

mixi 内のネットワークを対象としたスケールフリー性の検証

Inspection of scale-free for a network in mixi

<http://www.tani.cs.chs.nihon-u.ac.jp/g-2007/kaname/>

谷 研究室 安部 要

Kaname Abe

概要

近年、インターネットはスケールフリー構造であることが発見された。他にも現実世界の様々なネットワークがスケールフリーであることが発見されているが、本研究では、急激に発展した SNS のひとつである mixi のユーザーをノード、マイミクシィをリンクとして構成されるネットワークはスケールフリーであるか検証する。

1 はじめに

長年にわたりインターネットのような複雑なネットワークは、ランダムネットワークであると考えられてきた。ランダムネットワークでは、リンクがランダムに存在するならば、各ノードにはほぼ同数のリンクを持つとされている。しかし、近年スケールフリーといわれるネットワーク構造が発見され、現実世界で数多くのスケールフリー・ネットワークの実例が見つかっている。そこで、本研究では日本で最も発展している SNS である mixi 内の友人同士の繋がり（マイミクシィ）をリンクとするネットワークはスケールフリーであるかを検証する。

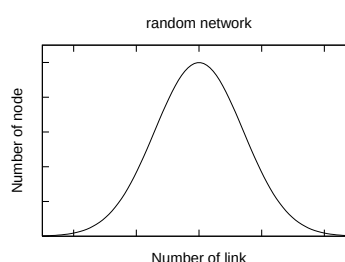
2 章ではスケールフリー・ネットワークについて、3 章では mixi について、4 章では検証方法について、5 章では検証結果・考察を述べてまとめとする。

2 スケールフリー・ネットワークについて

今回は代表的なネットワーク形態のひとつであるランダムネットワークと比較してスケールフリー・ネットワークを説明する。

2.1 ランダムネットワーク

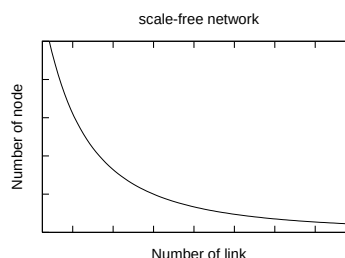
ランダムネットワークとは、リンクがランダムに貼られるネットワーク。このネットワークは、大半のノードはほぼ同数のリンクをもち、非常に多くのリンクを持つノードは少ない。ノードの次数分布で表すと、下のようになり正規分布になる。



2.2 スケールフリー・ネットワーク

スケールフリー・ネットワークとは、アルバート＝ラズロ・バラバシによって発見されたネットワーク。一部のノードが膨大なリンクをもつ一方で、ほとんどはごくわずかなノードとしか繋がっていないようなネットワーク構造である。膨大なリンクを持つノードをハブという。

このネットワークの最大の特徴は、新しいノードが次々に参入しても、ネットワークの形状が変化しないフラクタル性をもっているところにある。ノードの次数分布で表すと下図のようにべき法則になるのも特徴的である。



2.3 定義

あるネットワークの頂点の次数分布 $P(k)$ がある定数 γ によるべき法則分布に従うとき、このネットワークはスケールフリーであるという。

式で書くと

$$P(k) \propto k^{-\gamma}$$

が成り立つとき、そのスケールフリーであるという。

3 mixi について

3.1 SNS(Social Networking Service) とは

SNS とは、人と人とのつながりを促進・サポートする、コミュニティ型の web サイトである。趣味、居住地、出身校、「友人の友人」といったつながりを通して新たな人間関係を構築する場を提供する会員制のサービスである。数多くの SNS が存在するが、本研究では日本で最も広まっている「mixi」を研究対象にする。

3.2 mixi とは

日本では最も早い時期からサービスを開始していた SNS のひとつで、すでに入会している登録ユーザーから招待を受けないと利用登録ができない、完全招待制を採用している。利用者数は 2006 年 10 月現在で 520 万人。平均利用時間は日本ドメインでは yahooJapan に次ぐ 2 位で日本で社会現象となった。mixi では基本的には、日記を買いたり、他のユーザーの日記にコメントを書いたりしながらユーザー同士で交流している。

3.3 マイミクシィ

ユーザーごとに持っているリストの名前、またそのリストに登録された他のユーザー。mixi 事務局が想定しているのは、すでに友人や知人の関係にあるユーザー同士をリンクする使用方法である。

4 検証方法

4.1 対象のネットワーク

今回の検証では、「mixi」内でのユーザーをノード、ユーザー同士のマイミクシィのつながりをリンクとしてできるネットワークを対象とする。

4.2 検証方法

検証方法は WWW::Mixi という PERL の API があるので、これを使って Mixi 内のデータを取得していく。今回は安部のユーザーからマイミクシィを順々に取得し

ていく幅優先探索でデータを取得していくことにする。

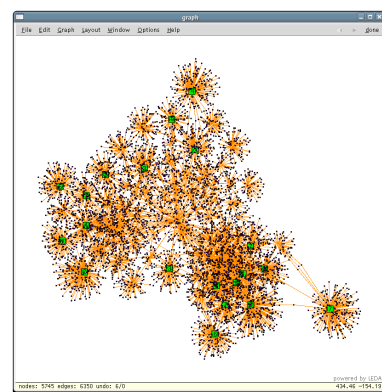
1. WWW::Mixi を使い、各ユーザー（ノード）のマイミクシィ数（リンク数）をデータベースに蓄積するプログラム (parse_my_mixi_data.pl) を作成し、データを取得する。Mixi の全ユーザーは 2006 年 10 月現在 520 万人と言われているが、今回は約 40 万のユーザーの情報を取得する。
2. 収集したデータから、前処理を行い、各リンク数ごとのノード数を集計する。
3. そのデータを元に分布図を作成し、得られた図がべき法則分布であれば、スケールフリー・ネットワークであるといえる。

という方法で検証する。

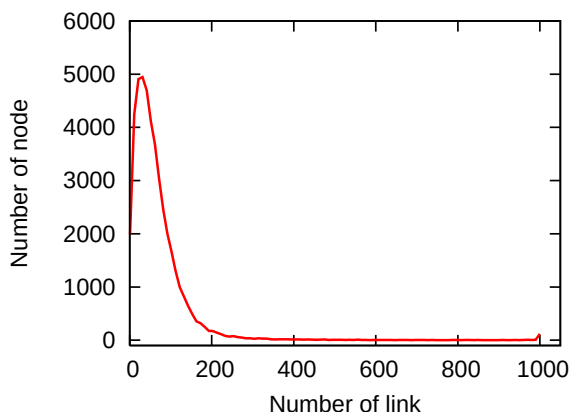
5 検証結果・考察

5.1 検証結果

結果としては、次のようなネットワークが得られた。取得した全てのデータでネットワークを作ると大きくなり過ぎてしまうので、データ数約 4000 でネットワークを作成した。

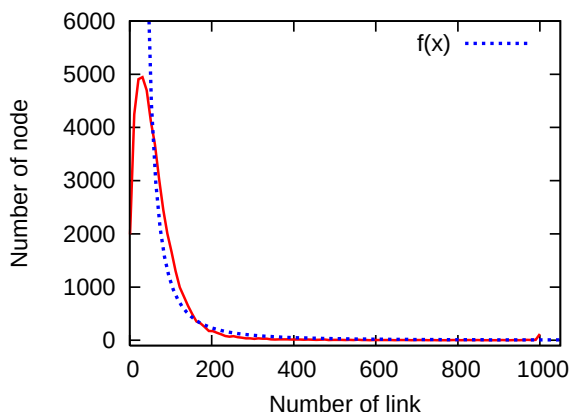


緑のノードはマイミクシィが 100 以上、紫のノードはマイミクシィ 20 以下、オレンジはリンクを表した。またさきほどのネットワークを度数分布図で表したのが下図である。



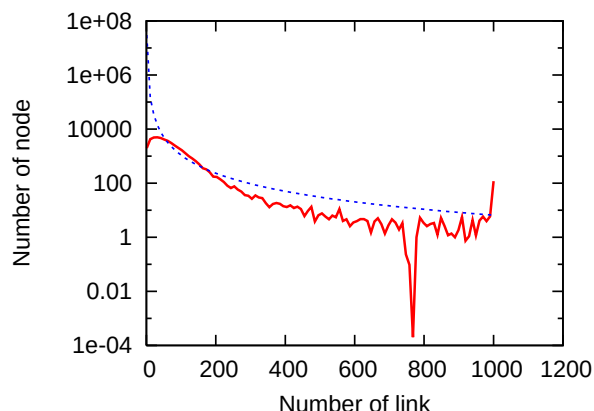
この図はリンク数が 1000 のところで切れているが、これはマイクシィ登録数が 1000 までと制限されているからである。図を見ると、小数のリンクを持つノードが大多数あり、数多くのリンクを持ついくつかのノードが混在している。

5.2 考察



この図では、べき乗分布を青の点線で表している。取得したデータからなる赤線と、べき乗分布の青線はほぼ近

似しているといえる。



また Y 軸を対数で表示した図が上の図になる。この図では取得データのグラフが一時的に下がっている箇所があり、取得データに偏りがあったのかもしれないが、ほぼべき乗分布と近似しているので、このネットワークはスケールフリー・ネットワークであるといえる。

6 今後の課題

- ・対数表示した図で一時的に下降しているのはなぜかを調べる
- ・新しいネットワークの対象としてコミュニティをノード、そのコミュニティに入会ユーザーが入会している他のコミュニティとリンクで繋いだネットワークはどのようなネットワークかを検証する

参考文献

- [1] アルバート＝ラズロ・バラバシ (著) 青木薫 (翻訳) 新ネットワーク思考－世界のしくみを読み解く