

NII市民講座: 複雑科学とネットワーク ～つながり方の科学～

国立情報学研究所

福田 健介

kensuke@nii.ac.jp

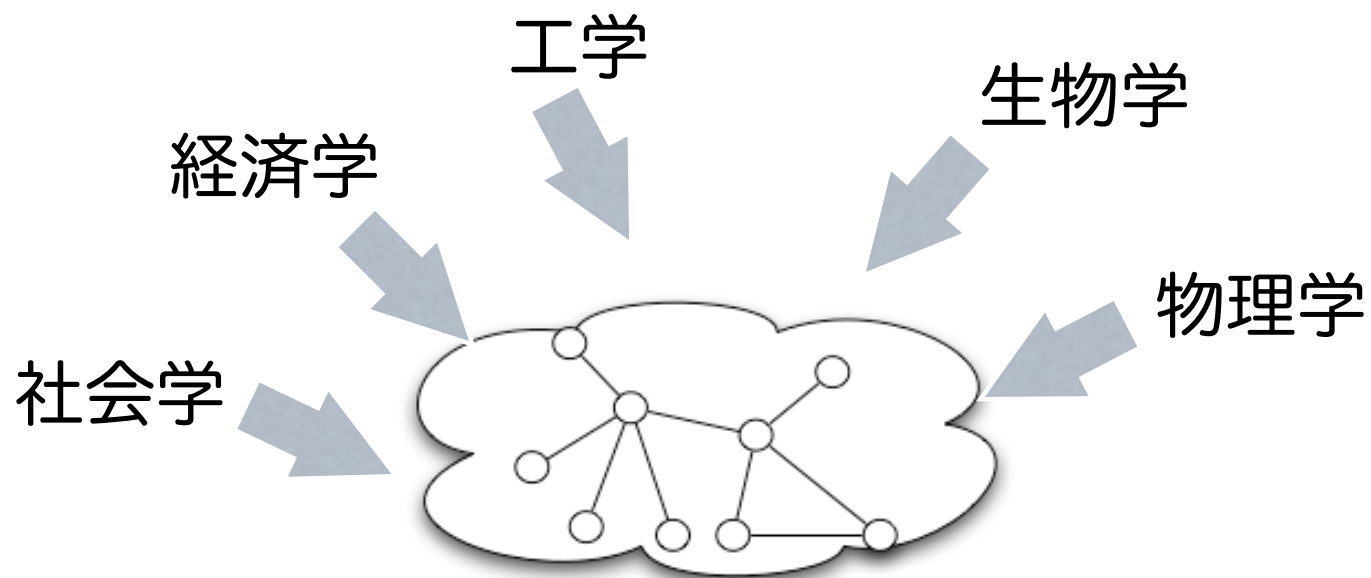
今日の話題

- 友達の輪!
- 転職を頼むには誰にお願いするのが良い？
- なんでxxxは流行るのか？
- ケビンベーコンって知ってます？
- インターネットやWebの直径はどのくらい？
- HIV/SARSの感染
- 漢方薬と西洋薬

今日の話題

- 世の中にはたくさんのネットワークが存在
 - 人間関係のネットワーク
 - 自然界に見られるネットワーク
 - 人工物に見られるネットワーク
- 共通する規則性の存在
 - スケールフリー
 - スモールワールド

“ネットワーク”の研究



ネットワークとは？

- 頂点(ノード)と辺(リンク)から構成される構造
 - 人間(ノード)と人間関係(リンク)
 - 空港(ノード)と航空路線(リンク)
 - ウェブページ(ノード)とページリンク(リンク)
- “グラフ”と言ってもいい
- (注) 今日のお話では、インターネットはネットワークの一つの例

(例)人間関係ネットワーク

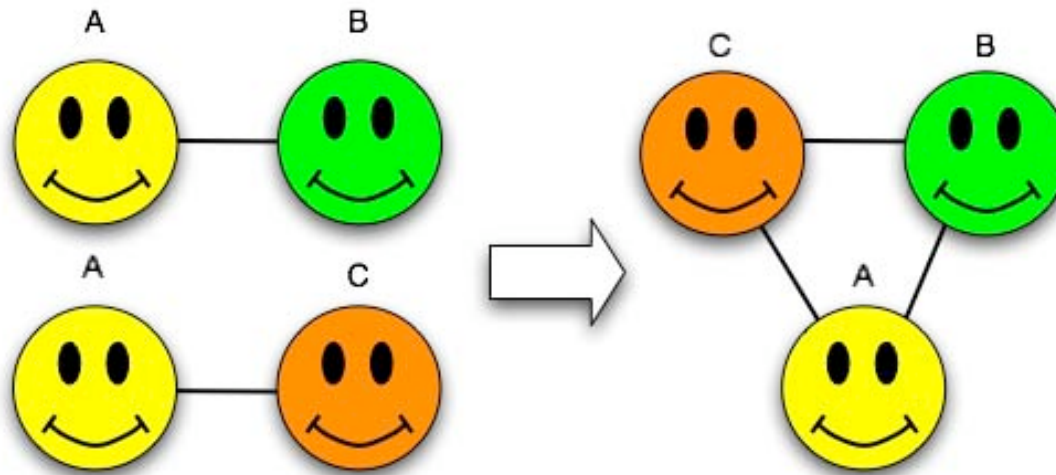
- ノード: 人間
- リンク: 二人の間関係
 - 関係あればリンクが存在, なければなし
 - 通常は重みなし・無向で考える
 - 実は“関係”を定義することが難しい



友人関係？

- 色々な友人関係
 - 義兄弟
 - お金貸してくれる人
 - 同僚
 - 同じ趣味の人
- 関係の定義によって作られるネットワークは全く異なる
- 時間の経過によっても変化

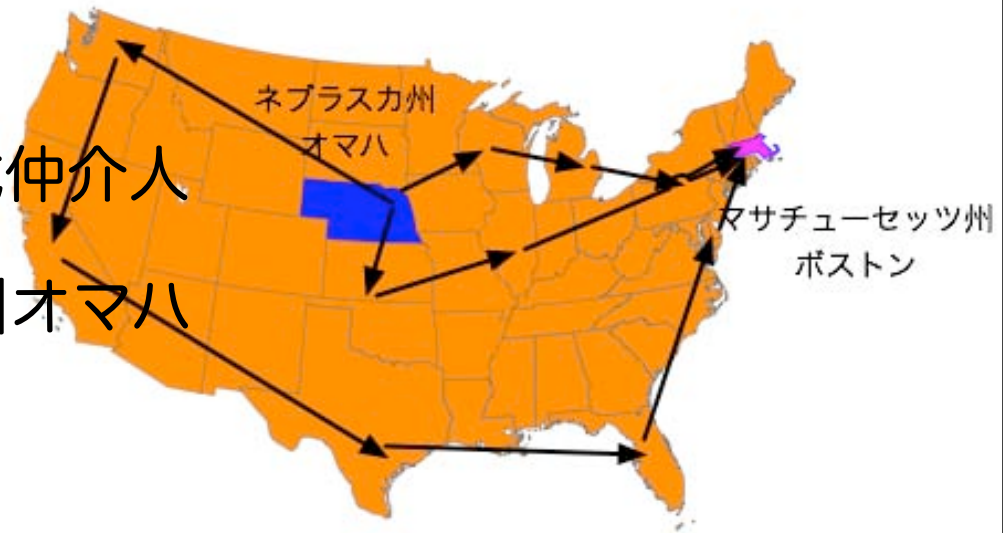
It's a small world



- 初対面のBさんとCさんが話をしていると、Aさんという共通の知人がいることが判明
- 三角形(クラスタ)で表現
- 世の中本当に狭いのか？

友人世界のサイズはどのくらい？

- スタンレー・ミルグラムの実験 (1967)
- 手紙をスタートからゴールまで届ける
- 中継できるのはファーストネームで呼び合う仲の人とだけ
- ゴール: ボストンの株式仲介人
- スタート: ネブラスカ州オマハ



ミルグラムの実験結果

- 160通のうち42通が到着
- 平均の知人の輪: 5.5人 (→ 6次の隔たり)
- 当時のアメリカの人口: 2億人
- 人間関係は思っているよりも、ずっとコンパクトなネットワーク

J.Travers and S.Milgram,
An Experimental Study of the Small World Problem,
Sociometry, vol.32, no.4, pp.425-443, 1969

ミルグラムの実験再び(2003)

- 実験を世界規模へ拡張
 - ゴール: 13人(18ヶ国)
 - スタート: 61668人(166ヶ国)
 - ただし今回はemail
- <http://smallworld.columbia.edu/>

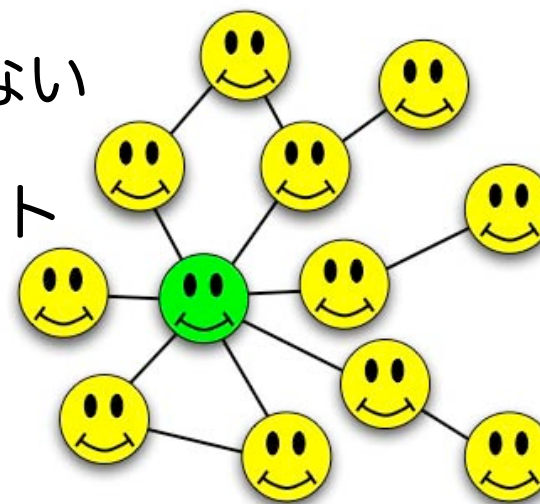
P.S.Dodds, R.Muhamad, and D.Watts,
An Experimental Study of Search in Global Social Networks,
Science, vol.31, no.5634, pp.827-829, 2003

email実験の結果

- 32%のemailが転送
- 友人の輪: 5人(同じ国の人), 7人(異なる国の人)
- やはりコンパクトなネットワーク

xxxはなぜ流行るのか？

- 始めは少数のマニア内での人気
- 影響力のある「コネクター」に伝わると人気
が急上昇
- 人間の他者への影響力は均一ではない
- 流行する閾値: ティッピングポイント

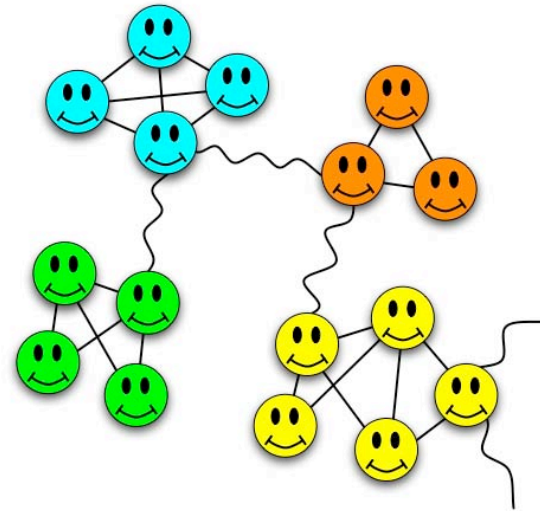


転職はどうやって決める？

- 転職に関するグラノヴェッターの研究 (1969)
 - 直接応募: 20%
 - 転職会社: 18.8%
 - 知り合いの紹介: 56%
 - しばしば会う人: 16.7%
 - たまに会う人: 55.6%
 - ほとんど会わない人: 28%

転職はどうやって決める？

- 助けてくれるのは少し離れた人
 - 親しい友人間だとネットワークが孤立 (強いつながり)
 - 異なるコミュニティを結びつける人の存在 (弱いつながり)
- 「弱い紐帯の強さ」



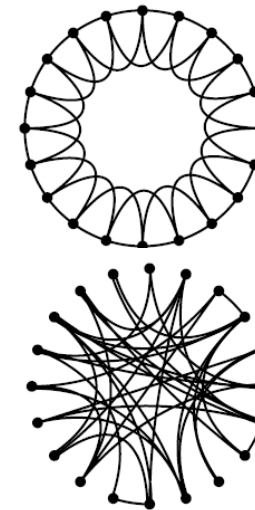
M.Granovetter
The Strength of Weak Ties
American Journal of Sociology, Vol.78, pp.1360-1380, 1973

ネットワークの特徴

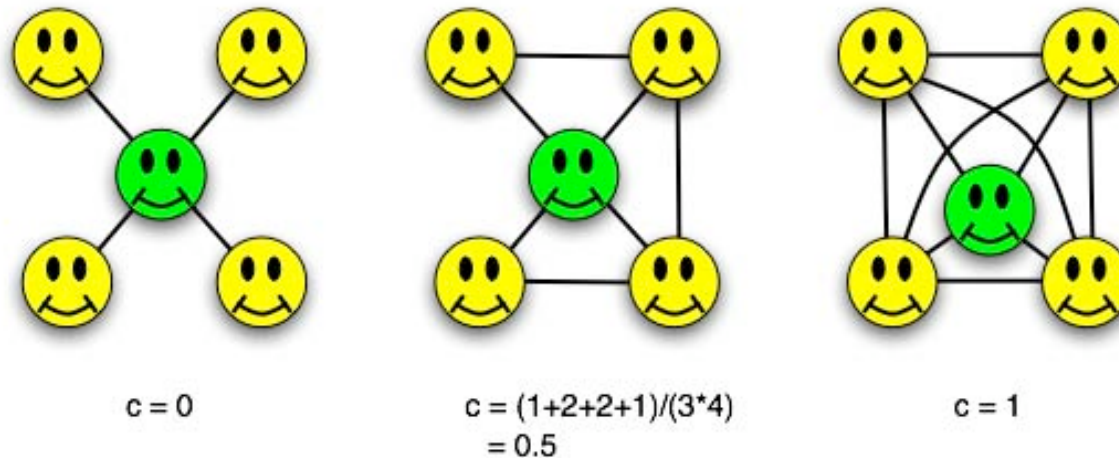
- 世界は狭い (→ スモールワールド)
- コンパクト (任意の二人の距離は短い)
- クラスタ (友達の友達は友達)
- ノードの持つリンク数に偏り (→ スケールフリー)
 - 多数のノードのリンク数は小さく, 少数のノードのリンク数は大きい

スモールワールド

- 統計的性質
 - **小直径**: ショートカットが存在
 - **高クラスタリング係数**: 三角形がたくさん
- $\Leftrightarrow$ 規則的なネットワーク:
 - 大直径, クラスタ化
- $\Leftrightarrow$ ランダムなネットワーク:
 - 小直径, クラスタなし



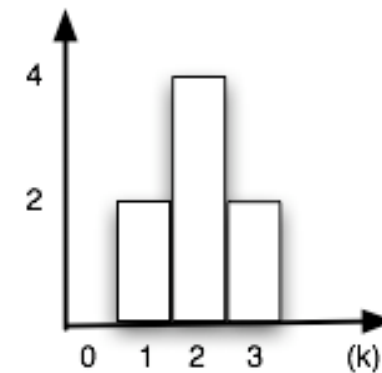
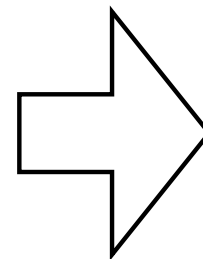
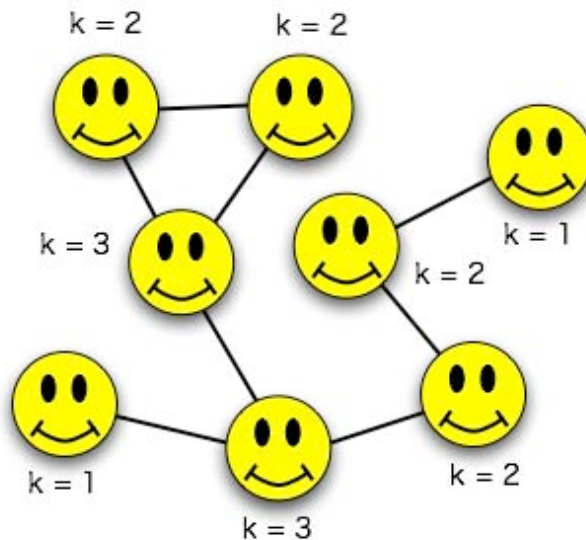
クラスタリング係数



- 自分の友達同士が友達である割合 (三角形のできやすさ)
- $c = \text{友達間のリンク数} / \text{友達間に存在しうるリンク数}$
 - $c = 0$; 友達なし
 - $c = 1$; 全員が密に結合

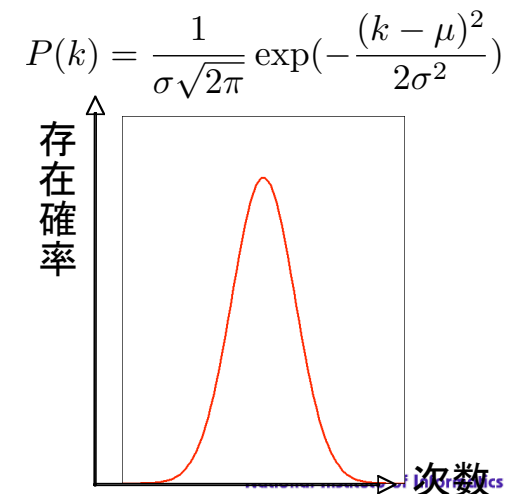
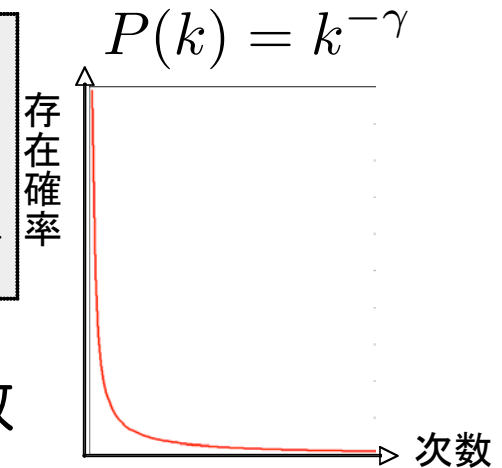
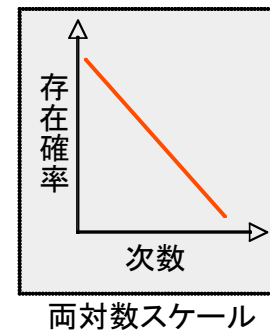
次数

- ノードの持つリンク数
 - 「入」と「出」を分ける場合もあり
- 次数の分布に着目



スケールフリーネットワーク

- スケールフリーネットワーク
 - 次数分布が**ベキ分布**で表現
 - 大多数は小さい次数，少数の大きな次数
 - 平均的なサイズなし
- $\Leftrightarrow$ ランダムネットワーク
 - 次数分布が**正規分布**で表現
 - 平均値の周りに分布



様々なデータベース

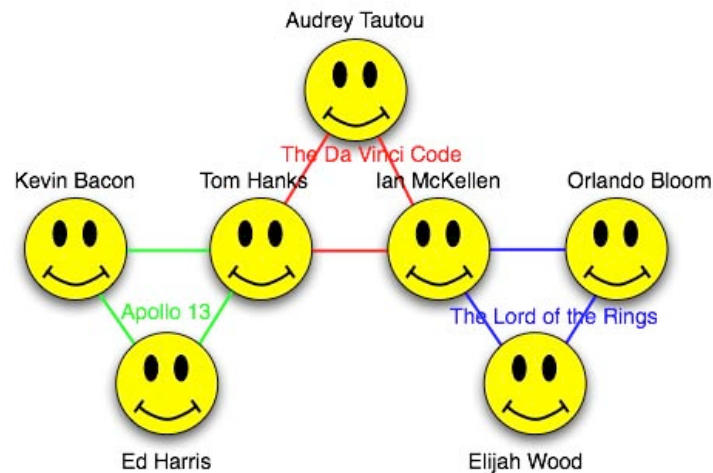
- 映画の競演関係
- 論文の共著者関係
- インターネットの接続
- バイオ関係
- 特許関係
- 株式の持ち合い

Internet Movie Database

- <http://www.imdb.com>
- 役者数: 875711

映画俳優の競演の輪

- Ex) Kevin BaconからOrlando Bloomまでの友達の輪
 - 「アポロ13」 Kevin Bacon, Tom Hanksが競演
 - 「ダヴィンチコード」 Tom Hanks, Ian McKellenが競演
 - 「ロードオブザリング」 Ian McKellenとOrlando Bloomが競演
- 友達の輪は3



競演ネットワーク

- ノード数: 225226, リンク数: 13738786
- 平均距離: 3.65
- クラスタリング係数: 0.79
- ベキ指数: 2.3

規則的なグラフ

平均距離: 1817

クラスタリング係数: 0.999

ランダムグラフ

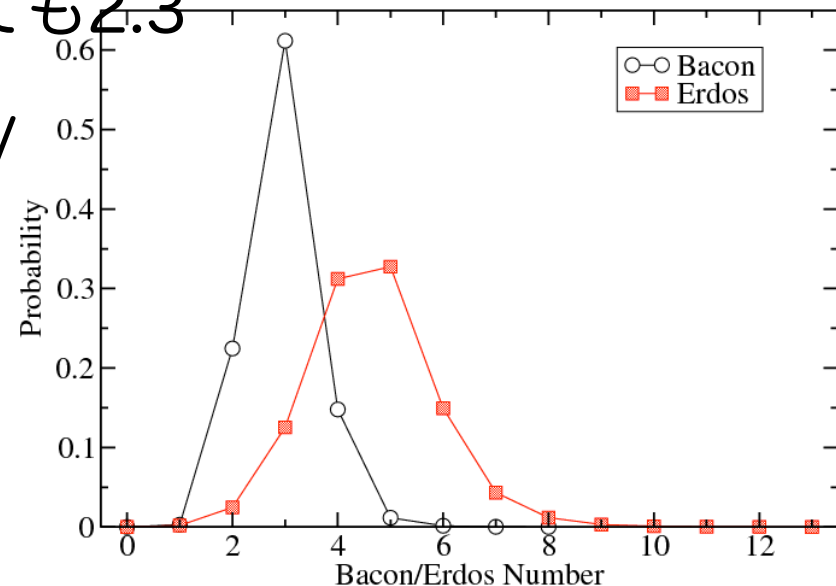
平均距離: 2.99

クラスタリング係数: 0.00027

D.Watts, S.Strogatz,
Collective dynamics of 'small-world' networks,
Nature, vol.393, pp.440-442, 1998

ケビンベーコンゲーム

- Kevin Baconからの距離(Bacon数)を計算
- Tom Hanks: 1, Ian Mckellen: 2
- 平均値: 2.9
- Sean Conneryを基準としても2.3
- <http://oracleofbacon.org/>



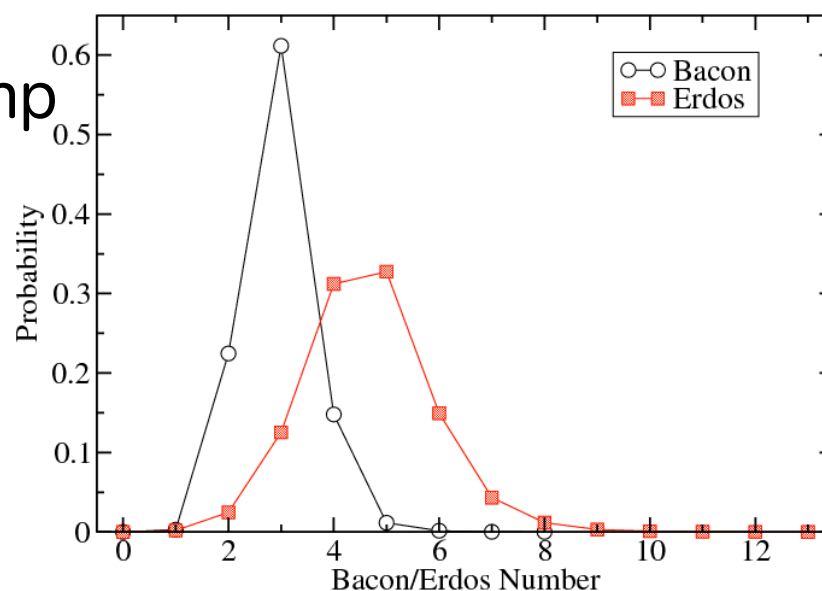
論文の共著者関係

- 「論文」を「映画」だと考える
- 190万の論文, 401000人の著者 (MR 2004)
- 平均距離: 7.64, 直径: 23
- クラスタ係数: 0.14
- ベキ指数: 2.4

<http://www.oakland.edu/enp/>
<http://www.orgnet.com/Erdos.html>

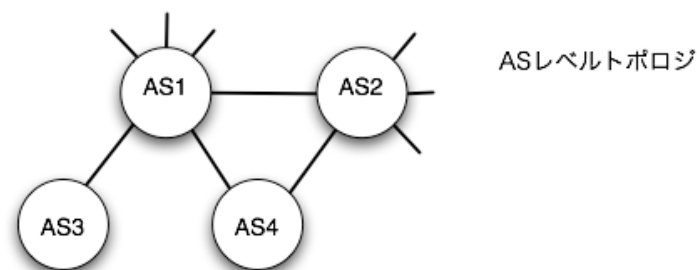
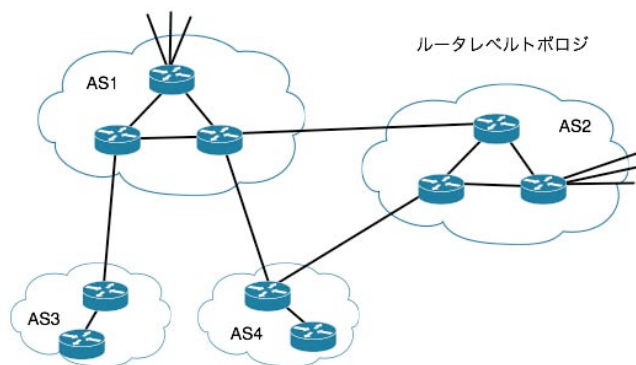
Erdos数

- Paul Erdosを0とした時の共著者の輪
- Erdos数 1: 472人
- 平均Erdos数: 5
- <http://www.oakland.edu/enp>

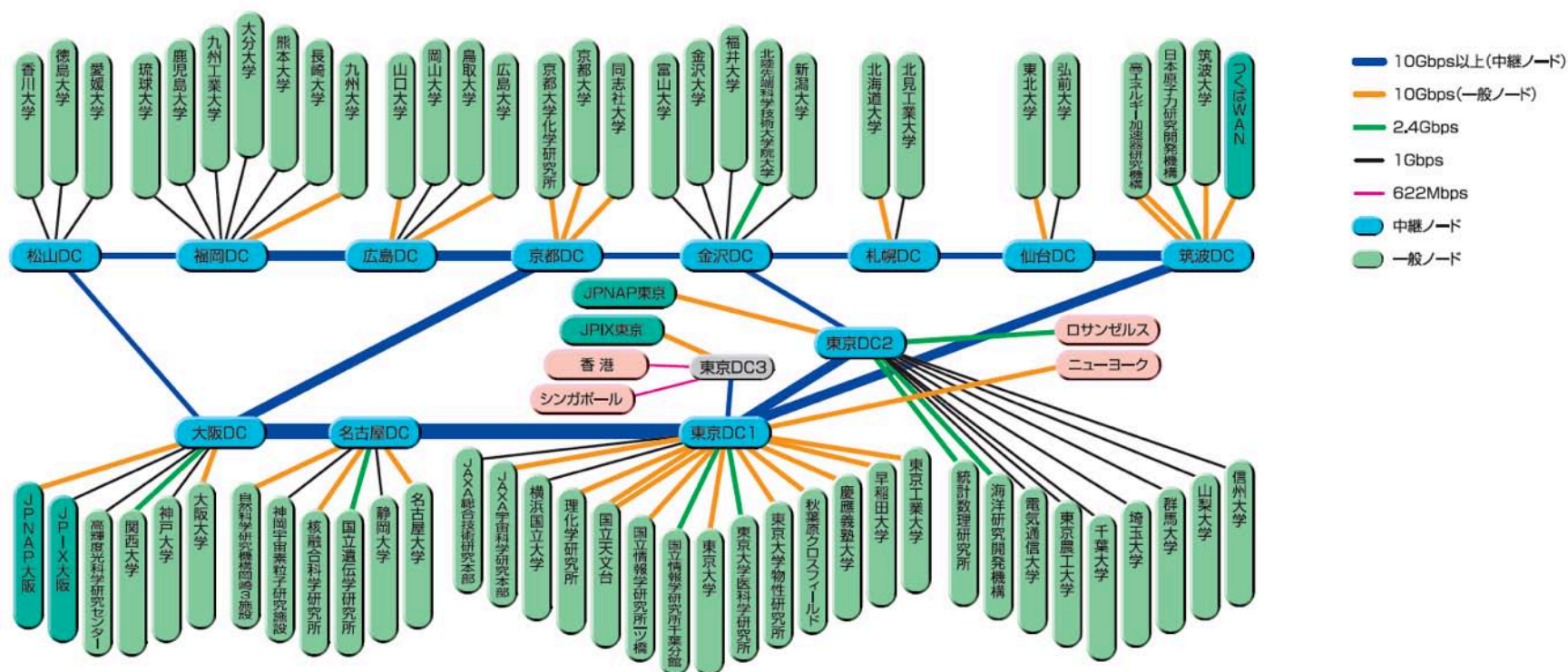


インターネット

- 人間が作った世界規模のネットワーク
- 「ネットワーク」が複数接続されたネットワークのネットワーク
- 中央制御のない自律分散システム



ネットワークの例



<http://www.sinet.ad.jp/>

インターネットトポロジ

- ASレベルトポロジ
 - ノード数: 17496, リンク数: 40805
 - 平均距離: 3.69, 直径10程度
 - クラスタリング係数: 0.02
 - ベキ指数: 2.16

M.Faloutsos, P.Faloutsos, and C.Faloutsos,
On power-law relationships of the Internet topology,
Computer Communication Review, vol.2, pp.251-262, 1999

R.Pastor-Satorras, E.Smith, and R.V.Sole,
Dynamical and correlation properties of the Internet
Phys.Rev.Lett., vol.87, 258701, 2000

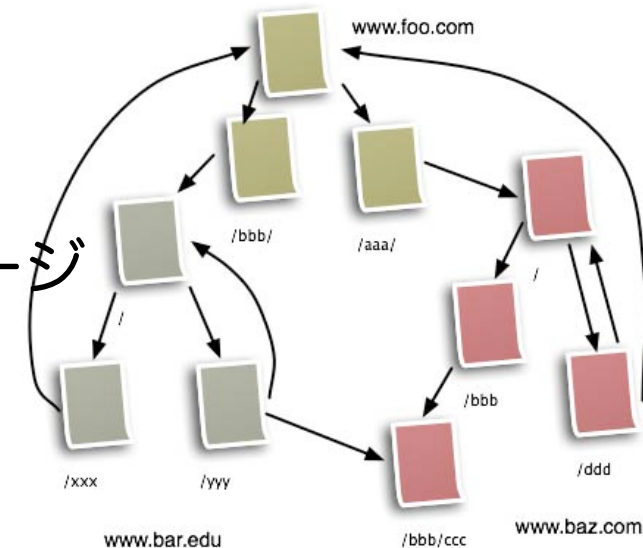
インターネットの耐性

- ランダムな故障でノードが取り除かれてもネットワークは分断されない
- スケールフリーネットワークではほとんどのノードの次数は小さい
- 80%のノードが壊れても動き続ける
- 次数の大きなノードへの意図的な攻撃には弱い

R.Albert, H.Jeong, and A.-L.Barabasi,
Error and attack tolerance of complex networks
Nature, vol.406, pp.378-382, 2000

WWWの構造

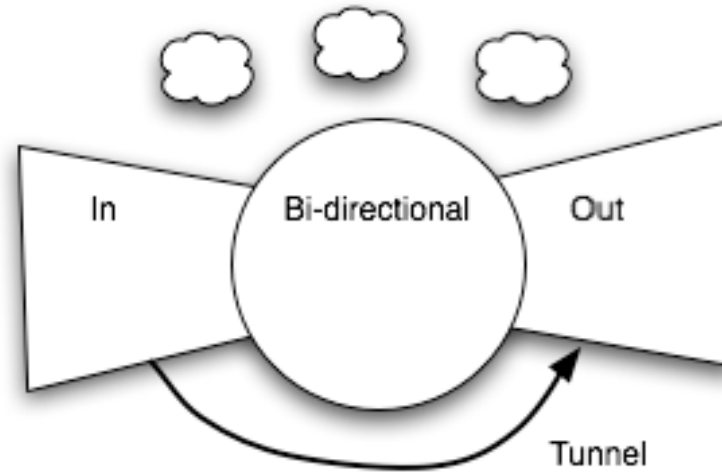
- 指数2.45(out), 2.1(in)のスケールフリー
- 1999年の推定ページ数: 8億ページ
 - 推定直径: 19ページ
- 2006年の推定ページ数: 350億ページ
 - 推定直径: 20-21ページ



R.Albert, H.Jeong, and A.-L.Barabasi,
Diameter of the world wide web,
Nature, vol.401, pp.130-131, 1999

加藤真, 山名早人,
Fact of the Web
Proc. DEWS, 2006

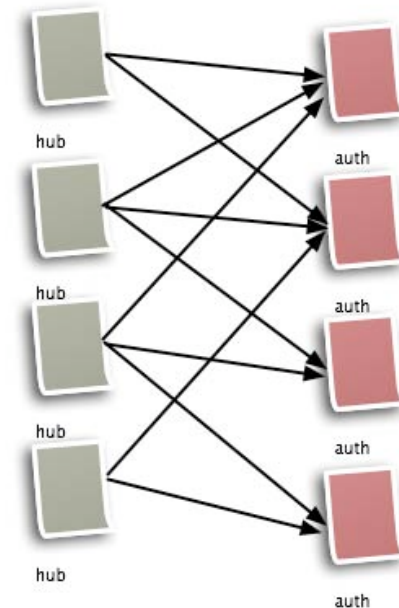
WWWの構造



- Web構造は蝶ネクタイ構造
- 密な領域: 30%, 一方向の領域: 22%ずつ

WWWの構造

- Webページの有用性の評価基準
 - ハブ: あるトピックに関してたくさんのリンクを持つページ
 - オーソリティー: 多くのページから参照される, 専門的なページ
- HITS (Hypertext Induced Topic Search) アルゴリズム



J.Kleinberg,
Authoritative sources in a hyperlinked environment,
Proc. ACM-SIAM Symp. on Discrete Algorithms, 1998

代謝ネットワーク

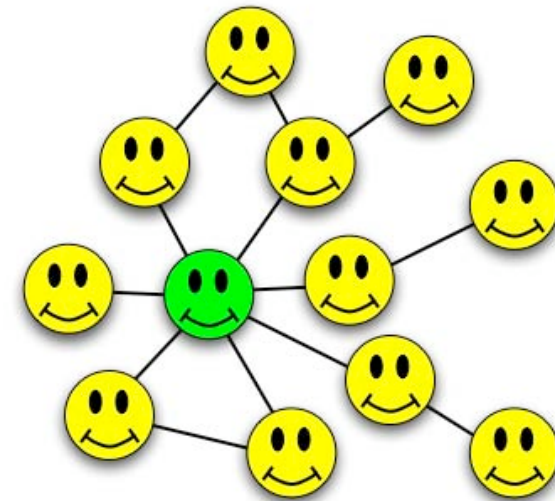
- 頂点: 化合物, 辺: 代謝反応
 - 例) アミノ酸の合成・分解
- ノード数: 275, リンク数: 1309
- ベキ指数: 2.2
- クラスタリング係数: 0.44
- 平均距離: 3.9

H.Jeong, B.Tombor, R.Albert, Z.N.Oltvai, and A.-L.Barabasi
The large-scale organization of metabolic networks,
Nature, vol.407, pp.651-654, 2000

D.Fell and A.Wagner
The small world of metabolism,
Nature Biotechnology, vol.18, pp.1121-1122, 2000

性的パートナー数の分布

- 過去1年間の性的パートナー数の分布
- 分布はベキ的減衰
 - 90%の人は3人以下
 - 20人以上の人
- 「ハブ」の存在



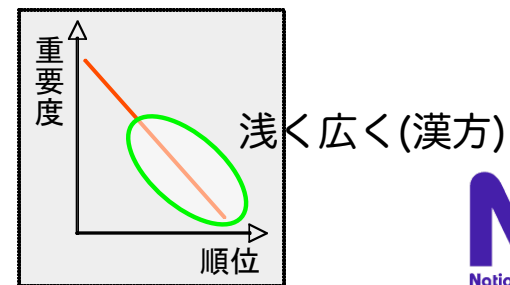
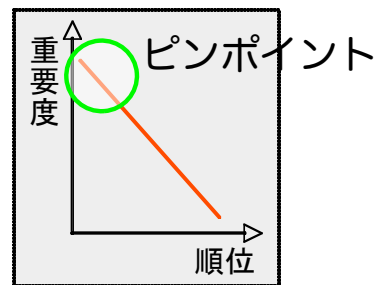
F.Liljeros, C.R.Edling, L.A.N.Amaral, H.E.Stanley, and Y.Åberg,
The web of human sexual contacts,
Nature, vol.411, pp.907-908, 2001

HIV/SARSの感染

- 次数の大きいノード(ハブ)が感染を拡散
- スケールフリーネットワークでは, ハブの数は少ない
- ハブを優先的に扱うことで拡散を抑制
 - スーパースプレッダへの治療
 - 大きな空港での検疫

創薬への期待

- 遺伝子の役割にも重みが存在
- 少数の遺伝子进行操作することで治療が可能
 - 効き目は強いが、他の機能にも副作用
- 機能の少ない複数の遺伝子进行操作
 - 効き目は弱いが、体に優しい



ネットワークの生成モデル

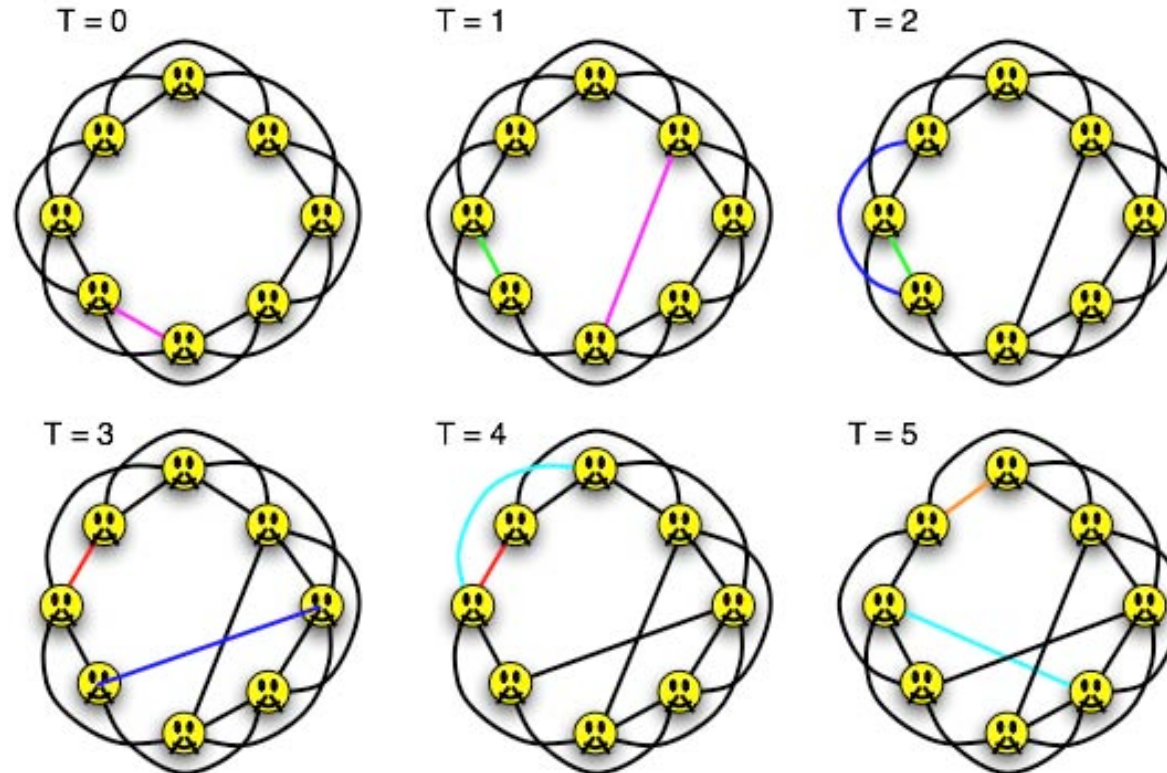
- 実ネットワークの諸性質を再現できるモデル (=アルゴリズム)が存在するのか？
- スモールワールド
 - β モデル
- スケールフリー
 - 優先的選択

β モデル

- スモールワールドを生成するモデル
- 規則的なネットワークから, リンクを確率的に張り替え
- 張り替え確率 p
 - $p = 0$: ラティスネットワーク(規則的)
 - $p = 1$: ランダムネットワーク(ランダム)

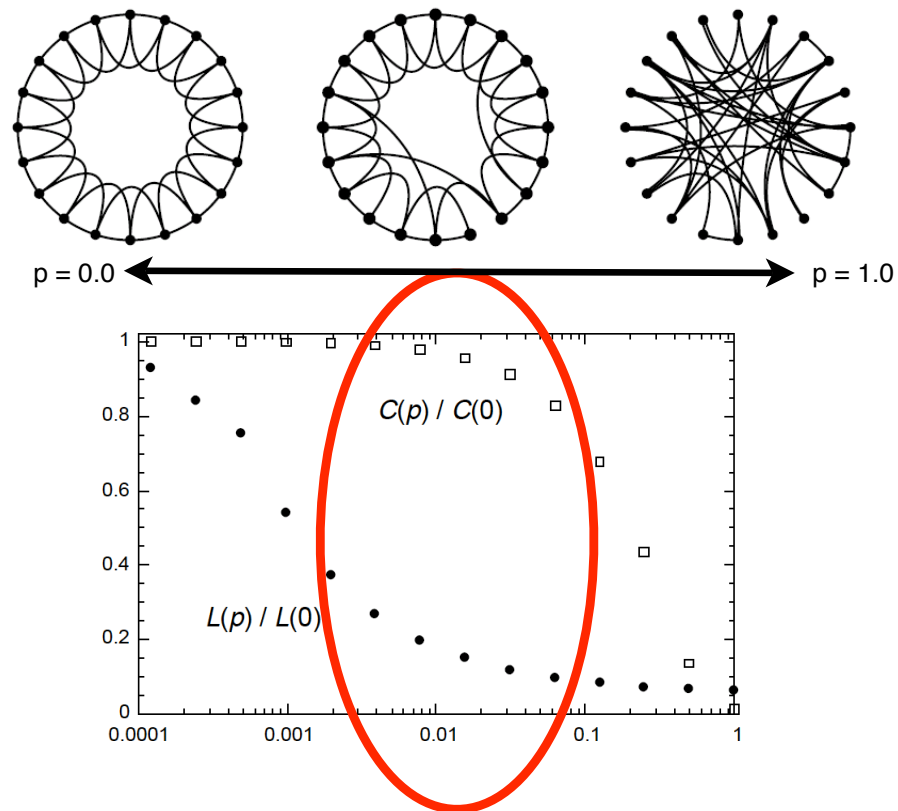
D.J.Watts, and S.H.Strogatz,
Collective dynamics of 'small-world network',
Nature, vol.393, pp.440-442, 1998

リンクの張り替え



- ノードは, リンクごとにサイコロを1回振って1,2,3
だったら張り替え ($p=0.5$)
- マゼンダ, 青, シアンで張り替え

β モデル



- 直径が小さくクラスタ化された領域が存在
→ スモールワールドネットワーク

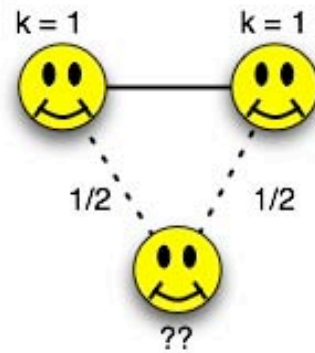
優先的選択

- スケールフリーを生み出すモデル
- rich-get-richer
 - 新規ノードは、次数の大きなノードに高い確率で接続
 - たまには次数の小さいものにも接続
- 優先的選択 $\Leftrightarrow$ ランダム選択

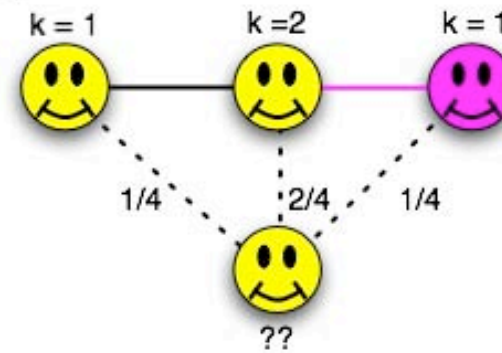
A.-L.Barabasi, and R.Albert,
Emergence of scaling in random networks,
Science, vol.286, pp.509-512, 1999

優先的選擇

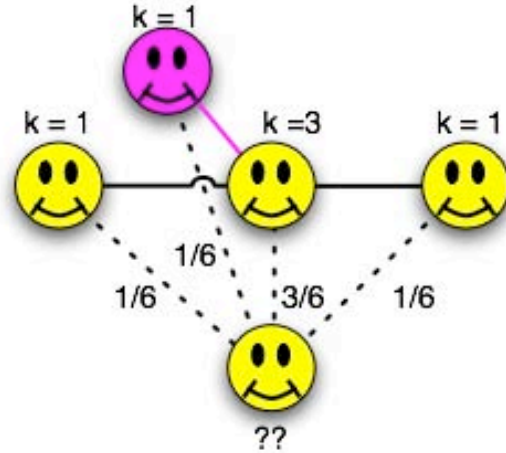
T = 0



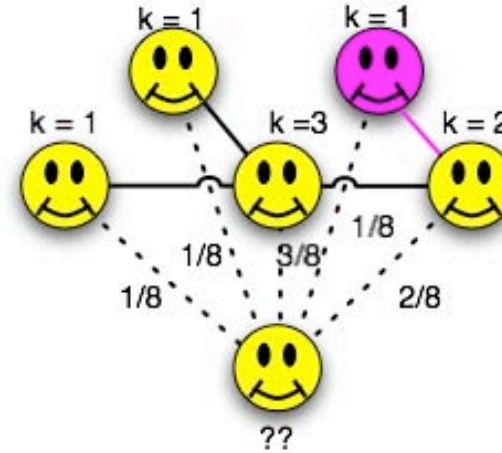
T = 1



T = 2



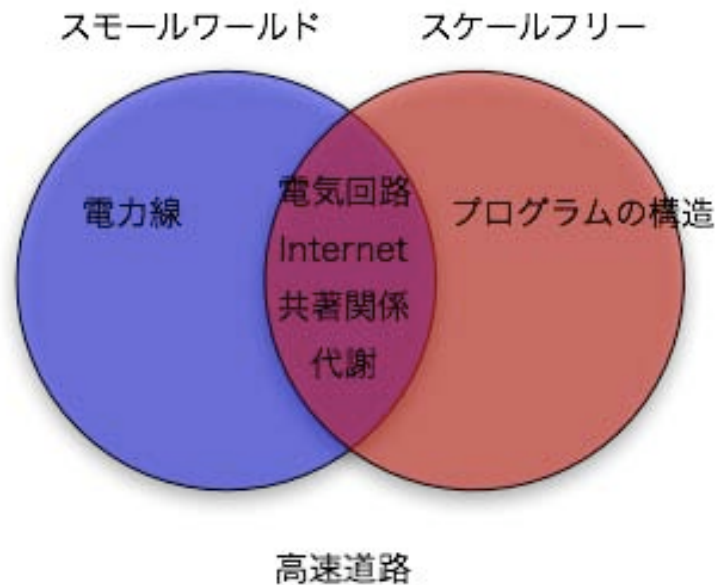
T = 3



次数相関

- 友達の多い人は友達の多い人と結びつくのか？
- 正の相関: 多い人同士がつながる
 - 人間関係
- 負の相関: 多い人は少ない人とつながる
 - インターネット, web
 - 代謝ネットワーク, 神経網

スケールフリーとスモールワールド



- 必ずしもどちらも成立するわけではない
- どちらにも属さないものも

まとめ

- 世の中には様々なネットワークが存在
- 特徴的な性質
 - スモールワールド
 - スケールフリー
- 発見→モデル→応用

参考文献

- 「新ネットワーク思考」 アルバート＝ラズロ バラバシ(著), 青木(訳), NHK出版, 2002, ISBN: 4140807431
- 「スモールワールド・ネットワーク」 ダンカン ワッツ(著), 辻, 友近(訳), 阪急コミュニケーションズ, 2004, ISBN: 4484041162
- 「複雑な世界, 単純な法則」 マーク ブキャナン(著), 阪本(訳), 草思社, 2005, ISBN: 4794213859
- 「複雑ネットワークの科学」 増田直紀, 今野紀雄(著), 産業図書, 2005, ISBN: 4782851517
- 「スモールワールド」 ダンカン ワッツ(著), 栗原, 佐藤, 福田(訳), 東京電機大学出版局, 2006, ISBN: 4501540702
- 「数理科学」 リレー連載: ネットワーク科学最前線. 2006.10-, サイエンス社