

蔓延する利己主義と **Scale-Free** 則

林 幸雄

北陸先端科学技術大学院大学

改訂第 2 版 2019

1. Best Business Books '02

複雑な世界、

NEXUS: Small Worlds and the Groundbreaking Science of Networks

ネットワーク科学の最前線

単純な法則

マーク・ブキャナン

阪本芳久=訳



エイズの蔓延から貧富の格差まで、この発見が数々の難問を解き明かす!

脳細胞、インターネット、
食物連鎖、人間社会……。
どれも同じ法則に従って、
つながっていた!

草思社



アルバート=ラズロ・バラバシ[著]

Albert-László Barabási

青木 薫[訳]

Translated by Kaoru Aoki

LINKED: The New Science of Networks
新ネットワーク思考
世界のしくみを読み解く

すべてはリンクしている。

インターネットの弱点、 エイズの急速な広がり、

すべてを説明するルールがあった

マイクロソフトのひとり勝ち、 アルカイダの組織など、

ついに複雑系の姿をとらえた話題の書!

複雑ネットワーク科学の誕生

日々の生活や経済活動, 物流や通信手段, 我々自身の体内メカニズムに至るまで, どれも **複雑なネットワーク** システムの上で成り立っている.

例えば,

- 人や企業間の関係: 口コミ効果, 連鎖倒産の阻止
- 自律分散システム: インターネット, 移動体通信
- バイオ技術: 分子モータ, 遺伝子治療薬

全く異なる対象や構成要素であるが...

⇒ 新発見はまさに世紀末に起こった!!

Many Real Networks

全く異なる対象や構成要素であるにもかかわらず、現実の多くのネットワークには共通性が存在 R.Albert,A.-L.Barabási, Rev.Med.Phys. 74, 2002.

$$P(k) \sim k^{-\gamma}, \quad 2 < \gamma < 3.$$

社会的関係

知人関係, 企業間取引, 映画の共演, 論文引用, 性的関係, 言語

インフラ技術

インターネット (ルータおよび ASレベル), WWW, 航空路線, 電力網, 電子メール送受信

生物系

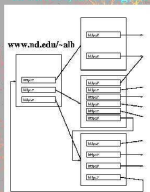
神経回路網, 遺伝子やエネルギー代謝の反応系, 食物連鎖

⇒ 何か普遍的な生成原理が存在するのでは

Technological Nets (Infra.)

World Wide Web

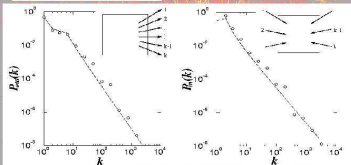
R. Albert, H. Jeong, K.-L. Barabási, *Nature* (in press)



nodes: WWW documents
edges: URL links

800 million documents
(S. Lawrence, 1999)

ROBOT: adds to its database all URL's found in a document and follows them recursively



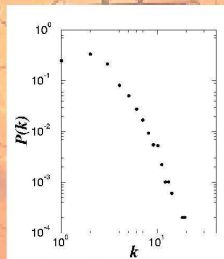
$$P_{\text{out}}(k) \sim k^{-\gamma_{\text{out}}} \quad \gamma_{\text{out}} = 2.45 \quad P_{\text{in}}(k) \sim k^{-\gamma_{\text{in}}} \quad \gamma_{\text{in}} = 2.1$$

(see also Huberman *et al.*, Kleinberg *et al.*, Faloutsos *et al.*)

DIAMETER OF THE WWW : 19 clicks...

ELECTRICAL POWERGRID

nodes: generators, transformers
edges: transmission lines

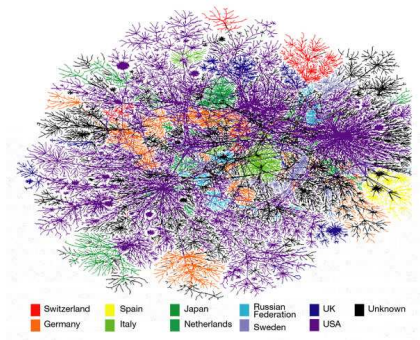


The electrical power grid of the western United States

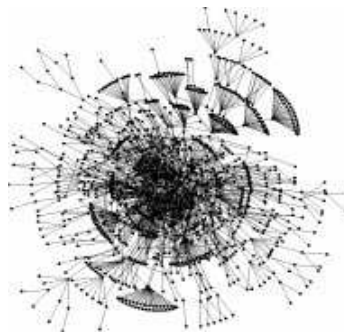
$$\gamma \approx 4$$

<http://www.nd.edu/~networks/ppt/SanDiego.ppt/>

Heterogeneous Structure

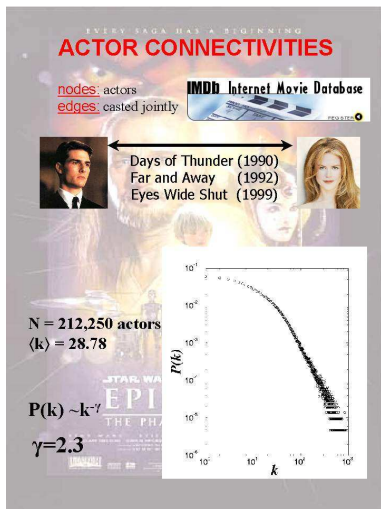


Internet connectivity with selected
backbone ISPs
Nature 406 (CAIDA) 2000

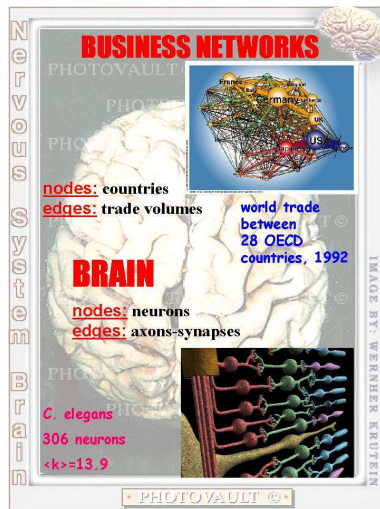


Java Class Components (JDK1.2)
Europhysics Letters 2002

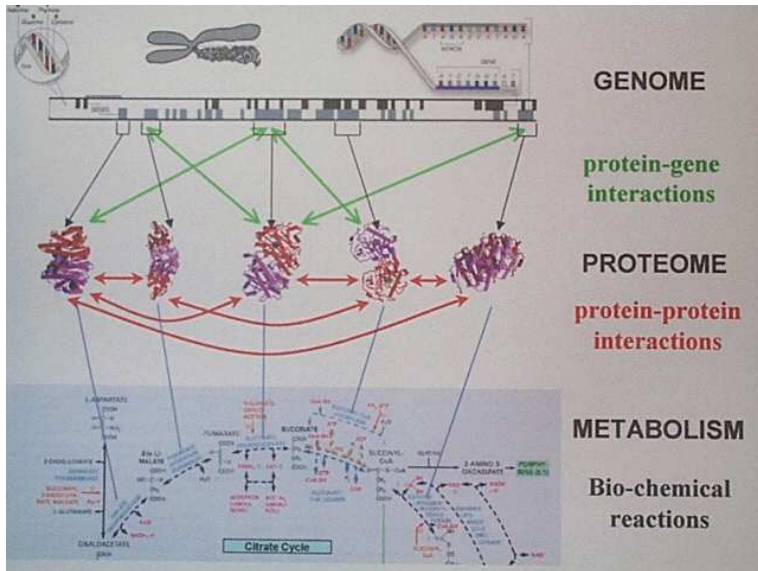
Social Nets



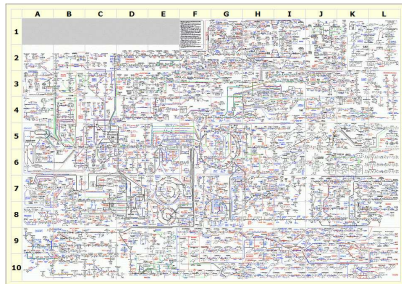
<http://www.nd.edu/~networks/ppt/SanDiego.ppt/>



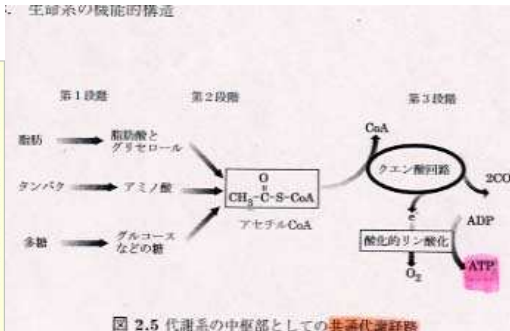
Bio-Chemical Nets



Metabolic Pathways



生体系の機能的構造

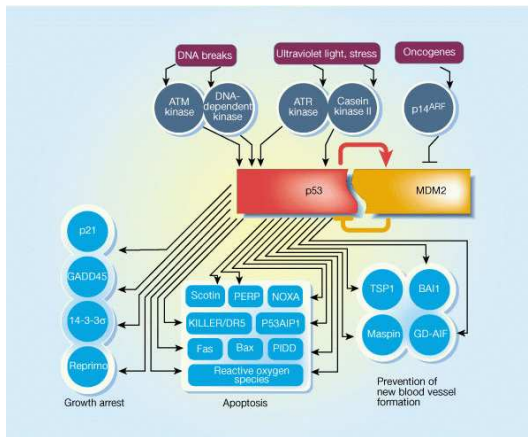


田中著, 生命と複雑系, 培風館, 2002

ATP(アデノシン三リン酸) がエネルギー代謝反応系のハブ

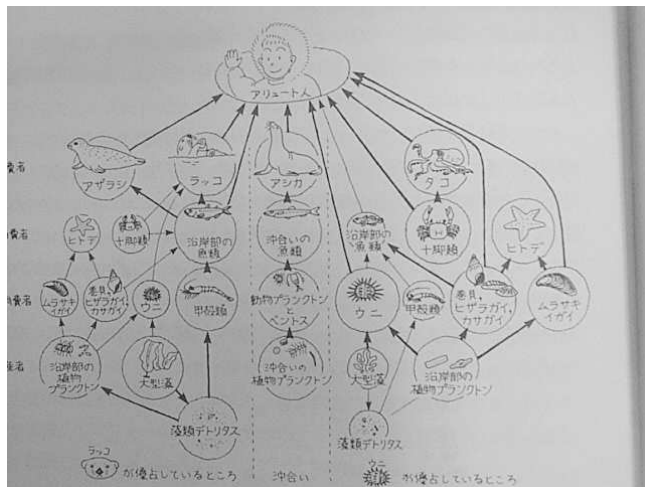
Surfing the p53 Network

癌発病に深く 関与する p53 ハブ遺伝子 ⇒ 新薬の開発



K.W. Kohn, Mol.Bio.Cell 10, 1999 & B. Vogelstein et al., Nature 408, 2000

Food Web



松本著, 生態と環境 (C.A. Simenstand et al., Science 200, 1978.)

2. 優先的選択：次数に比例したリンク確率

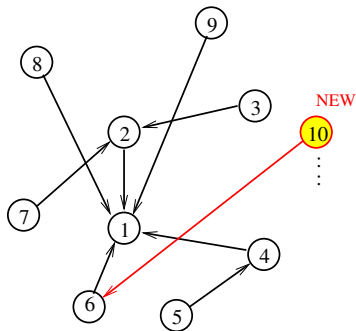
なぜ共通な構造が存在？ → リンク先選択に何か原理があるのでは

- 国内航空網において新規航路を開設する際、全国各地への多くの乗り継ぎ便を持ったハブである羽田空港に乗り入れると便利で自社の利用客が増える。
- 多くの回線が接続する通信拠点に繋ぐ方が少ない経由数で何処にでも辿り着けそう。
- 貧乏人より金持ちと付き合った方が金回りが良い。

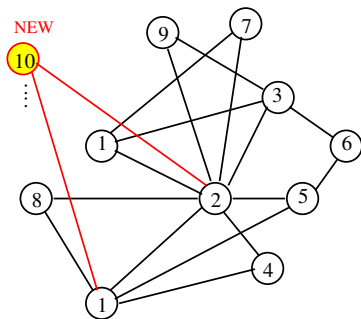
ノードを人に、リンクを経済取引から得るお金に対応付ければ、優先的選択は、"金持ちはより金持ちになる (rich get richer)" 法則に読み替えられる。

優先的選択は単なるネットワーク生成モデルの基本的規則に留まらず、ある種の利己原理を個々人が持てば、例え全体への影響を意図しなくても不平等な世界が生じうる！

GN 木と BA モデル



Krapivsky-Redner's GN model



BA model

P.L. Krapivsky, S.Redner, F.Leyvraz, *Phy.Rev.Lett.* 85, 2000, A.-L.Barabási, R.Albert, H.Jeong, *Physica A* 272, 1999

マスター方程式による次数分布の導出

時刻 t_i での追加ノード i が時刻 t で次数 k を持つ確率

$$p(k, t_i, t+1) = \frac{k-1}{2t} p(k-1, t_i, t) + \left(1 - \frac{k}{2t}\right) p(k, t_i, t),$$

$$\sum_{t_i=1}^{t+1} p(k, t_i, t+1) = \frac{k-1}{2t} \sum_{t_i=1}^t p(k-1, t_i, t) + \left(1 - \frac{k}{2t}\right) \sum_{t_i=1}^t p(k, t_i, t).$$

$p(k) = \lim_{t \rightarrow \infty} \sum_{t_i} p(k, t_i, t)/t \approx \sum_{t_i=1}^t p(k, t_i, t)/t$ を代入,

$$(t+1)p(k) = \frac{k-1}{2t} t p(k-1) + \left(1 - \frac{k}{2t}\right) t p(k),$$

$$p(k) = \frac{k-1}{k+2} p(k-1) = \frac{4}{k(k+1)(k+2)} \sim k^{-3}.$$

利己性の度合いに従った $P(k)$ の変化

Step 0: 次数 0 の孤立ノードがない, N_0 個のノードが連結した初期構成を考える.

Step 1: $t = 1, 2, 3, \dots$ の毎時刻に新ノードを 1 個追加して, 新ノードから既存ノードに m 本のリンクを張る. 各ノード i は確率 $\Pi_i \propto k_i^\nu$ で選択される.

Step 2: 所望のノード数 N まで, Step 1 を繰り返す.

利己性	無 $\nu = 0$	弱 $\leftarrow$ $0 < \nu < 1$	優先的選択 $\nu = 1$	$\rightarrow$ 強 $\nu > 1$
分布 $P(k)$ 核 A_k	指数	カット オフ 付べき乗 sublinear	べき乗 linear	独占状態 superlinear
ハブ の有無	ハブ無	最大次数が抑制	ハブ創出 BA 木	巨大ハブ

度数に比例する優先的選択: ベキ乗分布

時刻 t における次数 k のノード数を $N_k(t)$, 次数 k のノードにリンクが追加される確率を $A_k/A(t)$ として,

$$\frac{dN_k(t)}{dt} = \frac{A_{k-1}N_{k-1}(t) - A_kN_k(t)}{A(t)} + \delta_{k,1},$$

を考える. 優先的選択の $A_k = k$ の時は, $A(t) = \sum_{j \geq 1} jN_j(t) = 2t$: 総リンク数の倍となり, 上式から $N_1 = 2t/3$ と $N_2 = t/6, \dots$ を得る. 従って, より一般に $N_k = t \times p_k$ となる. これらを整理すると,

$$p_k = \frac{k-1}{k+2} \times p_{k-1} = \frac{4}{k(k+1)(k+2)} \approx k^{-3},$$

のベキ乗分布が得られる ($m=1$ の BA 木に相当).

利己性がないとき：指数分布

Growing Exponential Network(GEN) モデルと呼ばれる, $\nu = 0$ のランダム選択では $A_j = 1$, 毎時刻 t に1 個の新ノードが追加されるので $A(t) = \sum_j N_j(t) = t$ より,

$$N_k(t+1) - N_k(t) = \frac{N_{k-1}(t) - N_k(t)}{t} + \delta_{k,1}.$$

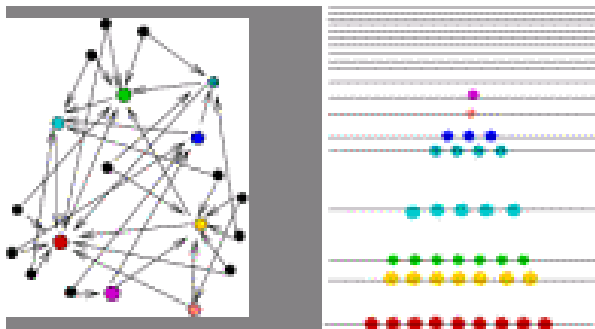
$P(k, t) = N_k(t)/t$ を用いて連続時間の近似をすると,

$$\frac{\partial P(k, t)}{\partial t} = P(k-1, t) - P(k, t) + \delta_{k,1}.$$

$t \rightarrow \infty$ における定常分布 $\frac{\partial P(k, t)}{\partial t} = 0$ を考えると,
 $2P(k) - P(k-1) = \delta_{k,1}$ より, 指数分布 $P(k) = 2^{-k}$.

S.N.Dorogovtsev, J.F.F.Mendes, Evolution of Networks, Oxford University Press, 2003

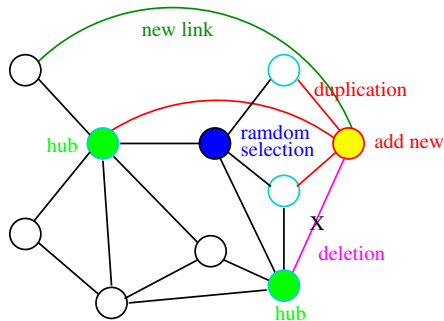
Fitness Model



G. Bianconi et al., PRL 86, 2001.

Bose-Einstein 凝縮と同様なメカニズムによって、リンクの独占が起こることを示唆 (MS の市場独占に対応)

Duplication-Divergence Model



ランダムに選ばれた頂点に隣接するハブにはリンク複写のチャンス大!

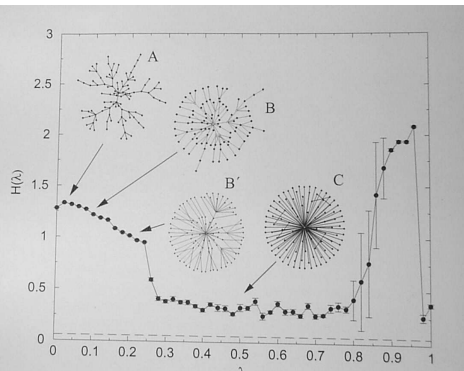
⇒ 生物は、選択不要のより単純な機構で、優先的結合を実現!

R.V.Solé et al., SantaFe Inst. Working Paper, 01-08-041, 2001

3. Optimal Topology: SFはSW

経済性: リンク数 $\leftarrow 0 < \lambda < 1 \rightarrow$ 通信効率: 距離

Random (tree) - Pref. (SF) - Forced (star, clique)



$$\min E(\lambda) = \lambda d + (1 - \lambda)\rho,$$

$$\text{距離 } d \stackrel{\text{def}}{=} \frac{\sum_{i < j} D_{ij}}{n C_2} / D_{\max},$$

$$\text{リンク数 } \rho \stackrel{\text{def}}{=} \frac{\sum_{i < j} a_{ij}}{n C_2},$$

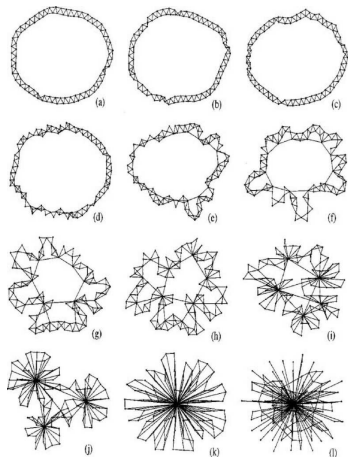
次数 k の頂点頻度 p_k のエントロピーで評価

$$H \stackrel{\text{def}}{=} - \sum_{k=1}^{n-1} p_k \log p_k$$

R.F.iCancho, R.V.Solé, SantaFe Inst. Working Paper 01-11-068, 2001

Hub Evolution

直感的に、ハブ経由のパスはホップ数的に短くなる



最短パスのリンク数 L と頂点間距離 W に関して

$$\min E = \lambda L + (1 - \lambda)W,$$

を Metropolis の SA 法で最適化 距離重視

$\lambda \approx 0$ からリンク数重視 $\lambda \approx 1$ にかけて,
(a) regular $\rightarrow$ (g)~(j) SF $\rightarrow$ (l) random

$\Rightarrow$ SF は経済性と効率を両立する最適構造!

N.Mathias, V.Gopal, Phys.Rev. E 63, 2001

4. K_{max} の推定値: ベキ乗分布

$P(k) = Ck^{-\gamma}$, C は正規化定数で,

$$1 = \int_{K_{min}}^{\infty} P(k) dk = \frac{C}{\gamma - 1} K_{min}^{1-\gamma},$$

より, ベキ指数 $2 \leq \gamma \leq 3$ で $C = (\gamma - 1)K_{min}^{\gamma-1}$.

ネットワーク全体の N 個のノードの中でカット オフ次数 K_{max} 以上の次数のノード 和が高々 1 個と定義され,

$$\frac{1}{N} = \int_{K_{max}}^{\infty} P(k) dk = \frac{C}{\gamma - 1} K_{max}^{1-\gamma},$$

$$K_{max} = K_{min} N^{\frac{1}{\gamma-1}},$$

と表される. BA モデルでは $\gamma = 3$ で最大次数は $O(\sqrt{N})$

K_{max} の推定値: 指数分布

パラメータ $\alpha > 0$ の指数分布 $P(k) = Ce^{-\alpha k}$ の時は,

$$1 = C \int_{K_{min}}^{\infty} e^{-\alpha k} dk = \frac{C}{\alpha} e^{-\alpha K_{min}},$$

より, $C = \alpha e^{\alpha K_{min}}$ で,

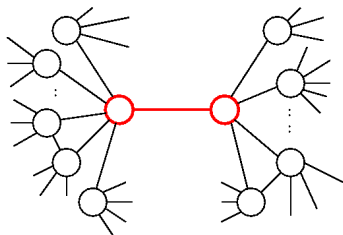
$$\frac{1}{N} = C \int_{K_{max}}^{\infty} e^{-\alpha k} dk = \frac{C}{\alpha} e^{-\alpha K_{max}},$$

$$K_{max} = \frac{1}{\alpha} (\alpha K_{min} + \log N),$$

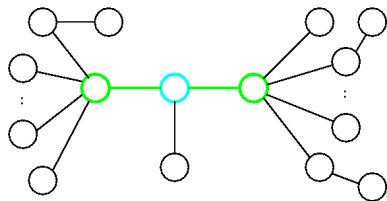
と表され, 最大次数は $O(\log N)$.

5. Degree-degree correlation

$$\bar{k}_{nn}(k) \stackrel{\text{def}}{=} \sum_{k'} k' P(k'|k) = \sum_{i=1}^N \delta(k_i, k) \frac{\sum_{j \in \partial i} k_j}{k N P(k)}.$$



Assortative 結合 ハブ同士など、同程度の次数ノードが結合しがちな『正の次数相関』の傾向
社会システム



Disassortative 結合 ハブと低次数ノードなど、次数が異なるノードが結合しがちな『負の次数相関』の傾向
技術・生物システム

⇒ 結合傾向に依存した伝搬特性, それぞれの長短所を!

M.E.J.Newman, PRL 89, 2002 & PRE 67, 2003

Assortativity

次数相関の強さを表す(次数に対する Pearson 相関係数に相当する)
Assortativity:

$$r \stackrel{\text{def}}{=} \frac{4M \sum_e (k_e k'_e) - [\sum_e (k_e + k'_e)]^2}{2M \sum_e (k_e^2 + k_e'^2) - [\sum_e (k_e + k'_e)]^2},$$

ここで, k_e と k'_e はリンク $e = (i, j)$ の両端ノード i と j の次数を表し,
 $-1 \leq r \leq 1$ である. $r > 0$ が正の次数相関を示す.
上式と等価な r 値の計算式として,

$$r \stackrel{\text{def}}{=} \frac{S_1 S_e - S_2^2}{S_1 S_3 - S_2^2},$$

などもある. $S_e \stackrel{\text{def}}{=} \sum_{i,j} A_{ij} k_i k_j$, $S_1 \stackrel{\text{def}}{=} \sum_i k_i$, $S_2 \stackrel{\text{def}}{=} \sum_i k_i^2$, $S_3 \stackrel{\text{def}}{=} \sum_i k_i^3$ とした. 但し, データ構造に依存して, あるいは(3 乗などによる) 数値的発散なども考慮して使い分ける.

M.E.J.Newman, Networks -Introduction, Oxford Univ. Press, 2010.

相関係数との等価性

次数 k_i に対する平均 μ と共分散 $\text{cov}(k_i, k_j)$ を用いて, 相関係数 r は以下のように導出される.

$$\mu \stackrel{\text{def}}{=} \frac{\sum_{ij} A_{ij} x_i}{\sum_{ij} A_{ij}} = \frac{\sum_i k_i x_i}{\sum_i k_i} = \frac{1}{2M} \sum_i k_i x_i,$$

$$\begin{aligned} \text{cov}(x_i, x_j) &\stackrel{\text{def}}{=} \frac{\sum_{ij} A_{ij} (x_i - \mu)(x_j - \mu)}{\sum_{ij} A_{ij}} \\ &= \frac{1}{2M} \sum_{ij} A_{ij} (x_i x_j - \mu x_i - \mu x_j + \mu^2) \\ &= \frac{1}{2M} \sum_{ij} A_{ij} x_i x_j - \mu^2 = \frac{1}{2M} \sum_{ij} A_{ij} x_i x_j - \frac{1}{(2M)^2} \sum_{i,j} k_i k_j x_i x_j \\ r &\stackrel{\text{def}}{=} \frac{\text{cov}(k_i, k_j)}{\sqrt{\text{cov}(k_i, k_i)} \sqrt{\text{cov}(k_j, k_j)}} = \frac{\sum_{ij} (A_{ij} - k_i k_j / 2M) k_i k_j}{\sum_{ij} (k_i \delta_{ij} - k_i k_j / 2M) k_i k_j}. \end{aligned}$$

M.E.J.Newman, Networks -Introduction, Oxford Univ. Press, 2010.