スケールフリーランダムグラフの直径に関する考察 II

浜 口 幸 弘

1. はじめに

近年、複雑な現実世界のネットワーク (complex real-world networks) は自然科学のみならず各分野で注目されておき、とりわけスモールワールド (small-world) 現象は多くの研究が行われてきている。例えば最初の頃では、Albert, Jeong and Barabási (1999, 2000) による研究が有名である。この現象では、観測データやコンピュータシミュレーションによって、頂点数 n のグラフの直径は、ほぼ $\log n$ に近くなることが指摘されている。Barabási and Albert (1999) は最初にその現象のモデル化を行っているが、さらに Bollobás and Riordan (2004) は、それにランダムグラフの理論を適用してモデル化し、その直径を理論的に与えている。その定理によれば、グラフの直径はほぼ $\log n / \log \log n$ であり、観測データやシミュレーション結果より、かなり小さいものとなっている。

さて前稿 (浜口 (2017)) では、Bollobás and Riordan (2004) のモデルを検討していくつかの課題を提示した。今回は第 2 段階として、このランダムプロセスから生成されるグラフ G_1^n を対象にして、 G_1^n 上の単調パスの存在確率および G_1^n に含まれる最大パスの長さを近似評価する。結果として、 G_1^n (および一般的な木も含む) の特徴的な大きさを表すような定理 1 および定理 2 を導出した。

2. 概念とモデルの定義

Bollobás and Riordan (2004) はスモールワールドをランダムグラフでモデル化して、その直径の下限と上限を理論的に導出している。そのモデル化についての概略は、以下のとおりである。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。

さて、 $m=1$ の場合について、このことを厳密に定式化すると以下ようになる。グラフに追加していく頂点列を $v_1, v_2, \dots$ 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。そしてランダム

グラフのプロセス $(G_1^t)_{t \geq 0}$ 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 $v_i \in V(G_1^{t-1})$ を 1 本の辺で結ぶ。このとき頂点 i (確率変数である) は次のようにランダムに選択される。

$$P(i=s) = \begin{cases} d_{G_1^{t-1}(s)} / (2t-1) & 1 \leq s \leq t-1 \\ 1 / (2t-1) & s = t \end{cases} .$$

これは、Barabási and Albert (1999) のモデルを定式化したものとなっている。また、preferential attachment として、ネットワーク理論の基礎になる式である。

$m > 1$ の場合については、一度に v_t から m 本の辺が 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。また、以下では、頂点 v_i をその添え字の番号 i で表し、数直線上の閉区間 $[1, n]$ 内に位置すると考える。そして、このグラフにおける i_0 を始点とし、 i_k を終点とするパスを $i_0 i_1 \dots i_k$ と記す (パス $i_0 \rightarrow i_k$ とも記す)。このとき、始点からの頂点の並びが $i_0 > i_1 > \dots > i_k$ となるような、頂点番号が単調に減少するパスを、本稿では単調パスと記す。ここで、頂点選択の試行の順番とは逆の並びになることに注意する。

さて、 G_1^t は $[t] = \{1, 2, \dots, mn\}$ 上のランダムグラフとする。また、追加する頂点 j が接続されるグラフ内の頂点を g_j として表す。そして、ランダムグラフの頂点 i と j が辺で接続される事象を $A_{i,j}$ と記す。今回は頂点数 mn の木を対象にするので、 mn を n と置き換えて考える。すなわち、グラフ G_1^n を対象にする。

3. ランダムモデルによって生成される木の性質

Bollobás and Riordan (2004) は、確率 $P(g_j = i)$ の大きさを評価する補題を導出している。本稿では、その評価量をもう少し詳細に導くことにする。なお、補題の (2) に関しては、 $i < j < k$ という条件から、 $g_j = i$ か否かは G_1^{k-1} のもとで確定するので、浜口 (2017) の指摘は誤りである。よって、Bollobás and Riordan (2004) の証明方法を踏襲する。

補題 1

(1) $1 \leq i < j$ のとき、ある定数 $c_1 > 0$ および $c_2 > 0$ に対して、以下の式が成り立つ。

$$\frac{c_2}{\sqrt{ij}} \leq P(g_j = i) \leq \frac{c_1}{\sqrt{ij}} .$$

また、以下のようにも表せる。

$$P(g_j = i) = \frac{1}{2\sqrt{ij}} \left(1 + O\left(\frac{1}{i}\right) \right) .$$

(2) $1 \leq i < j < k$ のとき、ある定数 $c_1 > 0$ および $c_2 > 0$ に対して、以下の式が成り立つ。

$$\frac{c_2}{i\sqrt{jk}} \leq P(g_j = i, g_k = i) \leq \frac{c_1}{i\sqrt{jk}} .$$

また、以下のようにも表せる.

$$P(g_j=i, g_k=i) = \frac{1}{2i\sqrt{jk}} \left(1 + O\left(\frac{1}{i}\right)\right).$$

証明

(1) グラフの構成プロセス (G_1^t) を考える. G_1^t は確率変数である. また, グラフ G_1^t における頂点 i の次数を $d_{t,i} = d_{G_1^t}(i)$ と記す. プロセス (G_1^t) の定義から, 以下となる.

$$P(g_t=i | G_1^{t-1}) = \begin{cases} d_{t-1,i}/(2t-1) & t > i \\ 1/(2i-1) & t = i \end{cases}.$$

このとき, $t > i$ に対して, $d_{t,i} = d_{t-1,i} + \mathbf{I}_{\{g_t=i\}}$ となるから,

$$E(d_{t,i} | G_1^{t-1}) = d_{t-1,i} + d_{t-1,i}/(2t-1) = \left(1 + \frac{1}{2t-1}\right) d_{t-1,i}.$$

上式の両辺の期待値をとると,

$$E(d_{t,i}) = \left(1 + \frac{1}{2t-1}\right) E(d_{t-1,i}).$$

ここで, $E(d_{t,i})$ の代わりに, $\mu_{t,i}$ と記すものとする. さて, $d_{i,i}$ は 1 または 2 をとるので, $\mu_{i,i} = \frac{2i}{2i-1}$

となり, $t \geq i$ に対して,

$$\mu_{t,i} = \prod_{r=i}^t \left(1 + \frac{1}{2r-1}\right)$$

となる. そして, $\mu_{t,i}$ の大きさを近似評価するために, 以下の不等式を利用する.

$$0 \leq x \leq 1 \text{ に対して, } x - \frac{x^2}{2} \leq \log(1+x) \leq x.$$

$$0 \leq x \leq 1 \text{ に対して, } \log(1-x) \leq -x \text{ および}$$

$$0 \leq x \leq 1/2 \text{ に対して, } -x - x^2 \leq \log(1-x).$$

ここで, 関数 $y = \log\left(1 + \frac{1}{2x-1}\right)$ ($x \geq 1$) は単調減少するから, 以下の式が成り立つ.

$i \geq 2$ の場合,

$$\begin{aligned} \log \mu_{t,i} &= \sum_{r=i}^t \log\left(1 + \frac{1}{2r-1}\right) \leq \int_{i-1}^t \log\left(1 + \frac{1}{2x-1}\right) dx \\ &= \int_{i-1}^t \log 2x \, dx - \int_{i-1}^t \log(2x-1) \, dx \\ &= \frac{1}{2} \log \frac{t}{i} - (i-1) \log\left(1 - \frac{1}{i}\right) - \frac{2t-1}{2} \log\left(1 - \frac{1}{2t}\right) + \frac{2i-3}{2} \log\left(1 - \frac{3}{2i}\right) \end{aligned}$$

$$\leq \log\left(\left(\frac{t}{i}\right)^{1/2} e^{9/(4i) - 1/i^2 - 1/(8t^2)}\right).$$

$i=1$ の場合,

$$\begin{aligned} \log \mu_{t,1} &= \sum_{r=1}^t \log\left(1 + \frac{1}{2r-1}\right) \leq \log 2 + \int_1^t \log\left(1 + \frac{1}{2x-1}\right) dx \\ &= \frac{\log 2}{2} + \frac{\log t}{2} - \frac{2t-1}{2} \log\left(1 - \frac{1}{2t}\right) \\ &\leq \log(2^{1/2} t^{1/2} e^{1/2 - 1/(8t^2)}). \end{aligned}$$

一方, $i \geq 1$ に対して,

$$\begin{aligned} \log \mu_{t,i} &\geq \int_i^{t+1} \log\left(1 + \frac{1}{2x-1}\right) dx \\ &= \frac{1}{2} \log \frac{t}{i} + (t+1) \log\left(1 + \frac{1}{t}\right) - \frac{2t+1}{2} \log\left(1 + \frac{1}{2t}\right) + \frac{2i-1}{2} \log\left(1 - \frac{1}{2i}\right) \\ &\geq \log\left(\left(\frac{t}{i}\right)^{1/2} e^{1/(4t) - 1/(2t^2) + 1/(8i^2)}\right). \end{aligned}$$

まとめると, $i \geq 2$ の場合,

$$\left(\frac{t}{i}\right)^{1/2} e^{1/(4t) - 1/(2t^2) + 1/(8i^2)} \leq \mu_{t,i} \leq \left(\frac{t}{i}\right)^{1/2} e^{9/(4i) - 1/i^2 - 1/(8t^2)}.$$

また, $i=1$ の場合,

$$t^{1/2} e^{1/8 + 1/(4t) - 1/(2t^2)} \leq \mu_{t,1} \leq 2^{1/2} t^{1/2} e^{1/2 - 1/(8t^2)}.$$

したがって, 頂点選択における定義式の両辺の期待値をとり, t を $j-1$ として上述の結果を適用すると, ある定数 $c_1 > 0$ および $c_2 > 0$ に対して, 以下が成り立つ.

$$\frac{c_2}{\sqrt{ij}} \leq \mathbf{P}(g_j = i) = \frac{\mu_{j-1,i}}{2j-1} \leq \frac{c_1}{\sqrt{ij}}.$$

また, 以下のようにも表せる.

$$\mathbf{P}(g_j = i) = \frac{1}{2\sqrt{ij}} \left(1 + O\left(\frac{1}{i}\right)\right).$$

(2) $t > j > i$ とする. まず, $d_{t,i} = d_{t-1,i} + \mathbf{I}_{\{g_t = i\}}$ であるから, 以下を得る.

$$\mathbf{E}(d_{t,i}^2 | G_1^{t-1}) = d_{t-1,i}^2 \left(1 + \frac{2}{2t-1}\right) + \frac{d_{t-1,i}}{2t-1}.$$

そして両辺の期待値をとると, 次のようになる.

$$\mathbf{E}(d_{t,i}^2) = \left(1 + \frac{2}{2t-1}\right) \mathbf{E}(d_{t-1,i}^2) + \frac{\mathbf{E}(d_{t-1,i})}{2t-1}.$$

また, $\mathbf{E}(d_{t,i}) = \left(1 + \frac{1}{2t-1}\right) \mathbf{E}(d_{t-1,i})$ であるから, これを上式の辺々に加えて, $\mu_{t,i}^{(2)} = \mathbf{E}(d_{t,i}^2) + \mathbf{E}(d_{t,i})$

と記せば、以下を得る.

$$\mu_{t,i}^{(2)} = \left(1 + \frac{2}{2t-1}\right) \mu_{t-1,i}^{(2)}.$$

ここで、 $d_{i,i}$ は 1 または 2 をとるので、 $\mu_{i,i}^{(2)} = \frac{4i+2}{2i-1}$ となり、以下を得る.

$$\mu_{t,i}^{(2)} = \left(1 + \frac{2}{2t-1}\right) \mu_{t-1,i}^{(2)} = \prod_{r=i+1}^t \left(1 + \frac{2}{2r-1}\right) \mu_{i,i}^{(2)} = \frac{2t+1}{2i+1} \frac{4i+2}{2i-1}.$$

さて、頂点選択における定義式から、

$$\mathbb{E}\left(d_{j,i} \mathbf{I}_{\{g_j=i\}} \middle| G_1^{j-1}\right) = (d_{j-1,i} + 1) \frac{d_{j-1,i}}{2j-1}$$

となるので、両辺の期待値をとると、

$$\mathbb{E}\left(d_{j,i} \mathbf{I}_{\{g_j=i\}}\right) = \frac{\mu_{j-1,i}^{(2)}}{2j-1} = \frac{4i+2}{(2i-1)(2i+1)}.$$

ところで、

$$\mathbb{E}\left(d_{t,i} \mathbf{I}_{\{g_t=i\}} \middle| G_1^{t-1}\right) = \left(d_{t-1,i} + \frac{d_{t-1,i}}{2t-1}\right) \mathbf{I}_{\{g_t=i\}} \text{ となるので、}$$

両辺の期待値をとると、

$$\begin{aligned} \mathbb{E}\left(d_{t,i} \mathbf{I}_{\{g_t=i\}}\right) &= \left(1 + \frac{1}{2t-1}\right) \mathbb{E}\left(d_{t-1,i} \mathbf{I}_{\{g_{t-1}=i\}}\right) \\ &= \cdots = \prod_{r=j+1}^t \left(1 + \frac{1}{2r-1}\right) \mathbb{E}\left(d_{j,i} \mathbf{I}_{\{g_j=i\}}\right). \end{aligned}$$

$\prod_{r=j+1}^t \left(1 + \frac{1}{2r-1}\right)$ は、(1) と同様に評価できるので、 $i \geq 1$ に対して、以下の式を得る.

$$\begin{aligned} &\frac{4i+2}{(2i-1)(2i+1)} \left(\frac{t}{j}\right)^{1/2} e^{1/(4t) - 1/(2t^2) - 7/(8j) - 1/(16j^2)} \leq \mathbb{E}\left(d_{t,i} \mathbf{I}_{\{g_t=i\}}\right) \\ &\leq \frac{4i+2}{(2i-1)(2i+1)} \left(\frac{t}{j}\right)^{1/2} e^{1/(4j) - 1/(8t^2)}. \end{aligned}$$

一方、 $k > j > i$ に対して、

$$\mathbb{P}(g_j=i, g_k=i \mid G_1^{k-1}) = \mathbf{I}_{\{g_j=i\}} \mathbb{P}(g_k=i \mid G_1^{k-1}) = \mathbf{I}_{\{g_j=i\}} \frac{d_{k-1,i}}{2k-1} \text{ となる. このとき、両辺の期待値を}$$

とると、

$$\mathbb{P}(g_j=i, g_k=i) = \frac{1}{2k-1} \mathbb{E}\left(d_{k-1,i} \mathbf{I}_{\{g_j=i\}}\right).$$

そして上述の $\mathbb{E}\left(d_{t,i} \mathbf{I}_{\{g_t=i\}}\right)$ において、 t を $k-1$ とすれば、主張を得る. $\square$

次に、辺が生成される事象間の同時確率に関する性質を表した Bollobás and Riordan (2004) の補題

を示す. なお, 本稿ではその証明ではなく, 補題の意味を直観的に理解できる説明を与えることにする.
補題 2 (Bollobás and Riordan (2004))

事象 E と E' を次のように設定する.

$$E = \bigcap_{t=1}^r \{g_{j_t} = i_t\}, \quad E' = \bigcap_{t=1}^{r'} \{g_{j'_t} = i'_t\}.$$

このとき, すべての t に対して, $i_t < j_t$ および $i'_t < j'_t$ である. そして集合 $\{i_1, \dots, i_r\}$ と $\{i'_1, \dots, i'_{r'}\}$ が共通要素をもたないとき, 以下の式が成り立つ.

$$P(E \cap E') \leq P(E)P(E').$$

解説

事象 F を $F = \bigcup_{t=1}^r \{g_{j_t} \in \{i'_1, \dots, i'_{r'}\}\}$ とし, 頂点の集合 V_t を $V_t = V(G_1^{j_t-1})$ とする. 確率の基本法則から, 以下の式が成り立つ.

$$P(E') = P(F)P(E'|F) + (1 - P(F))P\left(E' \bigcap_{t=1}^r \{g_{j_t} \in V_t \setminus \{i'_1, \dots, i'_{r'}\}\}\right).$$

また, 以下の式が成り立つ.

$$P\left(E' \bigcap_{t=1}^r \{g_{j_t} \in V_t \setminus \{i'_1, \dots, i'_{r'}\}\}\right) = P(E'E).$$

これは, どちらの条件付確率も j_t は $\{i'_1, \dots, i'_{r'}\}$ 以外の頂点に接続するという条件は同じなので, j'_t ($1 \leq t \leq r'$) の接続する頂点を選択する際, 両式で i'_t の次数の確率分布は同じになるということによる. したがって, 以下の式を得る.

$$P(E') - P(E'E) = P(F) \left(P(E'|F) - P\left(E' \bigcap_{t=1}^r \{g_{j_t} \in V_t \setminus \{i'_1, \dots, i'_{r'}\}\}\right) \right).$$

このとき, $P(E'|F) \geq P\left(E' \bigcap_{t=1}^r \{g_{j_t} \in V_t \setminus \{i'_1, \dots, i'_{r'}\}\}\right)$ ということが予想されるので, 主張は成り立つと言える. $\square$

ここで, プロセス $(G_t^i)_{t \geq 0}$ において, 頂点 i が自分自身を選択しない, すなわち, ループを持たないように頂点選択のモデルを若干変更して, 単調パス $i_0 i_1 \dots i_k$ の存在する確率 $P(A_{i_0, i_1} \cap \dots \cap A_{i_{k-1}, i_k})$ を考える. このとき, 事象 $A_{i_0, i_1} \cap \dots \cap A_{i_{t-2}, i_{t-1}}$ と事象 A_{i_{t-1}, i_t} ($2 \leq t \leq k$) は独立となるので, $P(A_{i_0, i_1} \cap \dots \cap A_{i_{k-1}, i_k}) = P(A_{i_0, i_1}) \dots P(A_{i_{k-1}, i_k})$ が成り立つ. このモデルの場合については, 次回に議論する.

次に以下では, 各頂点を数直線上の閉区間 $[1, n]$ 上に位置付けて考える. そして $[1, n]$ 上の区間 I_k を $[a + \Delta(k-1) + 1, a + \Delta k]$ と定義する. ここで, k は $1 \leq k \leq (n-a)/\Delta$ を満たす整数である.

さて補題 1 と補題 2 を適用して, 単調パスに関する基本的な補題を以下に与える.

補題 3

(1) $a > 0$ とし, $\Delta^2 = o(a)$ かつ k は $2 \leq k \leq (n-a)/\Delta$ を満たす整数とする. このとき, 区間 I_k に含まれる任意の頂点は, I_{k-1} に含まれるどの頂点にも接続しない.

- (2) $a > 0$ とし, m は 2 以上の整数で, かつ $\Delta = o(a^{(m-1)/m})$ のとき, 長さ m 以上の単調パスは, $1 \leq k \leq (n-a)/\Delta$ を満たすほとんどすべての I_k 内には存在しない.

証明

- (1) 区間 I_k に含まれる頂点 j が, I_{k-1} に含まれる頂点 i に接続する事象を $A_{i,j}^{(k)}$ とする. このとき, 区間 I_k に含まれる頂点が, I_{k-1} に含まれる頂点に接続する確率は, 補題 1 を用いて以下のように評価できる. ある正の定数 c に対して,

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i \in I_{k-1}} \bigcup_{j \in I_k} A_{i,j}^{(k)}\right) &\leq \sum_{i \in I_{k-1}} \sum_{j \in I_k} \mathbb{P}\left(A_{i,j}^{(k)}\right) \leq c \sum_{i \in I_{k-1}} \sum_{j \in I_k} \frac{1}{\sqrt{ij}} \\ &\leq c \int_{a+\Delta(k-1)}^{a+\Delta k} \frac{1}{\sqrt{x}} dx \int_{a+\Delta(k-2)}^{a+\Delta(k-1)} \frac{1}{\sqrt{x}} dx \leq \frac{c\Delta^2}{a+\Delta(k-2)} = o(1). \end{aligned}$$

- (2) 確率変数 $X_{i_0}^{(k)}$ を i_0 を始点とする I_k 内の長さ m の単調パスの個数とする (なお, 単調パスはコンポーネントとは限らない). また, $X^{(k)}$ を I_k 内の長さ m の単調パスの個数とし, $X_{i_0 i_1 \dots i_m}^{(k)}$ を I_k 内の頂点 $i_0 > i_1 > \dots > i_m$ に対して, 単調パス $i_0 i_1 \dots i_m$ が存在すれば 1, そうでなければ 0 をとる確率変数とする. さらに単調パス $i_0 i_1 \dots i_m$ が I_k に含まれることを $i_0 i_1 \dots i_m \subset I_k$ と表し, $A_{x,y}^{(k)}$ を I_k 内の頂点 x と y が接続する事象とすれば, 補題 1 と 2 を用いて以下を得る.

$X_{i_0}^{(k)} = \sum_{i_0 i_1 \dots i_m \subset I_k} X_{i_0 i_1 \dots i_m}^{(k)}$ であり, ある正の定数 c に対して,

$$\begin{aligned} \mathbb{E}\left(X^{(k)}\right) &= \sum_{i_0 \in I_k} \mathbb{E}\left(X_{i_0}^{(k)}\right) = \sum_{i_0 \in I_k} \sum_{i_0 i_1 \dots i_m \subset I_k} \mathbb{E}\left(X_{i_0 i_1 \dots i_m}^{(k)}\right) \\ &= \sum_{i_0 \in I_k} \sum_{i_0 i_1 \dots i_m \subset I_k} \mathbb{P}\left(A_{i_0, i_1}^{(k)} \cap \dots \cap A_{i_{m-1}, i_m}^{(k)}\right) \\ &\leq \sum_{i_0 \in I_k} \sum_{i_0 i_1 \dots i_m \subset I_k} \mathbb{P}\left(A_{i_0, i_1}^{(k)}\right) \dots \mathbb{P}\left(A_{i_{m-1}, i_m}^{(k)}\right) \leq \Delta \left(\frac{c\Delta}{a+\Delta(k-1)}\right)^m. \end{aligned}$$

ここで, $X = \sum_{k=1}^K X^{(k)}$ とする. ただし, $K = (n-a)/\Delta$ である. このとき, 以下を得る.

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^K \mathbb{E}(X^{(k)}) = \sum_{k=1}^K \Delta \left(\frac{c\Delta}{a+\Delta(k-1)}\right)^m \\ &\leq c^m \Delta \int_0^K \frac{1}{\left(x-1+\frac{a}{\Delta}\right)^m} dx \\ &= \frac{c^m \Delta}{m-1} \left(\frac{1}{\left(\frac{a}{\Delta}-1\right)^{m-1}} - \frac{1}{\left(\frac{n}{\Delta}-1\right)^{m-1}}\right) = o(1). \end{aligned}$$

よって, Markov の不等式から主張を得る. $\square$

例えば, $a = \log n$ とすれば, 補題 3 の (1) の条件は, $\Delta^2 = o(\log n)$, (2) の条件は, $\Delta = o((\log n)^{(m-1)/m})$ となる.

次に、本稿の目的である定理 1 および定理 2 を導くために、頂点 i_{k+1} のみ区間 $I_\sigma = [1, \sigma]$ に入るような、 i_0 を始点とする長さ $k+1$ の単調パス $i_0 \cdots i_{k+1}$ が存在する確率の上限を評価する不等式を与える。なお、これに関連して、単調パスの存在する確率の上限および下限を評価する別の補題を次回でも議論する。

補題 4

c_0 および c_1 を正の定数とし、 $\sigma \rightarrow \infty$ とする。頂点 i_{k+1} のみ区間 $I_\sigma = [1, \sigma]$ に入るような、 i_0 を始点とする長さ $k+1$ の単調パス $i_0 \cdots i_{k+1}$ が存在する確率 $P_{i_0 \rightarrow I_\sigma}^{k+1}$ は以下の不等式を満たす。

$$P_{i_0 \rightarrow I_\sigma}^{k+1} \leq \frac{2c_1\sqrt{\sigma}}{\sqrt{i_0}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma} \right) \right)^k \sum_{r=0}^k a_r \frac{(\log i_0)^{k-r}}{(k-r)!}.$$

このとき、 $a_0 = 1$, $a_1 = 0$, $a_r = \left(\frac{e+1}{e} \right)^{r-1}$ ($2 \leq r \leq k$) である。

証明

G_1^n において、頂点 x と y が接続する事象を $A_{x,y}$ とする。そして、頂点 i_{k+1} のみ区間 $I_\sigma = [1, \sigma]$ に入るような、始点を i_0 、終点を i_{k+1} とする長さ $k+1$ の単調パス $i_0 \cdots i_{k+1}$ が存在する確率を $P_{i_0 \rightarrow I_\sigma}^{k+1}$ とする。そして補題 1 と補題 2 を用いて以下を得る。

ある正の定数 c_0 および c_1 に対して、 $k=0$ の場合、主張は明らかに成り立つ。 $k \geq 1$ の場合は、

$$\begin{aligned} P_{i_0 \rightarrow I_\sigma}^{k+1} &= P \left(\bigcup_{n \geq i_0 > \cdots > i_k > \sigma \geq i_{k+1} \geq 1} (A_{i_0, i_1} \cap A_{i_1, i_2} \cap \cdots \cap A_{i_k, i_{k+1}}) \right) \\ &= \sum_{n \geq i_0 > \cdots > i_k > \sigma \geq i_{k+1} \geq 1} P (A_{i_0, i_1} \cap A_{i_1, i_2} \cap \cdots \cap A_{i_k, i_{k+1}}) \\ &\leq \sum_{i_1 = \sigma+k}^{i_0-1} \sum_{i_2 = \sigma+k-1}^{i_1-1} \cdots \sum_{i_k = \sigma+1}^{i_{k-1}-1} \sum_{i_{k+1}=1}^{\sigma} P (A_{i_0, i_1}) P (A_{i_1, i_2}) \cdots P (A_{i_k, i_{k+1}}) \\ &\leq c_1 \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma} \right) \right)^k \sum_{i_1 = \sigma+k}^{i_0-1} \sum_{i_2 = \sigma+k-1}^{i_1-1} \cdots \sum_{i_k = \sigma+1}^{i_{k-1}-1} \sum_{i_{k+1}=1}^{\sigma} \frac{1}{\sqrt{i_0 i_1}} \frac{1}{\sqrt{i_1 i_2}} \cdots \frac{1}{\sqrt{i_k i_{k+1}}} \\ &= \frac{c_1}{\sqrt{i_0}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma} \right) \right)^k \sum_{i_1 = \sigma+k}^{i_0-1} \sum_{i_2 = \sigma+k-1}^{i_1-1} \cdots \sum_{i_k = \sigma+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k} \sum_{i_{k+1}=1}^{\sigma} \frac{1}{\sqrt{i_{k+1}}} \\ &\leq \frac{c_1}{\sqrt{i_0}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma} \right) \right)^k \sum_{i_1 = \sigma+k}^{i_0-1} \sum_{i_2 = \sigma+k-1}^{i_1-1} \cdots \sum_{i_k = \sigma+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k} \left(1 + \int_1^{\sigma} \frac{1}{\sqrt{x}} dx \right) \end{aligned}$$

$$= \frac{2c_1(\sqrt{\sigma}-1)}{\sqrt{i_0}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma} \right) \right)^k \sum_{i_1=\sigma+k}^{i_0-1} \sum_{i_2=\sigma+k-1}^{i_1-1} \cdots \sum_{i_k=\sigma+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k}$$

このとき、以下の不等式が成り立つことを j に関する帰納法で示す。

$1 \leq j \leq k$ に対して、

$$\sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \cdots \sum_{i_k=\sigma+1}^{i_{k-1}^{-1}} \frac{1}{i_{k-(j-1)}} \cdots \frac{1}{i_k} \leq \sum_{i=0}^j a_i \frac{(\log i_{k-j})^{j-i}}{(j-i)!}. \quad \text{このとき, } a_0 = 1, \quad a_1 = 0,$$

$$a_i = \left(\frac{e+1}{e} \right)^{i-1} \quad (2 \leq i \leq k) \quad \text{である.}$$

$j=1$ の場合、以下となる。

$$\sum_{i_k=\sigma+1}^{i_{k-1}^{-1}} \frac{1}{i_k} \leq \int_1^{i_{k-1}} \frac{1}{x} dx = \log i_{k-1} \leq a_0 \log i_{k-1}.$$

$j=2$ の場合、以下となる。

$$\sum_{i_{k-1}=\sigma+2}^{i_{k-2}^{-1}} \sum_{i_k=\sigma+1}^{i_{k-1}^{-1}} \frac{1}{i_{k-1}} \frac{1}{i_k} \leq \sum_{i_{k-1}=\sigma+2}^{i_{k-2}^{-1}} \frac{\log i_{k-1}}{i_{k-1}} \leq \frac{1}{2} (i_{k-2})^2 \leq a_0 \frac{(i_{k-2})^2}{2!} + a_2.$$

よって、 $j=1$ および $j=2$ の場合、主張は成り立つ。次に、 $j-1$ ($j \geq 3$) の場合、主張が成り立つと

仮定する。このとき予め、 $\sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \frac{(\log i_{k-(j-1)})^m}{i_{k-(j-1)}}$ の大きさを評価しておく（ただし、 m は負でない

整数）。そして、 $f(x) = \frac{(\log x)^m}{x}$ ($x \geq 1$) とする。 $f'(x)$ は、 $f'(x) = \frac{(\log x)^{m-1}(m - \log x)}{x^2}$ となる

から、 $1 \leq x \leq e^m$ のとき、 $f'(x) \geq 0$ であり、 $x \geq e^m$ のとき、 $f'(x) \leq 0$ である。そして、 $x = e^m$ のとき $f(x)$ は最大値をとる。このとき、 $t \leq e^m < t+1$ を満たす整数 t をとり、 t が区間 $[\sigma+j, i_{k-j}-1]$ に含まれる場合を考えると、 $f(t) \geq f(t+1)$ の場合、以下ようになる（ $f(t) < f(t+1)$ の場合も同様に評価できる）。

$$\begin{aligned} \sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \frac{(\log i_{k-(j-1)})^m}{i_{k-(j-1)}} &\leq \int_1^t \frac{(\log x)^m}{x} dx + \left(\frac{m}{e} \right)^m + \int_{t+1}^{i_{k-j}} \frac{(\log(x-1))^m}{x-1} dx \\ &\leq \frac{1}{m+1} (\log i_{k-j})^{m+1} + \left(\frac{m}{e} \right)^m. \end{aligned}$$

ただし、 $m=0$ のとき、 $\left(\frac{m}{e} \right)^m = 1$ とする。よって、帰納法の仮定および上の結果を用いて以下を得る。

$$\begin{aligned}
& \sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \cdots \sum_{i_k=\sigma+1}^{i_{k-1}^{-1}} \frac{1}{i_{k-(j-1)}} \cdots \frac{1}{i_k} = \sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \frac{1}{i_{k-(j-1)}} \sum_{i_{k-(j-2)}=\sigma+j-1}^{i_{k-(j-1)}^{-1}} \cdots \sum_{i_k=\sigma+1}^{i_{k-1}^{-1}} \frac{1}{i_{k-(j-2)}} \cdots \frac{1}{i_k} \\
& \leq \sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \frac{1}{i_{k-(j-1)}} \sum_{i=0}^{j-1} a_i \frac{(\log i_{k-(j-1)})^{j-i-1}}{(j-i-1)!} \\
& \leq \sum_{i=0}^{j-1} \frac{a_i}{(j-i-1)!} \sum_{i_{k-(j-1)}=\sigma+j}^{i_{k-j}^{-1}} \frac{(\log i_{k-(j-1)})^{j-i-1}}{i_{k-(j-1)}} \\
& \leq \sum_{i=0}^{j-1} \frac{a_i}{(j-i-1)!} \left(\frac{1}{j-i} (\log i_{k-j})^{j-i} + \left(\frac{j-i-1}{e} \right)^{j-i-1} \right) \\
& = \sum_{i=0}^{j-1} \frac{a_i}{(j-i)!} (\log i_{k-j})^{j-i} + \sum_{i=0}^{j-1} \frac{a_i}{(j-i-1)!} \left(\frac{j-i-1}{e} \right)^{j-i-1} \\
& \leq \sum_{i=0}^j a_i \frac{(\log i_{k-j})^{j-i}}{(j-i)!}.
\end{aligned}$$

ここで最後の不等号は、 $n! \geq e \left(\frac{n}{e} \right)^n$ を用いて、以下となることによる。

$$\begin{aligned}
\sum_{i=0}^{j-1} \frac{a_i}{(j-i-1)!} \left(\frac{j-i-1}{e} \right)^{j-i-1} & \leq \frac{1}{e} \sum_{i=0}^{j-1} a_i \leq \frac{1}{e} + \frac{1}{e} \sum_{i=2}^{j-1} \left(\frac{e+1}{e} \right)^{i-1} \\
& = \frac{1}{e} \sum_{i=1}^{j-1} \left(\frac{e+1}{e} \right)^{i-1} = \left(\frac{e+1}{e} \right)^{j-1} - 1 \leq a_j.
\end{aligned}$$

t がこの区間に含まれない場合についても、 $f(x)$ はこの区間で単調に変化するので、前述の不等式は成り立つ。したがって、 j の場合も主張は成り立ち、この補題は示された。□

さて、ランダムに生成される木について、その最大の深さはおよそ $\log n$ として、すでに Pittel (1994) によって導出されている。しかし、ブランチングプロセス (branching processes) による評価方法で、パスの存在確率を直接評価する方法はとっていない。これに対して本稿では、補題 4 を用いて単調パスの存在する確率を直接評価することにより、以下の定理を導出する。ここでは、前述のようにループをもつ G_1^n (すなわち 1 つの端点がループを持つ木の集まりとも言える) を対象にするが、頂点選択の確率を若干変更すれば、以下の定理 1 と 2 は一般的な木にも対応できる。この場合、Pittel (1994) の定理にある結果とは若干係数の大きさが違うものとなるが、これは検討課題となる。

定理 1

定数 ε を $0 < \varepsilon < 1$ とする。十分大きな n に対して、 G_1^n における最長の単調パスの長さは $\frac{(1-\varepsilon)\log n}{2}$

以上かつ $\frac{(1+\varepsilon)\log n}{2}$ 以下である.

証明

G_1^n において, i_0 を始点とする長さ $k+1$ の単調パス $i_0 i_1 \cdots i_{k+1}$ を考える ($k \geq 0$). そして $i_0 = n$ とし, 十分小さな定数 $0 < \alpha < 1$ に対して, $\sigma = \alpha \log n$ とする. また, 区間 $I_\sigma = [1, \sigma]$ とする. このとき, 補題 4 から以下を得る. 正の定数 c_0 および c_1 に対して,

$$P_{i_0 \rightarrow I_\sigma}^{k+1} \leq \frac{2c_1 \sqrt{\sigma}}{\sqrt{n}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\alpha \log n} \right) \right)^k \sum_{r=0}^k a_r \frac{(\log n)^{k-r}}{(k-r)!}. \quad \text{このとき, } a_r \text{ は補題 4 の定義に従う.}$$

ここで, $F_k(r)$ ($0 \leq r \leq k$) を以下のように定める.

$$F_k(0) = \left(\frac{1}{2} \left(1 + \frac{c_0}{\alpha \log n} \right) \right)^k \frac{(\log n)^k}{k!} \quad \text{および, } F_k(1) = 0 \text{ とし, } 2 \leq r \leq k \text{ に対して,}$$

$$F_k(r) = \left(\frac{1}{2} \left(1 + \frac{c_0}{\alpha \log n} \right) \right)^k \left(\frac{e+1}{e} \right)^{r-1} \frac{(\log n)^{k-r}}{(k-r)!} \text{ とする.}$$

次に, 以下のように各事象を設定する. $k \leq \frac{(1-\varepsilon)\log n}{2} - 2$ であり, 頂点 i_{k+1} (終点とは限らない)

のみ区間 I_σ に入るような単調パス $i_0 i_1 \cdots i_{k+1}$ が存在する事象を $E_{1,k}$ とし, $k \geq \frac{(1+\varepsilon)\log n}{2}$ であり, 頂

点 i_{k+1} (終点とは限らない) のみ区間 I_σ に入るような単調パス $i_0 i_1 \cdots i_{k+1}$ が存在する事象を $E_{2,k}$ とし,

$k \leq \frac{(1-\varepsilon)\log n}{2} - 2$ であり, 終点 i_{k+1} が区間 I_σ に入らずループ辺を持つような単調パス $i_0 i_1 \cdots i_{k+1}$ が

存在する事象を $E_{3,k}$ とし, また, $k \geq \frac{(1+\varepsilon)\log n}{2}$ であり, 終点 i_{k+1} が区間 I_σ に入らずループ辺を持

つような単調パス $i_0 i_1 \cdots i_{k+1}$ が存在する事象を $E_{4,k}$ とする. さらに, $\frac{(1-\varepsilon)\log n}{2} - 1 \leq k$

$\leq \frac{(1+\varepsilon)\log n}{2} - 1$ であり, 終点 i_{k+1} が区間 I_σ に入らずループ辺を持つような単調パス $i_0 i_1 \cdots i_{k+1}$ が

存在する事象を $E_{5,k}$ とする (注: この証明では, ループ辺は長さ 0 と見なす). そして, $1 \leq i \leq 5$ に

対して, $E_i = \bigcup_k E_{i,k}$ とする. ただし, E_3 は i_0 でループ辺を持つ事象も含むとする.

以下では, $P(E_1 \cup E_2 \cup E_3 \cup E_4) \rightarrow 0$ を示す. 各事象は互いに排反であるから, $P(E_1 \cup E_2 \cup E_3 \cup E_4) = P(E_1) + P(E_2) + P(E_3) + P(E_4)$ である. そして各確率の評価をそれぞれ行うが, そのために, まず $F_k(r)$ ($0 \leq r \leq k$) の大きさを以下のように評価比較する.

$2 \leq r \leq k-1$ に対して,

$$\frac{F_k(r+1)}{F_k(r)} = \frac{(e+1)(k-r)}{e} \frac{1}{\log n}.$$

よって、 $\frac{F_k(r+1)}{F_k(r)} \leq 1$ を満たす最小の整数 r は、 $k - \left\lfloor \frac{e}{e+1} \log n \right\rfloor$ である。

$k \geq \frac{e}{e+1} \log n + 2$ の場合、 $r = k - \left\lfloor \frac{e}{e+1} \log n \right\rfloor$ で $F_k(r)$ ($2 \leq r \leq k$) は最大となる。すなわち、

以下を得る。

$$F_k(1) \leq \cdots \leq F_k\left(k - \left\lfloor \frac{e}{e+1} \log n \right\rfloor - 1\right) \leq F_k\left(k - \left\lfloor \frac{e}{e+1} \log n \right\rfloor\right) \geq \cdots \geq F_k(k).$$

$m = \left\lfloor \frac{e}{e+1} \log n \right\rfloor$ とすると、 $\frac{e}{e+1} \log n - 1 < m \leq \frac{e}{e+1} \log n$ 。このとき、

$$\begin{aligned} \max_{r \geq 2} F_k(r) &= F_k(k-m) \\ &= \left(\frac{1}{2} \left(1 + \frac{c_0}{a \log n}\right)\right)^k \left(\frac{e+1}{e}\right)^{k-m-1} \frac{(\log n)^m}{(m)!} \\ &\leq \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right)\right)^k \left(\frac{e+1}{e}\right)^{-m-1} \frac{(\log n)^m}{e \left(\frac{m}{e}\right)^m} \\ &\leq \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right)\right)^k \left(\frac{e}{e+1}\right)^{m+1} \frac{\left(\frac{e \log n}{m}\right)^m}{e} \\ &\leq \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right)\right)^k \frac{1}{e+1} \left(\frac{e^2 \log n}{(e+1)m}\right)^m \\ &\leq \frac{e^2}{e+1} \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right)\right)^k n^{e/(e+1)} \end{aligned}$$

これは、 $e^{-2x} \leq 1-x$ ($0 \leq x \leq \frac{1}{4}$) を用いて、

$$\begin{aligned} \left(\frac{e^2 \log n}{(e+1)m}\right)^m &\leq \left(\frac{e^2 \log n}{(e+1)\left(\frac{e}{e+1} \log n - 1\right)}\right)^{(e/(e+1)) \log n} \\ &= \left(\frac{e}{1 - \frac{e+1}{e \log n}}\right)^{(e/(e+1)) \log n} = n^{e/(e+1)} \frac{1}{\left(1 - \frac{e+1}{e \log n}\right)^{(e/(e+1)) \log n}} \leq e^2 n^{e/(e+1)} \end{aligned}$$

となることによる。このとき、 $\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right) < 1$ となるような十分大きな n で考える。よって、

$\max_{r \geq 2} F_k(r)$ は $k = m + 2$ または、 $k = m + 3$ のとき最大となり、以下を得る。

$$\max_k \left(\max_{r \geq 2} F_k(r)\right) \leq \frac{e^2}{e+1} \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{a \log n}\right)\right)^{m+2} n^{e/(e+1)}$$

$$\begin{aligned}
 &\leq \frac{e^2}{e+1} \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n} \right) \right)^{(e/(e+1)) \log n + 1} n^{e/(e+1)} \\
 &\leq e \left(\frac{e+1}{2e} e^{c_0/(\alpha \log n)} \right)^{(e/(e+1)) \log n} n^{e/(e+1)} \\
 &\leq c'n^{e/(e+1)} n^{(e/(e+1)) \log((e+1)/(2e))} \leq c'n^{0.454}.
 \end{aligned}$$

ただし, $c' = e^{c_0 e/(\alpha(e+1)+1)}$ である.

$$\text{また, } \frac{\max_{r \geq 2} F_{k+1}(r)}{\max_{r \geq 2} F_k(r)} = \frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n} \right) < 1 \text{ となる.}$$

一方, $k \leq \frac{e}{e+1} \log n + 1$ の場合, $r \geq k - \frac{e}{e+1} \log n$ は常に成り立つので, $\frac{F_k(r+1)}{F_k(r)} \leq 1$ を満たす最小の r は 2 である. すなわち以下となる.

$$F_k(2) \geq F_k(3) \geq \dots \geq F_k(k).$$

また, $F_k(0)$ の大きさを以下のように評価比較する.

$$\frac{F_k(2)}{F_k(0)} = \frac{e+1}{e} \frac{k(k-1)}{(\log n)^2} = \frac{e+1}{e} \left(\frac{k}{\log n} \right)^2 \left(1 - \frac{1}{k} \right). \text{ および,}$$

$$\frac{F_{k+1}(0)}{F_k(0)} = \left(1 + \frac{c_0}{\alpha \log n} \right) \frac{\log n}{2(k+1)} \text{ すなわち } \left(1 + \frac{c_0}{\alpha \log n} \right) \frac{\log n}{2} - 1 \geq k \text{ ならば, } F_k(0) \leq F_{k+1}(0).$$

さて, $k \leq \frac{(1-\varepsilon) \log n}{2} - 2$ の場合, 上述のことから, $F_k(r) \leq F_k(0) \leq F_{k+1}(0)$ ($1 \leq r \leq k$) となるので, $P(E_1)$ の大きさを評価すると補題 1, 2 および 4 から以下を得る.

$$\begin{aligned}
 P(E_1) &= P \left(\bigcup_{k=0}^{((1-\varepsilon) \log n)/2-2} E_{1,k} \right) = \sum_{k=0}^{((1-\varepsilon) \log n)/2-2} P(E_{1,k}) \\
 &= \sum_{k=0}^{((1-\varepsilon) \log n)/2-2} P_{i_0 \rightarrow I_\sigma}^{k+1} \leq \frac{2c_1 \sqrt{\sigma}}{\sqrt{n}} \sum_{k=0}^{((1-\varepsilon) \log n)/2-2} \sum_{r=0}^k F_k(r) \\
 &\leq \frac{2c_1 \sqrt{\sigma}}{\sqrt{n}} \sum_{k=0}^{((1-\varepsilon) \log n)/2-2} k F_{((1-\varepsilon) \log n)/2}(0) \leq \frac{2c_1 \sqrt{\sigma}}{\sqrt{n}} \left(\frac{(1-\varepsilon) \log n}{2} \right)^2 F_{((1-\varepsilon) \log n)/2}(0) \\
 &= \frac{2c_1 \sqrt{\sigma}}{\sqrt{n}} \left(\frac{(1-\varepsilon) \log n}{2} \right)^2 \left(1 + \frac{c_0}{\alpha \log n} \right)^{((1-\varepsilon) \log n)/2} \frac{1}{2^{((1-\varepsilon) \log n)/2}} \frac{(\log n)^{((1-\varepsilon) \log n)/2}}{\left(\frac{(1-\varepsilon) \log n}{2} \right)!}
 \end{aligned}$$

$$\begin{aligned}
 &\leq \frac{c_1(1-\varepsilon)^2}{2} e^{((1-\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{\sqrt{n}} \frac{1}{2 \log n^{(1-\varepsilon)/2}} \frac{1}{e\left(\frac{1-\varepsilon}{2e} \log n\right)^{\log n^{(1-\varepsilon)/2}}} (\log n)^{\log n^{(1-\varepsilon)/2}} \\
 &\leq \frac{c_1(1-\varepsilon)^2}{2} e^{((1-\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{\sqrt{n}} \left(\frac{e}{1-\varepsilon}\right)^{\log n^{(1-\varepsilon)/2}} \\
 &= \frac{c_1(1-\varepsilon)^2}{2} e^{((1-\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{\sqrt{n}} n^{((1-\varepsilon)/2)\log(e/(1-\varepsilon))} \\
 &\leq \frac{c_1(1-\varepsilon)^2}{2} e^{((1-\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{n^{\varepsilon^2/4}} = O\left(\frac{\sqrt{\sigma}(\log n)^2}{n^{\varepsilon^2/4}}\right).
 \end{aligned}$$

次に, $P(E_2)$ の大きさを評価する.

$$P(E_2) = P\left(\bigcup_{k=(1+\varepsilon)\log n/2}^{\infty} E_{2,k}\right) = \sum_{k=(1+\varepsilon)\log n/2}^{\infty} P(E_{2,k}) = \sum_{k=(1+\varepsilon)\log n/2}^{\infty} P_{i_0 \rightarrow I_\sigma}^{k+1}.$$

このとき, 以下のように場合に分けて考える. まず $\frac{(1+\varepsilon)\log n}{2} \leq k \leq \log n - 1$ の場合, 前述のこと

から, $F_{((1+\varepsilon)\log n)/2}(0) \geq F_k(r)$ ($0 \leq r \leq k$) となるので, 以下を得る.

$$\begin{aligned}
 \sum_{k=(1+\varepsilon)\log n/2}^{\log n-1} P_{i_0 \rightarrow I_\sigma}^{k+1} &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \sum_{k=(1+\varepsilon)\log n/2}^{\log n-1} \sum_{r=0}^k F_k(r) \\
 &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \sum_{k=(1+\varepsilon)\log n/2}^{\log n-1} (k+1) F_{((1+\varepsilon)\log n)/2}(0) \\
 &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \frac{(\log n)^2}{2} F_{((1+\varepsilon)\log n)/2}(0) \\
 &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \frac{(\log n)^2}{2} \left(1 + \frac{c_0}{\alpha \log n}\right)^{((1+\varepsilon)\log n)/2} \frac{1}{2^{((1+\varepsilon)\log n)/2}} \frac{(\log n)^{((1+\varepsilon)\log n)/2}}{\left(\frac{(1+\varepsilon)\log n}{2}\right)!} \\
 &\leq c_1 e^{((1+\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{\sqrt{n}} \left(\frac{e}{1+\varepsilon}\right)^{\log n^{(1+\varepsilon)/2}} \\
 &= c_1 e^{((1+\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{\sqrt{n}} n^{((1+\varepsilon)/2)\log(e/(1+\varepsilon))} \\
 &\leq c_1 e^{((1+\varepsilon)c_0)/(2\alpha)} \frac{\sqrt{\sigma}(\log n)^2}{n^{\varepsilon^2/8}}.
 \end{aligned}$$

$k \geq \log n$ の場合, 前述のことから以下の結果を得ている.

$$\max_{r \geq 0} F_k(r) = F_k\left(k - \frac{e}{e+1} \log n\right) \leq \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right)\right)^k n^{e/(e+1)}.$$

よって, $0.68 < \frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right) < 0.7$ を満たすような十分大きな n に対して, 以下を得る.

$$\begin{aligned} \sum_{k=\log n}^{\infty} \mathbb{P}_{i_0 \rightarrow I_\sigma}^{k+1} &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \sum_{k=\log n}^{\infty} \sum_{r=0}^k F_k(r) \\ &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} \sum_{k=\log n}^{\infty} k \max_{r \geq 0} F_k(r) \\ &\leq \frac{2c_1\sqrt{\sigma}}{\sqrt{n}} n^{e/(e+1)} \sum_{k=\log n}^{\infty} k \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right)\right)^k \\ &\leq \frac{200}{27} c_1 e^{c_0/\alpha} n^{e/(e+1)} n^{\log((e+1)/(2e))} (\log n) \sqrt{\frac{\sigma}{n}} \\ &= O\left(\frac{\sqrt{\sigma} \log n}{n^{0.14}}\right) \end{aligned}$$

最後の不等号の評価は, $0.68 < \frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right) < 0.7$ を用いて次の式が得られることによる.

$$\begin{aligned} \sum_{k=\log n}^{\infty} k \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right)\right)^k &\leq \frac{100}{27} \left(\frac{e+1}{2e} \left(1 + \frac{c_0}{\alpha \log n}\right)\right)^{\log n} \log n \\ &\leq \frac{100}{27} e^{c_0/\alpha} \left(\frac{e+1}{2e}\right)^{\log n} \log n. \end{aligned}$$

したがって, 以下を得る.

$$\mathbb{P}(E_2) = O\left(\frac{\sqrt{\sigma} (\log n)^2}{n^{\varepsilon^2/8}}\right).$$

次に, $\mathbb{P}(E_3)$ の大きさを評価する. i_{k+1} は区間 $J_\sigma = (\sigma, n]$ に含まれ, i_0 がループ辺を持つ場合も考慮すると, 補題 1 と 2 から以下となる.

$$\begin{aligned} \mathbb{P}(E_3) &= \mathbb{P}\left(\bigcup_{k=0}^{((1-\varepsilon)\log n)/2-2} E_{3,k}\right) + \frac{1}{2n-1} \\ &= \sum_{k=0}^{((1-\varepsilon)\log n)/2-2} \mathbb{P}(E_{3,k}) + \frac{1}{2n-1} \\ &= \sum_{k=0}^{((1-\varepsilon)\log n)/2-2} \mathbb{P}_{i_0 \rightarrow J_\sigma}^{k+1} + \frac{1}{2n-1}. \end{aligned}$$

ここで, $P_{i_0 \rightarrow J_\sigma}^{k+1}$ は以下のようになる.

$$\begin{aligned} P_{i_0 \rightarrow J_\sigma}^{k+1} &= P\left(n \geq i_0 > \dots > i_k > i_{k+1} > \sigma \left(A_{i_0, i_1} \cap A_{i_1, i_2} \cap \dots \cap A_{i_k, i_{k+1}} \cap A_{i_{k+1}, i_{k+1}} \right)\right) \\ &\leq \sum_{i_1=\sigma+k+1}^{i_0-1} \sum_{i_2=\sigma+k}^{i_1-1} \dots \sum_{i_k=\sigma+2}^{i_{k-1}-1} \sum_{i_{k+1}=\sigma+1}^{i_k-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \dots P(A_{i_k, i_{k+1}}) P(A_{i_{k+1}, i_{k+1}}) \\ &\leq \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma}\right)\right)^{k+1} \sum_{i_1=\sigma+k+1}^{i_0-1} \sum_{i_2=\sigma+k}^{i_1-1} \dots \sum_{i_k=\sigma+2}^{i_{k-1}-1} \sum_{i_{k+1}=\sigma+1}^{i_k-1} \frac{1}{\sqrt{i_0 i_1}} \frac{1}{\sqrt{i_1 i_2}} \dots \frac{1}{\sqrt{i_k i_{k+1}}} \frac{1}{2i_{k+1}-1} \\ &\leq \frac{1}{\sqrt{i_0}} \left(\frac{1}{2} \left(1 + \frac{c_0}{\sigma}\right)\right)^{k+1} \sum_{i_1=\sigma+k+1}^{i_0-1} \sum_{i_2=\sigma+k}^{i_1-1} \dots \sum_{i_k=\sigma+2}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \dots \frac{1}{i_k} \sum_{i_{k+1}=\sigma+1}^n \frac{1}{i_{k+1}^{3/2}}. \end{aligned}$$

上の式において, $P(A_{i_{k+1}, i_{k+1}}) = \frac{1}{2i_{k+1}-1}$ に注意する. したがって, $P(E_1)$ の場合と同様にして

以下を得る.

$$P(E_3) = O\left(\frac{(\log n)^2}{\sqrt{\sigma} n^{\varepsilon^2/4}}\right).$$

$P(E_4)$ の大きさを評価すると, $P(E_2)$ および $P(E_3)$ の場合と同様に考えて,

$$P(E_4) = P\left(\bigcup_{k=\lceil(1+\varepsilon)\log n\rceil/2}^{\infty} E_{4,k}\right) = \sum_{k=\lceil(1+\varepsilon)\log n\rceil/2}^{\infty} P(E_{4,k}) = O\left(\frac{(\log n)^2}{\sqrt{\sigma} n^{\varepsilon^2/8}}\right).$$

以上のことから, $P(E_1 \cup E_2 \cup E_3 \cup E_4) = O\left(\frac{\sqrt{\sigma} (\log n)^2}{n^{\varepsilon^2/8}}\right)$ を得る.

次に, 頂点 i_{k+1} (終点とは限らない) のみ区間 I_σ に入るような単調パスに関して, i_{k+1} を始点 i_0 と見なし, $i_0 = O(\log n)$ として I_σ におけるパスの長さを評価すると, 上記の結果から, このパスは I_σ において $O(\log \log n)$ の長さになる. よって, $i_0 = n$ を始点とする G_1^n の単調パス $i_0 i_1 \dots i_{k+1}$ の長さは $\frac{(1-\varepsilon)\log n}{2}$ 以上かつ $\frac{(1+\varepsilon)\log n}{2}$ 以下である. さらに, $i_0 < n$ の場合についても上記と同様に考えると,

$i_0 = n$ を始点とする G_1^n の単調パスが最長になることは明らかであり, 定理の主張は示された. $\square$

次に, 定理 1 に関連して以下の定理を示す.

定理 2

定数 ε を $0 < \varepsilon < 1$ とする. G_1^n において, 区間 $[n - n^{\varepsilon^3/8}, n]$ に含まれるほとんどすべての 2 頂点間の距離は, 少なくとも $\frac{(1-\varepsilon)\log n}{2}$ である.

証明

区間 $[n - n^{\varepsilon^{3/8}}, n]$ に含まれる任意の異なる 2 頂点を i_0 および j_0 とする. そして, i_0 を始点とする長さ $k+1$ の単調パス $i_0 \cdots i_{k+1}$ と j_0 を始点とする長さ $k'+1$ の単調パス $j_0 \cdots j_{k'+1}$ が頂点 i_{k+1} と $j_{k'+1}$ で最初に交わる確率の大きさを評価すると, 補題 1 と 2 から以下ようになる. なお, 共有点 i_{k+1} と $j_{k'+1}$ は同一なので, これを v とする. また, i_0 と j_0 は固定する.

$$\begin{aligned}
 & \mathbb{P} \left(\bigcup_{v=1}^{n-2} \bigcup_{k=0}^{((1-\varepsilon)\log n)/2} \bigcup_{k'=0}^{((1-\varepsilon)\log n)/2} \bigcup_{\substack{n \geq i_0 > \cdots > i_k > v, \\ n \geq j_0 > \cdots > j_{k'} > v}} \left(A_{i_0, i_1} \cap \cdots \cap A_{i_k, v} \cap A_{j_0, j_1} \cap \cdots \cap A_{j_{k'}, v} \right) \right) \\
 & \leq \sum_{v=1}^{n-2} \sum_{k=0}^{((1-\varepsilon)\log n)/2} \sum_{k'=0}^{((1-\varepsilon)\log n)/2} \sum_{\substack{n \geq i_0 > \cdots > i_k > v, \\ n \geq j_0 > \cdots > j_{k'} > v}} \mathbb{P} \left(A_{i_0, i_1} \cap \cdots \cap A_{i_k, v} \cap A_{j_0, j_1} \cap \cdots \cap A_{j_{k'}, v} \right) \\
 & \leq \sum_{v=1}^{n-2} \sum_{k=0}^{((1-\varepsilon)\log n)/2} \sum_{k'=0}^{((1-\varepsilon)\log n)/2} \sum_{\substack{n \geq i_0 > \cdots > i_k > v, \\ n \geq j_0 > \cdots > j_{k'} > v}} \mathbb{P} \left(A_{i_0, i_1} \right) \cdots \mathbb{P} \left(A_{i_{k-1}, i_k} \right) \\
 & \quad \times \mathbb{P} \left(A_{j_0, j_1} \right) \cdots \mathbb{P} \left(A_{j_{k'-1}, j_{k'}} \right) \mathbb{P} \left(A_{i_k, v} \cap A_{j_{k'}, v} \right) \\
 & \leq \sum_{v=1}^{n-2} \sum_{k=0}^{((1-\varepsilon)\log n)/2} \sum_{k'=0}^{((1-\varepsilon)\log n)/2} \sum_{\substack{n \geq i_0 > \cdots > i_k > v, \\ n \geq j_0 > \cdots > j_{k'} > v}} \mathbb{P} \left(A_{i_0, i_1} \right) \cdots \mathbb{P} \left(A_{i_{k-1}, i_k} \right) \\
 & \quad \times \mathbb{P} \left(A_{j_0, j_1} \right) \cdots \mathbb{P} \left(A_{j_{k'-1}, j_{k'}} \right) \frac{c_0}{v \sqrt{i_k j_{k'}}} \\
 & \leq \frac{c_0}{\sqrt{i_0 j_0}} \sum_{v=1}^{n-2} \frac{1}{v} \sum_{k=1}^{((1-\varepsilon)\log n)/2} c_1^k \sum_{i_1=v+k}^{i_0-1} \cdots \sum_{i_k=v+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k} \\
 & \quad \times \sum_{k'=1}^{((1-\varepsilon)\log n)/2} c_1^{k'} \sum_{j_1=v+k'}^{j_0-1} \cdots \sum_{j_{k'}=v+1}^{j_{k'-1}-1} \frac{1}{j_1} \frac{1}{j_2} \cdots \frac{1}{j_{k'}}
 \end{aligned}$$

ここで補題 1 の (1) の証明によれば, 任意の定数 $\delta > 0$ に対して, 2 つの頂点 $i < j$ の頂点番号がともにある正の定数 a 以上ならば, $\mathbb{P}(g_j = i) \leq \frac{c_1}{\sqrt{ij}}$ における定数 $c_1 > \frac{1}{2}$ は, $\frac{1+\delta}{2}$ より小さくなる. よって,

長さ $k+1$ の単調パスならば, 少なくとも $k-a+2$ 個の辺について, 上式の c_1 は $\frac{1+\delta}{2}$ より小さくなる.

なお, このパスはもう 1 つのパスと最後の頂点が共有されると考えるので, 最後の辺に関する定数 1 個は計算から除外する. このことから上の最後の式は以下のように変更できる.

$$\begin{aligned}
& \frac{c_0}{\sqrt{i_0 j_0}} \sum_{v=1}^{n-2} \frac{1}{v} \left(\sum_{k=1}^{a-2} c_1^k \sum_{i_1=v+k}^{i_0-1} \cdots \sum_{i_k=v+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k} \right. \\
& + \sum_{k=a-1}^{((1-\varepsilon)\log n)/2} c_1^{a-2} \left(\frac{1+\delta}{2} \right)^{k-a+2} \sum_{i_1=v+k}^{i_0-1} \cdots \sum_{i_k=v+1}^{i_{k-1}-1} \frac{1}{i_1} \frac{1}{i_2} \cdots \frac{1}{i_k} \Big) \\
& \times \left(\sum_{k'=1}^{a-2} c_1^{k'} \sum_{j_1=v+k'}^{j_0-1} \cdots \sum_{j_{k'}=v+1}^{j_{k'-1}-1} \frac{1}{j_1} \frac{1}{j_2} \cdots \frac{1}{j_{k'}} + \sum_{k'=a-1}^{((1-\varepsilon)\log n)/2} c_1^{a-2} \left(\frac{1+\delta}{2} \right)^{k'-a+2} \sum_{j_1=v+k'}^{j_0-1} \cdots \sum_{j_{k'}=v+1}^{j_{k'-1}-1} \frac{1}{j_1} \frac{1}{j_2} \cdots \frac{1}{j_{k'}} \right).
\end{aligned}$$

ここで、 $\delta \leq \varepsilon^3/4$ となるように a を設定すれば、定理 1 の証明における $P(E_1)$ の大きさを評価する部分を参照にして以下を得る。

$$\begin{aligned}
& P \left(\bigcup_{v=1}^{n-2} \bigcup_{k=0}^{((1-\varepsilon)\log n)/2} \bigcup_{k'=0}^{((1-\varepsilon)\log n)/2} \bigcup_{\substack{n \geq i_0 > \cdots > i_k > v, \\ n \geq j_0 > \cdots > j_{k'} > v}} \left(A_{i_0, i_1} \cap \cdots \cap A_{i_k, v} \cap A_{j_0, j_1} \cap \cdots \cap A_{j_{k'}, v} \right) \right) \\
& = O \left(\frac{(\log n)^3}{n^{\varepsilon^2/4 - \delta/2}} \right) O \left(\frac{(\log n)^3}{n^{\varepsilon^2/4 - \delta/2}} \right) = O \left(\frac{(\log n)^6}{n^{\varepsilon^2/2 - \delta}} \right).
\end{aligned}$$

区間 $[n - n^{\varepsilon^3/8}, n]$ に含まれる 2 頂点 i_0 および j_0 の組み合わせの個数は高々 $n^{\varepsilon^3/4}$ なので、主張は示されたことになる。□

この定理では、ループを認めるグラフ G_1^n を対象にしているので、その主張は区間 $[n - n^{\varepsilon^3/8}, n]$ に含まれる 2 頂点の間にパスがない場合、すなわち距離が無限大の場合も含むことになる。ループを認めない G_1^n 、すなわち一般的な木については、定理 1 と 2 から、最長のパスは $(1 - \varepsilon)\log n$ と予想できる。これについては次回議論する予定である。

4. 次回に向けて

次回は、補題 4 に関連する議論を行い、プロセス (G_m^t) で得られるランダムグラフを考察する。

参考文献

- A.-L. Barabási and R. Albert, "Emergence of scaling in random networks", *Science* (1999)
- R. Albert, H. Jeong and A.-L. Barabási, "Diameter of the world-wide web", *Nature* (1999)
- A.-L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: the topology of the world-wide web", *Physica A* (2000)
- B. Bollobás and O. Riordan, "The diameter of a scale-free random graph", *Combinatorica* (2004)
- B. Pittel, "Note on the heights of random recursive trees and random m-ary zsearch trees", *Random Structures*

スケールフリーランダムグラフの直径に関する考察Ⅱ

and Algorithms 5 (1994)

浜口幸弘, スケールフリーランダムグラフの直径に関する考察Ⅰ, 明治学院大学経済研究 153 (2017)

(2018年3月16日 受理)