

遺伝的アルゴリズムを用いた複雑ネットワークの生成

佐藤 浩, 久保正男, 生天目章

防衛大学校電気情報学群情報工学科

和文抄録

複雑ネットワークの研究においてはその構造に着目することが重要である。ネットワークの構造を掴むために、さまざまな指標が用いられるが、単一の指標は性質の一面しか表現できないため、同一の指標を持つネットワークであっても、構造がまったく異なる場合がある。本研究では、指標と構造の関係を明らかにすることを目的とする。与えられた指標がどれだけ達成されたかを目的関数とした遺伝的アルゴリズムを用い、最適化を通じたネットワークの生成を行う。実験により、問題空間には非対称性がみられ、最適化の指標により解空間の構造が大きく変わること、また、同じ指標を持ちながら異なる構造のネットワークが見出されることを明らかにした。

GA-based Complex Network Generating System and its Characteristics

Hiroshi Sato, Masao Kubo, Akira Namatame

Dept. of Computer Science,
National Defence Academy in Japan

Abstract

In the research of complex networks, it is important to focus on the structure of the networks. Various measures such as centrality or radius are used to analyze the networks. One measure, however, can only capture the one aspect, so two networks that have same measure sometimes show completely different structure. The purpose of this study is to figure out the relation between measurements and structures. Genetic Algorithm is used to find networks that have specific measurement value. We show that GA can find the networks that are different in view of structure, but the same in view of measurement value.

1 はじめに

複雑ネットワークの研究が盛んに行われるようになってつれ、人間社会や神経細胞のつながりといった、一見何の関係もないような実際のネットワークにおいて、スケールフリー性やスモールワールド性という共通した性質があることが分かってきた [Albert 02][Watts 98]。

ただし、一口にスケールフリーやスモールワールドといっても、ある指標で見れば同じような値を持つネッ

トワークであっても、他の指標で見た場合にまったく異なる構造に見える場合がある。たとえば映画俳優の共演関係といった社会的なネットワークとソフトウェアパッケージの依存関係図といったような人工的なネットワークの双方ともベキ分布に従うが、それらの次数相関係数は前者が正であり後者が負である、というまったく反対の結果を得ることが知られている [Newman 02]。

また、同じネットワークも視点によって構造が変わっ

てしまう例もある。たとえば、複雑ネットワークの中でもインターネットは人間が作り出した最も大規模なネットワークの一つであるが、対象が巨大過ぎて直接構造を明らかにすることは非常に難しい。これまで数々の研究が報告されているが、どのレベルでネットワークを見るかによって、得られるネットワークの構造が異なることが分かっている [Siganos 03][Sprint 04][Li 04]。

これらはネットワークの構造とその特徴量の非可逆性のために生じる問題である。ある一つの構造を定めたとき、それを測る指標の値はただ一つであるが、ある指標の値を定めたときに取り得る構造というのは複数存在することもあるし、存在さえしないこともある。

このインターネットのように大規模なネットワークのある一面をシミュレーションしようという場合には、ある特徴量の組を持ったネットワークを自由に生成できることが必要である。本研究では、特徴量からネットワーク構造を生成するシステムの構築とそのシステムが持つ性質を明らかにすることを目的とする。

2 遺伝的アルゴリズムによる複雑ネットワーク生成モデル

ネットワークの特徴量からその構造を生成するシステムとして、佐藤、花田らによるモデルがある [佐藤 05][花田 07][廣安 04]。このシステムはネットワークの生成を遺伝的アルゴリズム (Genetic Algorithm: GA) による最適化として実現している。

本研究で作成したシステムは、特に佐藤によるものに基づいたものである [佐藤 05]。モデルの詳細を以下に示す。

対象とするネットワーク 本研究において扱うネットワークは、簡単のため無向な連結グラフであり、自身へのリンクは持たないとする。ノード数は固定であり、エッジを張り替えることで新しいネットワークを生成する。GA の初期世代として、ある存在確率でエッジを生成したものをを用いる。

コード化 本モデルにおいては、ネットワークの接続行列を基本的な遺伝子表現として用いる。より具体的には、接続行列の上三角成分を直線状に並べたものを

染色体とする。このため、行列は対称であり、要素は 0 または 1 の値のみを取る。また、対角成分は 0 となる。

交叉 交叉は一点交叉を用いる。上記コード化により、接続行列を直線状に並べたものであるため、バイナリ GA における交叉はどんなものでも使えるという利点を持つ。

突然変異 ある確率で、接続行列の一部を反転させる。本論文で扱うネットワークは無向グラフであるため、ある成分を反転させるときは対称成分も同時に反転させるものとする。

修正操作 上記の交叉、突然変異では孤立点が生じてしまう可能性がある。ここでは修正操作として、新たに生成されたネットワークに孤立点が発見されたら、自身以外のある一点へリンクを張るという操作を考える。

適応度 ネットワークがどの程度特徴量を達成したかを適応度とし、最適化を行うことで所望のネットワークを探索する。

世代交代モデル 本研究においては、世代交代法として MGG [佐藤 97] と呼ばれるモデルに基づいたモデルを用いる。ただし、生存選択においては、最良の 2 個体を選択するものとする。

3 実験

前章で紹介したモデルにより、ネットワークの生成実験を行った。実験に用いたネットワークおよび GA のパラメータは以下のようなものである：

- ノード数：20
- 初期集団：存在確率=0.3 でエッジを生成したネットワーク
- 集団サイズ：10
- 生成個体数：1000
- 目的関数：中心性、グラフの直径、エッジ数の最大化/最小化の計 6 通り

上記の6通りの目的関数について実験を行った結果をエッジ数に基づいて整理すると以下の2つに大別できる：

- エッジ数が増大するもの（完全グラフ型）：中心性最小化、直径最小化、エッジ数最大化
- エッジ数が減少するもの→
 - －（直線型）：中心性最大、直径最大
 - －（スター型）：エッジ数最小

各々の場合について得られたネットワークを図1, 2に示す。

4 考察

4.1 ネットワークの構造と評価関数について

ネットワークのエッジ数で整理した図1, 2より目的関数による生成ネットワークの違いを見てみる。

グラフのエッジ数が増大する側への変化においては目的関数の変化による違いは見られない。これは次数がある程度大きくなると完全グラフに近づくためであり、この方向に対するGAのランドスケープは単調であるといえる。

一方で、エッジ数が減少する場合には、スター型と直線型といった明らかな構造の違いが見られる。この2つは同一のエッジ数であるにも関わらず他の指標においては大きな違いを持つ例である。よって中心性あるいは直径を対象とした場合には、スター型は適応度が低くなるため現れることはない。

興味深いのは、エッジ数最小化を目的関数とした場合では、両者に適応度の違いはないにも関わらず、その出現頻度には大きな差がある点である。

これは、各々の最適解を構成する部分解が共通する部分を持つかどうかによる違いによると思われる。直線型は構成要素が入れ替わってもエッジを共用できるが、スター型の場合、中心となるノードが入れ替わってしまうと、それまでの部分解が新しい部分解としてまったく使うことができない。

エッジ数が減少する側には2つの山があるが、その到達しやすさには大きな違いがあるため、その出現率に大きな違いが生じてしまうのである。これは二重円

巡回セールスマン問題におけるO型とC型の解空間の構造 [山村 92] と類似したものである。



図1: 次数が増大する変化：(a) 初期世代の最良解、(b) 最終世代の最良解

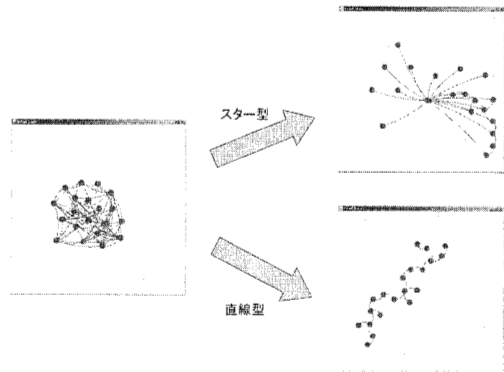


図2: 次数が減少する変化：(a) 初期世代の最良解、(b) 最終世代の最良解

4.2 修正操作の効果について

本研究では生成するネットワークは、連結グラフであることを条件としているが、接続行列を一点交差させるという方法では、グラフの連結が切れてしまう場合がある。これを防ぐために、修正操作を施しているが、それが良い方向に働く場合と悪い方向に働く場合があることが分かった。

まず、結果としてエッジ数が増大する場合のネットワーク生成においては、ある点が孤立してしまうような状況はまず起こらないため、修正操作のありなしで結果にほとんど影響しない。

一方で、結果としてエッジ数が減少する場合のネットワーク生成においては、修正操作が必要となること

が多い。ここで、目的関数をどのようにとるかで、修正操作が結果にまったく正反対の影響を与えることが分かった。結果を以下に示す：

- 修正操作が正の影響を与える場合：エッジ数の最小化問題
- 修正操作が負の影響を与える場合：中心性最大化問題、グラフの直径を最大化問題

これも前節同様、得られるグラフの形状により引き起こされる違いである。スター型のグラフは1つのリンクが切れてしまうだけで、孤立点が生じてしまうため修正操作がないとすぐに実行不可能解となってしまうが、線状のグラフは1点に平均2つのリンクがあるため、双方が切れるまでは孤立した点にならない。このため、修正操作に頼らずとも探索を続けることが可能である。

5 おわりに

本研究では、指標からネットワーク構造を発見することを目的としたシステムを作成し、その性質を調べた。作成したシステムは、佐藤、花田らによるシステム [佐藤 05] [花田 07] [廣安 04] に基づいたもので、ネットワークの接続行列を遺伝子表現とし、所与の指標値からの近さを適応度とした、遺伝的アルゴリズムによる最適化システムである。

本システムではコード化として接続行列を、交叉として接続行列の三角成分の一点交叉を採用した。本コード化／交叉により規定された問題空間は、ネットワークの次数が増大する方向には単峰であるが、減少する方向においては多峰性のあるランドスケープであることが分かった。また、どのような指標を最適化のために用いるかにより、実現される峰が異なることが分かった。

今後の課題としては、コード化／交叉の工夫による問題空間の平滑化や、最適化の際の指標のバランスによる最適化経路の制御などがあげられる。

参考文献

[Albert 02] Albert, R., and Barabási, A.-L., Statistical mechanics of complex networks, *Reviews of Modern Physics* 74, pp. 47-97 (2002)

[Li 04] L. Li, D. Alderson, W. Willinger, and J. Doyle, A first principles approach to understanding the Internet's router level topology, *ACM SIGCOMM Computer Communication Review*, vol. 34, pp. 3-14 (2004)

[Newman 02] M. E. J. Newman, Assortative mixing in networks, *Phys. Rev. Lett.* 89, 208701 (2002)

[Siganos 03] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, Power laws and the AS level Internet topology, *IEEE/ACM Transactions on Networking*, vol. 11, pp. 514-524 (2003)

[Sprint 04] N. Sprint, R. Mahajan, D. Wetherall, and T. Anderson, Measuring ISP topologies with rocketfuel, *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2-16 (2004)

[Watts 98] Watts, D.J., and Strogatz, S.H., Collective dynamics of small-world networks, *Nature* 393, pp. 440-442 (1998)

[佐藤 05] 佐藤史隆, 廣安知之, 三木光範, 鈴木泰博, 遺伝的アルゴリズムによるスケールフリーネットワークの設計, FCS/テクノシンポ名大/MPS シンポジウム 2005 講演資料 (2005)

[佐藤 97] 佐藤、小野、小林、遺伝的アルゴリズムにおける世代交代モデルの提案と評価、*人工知能学会誌*, Vol. 12, No. 5, pp. 734-744 (1997)

[花田 07] 花田良子, 佐藤史隆, 廣安知之, 三木光範, 鈴木泰博, 遺伝的アルゴリズムによるネットワーク特性量に着目したネットワーク設計法, *コンピュータソフトウェア* Vol. 24, No. 1, 91-100 (2007)

[廣安 04] 廣安知之, 三木光範, 佐藤史隆, 鈴木泰博, 遺伝的アルゴリズムによる複雑ネットワークの解法, 第11回 MPS シンポジウム講演資料 (2004)

[山村 92] 山村、小野、小林、形質の遺伝を重視した遺伝的アルゴリズムに基づく巡回セールスマン問題の解法、*人工知能学会誌*, Vol. 7, No. 6, pp. 117-127 (1992)