『経済研究』（明治学院大学）第 145 号　2012 年

スケールフリーネットワーク評価のための平均距離の分析

浜　口　幸　弘

1. 本稿の概要

本稿は浜口（2010）の続編に相当する。浜口（2010）では，Li et al.（2005）の提案したネットワークを識別するための新たな基準量 *s*- 値に基づき，次数列が与えられたときに *s*- 値を最大にする木 T_{max} とその中の特徴的な T^*_{max} について考察した。本稿では，その考察をさらに深め，次数列が与えられたとき，ネットワークの主要基準量の 1 つである 2 頂点間の平均距離に関して，それをできる限り小さくするような木 T_{max} ついて考察する。

2. 定義と基本的概念

Caldarelli（2007）はネットワークを特徴づけるための基準量の重要性を指摘しているが，本稿では主要な基準量の 1 つである 2 頂点間の平均距離を用いて，Li et al.（2005）の提案する *s*- 値を最大にする木 T_{max} を対象に分析を行う。*s*- 値を最大にする木 T_{max} およびその中の特徴的な T^*_{max} については，浜口（2010）でその構成方法および性質を考察し，幾つかの命題を与えている。本稿では，さらに平均距離という基準量の視点からより詳細な分析を試みる（本文中の記号の表記および概念の定義については浜口（2010）に従う）。

まず平面上における木 T の描画として，木の最大次数の頂点の 1 つを根 r として最上レベル 0（l_0 と記す）に置き，r から下方に向けて距離が 1 にある頂点の位置をレベル 1（l_1），距離が 2 にある頂点の位置をレベル 2（l_2），…とする。さらに議論をしやすくするために，次のような定義を与える。

定義 1（平面上における木の頂点の位置づけ）

平面上の木 T に対して，その l_1 において頂点を左端から右端に次数の大きいものから並べる。任意の頂点 v の下流の隣接点は，次数の大きい順に左から並べる。また，任意の頂点 u と v が同じレベルで

左からこの順に並ぶとき，それぞれの下流の任意の隣接点 α と β は，α が β より左になるようにする（このような位置づけによって得られる木は元の木と同型である）。また頂点の位置を数字で表すために，根 r を 1 とし，次に l_1 の左端から 2，3，…として，以降各レベルにおいて左端から右端に進みながら番号を順次与え，すぐ下のレベルに進んで同様に番号をつけていく。

さて，グラフの頂点次数列（非増加列とする）が与えられたとき，それを満たす木の存在条件について述べると，次の簡単な命題が成り立つ。

命題 1

グラフの頂点次数列 $D=(d_1, d_2, \cdots, d_n)$ に対して，$\sum_i d_i = 2(n-1)$ が成り立つならば，この次数列を満たす木が存在する。ただし，次数列は非増加列とする。

証明

n に関する帰納法による。$D=(d_1, d_2, \cdots, d_n)$（$d_n=1$）に対して，$\sum_i d_i = 2(n-1)$ とする（次数列は大きさの順）。$n-1$ に対して，主張が成り立つと仮定する。

このとき，次数列 $D'=(d_1-1, d_2, \cdots, d_{n-1})$ に対して，$(d_1-1)+d_2+\cdots+d_{n-1}=2((n-1)-1)$ となるから，帰納法の仮定より D' を満たす木が存在する。そこで，次数 d_1-1 の頂点に次数 $d_n=1$ の頂点を接続すれば，n の場合も成り立つ。□

以下の議論では，頂点次数列 D が与えられ，この前提条件が満たされているとする。

頂点数 n の木 T に対して，任意の 2 頂点間の道 P の長さの平均距離は，一般に次式のように与えられる（任意の 2 頂点間の道の数は 1 つである）。

$$\bar{l}(T) = (T \text{ 上の任意の 2 頂点間の道 } P \text{ の長さの総和}) \Big/ \binom{n}{2}$$

次に，本稿で必要とする T^*_{max} の特徴について以下に述べる（詳細は浜口（2010）参照）。

頂点の位置づけに従う任意の木 T に対して，f_i をレベル i の頂点の個数（$i \geq 1$）とし，$f_0=1$ とする。n_i をレベル i の左端の頂点の番号（$i \geq 1$）とし，$n_0=1$ とする。さらに，$d(j)$ を頂点番号 j の次数（$2 \leq j \leq n$）とし，$d(1)=d(n_0)=f_1$ とする。このとき，

$$n_i = f_0+f_1+\cdots+f_{i-1}+n_0 = n_{i-1}+f_{i-1}$$

であり，また，

$$f_i = (d(n_{i-1})-1)+(d(n_{i-1}+1)-1)+\cdots+(d(n_i-1)-1)$$

である。よって，

$$n_1 = 2, \qquad n_i = 2+d(1)+\sum_{j=2}^{n_{i-1}-1}(d(j)-1) \quad (i \geq 2) \text{ となる。}$$

ここで，T^*_{max} の n_i を n^*_i とおき，$d(j)$ を $d^*(j)$ とおき，f_i を f^*_i とおく。このとき，$n^*_i \geq n_i$ である。また，m を T_{max} の深さ（根から端点までの最大距離），m' を T^*_{max} の深さとすれば，明らかに $m' \leq m$ である。

3. 特定の次数列を満たす $\bar{l}\,(T^*_{max})$ の評価

一般に，木 T の 2 点間の道の平均距離を求めるには，かなりの計算量を要するし，また，その効率的なアルゴリズムも知られていない。そこで，以下のような次数列の簡単な例を与え，T^*_{max} と一般的 T_{max} の平均距離を比較することで，一般的な次数列を満たす T_{max} の場合を予想評価する。

$d_1=d\ (\geq 3)$，$d_2=2$，…，$d_{n-d}=2$，$d_{n-d+1}=1$，…，$d_n=1\ (n\geq 4)$ とする。また，便宜上，$n\equiv 1\ (mod\ d)$ とし，この木を T_0 とする（この次数列を満たす任意の木は，自然に T_{max} となる）。

木 T_0 の根 r に接続する部分木を左から T_1，T_2，…，T_d とする（根は除く）。そして，その頂点数をそれぞれ x_1，x_2，…，x_d として，$1\leq x_1\leq\cdots\leq x_d\leq n-d$ とする。このとき，$x_1+\cdots+x_d=n-1$ となる。

$$f(x)=\sum_{i=1}^{x} i\,(x-(i-1)),\qquad g(x,\ y)=\sum_{i=1}^{x}\ \sum_{j=1}^{y}\ (i+j)$$

とすれば，任意の $1\leq s<t\leq d$ に対して，$f(x)$ は 2 個の端点がともに $\{r\}\cup T_s\ (x:=|T_s|)$ にあるときの道の距離の総和で，$g(x,\ y)$ は 2 個の端点がそれぞれ T_s と $T_t\ (x:=|T_s|,\ y:=|T_t|)$ にあるときの道の距離の総和を表すものとする。よって，

$$\begin{aligned}\bar{l}\,(T_{max})\binom{n}{2} &= \sum_{s=1}^{d} f(x_s)+\sum_{1\leq s<t\leq d} g(x_s,\ x_t)\\ &= \sum_{s=1}^{d}\left(\frac{x_s^3}{6}+\frac{x_s^2}{2}+\frac{x_s}{3}\right)+\sum_{1\leq s<t\leq d}\left(\frac{x_s x_t\,(x_s+x_t+2)}{2}\right)\\ &= \frac{1}{6}\sum_{s=1}^{d} x_s^3+\frac{1}{2}\sum_{1\leq s<t\leq d} x_s x_t (x_s+x_t)+\frac{(n-1)^2}{2}+\frac{n-1}{3}\end{aligned}$$

ここで，

$$h(x_1,\cdots,x_d)=\sum_{s=1}^{d} x_s^3+3\sum_{1\leq s<t\leq d} x_s x_t\,(x_s+x_t)$$

とおくと，

$$\begin{aligned}&h(x_1,\cdots,\ x_d)-\frac{(3d-2)(n-1)^3}{d^2}\\ &= h(x_1,\ \cdots,\ x_d)-\frac{(3d-2)(x_1+\cdots+x_d)^3}{d^2}\\ &= \frac{d-2}{d^2}\sum_{1\leq s<t\leq d}(x_s-x_t)^2\left(x_s+x_t+\frac{3d-2}{d-2}\sum_{1\leq i\neq s,\,t\leq d} x_i\right)\geq 0\end{aligned}$$

となるから，$\bar{l}\,(T_{max})$ の最小値は，$x_1=\cdots=x_d=\dfrac{n-1}{d}$ のとき，すなわち，T^*_{max} の場合である。

さらに別の視点から考え，$n+1$ 番目の頂点 v を T_i または $T_j\ (x_i\leq x_j)$ に接続した場合の，距離の総和の増加分をそれぞれ ΔL_i，ΔL_j とする。このとき，

$$\Delta L_i-\Delta L_j=(x_i-x_j)\,(n-1-x_i-x_j)\leq 0$$

である。すなわち，木 T_0 に頂点を 1 個接続するとき，深さの最も小さい部分木に加えると，距離の総和の増加分は最も小さくなる。

以上のことから，一般的な次数列を満たす T^*_{max} の場合も，他の T_{max} に比べて 2 点間の道の平均距離は小さいと予想できる。

4. $\bar{l}\,(T^*_{max})$ の評価

前述のように，T^*_{max} の 2 点間の道の平均距離は小さくなると予想される。ここでは，それを一般的な次数列を満たす T^*_{max} に拡張して評価を試みる。

まず，一般木 T において，A を任意の 2 頂点間の道 P の集合とし，B_i を端点の最高レベルが l_i である P と内点の最高レベルが l_i である P を合わせた集合とすれば，$A = \cup_i B_i$ および $B_i \cap B_j = \emptyset\ (i \neq j)$ が成り立つ。よって，各レベルごとに任意の道を扱えば，重複なくすべての道を扱えることに注意する。

次に，T_{max} 上のレベル i（l_i と表記）における j 番目の頂点に対して，以下のように記号を定義する。

$N_1(i)$ ：端点の最高レベルが l_i である P の総数

$N_2(i)$ ：内点の最高レベルが l_i である P の総数

$L_{1,\,i}(j)$：j 番目の頂点を端点とし，その最高レベルが l_i である P の長さの和

$L_{2,\,i}(j)$：j 番目の頂点を内点とし，その最高レベルが l_i である P の長さの和

$$L_1(i) = \sum_j L_{1,i}(j) \quad , \qquad L_2(i) = \sum_j L_{2,i}(j)$$

$$L_1 = \sum_i L_1(i) \qquad , \qquad L_2 = \sum_i L_2(i)$$

とする。また，T^*_{max} の場合については，記号に“＊”をつけることにする。このとき，

$$N_1(i-1) = f_i + N_1(i) = f_i + f_{i+1} + \cdots + f_m = \sum_{k=i}^{m} f_k$$

であり，

$$N_1 = \sum_i N_1(i) \ \ ,\ \ N_2 = \sum_i N_2(i) \ \ ,\ \ N_1 + N_2 = \frac{n(n-1)}{2}$$

である。また，

$$L_1(i-1) = f_i + N_1(i) + L_1(i) = f_i + 2f_{i+1} + \cdots + (m-i+1)f_m = \sum_{k=i}^{m} (k-i+1) f_k$$

である。ただし，m は T_{max} の深さとし，長さ 0 の道は考えない。

T^*_{max} は任意のレベル i に対して，$\displaystyle\sum_{k=i}^{m} f_k$ すなわち，レベル i 以下の頂点数が最小となる（浜口（2010）の命題 8 −（b）の証明を参照）。よって，

$N^*_1(i) \leq N_1(i)$ すなわち $N^*_1 \leq N_1$ かつ $L^*_1(i) \leq L_1(i)$ を得る。また，$\dfrac{L_1(i)}{N_1(i)}$ は端点の最高レベルが l_i

である P の長さの平均値を表し，

$$\frac{L_1}{N_1}=\frac{\displaystyle\sum_{k=1}^{m}\frac{k(k+1)}{2}f_k}{\displaystyle\sum_{k=1}^{m}k\,f_k}$$

は木全体におけるその平均値を表す。さて，これまでの定義に基づき，以下の命題が成り立つ。

命題 2

m' は T^*_{max} の深さとする（$m'\geq 1$）。任意の $1\leq i\leq m'$ に対して，$\dfrac{L^*_1(i-1)}{N^*_1(i-1)}>\dfrac{L^*_1(i)}{N^*_1(i)}$ が成り立つ。

証明

$i=m'$ の場合は，明らかに成り立つ。そこで，$1\leq i\leq m'-1$ の場合について考える。

まず，以下のように帰納法を適用する。明らかに，$f^*_0\leq f^*_1$ である。l_{i-1}（$i\geq 2$）に対して，$f^*_0\leq f^*_1\leq\cdots\leq f^*_{i-1}$ が成り立ち，l_{i-2} には次数 1 の頂点が存在しないと仮定する。l_{i-1} に次数 1 の頂点が存在しないならば，$m'\geq i$ であり，$f^*_0\leq\cdots\leq f^*_{i-1}\leq f^*_i$ となるので，引き続き l_i について最初から同様に考える。

一方，l_{i-1} に次数 1 の頂点が存在するならば，T^*_{max} の定義から l_i には次数 2 の頂点が存在しない。よって，$m'=i-1$ または，$m'=i$ となる。前者の場合，$f^*_0\leq f^*_1\leq\cdots\leq f^*_{m'}$ であり，後者の場合，$f^*_0\leq f^*_1\leq\cdots\leq f^*_{m'}$ または，$f^*_0\leq f^*_1\leq\cdots\leq f^*_{m'-1}>f^*_{m'}$ である。

この結果を利用すれば，以下のように式を評価できる。

$$\frac{L^*_1(i-1)}{N^*_1(i-1)}-\frac{L^*_1(i)}{N^*_1(i)}$$

$$=\frac{f^*_i\displaystyle\sum_{k=i+1}^{m'}(k-i)f^*_k-\sum_{k=i}^{m'}f^*_k\sum_{k=i+1}^{m'}f^*_k}{\displaystyle\sum_{k=i}^{m'}f^*_k\sum_{k=i+1}^{m'}f^*_k}$$

$$=\frac{\left(\displaystyle\sum_{k=i}^{m'}f^*_k-f^*_i\right)f^*_{i+1}+\left(\displaystyle\sum_{k=i}^{m'}f^*_k-2f^*_i\right)f^*_{i+2}+\cdots+\left(\displaystyle\sum_{k=i}^{m'}f^*_k-(m'-i)f^*_i\right)f^*_{m'}}{\displaystyle\sum_{k=i}^{m'}f^*_k\sum_{k=i+1}^{m'}f^*_k}$$

$$=\frac{\displaystyle\sum_{j=i+1}^{m'}\left(\sum_{k=i}^{m'}f^*_k-(j-i)f^*_i\right)f^*_j}{\displaystyle\sum_{k=i}^{m'}f^*_k\sum_{k=i+1}^{m'}f^*_k}>0$$

したがって主張を得る。□

ここで，以下の条件 A を設定する

条件 A：木 T のレベル i における頂点数 f_i に対して，$f_i = f_i^* \left(1 \leq i \leq \left\lfloor \frac{m'}{2} \right\rfloor - 1 \right)$ が成り立つ。

このとき，次の命題が成り立つ。

命題 3

条件 A のもとで，$\frac{L_1^*}{N_1^*}$ より小さい $\frac{L_1}{N_1}$ をもつ木はない。

証明

任意の木 T に対して，f_0，f_1，…，f_m を既述のように定義し，m を T_{max} の深さ，m' を T_{max}^* の深さとする。$m \geq 2$ で，$m > m'$ の場合を考えればよい。そこで，

$$h_0 := \frac{L_1}{N_1} = \frac{\sum_{k=1}^{m} \frac{k(k+1)}{2} f_k}{\sum_{k=1}^{m} k f_k}$$

とし，f_m を $f_m - 1$ および f_{m-1} を $f_{m-1} + 1$ と操作した場合に対して，同様に

$$h_1 := \frac{\sum_{k=1}^{m-2} \frac{k(k+1)}{2} f_k + \frac{m(m-1)}{2}(f_{m-1}+1) + \frac{m(m+1)}{2}(f_m - 1)}{\sum_{k=1}^{m-2} k f_k + (m-1)(f_{m-1}+1) + m(f_m - 1)}$$

とする。分母をそろえて整理すれば $h_0 - h_1$ の分子は次のようになる。

$$\sum_{k=1}^{m} \left(mk - \frac{k(k+1)}{2} \right) f_k > 0 \ 。$$

よって，$h_0 > h_1$ となる。以降，操作の回数 j に対して，h_j を同様に定義する。ここで表記を簡単にするため，f_1，…，f_{m-2}，$f_{m-1}+1$，$f_m - 1$ を最初から順に改めて，f_1，…，f_m と置き直し，$m = m'$ となるまで同様の操作を繰り返す（$f_m - 1$ が 0 になれば，これは除外するので，m は小さくなる）。そして，$m = m'$ となるとき，T_{max}^* は任意のレベル i に対して，$\sum_{k=i}^{m} f_k$ すなわち，レベル i 以下の頂点数が最小となるから，$f_m \geq f_{m'}^*$ である。また，これ以降は m の大きさを固定する（すなわち，$m = m'$）。

$i \leq m$ の場合も添え字の大きい順から考えて，$f_i > f_i^*$ ならば（上記と同じ理由から常に $f_i \geq f_i^*$ である），上述と同様に f_i を $f_i - 1$ かつ f_{i-1} を $f_{i-1} + 1$ と操作する。このとき，この f_i 以前に行った操作の回数を j とすれば，$h_j > h_{j+1}$ である。この操作を $f_i = f_i^*$ となるまで繰り返す。そして，$f_i = f_i^*$ となれば，i を 1 減らして同様に考える。これら一連の操作を続けていけば，ある t 回目の操作に，f_0^*，f_1^*，…，$f_{m'}^*$ となって操作は終了する。したがって，

$$h_0 = \frac{L_1}{N_1} > \cdots > h_t = \frac{L_1^*}{N_1^*}$$

となり主張を得る。□

さて，$\bar{l}(T_{max}^*)$ の最小性について評価するが，それを直接考えることは難しい。そこで，命題 3 を用いて近似して評価を試みる。

まず，T_{max} において，内点の最高レベルが l_i である道 P は端点の最高レベルが l_i である 2 つの道から構成されるので，それぞれについて全体の平均の長さを考えると，

$$\frac{L_2}{N_2} \approx 2\frac{L_1}{N_1}$$

と仮定できる。よって，大雑把に評価して $\bar{l}\,(T^*_{max})$ は最小に近い可能性を持つと言える。ここでは，N_1 と N_2 の比（すなわち重み）については考えていないが，それを考慮してもう少し厳密に扱うと次のようになる。上述の仮定を使えば，

$$\begin{aligned}
\bar{l}\,(T_{max}) &= \frac{L_1+L_2}{N_1+N_2} \\
&= \frac{N_1}{N_1+N_2}\,\frac{L_1}{N_1} + \frac{N_2}{N_1+N_2}\,\frac{L_2}{N_2} \\
&\approx \frac{N_1+2N_2}{N_1+N_2}\,\frac{L_1}{N_1} \\
&= \left(1+\frac{N_2}{N_1+N_2}\right)\frac{L_1}{N_1} \\
&= 2\left(1-\frac{N_1}{n\,(n-1)}\right)\frac{L_1}{N_1}
\end{aligned}$$

となる。以下，n を固定して考える。

$$\frac{m}{n} \geq \frac{N_1}{n(n-1)}$$

だから，T^*_{max} と同じ s- 値をもつ T_{max} において，その深さ m が $m \ll n$ となるならば（すなわち，次数の大きな頂点の比率が大きい場合），条件 A のもとでは，命題 3 より，$\bar{l}\,(T^*_{max}) \leq \bar{l}\,(T_{max})$ となる可能性が高い。よってこの場合には，$\bar{l}\,(T^*_{max})$ が最小になり得る。

また，$\dfrac{L_1}{N_1}$ を次のように変形して考える。

$$\frac{L_1}{N_1} = \frac{N_1(0)}{N_1}\,\frac{L_1(0)}{N_1(0)} + \cdots + \frac{N_1(m)}{N_1}\,\frac{L_1(m)}{N_1(m)}$$

このとき，$\displaystyle\sum_{i=0}^{m}\frac{N_1(i)}{N_1} = 1$ である。

よって，$\dfrac{L_1}{N_1}$ は各レベルを端点とする道 P の長さの平均値に重みを乗じた和となる。T_{max} において，明らかに $\dfrac{L_1(i)}{N_1(i)}$ はおおよそ減少傾向にある。よって，$m \ll n$ で複数の T_{max} が同じ深さ m であれば，i が小さいとき，重みの変化分 $\dfrac{N_1(i-1)}{N_1} - \dfrac{N_1(i)}{N_1} = \dfrac{f_i}{N_1}$ すなわち，f_i が大きいと，$\dfrac{L_1}{N_1}$ が小さくなるので，

$\bar{l}(T_{max})$ がより小さくなると予想される。

5. 一般木の平均距離を算出するアルゴリズム

前述のことからわかるように，一般的な木の平均距離を求めるには膨大な計算を要する。ここでは，一定の条件が与えられれば効率的に一般木の平均距離を求められる方法を示す。

頂点の次数列が与えられたとき，次数列を満たす頂点の位置づけに従う任意の木 T が構成できるとする。T に対して $d(j)$ を頂点番号 j の次数（$1 \leq j \leq n$）とすれば，以下のようなアルゴリズムを提案できる。

［アルゴリズム］

入力値：

$$d(j) \quad (1 \leq j \leq n)$$

手続き：

$n_0 = 1$ とし，n_i をレベル i の左端の頂点の番号（$i \geq 1$）とする。本稿 2 章より，

$$n_i = 2 + d(1) + \sum_{j=2}^{n_{i-1}-1} (d(j) - 1)$$

である。

そこで，頂点の位置づけに従う木 T のレベル i に対して，左端から $k(\geq 1)$ 番目（すなわち，根 r から数えて $n_i + k - 1$ 番目）の頂点の次数を $d_i(k)$ とし，その下流の隣接点（レベル $i+1$ に属す）の集合を $A_i(k)$ とする。よって，

$$d_i(k) = d(n_i + k - 1)$$

$$A_i(k) = \left\{ 1 + \sum_{j=1}^{k-1} (d_i(j) - 1),\ \cdots,\ 1 + \sum_{j=1}^{k-1} (d_i(j) - 1) + d_i(k) - 2 \right\}$$

である。ただし，下流に隣接点がない場合は，$A_i(k) = \emptyset$ である。

$T_i(k)$ をレベル i における k 番目の頂点から下流にある頂点の総数とする（k 番目の頂点も含む）と，以下を得る。

$$T_i(k) = \sum_{t=0}^{d_i(k)-2} T_{i+1}\left(1 + \sum_{j=1}^{k-1} (d_i(j) - 1) + t\right) + 1$$

$$L_{1,i}(k) = \sum_{s \in A_i(k)} T_{i+1}(s) + \sum_{s \in A_i(k)} L_{1,i+1}(s)$$

$$L_{2,i}(k) = 2 \sum_{s < t \in A_i(k)} T_{i+1}(s)\, T_{i+1}(t) + \sum_{s \in A_i(k)} \left(\sum_{s \neq t \in A_i(k)} T_{i+1}(t) \right) L_{1,i+1}(s)$$

$$L_1(i) = \sum_k L_{1,i}(k) \quad , \qquad L_2(i) = \sum_k L_{2,i}(k)$$

出力値：

$$L_1 = \sum_i L_1(i) \qquad , \qquad L_2 = \sum_i L_2(i)$$

$$\bar{l}(T) = \frac{2(L_1 + L_2)}{n(n-1)}$$

命題 4

頂点の位置づけに従う任意の木 T に対して，以下の式が成り立つ。

$$T_i(k) = \sum_{t=0}^{d_i(k)-2} T_{i+1}\Bigl(1 + \sum_{j=1}^{k-1}(d_i(j)-1)+t\Bigr)+1$$

$$L_{1,i}(k) = \sum_{s\in A_i(k)} T_{i+1}(s) + \sum_{s\in A_i(k)} L_{1,\,i+1}(s)$$

$$L_{2,i}(k) = 2\sum_{s<t\in A_i(k)} T_{i+1}(s)\,T_{i+1}(t) + \sum_{s\in A_i(k)}\Bigl(\sum_{s\neq t\in A_i(k)} T_{i+1}(t)\Bigr) L_{1,i+1}(s)$$

証明

明らかに，

$$T_i(k) = \sum_{t=0}^{d_i(k)-2} T_{i+1}\Bigl(1 + \sum_{j=1}^{k-1}(d_i(j)-1)+t\Bigr)+1$$

$$L_{1,i}(k) = \sum_{s\in A_i(k)} T_{i+1}(s) + \sum_{s\in A_i(k)} L_{1,i+1}(s)$$

は成り立つ。木 T のレベル i における k 番目の頂点に対して，上述のように $A_i(k)$ を定義する。ここで，任意の $s \in A_i(k)$ に対して，$f_r(s)$ を頂点 s の下流かつレベル r（$i+1 \leq r \leq m$）に位置する頂点の個数とする（ただし，$f_{i+1}(s) = 1$ とし，レベル $i+1$ 上の s の場合も含む）。表記を簡潔にするため，$A_i(k) = \{1,\ 2,\ \cdots,\ \alpha\}$ とする。そして，レベル i における k 番目の頂点を内点とし，その最高レベルが l_i である道 P の長さの総和を $L_{2,\,i}(k)$ とする。まず，$A_i(k)$ に属す点 1 を固定し，任意の点 s（$2 \leq s \leq \alpha$）に対して，1 と s を通る P の長さの総和を求めると，次のようになる（s はその範囲を動かす）。ただし，以下では形式上 m まで加算するが，頂点のないレベルでは，

$f_r(s) = 0$ とする。$\sum_{j=1}^{m} f_{i+j}(1) = T_{i+1}(1)$ かつ $L_1(i) = \sum_{k=i+1}^{m}(k-i)\,f_k$ だから（$L_1(i)$ は本稿 4 章による），

$$\sum_{s=2}^{\alpha}\sum_{j=1}^{m} f_{i+j}(1)\Bigl((j+1)T_{i+1}(s) + L_{1,i+1}(s)\Bigr)$$

$$= \sum_{s=2}^{\alpha}\Bigl(\sum_{j=1}^{m} f_{i+j}(1) + \sum_{j=1}^{m} j f_{i+j}(1)\Bigr) T_{i+1}(s) + T_{i+1}(1)\sum_{s=2}^{\alpha} L_{1,i+1}(s)$$

$$= \Bigl(2T_{i+1}(1) + L_{1,\,i+1}(1)\Bigr)\sum_{s=2}^{\alpha} T_{i+1}(s) + T_{i+1}(1)\sum_{s=2}^{\alpha} L_{1,\,i+1}(s)$$

$$= 2\sum_{s=2}^{\alpha} T_{i+1}(s)\,T_{i+1}(1) + \sum_{s=2}^{\alpha} T_{i+1}(s)\,L_{1,\,i+1}(1) + T_{i+1}(1)\sum_{s=2}^{\alpha} L_{1,\,i+1}(s)$$

同様の方法で，$A_i(k)$ に属す 1 つの点を固定して上記のような総和を求め，それらを加算する。1 つ

の道を 2 回数えるので，それを考慮すると以下を得る。

$$L_{2,i}(k) = 2\sum_{s<t\in A_i(k)} T_{i+1}(s)\,T_{i+1}(t) + \sum_{s\in A_i(k)}\left(\sum_{s\neq t\in A_i(k)} T_{i+1}(t)\right) L_{1,i+1}(s) \quad \square$$

この命題に従えば，一般的な木の平均距離を簡潔に算出できる。

6. 課題

本稿では木を対象としたが，より一般的な，すなわち閉路を持つグラフに対する理論の拡張が必要である。

参考文献

Li, L., Anderson, D., Doyle, J. and Willinger, W. (2005). “Towards a Theory of Scale-Free Graphs: Definition, Properties and Implications”, *Internet Math.* 2: 4, 431-523

Caldarelli, G. (2007). *Scale-Free Networks: complex webs in nature and technology,* Oxford University Press

浜口（2010），スケールフリーネットワーク評価のための基準量の提案，明治学院大学経済研究