

種々のスケールフリーネットワークの位相不変量を用いた解析

02232014 小田部研究室 大沢 慶祐

1. はじめに 1999年にBarabásiらはWebサイトのリンクを調査し、ほとんどのノードはリンクを少ししか持たないが、ごく少数の限られたノードは膨大なリンクを持っているということを見出し、このようなネットワークをスケールフリーネットワーク(SFN)と名付けた。SFNは、頻度 P とリンク数 k の関係がべき乗則 $P \propto k^{-\gamma}$ に従うという特徴を持っている¹⁾。このSFNは様々なネットワークに当てはまるものであると言われている。今回は、その中でも、WWWのネットワークと、二つの生体ネットワークであるタンパク質-タンパク質相互作用(PPI)ネットワークと代謝ネットワークを調査し、特徴付けや解析を行った。例えば生体ネットワークでは、特徴付けができれば、まだ検出されていないネットワークの推測をして新しいネットワークを発見したり、耐攻撃性を調べたりすることができ、薬学、医学への貢献も期待できる。

今回、解析の指標としたのは、平均経路長、直径、ベッチ数である。平均経路長とは任意の異なる2頂点間の最短経路長の平均、直径とはその最大値である。ベッチ数とは位相不変量の一つで、直観的な解釈としては、 β_0 は独立したネットワークの数、 β_1 はネットワーク内の独立の閉ループの数として捉えることができる。さらに、 β_1 が大きいと、耐攻撃性に優れているということがわかっている²⁾。

2. 解析 解析対象のネットワークは、九州工業大学のホームページ(<http://www.kyutech.ac.jp/>)の内部リンク、金沢星陵大学のホームページ(<http://www.seiryu-u.ac.jp/>)の内部リンク、これらのネットワークとそれぞれ頂点数を同じにしたSFNである。これら二つのサイトは頂点数がほとんど同じであり、良い比較対象となった。このSFNはBarabásiらが開発したプログラムで作成した。九州工業大学のホームページと頂点数が同じものをSFN-A、金沢星陵大学のホームページと頂点数が同じものをSFN-Bとした。金沢星陵大学のホームページは日経BPコンサルティング社が2004年に行った全国の大学サイトの使いやすさを調査した「全国大学サイトユーザビリティ調査」の1位である。「使いやすさ」と位相的な解析の関係を調べるために調査対象に加えた。さらに、KEGG(Kyoto Encyclopedia of Genes and Genomes)から入手したタンパク質-タンパク質相互作用(PPI)ネットワーク、同じくKEGGから入手した代謝ネットワークについても調査した。これらのネットワークの平均経路長、直径、 β_0 、 β_1 を調査した。

3. 結果及び検討 表1はそれぞれのネットワークの解析結果である。平均経路長、直径の違いがWWWの使いやすさに影響を与えていると考えたが、結果はほとんど同じであった。両サイトとも、多くても6回リンクをたどれば見たいページへ行ける。生体ネットワークにおいてこれらの値、特に直径は非常に大きくなった。これは、特定の組み合わせでしか反応しないということが影響を与えていると考えられる。

β_0 は、WWWがともに1で、生体ネットワークが非常に大きな値となった。これは、WWWはネット

ワークが分断されておらず、一つの大きなネットワークを形成し、生体ネットワークは特定の組み合わせでしか反応しないので、数多く分断されている。

β_1 は閉ループの数を表し、ネットワークの疎密さを表している。WWWの二つを比べて最も違いの大きかったのがこの値であり、これが使いやすさの指標に影響を与えている可能性が高い。また、生体ネットワークの値が他のネットワークと比べて大きい、これは生体ネットワークのノード数がWWWの10倍になっているからである。

今回の解析では、全てのネットワークにおいて、図1のような頻度分布となった。この分布は、スケールフリーの特性であるべき乗則に従っている部分と従っていない部分がある。従っていない部分は横軸に平行であり、これに移行するところは、頻度がノード数の逆数に近いところである。よって、この特性もスケールフリーの特性の一つと考えるのが自然である。これを利用すれば、まだ発見されていない生体ネットワークや、新たなハブの発見に役立つと期待できる。

以上のように、従来同一のスケールフリーネットワークとして捉えられていた様々なネットワークをベッチ数で解析することにより、よりネットワークの特徴を理解することができることがわかった。

表1. 解析結果

network	ノード数	平均経路長	直径	β_0	β_1
九工大	361	3.30	6	1	46
SFN-A	361	2.27	4	1	8
金沢星陵	351	3.47	6	1	159
SFN-B	351	2.30	5	1	6
PPI	3310	4.93	15	188	1174
代謝	3104	11.56	46	55	692

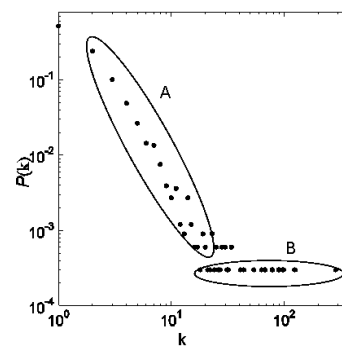


図1: タンパク質-タンパク質相互作用(PPI)ネットワークのリンクの分布。A: べき乗則に従うところ、B: べき乗則に従わないところ。

【参考文献】

- 1) R. Albert, H. Jeong, A-L. Barabási: Nature 401 (1999) 130.
- 2) 廣田裕二: 位相不変量によるネットワークの解析 [九州工業大学卒業論文 平成17年]