

スケールフリーランダムグラフの直径に関する考察Ⅲ

浜 口 幸 弘

1. はじめに

前稿の浜口（2018）では，Bollobás and Riordan（2004）のグラフ生成プロセスモデル $(G_1^t)_{t \geq 0}$ に基づき，それから生成されるグラフの大きさに関する定理を 2 つ導いた．今回は，ランダムプロセス $(G_1^t)_{t \geq 0}$ を若干変更したプロセス $(H_1^t)_{t \geq 0}$ から生成される 1 つの木を対象にして，その性質を考察する．このモデルは，グラフに新たな頂点を追加するとき，グラフ内の既存の頂点のみとその次数の大きさに応じて確率的に接続するというものである．したがって，追加する頂点そのものを接続先として選択することはないので，生成されるグラフは 1 本の木となる．したがって，そのアプローチの方法では異なるが，Pittel（1994）の考えるモデルと同じである．そして結果として，定理 3 において木上の長さ k の単調パスの存在確率の上限と下限を導き，定理 4 では，その長さ k の単調パスの始点から数えて k 番目の頂点の位置の平均値とその存在確率を示した．なお，文中で用いる用語と定義は前回（浜口 2017, 2018）と同じであり，補題および定理の番号は，前回からの続きとなる．

2. 概念とモデルの定義

本稿では，プロセス $(G_1^t)_{t \geq 0}$ において，頂点 v が自分自身を選択しないように Bollobás and Riordan（2004）の頂点選択のモデルを若干変更して議論する．すなわち，グラフ G に新しい頂点 v を 1 個ずつ加えるとき，その頂点 v から m 本の辺をグラフ G に接続する（よって， G はループを持たない）ものとする．そして，この新たなグラフ生成プロセスを $(H_1^t)_{t \geq 0}$ と呼ぶことにする．

さて， $m=1$ の場合について，厳密にこのことを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順番に $v_1, v_2, \dots$ と固定し，グラフ H における頂点 v の次数を $d_H(v)$ と記す．ランダムグラフのプロセス $(H_1^t)_{t \geq 0}$ 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 H_1^t が構成される．すなわち，まず 1 個の頂点 v_1 （次数 1 とする）を持つグラフ H_1^1 から出発する．そして， H_1^{t-1} ままで構成されたとき，次に追加する頂点 v_t と $v_i \in V(H_1^{t-1})$ を 1 本の辺で結ぶ．このとき頂点 $i (=v_i)$ は次

のようにランダムに選択されるものとする。

$$P(i=s) = d_{H_1^{t-1}}(s)/(2t-3) \quad 1 \leq s \leq t-1.$$

このモデルは、Pittel (1994) に従い、頂点 $v_i \in V(H_1^{t-1})$ の選択される確率が $D_{H_1^{t-1}}(v_i) + 1$ （ここで $D_{H_1^{t-1}}(v_i)$ は、 H_1^{t-1} において、 $i+1$ 番目以降に v_i に接続した辺の本数を意味する）に比例することから決定される。このとき構成されるグラフは 1 つの木である。

$m > 1$ の場合については、一度に v_i から m 本の辺が H_1^{t-1} に接続される。よって、プロセス (H_m^t) については、プロセス (H_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 (H_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 k 番目の頂点とするパスを $i_0 i_1 \cdots i_k$ 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 $i_0 > i_1 > \cdots > i_k$ 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、 H_1^t は $[t] = \{1, 2, \dots, mn\}$ 上のランダムグラフとする。このとき、 h_j をグラフに追加する頂点 j が接続されるグラフ内の頂点として表す。また、ランダムグラフの頂点 i と j が 1 本の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 H_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

プロセス $(H_1^t)_{t \geq 0}$ に対して、補題 1（浜口 2018）と同様の考え方により $P(h_j = i)$ の大きさを評価すると、以下の補題 5 を得る。

補題 5

$(H_1^t)_{t \geq 0}$ に対して、 $1 \leq i < j$ 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。

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

また、以下のようにも表せる。

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

証明

グラフの構成プロセス (H_1^t) を考える。 H_1^t は確率変数である。また、グラフ H_1^t における頂点 i の次数を $d_{t,i} = d_{H_1^t}(i)$ と記す。上述のプロセス (H_1^t) の定義から、以下となる。

$$P(h_t = i | H_1^{t-1}) = d_{t-1,i} / (2t-3) \quad t > i \geq 1.$$

このとき、 $d_{t,i} = d_{t-1,i} + I_{|h_t=i|}$ となるから、

$$E(d_{t,i} | H_1^{t-1}) = d_{t-1,i} + d_{t-1,i} / (2t-3) = \left(1 + \frac{1}{2t-3} \right) d_{t-1,i}.$$

上式の両辺の期待値をとると,

$$E(d_{t,i}) = \left(1 + \frac{1}{2t-3}\right) E(d_{t-1,i}).$$

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 をとるので, $\mu_{i,i} = 1$ となり, $t \geq i$ に対して,

$$\mu_{t,i} = \mu_{i,i} \prod_{r=i+1}^t \left(1 + \frac{1}{2r-3}\right) = \prod_{r=i}^{t-1} \left(1 + \frac{1}{2r-1}\right)$$

となる. よって, t を $t-1$ として置き換えれば, 補題 1 と同様の結果を得る. $\square$

あとの定理で前提条件となる $i \rightarrow \infty$ の場合は, 補題 1 の証明によって, 上記の補題は以下のようになる. $i < j$ のとき, ある定数 $c_1 > 0$ および $c_2 > 0$ に対して,

$$\frac{1}{2\sqrt{ij}} \left(1 + \frac{c_2}{i^2}\right) \leq P(h_j = i) \leq \frac{1}{2\sqrt{ij}} \left(1 + \frac{c_1}{i}\right).$$

補題 5 の結果は, プロセス $(G_1^t)_{t \geq 0}$ において得られる結果と同じである. したがって, 定理 1 と定理 2 (浜口 2018) はプロセス $(H_1^t)_{t \geq 0}$ に対しても成り立つことになる. また, プロセス $(H_1^t)_{t \geq 0}$ に対しては, 以下の補題が成り立つ.

補題 6

H_1^n における単調パス $i_0 i_1 \cdots i_k$ に対して, 以下の式が成り立つ.

$$P(A_{i_0 i_1} \cap \cdots \cap A_{i_{k-1} i_k}) = P(A_{i_0 i_1}) \cdots P(A_{i_{k-1} i_k}).$$

証明

プロセス (H_1^t) において単調パス $i_0 i_1 \cdots i_t i_{t+1} \cdots i_k$ が構成されているとする (パスの頂点は添え字の大きい方から順に接続されることに注意). そして, 最初の頂点 1 から頂点 $i_t - 1$ (i_t の直前) まで接続の順序に従って追加される頂点の集合を $\{1, 2, \dots, i_k, \dots, i_t - 1\}$ とする. i_t の追加までプロセスが実行されたとき, この集合を構成するすべての頂点の次数の和と i_t の次数は, どのようなプロセスの場合でも同じである. したがって, それ以降の頂点 $i_t + 1, \dots, i_{t-1}$ をグラフに追加するときの i_t の次数の確率分布は, i_t までは頂点同士がどのように接続されても同じである. よって事象 $A_{i_t i_{t+1}} \cap \cdots \cap A_{i_{k-1} i_k}$ と事象 $A_{i_{t-1} i_t}$ ($1 \leq t \leq k-1$) は独立となり, $P(A_{i_0 i_1} \cap \cdots \cap A_{i_{k-1} i_k}) = P(A_{i_0 i_1}) \cdots P(A_{i_{k-1} i_k})$ が成り立つ. $\square$

これらの補題をもとに, H_1^n 上で, i_0 を始点とする長さ k の単調パス $i_0 \cdots i_k$ (H_1^n 上で, i_k はこのパスの終点とは限らない) が存在する確率の上限および下限を評価する以下の定理 3 を導出する.

定理 3

H_1^n 上で, その長さが k ($k \geq 1$) であり, 始点が i_0 , 始点から k 番目の頂点が i であるような単調パスの存在する確率を $P_{i_0 \rightarrow i}^k$ とする.

(1) $0 < \varepsilon_1 < 1$ を任意の定数とし, $i \rightarrow \infty$ とする. このとき, ある正の整数を N とすると, 任意の整数 $n \geq N$ に対して, $P_{i_0 \rightarrow i}^k$ は以下の不等式を満たす.

$$P_{i_0 \rightarrow i}^k \leq \left(\frac{1 + \varepsilon_1}{2}\right)^k \frac{(\log i_0 - \log i)^{k-1}}{(k-1)! \sqrt{i_0 i}}.$$

(2) $0 < \varepsilon_2 < 1$ を任意の定数とし, 定数 $\delta = \max \left\{ \min_{x \text{ は正の整数}} \left\{ x : \frac{2}{3^{x-1}} + 2 \left(\frac{4}{9} \right)^x < \varepsilon_2 \right\}, 3 \right\}$ および

$k^{\delta+1} = o(i)$ とする. また, $i_0 \geq i + k^\delta$ とする. このとき, ある正の整数を N とすると, 任意の整数 $n \geq N$ に対して, $P_{i_0 \rightarrow i}^k$ は以下の不等式を満たす.

$$P_{i_0 \rightarrow i}^k \geq \left(\frac{1 - \varepsilon_2}{2} \right)^k \frac{(\log i_0 - \log i)^{k-1}}{(k-1)! \sqrt{i_0 i}}.$$

証明

(1) H_1^n において, 頂点 x と y が 1 本の辺で接続される事象を $A_{x,y}$ とする. このとき, 始点を i_0 とし, 始点から k 番目の頂点を i_k とするような長さ k の単調パス $i_0 i_1 \cdots i_k$ が存在する任意の事象 $A_{i_0, i_1} \cap \cdots \cap A_{i_{k-1}, i_k}$ 同士は, 互いに排反となり, また, 事象 $A_{i_t, i_{t+1}} \cap \cdots \cap A_{i_{k-1}, i_k}$ と事象 A_{i_{t-1}, i_t} ($1 \leq t \leq k-1$) は独立となるので, 補題 5 と補題 6 を用いて以下を得る. なお便宜上, パスの k 番目の頂点は $i_k (= i)$ と記し, i_k は固定して考える.

$$\begin{aligned} P_{i_0 \rightarrow i_k}^k &= P \left(\bigcup_{n \geq i_0 > i_1 > \cdots > i_k} (A_{i_0, i_1} \cap A_{i_1, i_2} \cap \cdots \cap A_{i_{k-1}, i_k}) \right) \\ &= \sum_{n \geq i_0 > i_1 > \cdots > i_k} P(A_{i_0, i_1} \cap A_{i_1, i_2} \cap \cdots \cap A_{i_{k-1}, i_k}) \\ &= \sum_{i_1 = i_k + k - 1}^{i_0 - 1} \sum_{i_2 = i_k + k - 2}^{i_1 - 1} \cdots \sum_{i_{k-1} = i_k + 1}^{i_{k-2} - 1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}). \end{aligned}$$

ここで, パスの長さ k に関する帰納法を適用する. 証明すべき式は以下ようになる. $0 < \varepsilon_1 < 1$ を任意の定数とし, $i_k \rightarrow \infty$ とする. このとき十分大きな n に対して,

$$\sum_{i_1 = i_k + k - 1}^{i_0 - 1} \sum_{i_2 = i_k + k - 2}^{i_1 - 1} \cdots \sum_{i_{k-1} = i_k + 1}^{i_{k-2} - 1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \leq \left(\frac{1 + \varepsilon_1}{2} \right)^k \frac{(\log i_0 - \log i_k)^{k-1}}{(k-1)! \sqrt{i_0 i_k}}.$$

まず, $k=1$ の場合, 補題 5 から主張は成り立つ. 次に, 帰納法の仮定として $k-1$ の場合, この不等式が成り立つとし, 始点が i_1 で始点から $k-1$ 番目の頂点が i_k である長さ $k-1$ の単調パスにこの仮定を適用すれば, c を正の定数として, 補題 5 から以下を得る.

$$\begin{aligned} &\sum_{i_1 = i_k + k - 1}^{i_0 - 1} \sum_{i_2 = i_k + k - 2}^{i_1 - 1} \cdots \sum_{i_{k-1} = i_k + 1}^{i_{k-2} - 1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \\ &= \sum_{i_1 = i_k + k - 1}^{i_0 - 1} P(A_{i_0, i_1}) \sum_{i_2 = i_k + k - 2}^{i_1 - 1} \cdots \sum_{i_{k-1} = i_k + 1}^{i_{k-2} - 1} P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{i_1=i_k+k-1}^{i_0-1} \frac{1}{2\sqrt{i_0 i_1}} \left(1 + \frac{c}{i_1}\right) \left(\frac{1+\varepsilon_1}{2}\right)^{k-1} \frac{(\log i_1 - \log i_k)^{k-2}}{(k-2)! \sqrt{i_1 i_k}} \\
 &\leq \frac{1}{2} \left(1 + \frac{c}{i_k}\right) \left(\frac{1+\varepsilon_1}{2}\right)^{k-1} \frac{1}{(k-2)! \sqrt{i_0 i_k}} \sum_{i_1=i_k+k-1}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1}.
 \end{aligned}$$

ここで, $\sum_{i_1=i_k+k-1}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1}$ を評価する. そこで, $a > 0$ および正の整数 m に対して,

$$f(x) = \frac{(\log x - \log a)^m}{x} \quad (x \geq 1) \text{ とする. } f'(x) \text{ は}$$

$$f'(x) = \frac{(\log x - \log a)^{m-1} (m - (\log x - \log a))}{x^2} \text{ となるから,}$$

$1 \leq x \leq ae^m$ のとき, $f'(x) \geq 0$ であり, $x \geq ae^m$ のとき, $f'(x) \leq 0$ である. そして, $x = ae^m$ のとき $f(x)$ は最大値 $m^m / (ae^m)$ をとる. このとき, $m = k-2$ および $a = i_k$ とし, 区間 $[i_k + k-1, i_0-1]$ で $f(x)$ を考える. そして, $t \leq ae^m < t+1$ を満たす整数 t をとる. 以下では, t がこの区間内にある場合を考えるが, t が区間内でない場合については, $f(x)$ は区間内で単調に変化するので, 以下の議論を簡単にして同様に扱える.

$$\text{よって, } \sum_{i_1=i_k+k-1}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1} \text{ の上限は } f(t) \geq f(t+1) \text{ の場合, 以下ようになる}$$

($f(t) < f(t+1)$ の場合も同様に評価できる).

$$\begin{aligned}
 &\sum_{i_1=i_k+k-1}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1} \\
 &\leq \int_{i_k+k-1}^t \frac{(\log x - \log i_k)^{k-2}}{x} dx + \int_{t+1}^{i_0} \frac{(\log (x-1) - \log i_k)^{k-2}}{x-1} dx + \frac{(\log t - \log i_k)^{k-2}}{t} \\
 &\leq \frac{(\log (i_0-1) - \log i_k)^{k-1}}{k-1} - \frac{(\log (i_k+k-1) - \log i_k)^{k-1}}{k-1} + \frac{1}{i_k} \left(\frac{k-2}{e}\right)^{k-2} \\
 &\leq \frac{(\log i_0 - \log i_k)^{k-1}}{k-1} \left(1 - \left(\frac{\log (i_k+k-1) - \log i_k}{\log i_0 - \log i_k}\right)^{k-1} + \frac{k-1}{i_k (\log i_0 - \log i_k)} \left(\frac{k-2}{e (\log i_0 - \log i_k)}\right)^{k-2}\right) \\
 &\leq \frac{(\log i_0 - \log i_k)^{k-1}}{k-1} \left(1 + \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right).
 \end{aligned}$$

最後の不等式は以下の理由による. t は区間 $[i_k + k-1, i_0-1]$ の中にあるので, 以下となる.

$$i_0 > i_k e^{k-2} \text{ すなわち, } \log i_0 - \log i_k > k-2.$$

そして、 $k \geq 3$ として考える。なお、 $k=2$ の場合については、最初の式に $k=2$ を代入して直接計算すれば、上の不等式を満たす。このことから、以下を得る。

$$\begin{aligned} & \sum_{i_1=i_k+k-1}^{i_0-1} \sum_{i_2=i_k+k-2}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \\ & \leq \frac{1}{2} \left(1 + \frac{c}{i_k}\right) \left(\frac{1+\varepsilon_1}{2}\right)^{k-1} \left(1 + \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right) \frac{(\log i_0 - \log i_k)^{k-1}}{(k-1)! \sqrt{i_0 i_k}}. \end{aligned}$$

よって、十分大きな n をとれば、 $i_k \rightarrow \infty$ なので、以下の不等式が成り立つ。

$$\frac{1}{2} \left(1 + \frac{c}{i_k}\right) \left(1 + \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right) \leq \frac{1+\varepsilon_1}{2}.$$

また、 t が区間内でない場合についても同じ結果を得られるので、以上のことから、(1) は示された。

(2) パスの長さ k に関する帰納法を適用する。証明すべき式は以下ようになる。 $0 < \varepsilon_2 < 1$ を任意の定数とし、定数 $\delta = \max \left\{ \min_{x \text{ は正の整数}} \left\{ x : \frac{2}{3^{x-1}} + 2 \left(\frac{4}{9} \right)^x < \varepsilon_2 \right\}, 3 \right\}$ および $k^{\delta+1} = o(i)$ とする。このとき (1) を参考にすれば、

$$\begin{aligned} P_{i_0 \rightarrow i_k}^k &= P \left(\bigcup_{n \geq i_0 > i_1 > \cdots > i_k} (A_{i_0, i_1} \cap A_{i_1, i_2} \cap \cdots \cap A_{i_{k-1}, i_k}) \right) \\ &= \sum_{i_1=i_k+k-1}^{i_0-1} \sum_{i_2=i_k+k-2}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \\ &\geq \sum_{i_1=i_k+(k-1)^\delta}^{i_0-1} \sum_{i_2=i_k+(k-2)^\delta}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1^\delta}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \end{aligned}$$

が成り立つ。よって、以下を示せばよい。

$$\begin{aligned} & \sum_{i_1=i_k+(k-1)^\delta}^{i_0-1} \sum_{i_2=i_k+(k-2)^\delta}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1^\delta}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \\ & \geq \left(\frac{1-\varepsilon_2}{2} \right)^k \frac{(\log i_0 - \log i_k)^{k-1}}{(k-1)! \sqrt{i_0 i_k}}. \end{aligned}$$

まず、 $k=1$ および $k=2$ の場合、補題 5 から主張は成り立つ。 $k \geq 3$ の場合、(1) と同様に考え、 c を正の定数として、帰納法の仮定から以下を得る。このとき、 $i_1 \geq i_k + (k-1)^\delta$ だから、帰納法を適用する際、その前提条件を満たすことに注意する。

$$\sum_{i_1=i_k+(k-1)^\delta}^{i_0-1} \sum_{i_2=i_k+(k-2)^\delta}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1^\delta}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k})$$

$$\begin{aligned}
 &= \sum_{i_1=i_k+(k-1)\delta}^{i_0-1} \mathbf{P}(A_{i_0, i_1}) \sum_{i_2=i_k+(k-2)\delta}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1\delta}^{i_{k-2}-1} \mathbf{P}(A_{i_1, i_2}) \cdots \mathbf{P}(A_{i_{k-1}, i_k}) \\
 &\geq \sum_{i_1=i_k+(k-1)\delta}^{i_0-1} \frac{1}{2\sqrt{i_0 i_1}} \left(1 + \frac{c}{i_1^2}\right) \left(\frac{1-\varepsilon_2}{2}\right)^{k-1} \frac{(\log i_1 - \log i_k)^{k-2}}{(k-2)! \sqrt{i_1 i_k}} \\
 &\geq \frac{1}{2} \left(1 + \frac{c}{i_0^2}\right) \left(\frac{1-\varepsilon_2}{2}\right)^{k-1} \frac{1}{(k-2)! \sqrt{i_0 i_k}} \sum_{i_1=i_k+(k-1)\delta}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1}
 \end{aligned}$$

ここで, $\sum_{i_1=i_k+(k-1)\delta}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1}$ の下限は(1)と同様に考えて, $f(t) \geq f(t+1)$ の場合,

以下のようになる. このとき, t が区間 $[i_k + (k-1)\delta, i_0 - 1]$ の中にある場合を考えるが, t が区間外の場合についても同様に扱える ($f(t) < f(t+1)$ の場合も同様に評価できる).

$$\begin{aligned}
 &\sum_{i_1=i_k+(k-1)\delta}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1} \\
 &\geq \int_{i_k+(k-1)\delta-1}^t \frac{(\log x - \log i_k)^{k-2}}{x} dx + \int_t^{i_0-1} \frac{(\log(x+1) - \log i_k)^{k-2}}{x+1} dx \\
 &= \int_{i_k+(k-1)\delta-1}^{i_0} \frac{(\log x - \log i_k)^{k-2}}{x} dx - \int_t^{t+1} \frac{(\log x - \log i_k)^{k-2}}{x} dx \\
 &\geq \int_{i_k+(k-1)\delta}^{i_0} \frac{(\log x - \log i_k)^{k-2}}{x} dx - \frac{1}{i_k} \left(\frac{k-2}{e}\right)^{k-2} \\
 &\geq \frac{(\log i_0 - \log i_k)^{k-1}}{k-1} \left(1 - \left(\frac{\log(i_k + (k-1)\delta) - \log i_k}{\log i_0 - \log i_k}\right)^{k-1} - \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right).
 \end{aligned}$$

ここで, 以下のような不等式の評価を行う. $i_0 \geq i_k + k^\delta$ だから,

$$\begin{aligned}
 &\left(\frac{\log(i_k + (k-1)\delta) - \log i_k}{\log i_0 - \log i_k}\right)^{k-1} \leq \left(\frac{\log(i_k + (k-1)\delta) - \log i_k}{\log(i_k + k^\delta) - \log i_k}\right)^{k-1} = \left(\frac{\log(1 + (k-1)\delta/i_k)}{\log(1 + k^\delta/i_k)}\right)^{k-1} \\
 &\leq \left(1 - \frac{1}{k}\right)^{(k-1)\delta} \left(1 + \frac{2k^\delta}{3i_k}\right)^{k-1} \leq \left(\frac{1}{(1+1/(k-1))^{k-1}}\right)^\delta \left(1 + \frac{k^{\delta+1}}{i_k}\right) \\
 &\leq \left(\frac{4}{9}\right)^\delta \left(1 + \frac{k^{\delta+1}}{i_k}\right).
 \end{aligned}$$

上記の2番目の不等式は, $0 \leq x \leq 1/2$ に対して,

$$x - \frac{x^2}{2} \leq \log(1+x) \leq x - \frac{x^2}{3}$$

が成り立つことから,

$$\frac{\log(1+y)}{\log(1+x)} \leq \frac{y}{x} \left(1 + \frac{2x}{3}\right)$$

を得ることによる.

したがって, $k \geq 3$ かつ $k^{\delta+1} = o(i_k)$ だから,

$$\begin{aligned} & \sum_{i_1=i_k+(k-1)^\delta}^{i_0-1} \sum_{i_2=i_k+(k-2)^\delta}^{i_1-1} \cdots \sum_{i_{k-1}=i_k+1}^{i_{k-2}-1} P(A_{i_0, i_1}) P(A_{i_1, i_2}) \cdots P(A_{i_{k-1}, i_k}) \\ & \geq \frac{1}{2} \left(1 + \frac{c}{i_0^2}\right) \left(\frac{1-\varepsilon_2}{2}\right)^{k-1} \frac{1}{(k-1)! \sqrt{i_0 i_k}} (\log i_0 - \log i_k)^{k-1} \left(1 - \left(\frac{4}{9}\right)^\delta \left(1 + \frac{k^{\delta+1}}{i_k}\right) - \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right) \\ & \geq \frac{1}{2} \left(1 + \frac{c}{i_0^2}\right) \left(\frac{1-\varepsilon_2}{2}\right)^{k-1} \frac{1}{(k-1)! \sqrt{i_0 i_k}} (\log i_0 - \log i_k)^{k-1} \left(1 - 2\left(\frac{4}{9}\right)^\delta - \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right). \end{aligned}$$

このとき, 十分大きな n をとれば, δ の定義から以下の不等式が成り立つ.

$$\frac{1}{2} \left(1 + \frac{c}{i_0^2}\right) \left(1 - 2\left(\frac{4}{9}\right)^\delta - \frac{2}{i_k} \left(\frac{1}{e}\right)^{k-2}\right) \geq \frac{1-\varepsilon_2}{2}$$

また, t が区間内でない場合も同様に考えられるが, $f(x)$ が区間内において単調増加する場合は, 以下のように評価する.

$$\begin{aligned} & \sum_{i_1=i_k+(k-1)^\delta}^{i_0-1} \frac{(\log i_1 - \log i_k)^{k-2}}{i_1} \\ & \geq \int_{i_k+(k-1)^\delta-1}^{i_0-1} \frac{(\log x - \log i_k)^{k-2}}{x} dx \\ & \geq \frac{(\log i_0 - \log i_k)^{k-1}}{k-1} \left(\left(1 + \frac{\log(1-1/i_0)}{\log i_0 - \log i_k}\right)^{k-1} - \left(\frac{\log(i_k + (k-1)^\delta) - \log i_k}{\log i_0 - \log i_k}\right)^{k-1} \right) \\ & \geq \frac{(\log i_0 - \log i_k)^{k-1}}{k-1} \left(1 - \frac{2}{k^{\delta-1}} - \left(\frac{4}{9}\right)^\delta \left(1 + \frac{k^{\delta+1}}{i_k}\right)\right). \end{aligned}$$

このとき, 最後の不等式は以下の評価による. $0 \leq x \leq 1/10$ に対して, $\log(1-x) \geq -\sqrt{2}x$ かつ

$\log(1+x) \geq \frac{1}{\sqrt{2}}x$ となるから,

$$\begin{aligned} & \left(1 + \frac{\log(1-1/i_0)}{\log i_0 - \log i_k}\right)^{k-1} \geq \left(1 - \frac{\sqrt{2}}{i_0(\log i_0 - \log i_k)}\right)^{k-1} \\ & \geq \left(1 - \frac{\sqrt{2}}{(i_k + k^\delta) \log(1 + k^\delta/i_k)}\right)^{k-1} \geq 1 - \frac{2}{k^{\delta-1}}. \end{aligned}$$

以上のことから, (2)は示された. □

次に、この定理を用いて、 i_0 を始点とする長さ k の単調パスの k 番目の頂点の位置の平均値およびその存在しない範囲を明らかにするための定理を導く。

定理 4

$0 < \mu < 1$ を任意の定数とし、 k は $1 \leq k \leq \frac{(1-\mu)\log n}{2}$ を満たす整数とする。そして、 H_1^n において、始点を $i_0 (=n)$ とし、始点から k 番目の頂点を i とするような長さ k の単調パスが存在する確率を $P_{i_0 \rightarrow i}^k$ とする。そして確率変数 X を、 i_0 を始点とする長さ k の単調パスの k 番目の頂点の番号とする。このとき、 $0 < \varepsilon < 1$ および $0 < \alpha < 1$ を任意の定数として、ある正の整数を N とすると、任意の整数 $n \geq N$ に対して、以下の不等式が成り立つ。

$$\left(\frac{1-\varepsilon}{3}\right)^k n(1+o(1)) \leq E(X) \leq \left(\frac{1+\varepsilon}{3}\right)^k n(1+o(1))$$

$$P\left(|X - E(X)| \geq \left(\frac{(1+\varepsilon)(1+\alpha)}{5}\right)^{k/2} n\right) \leq \frac{1}{(1+\alpha)^k}.$$

証明

ここでは定理 3 を適用するが、その際、定理において用いている独立した任意の定数 $0 < \varepsilon_1 < 1$ および $0 < \varepsilon_2 < 1$ に対して、 $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ とし、 ε_1 と ε_2 の代わりに、共通の ε を用いることにする。

この定数 ε に対して、定数 $\delta = \max\left\{\min_{x \text{ は正の整数}} \left\{x : \frac{2}{3^{x-1}} + 2\left(\frac{4}{9}\right)^x \leq \varepsilon\right\}, 3\right\}$ とする。また、 $\sigma = (\log n)^{\delta+2}$

として、区間を $I_\sigma = [1, \sigma]$ とする。便宜上、始点を $n = i_0$ とする。そして、確率変数 X を、 i_0 を始点とする長さ k の単調パスの始点から k 番目の頂点の番号とする。このとき、平均値 $E(X)$ は以下のうに 3 つの式から構成される。

$$E(X) = \sum_{i=1}^{i_0-k} i P_{i_0 \rightarrow i}^k = \sum_{i=1}^{\sigma} i P_{i_0 \rightarrow i}^k + \sum_{i=\sigma+1}^{i_0-k\delta} i P_{i_0 \rightarrow i}^k + \sum_{i=i_0-k\delta+1}^{i_0-k} i P_{i_0 \rightarrow i}^k.$$

まず、 $2 \leq k \leq \frac{(1-\mu)\log n}{2}$ の場合について、これらの 3 つの式を評価する。

$E(X)$ を構成する最初の式について、 i_0 を始点とする長さ k の単調パスの k 番目の頂点 i は区間 I_σ に入るので、この単調パスにおいて最初に I_σ に入る頂点は始点から k 番目以下の頂点である。よって定理 1 (浜口 2018) の証明を参照すれば、こうした単調パスの存在する事象は事象 E_1 に対応するので、この確率は $O(\sqrt{\sigma}(\log n)^2/n^{\mu^2/4})$ である。このことから以下を得る。

$$\sum_{i=1}^{\sigma} i P_{i_0 \rightarrow i}^k \leq \sigma \sum_{i=1}^{\sigma} P_{i_0 \rightarrow i}^k = O(\sigma^{3/2}(\log n)^2/n^{\mu^2/4}) = o(1).$$

次に、 $E(X)$ を構成する最後の式について、その大きさを評価すると補題 5 から以下ようになる。ただし、 c は正の定数とする。

$$\begin{aligned}
& \sum_{i=i_0-k^{\delta+1}}^{i_0-k} i P_{i_0 \rightarrow i}^k \leq (i_0 - k) \left(\frac{k^{\delta} - 1}{k} \right) \left(\frac{c}{\sqrt{i_0 - k^{\delta} + 1} \sqrt{i_0 - k^{\delta} + 2}} \right)^k \\
& = O \left(n \left(\frac{ck^{\delta}}{n} \left(1 + \frac{k^{\delta}}{n} \right) \right)^k \right) = o(1).
\end{aligned}$$

これらのことから、以下を得る.

$$E(X) = \sum_{i=1}^{n-k} i P_{i_0 \rightarrow i}^k = \sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i P_{i_0 \rightarrow i}^k + o(1).$$

よって、 $E(X)$ の大きさを評価するには、 $\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i P_{i_0 \rightarrow i}^k$ の大きさを評価すればよい. このとき、

定理 3 の前提条件はすべて満たされるので、定理 3 を適用できる. よって、以下を得る.

$$\begin{aligned}
& \left(\frac{1-\varepsilon}{2} \right)^k \frac{1}{(k-1)! \sqrt{n}} \sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1} \\
& \leq \sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i P_{i_0 \rightarrow i}^k \\
& \leq \left(\frac{1+\varepsilon}{2} \right)^k \frac{1}{(k-1)! \sqrt{n}} \sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1}.
\end{aligned}$$

ここで、 $\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1}$ の大きさを評価する. $f(x) = x^{1/2} (\log n - \log x)^{k-1}$ ($x \geq 1$)

とすると、 $f(x)$ は $x = n/e^{2(k-1)}$ において極大値をとる. そして、区間 $[(\log n)^{\delta+2}, n-k^{\delta}]$ で $f(x)$ を考えることにし、 $t \leq n/e^{2(k-1)} < t+1$ を満たす整数 t をとる. 以下では、 t がこの区間内にある場合を考えるが、 t が区間内でない場合については、 $f(x)$ は区間内で単調に変化するので、以下の議論を簡

単にして同様に扱える. よって、 $\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1}$ の上限は $f(t) \geq f(t+1)$ の場合、

以下のようになる ($f(t) < f(t+1)$ の場合も同様に評価できる).

$$\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1}$$

$$\begin{aligned}
 &\leq \int_{(\log n)^{\delta+2}}^t x^{1/2} (\log n - \log x)^{k-1} dx + t^{1/2} (\log n - \log t)^{k-1} + \int_{t+1}^{n-k^{\delta+1}} (x-1)^{1/2} (\log n - \log(x-1))^{k-1} dx \\
 &\leq \int_{(\log n)^{\delta+2}}^n x^{1/2} (\log n - \log x)^{k-1} dx + \frac{n^{1/2}}{e^{k-1}} (2(k-1))^{k-1}.
 \end{aligned}$$

ここで、以下のように a_k を定めて、漸化式を構成する.

$$\begin{aligned}
 a_k &= \int_{(\log n)^{\delta+2}}^n x^{1/2} (\log n - \log x)^{k-1} dx \\
 &= \left[\frac{2}{3} x^{3/2} (\log n - \log x)^{k-1} \right]_{(\log n)^{\delta+2}}^n + \frac{2(k-1)}{3} \int_{(\log n)^{\delta+2}}^n x^{1/2} (\log n - \log x)^{k-2} dx \\
 &= -\frac{2}{3} (\log n)^{3(\delta+2)/2} (\log n - (\delta+2) \log \log n)^{k-1} + \frac{2}{3} (k-1) a_{k-1}.
 \end{aligned}$$

このとき、 $k \leq \frac{(1-\mu) \log n}{2}$ に注意してこの漸化式を解くと、以下を得る.

$$a_k = \left(\frac{2}{3} \right)^k (k-1)! (n^{3/2} + O(n^{1.1})).$$

よって、

$$\begin{aligned}
 &\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1} \\
 &\leq a_k + \frac{n^{1/2}}{e^{k-1}} (2(k-1))^{k-1} \\
 &= \left(\frac{2}{3} \right)^k (k-1)! n^{3/2} \left(1 + O\left(\frac{1}{n^{0.4}} \right) \right).
 \end{aligned}$$

一方、 $\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1}$ の下限は以下のように評価される.

$$\begin{aligned}
 &\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2} (\log n - \log i)^{k-1} \\
 &\geq \int_{(\log n)^{\delta+2-1}}^t x^{1/2} (\log n - \log x)^{k-1} dx + \int_t^{n-k^{\delta}} (x+1)^{1/2} (\log n - \log(x+1))^{k-1} dx \\
 &\geq \int_{(\log n)^{\delta+2}}^{n-k^{\delta}} x^{1/2} (\log n - \log x)^{k-1} dx - f(n/e^{2(k-1)}).
 \end{aligned}$$

あとは上限と同様に評価できる. したがって、

$$\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{1/2}(\log n - \log i)^{k-1} = \left(\frac{2}{3}\right)^k (k-1)! n^{3/2} \left(1 + O\left(\frac{1}{n^{0.4}}\right)\right)$$

となることから, 以下を得る.

$$\left(\frac{1-\varepsilon}{3}\right)^k n(1+o(1)) \leq E(X) \leq \left(\frac{1+\varepsilon}{3}\right)^k n(1+o(1)).$$

次に, $E(X^2)$ の大きさを評価する. $E(X)$ の議論と同様にして以下が成り立つ.

$$E(X^2) = \sum_{i=1}^{n-k} i^2 P_{i_0 \rightarrow i}^k = \sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^2 P_{i_0 \rightarrow i}^k + o(1).$$

そして,

$$\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} i^{3/2}(\log n - \log i)^{k-1} = \left(\frac{2}{5}\right)^k (k-1)! n^{5/2} \left(1 + O\left(\frac{1}{n^{0.7}}\right)\right)$$

となるから, 以下を得る.

$$\left(\frac{1-\varepsilon}{5}\right)^k n^2(1+o(1)) \leq E(X^2) \leq \left(\frac{1+\varepsilon}{5}\right)^k n^2(1+o(1)).$$

よって, 分散 $V(X)$ の上限は以下のように評価できる.

$$V(X) \leq \left(\frac{1+\varepsilon}{5}\right)^k n^2(1+o(1)).$$

ここで, Chebyshev の不等式を適用すると, 任意の定数 $0 < \alpha < 1$ に対して, 以下を得る.

$$P\left(|X - E(X)| \geq \left(\frac{(1+\varepsilon)(1+\alpha)}{5}\right)^{k/2} n\right) \leq \frac{1}{(1+\alpha)^k}.$$

$k=1$ の場合も上記に含まれる. よって主張は示された. $\square$

この定理を用いれば, H_1^n において, n を始点とする長さ k の単調パスの k 番目の頂点 i は, $k \rightarrow \infty$ のとき, $\left(\frac{1+\varepsilon}{\sqrt{5}}\right)^k n$ から n までの範囲にはほとんど存在しないことがわかる.

4. 課題

まず, 定理 4 の条件下で, 定理 3 における確率の大きさを評価すると, 確率の下限に対しては, 頂点 i の存在し得る範囲でその合計をとると, 以下のようになる.

$$\sum_{i=(\log n)^{\delta+2}}^{n-k^{\delta}} \left(\frac{1-\varepsilon}{2}\right)^k \frac{1}{(k-1)!\sqrt{n}} (\log n - \log i)^{k-1} \sim (1-\varepsilon)^k.$$

一方, 確率の上限に対して同様の計算をすると, $(1+\varepsilon)^k$ となる. $k \rightarrow \infty$ のとき前者は 0 に近づき, 後者は 1 を超えるので, 上限と下限の確率を求める際の近似を改善する必要があると思われる. さらに, ランダムプロセス $(H_1^t)_{t \geq 0}$ を一般化したランダムプロセス (H_m^t) で生成されるグラフの性質を考察することが次の課題となる.

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 search trees", *Random Structures 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)
 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Ⅱ, 明治学院大学経済研究 156 (2018)