

遺伝的アルゴリズムによる最適ネットワークの生成

申 昇燁 佐藤 浩 生天目 章

防衛大学校情報工学科

Abstract: 現実にする多くの複雑ネットワークは、べき分布に従うといわれている。べき分布を持つネットワークを生成するために、いろいろなモデルが提案されてきた。だが、次数分布が同じであっても、生成されたネットワークは、多くの点で特性が異なることが多い。本研究では、遺伝的アルゴリズム(GA)を用いて、与えられた目的関数を最適化するネットワークを生成する方法を提案する。そして、進化的に生成されたネットワークの特性を分析する。

最適化のための指標として、ノード中心性に基づく混雑度とリンク密度を用いる。つまり、ネットワークの混雑とコスト制約の最小化を目的として、最適なネットワークを生成する。そのようなネットワークは、ノード中心性の最大値を下げて、トラヒックの混雑上のボトルネックをなくすという性質を持つことを示す。特に、ネットワークの混雑度回避を重視した場合、中心に円型のコアノードと階層構造を持つ最適ネットワークが生成されることを示す。

Evolutionary Design of Optimal Networks

Seung-Youp Shin Hiroshi Sato Akira Namatame

Dept. of Computer Science,
National Defense Academy of Japan

Abstract: Networks in real world often show the power-law distributions. A variety of models that can produce the power-law have been proposed, but they are different in many respects other than power-law. Recent study shows that the growth of network is interpreted as a result of optimization of various measures. In this paper, we use genetic algorithm (GA) for tool of optimization and investigate the properties of evolve networks. We designed networks that minimize congestion and constructional cost. From the simulation, we show that the optimized network has a circle of core nodes and hierarchical structure. It has a property that lose the traffic bottleneck because can decrease greatest betweenness.

Keywords: Betweenness, Optimal network, Backbone, Power-law, Genetic Algorithm.

1. はじめ

近年の研究により、インターネットの構造は、スケールフリー構造を持つことが知られている[2, 6]。スケールフリーネットワークは、各ノードが、ランダムにまたは均等につながっているのではなく、ハブノードという非常に多くつながっている極少数のノードが存在する。これらの少数のハブノードは、ネットワークの全体の性質を大きく左右している。この性質は、ネットワークの規模に関係なく、非常に多いリンクを持つノードと残りのネットワークの比率によるスケールフリー性として知られている。

一方で、ランダムネットワークには、多くつながっているノード存在がなく、すべてのノードがほぼ同じ

のリンク数を持つ。そして、リンク数のほとんどが全体の平均値の付近にある。このようなランダムネットワークは、あまり混雑しないという特性を持っている。

本研究の目的は、ネットワークの混雑に耐えられ、ネットワーク設計上のコストが少ない最適なネットワーク構造を求めることである。そのために、進化的アルゴリズムを使ってネットワークを生成する方法を提案する。

ネットワークは、トラヒックの送受信するグラフとして生成する。グラフのノードは、ネットワークの送受信するノードを、リンクはネットワークのケーブルを表している。トラヒックの生成はランダムなノードで発生して、ランダムなノードに送信することを仮定する。

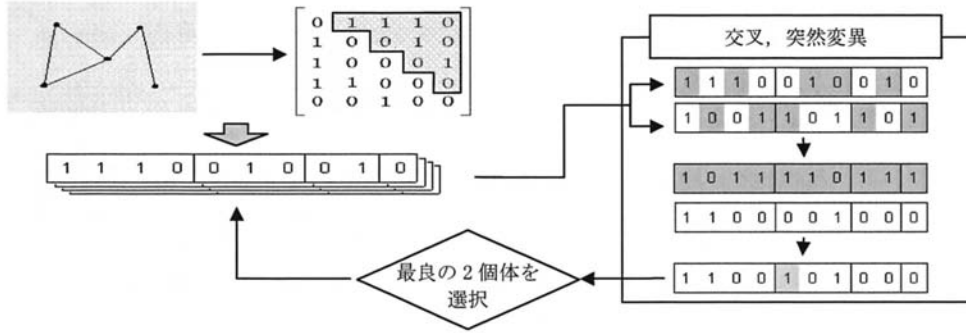


図.1. 適用する遺伝アルゴリズムの概念図.

2. 目的関数の設定

ネットワークは、一定のノード数を持つ無向グラフとして表される。ここで、隣接行列の $A=\{a_{ij}\}$ を定義する。ある2つのノード、 i と j がつながっていれば、 a_{ij} を1、つながっていなければ、0の値をとる。そして、 A は全体で $n \times n$ の要素を持つ。

2.1 リンク密度

ネットワークの多くの特性は、根底にあるネットワーク構造によって決まる[3, 5]。そして、それらの特性は隣接行列の A から把握することができる。

ここで、ネットワークの密度を以下の式で定義する。

$$\alpha = \frac{1}{n C_2} \sum_{i < j} a_{ij} \quad (1)$$

ネットワークが全て連結されていれば(全連結グラフ)、全部で $n C_2$ 本のリンクが存在する。その値を基準にして、ネットワークが持つリンク数の相対比により、リンク密度を定義する。

2.2 混雑度指数

ネットワークのトラフィックの混雑度を計るための指標として、各ノードに定義されるノード中心性 (betweenness) を用いる。ここでは、ネットワーク全体で予想されるキューの長さ $\lambda(\rho)$ を用いる[1]。それを、次式で定義する。

$$\lambda(\rho) = \sum_{i \in N} \frac{\rho \frac{\beta_i}{n-1}}{1 - \rho \frac{\beta_i}{n-1}} \quad (2)$$

(2)式において、 β_i はノード i の正規化した betweenness、また、ネットワーク全体の平均パケット生成率を ρ とする。その場合、ノード i に入ってくるパケット期待値は、 $\rho (\beta_i/n-1)$ として求まる。そして、(2)式は、ネットワーク全体のキューの長さを表すことか

ら、 $\lambda(\rho)$ を最小化することで、ネットワーク全体のトラフィックの混雑度を下げることになる。

2.3 適応関数の定義

進化アルゴリズムにより最適化する適応関数は、(1)と(2)の2つを重み付けした次式で与える。

$$E(\omega) = \omega \lambda(\rho) + (1 - \omega) \alpha \quad (0 < \omega < 1) \quad (3)$$

そして、(1) $\omega=0$ とした場合は、リンク密度、(2) $\omega=1$ とした場合は、ネットワークの混雑度だけを最適化することになる。

3. 進化的アルゴリズムの適用

本研究では、ネットワークの隣接行列 A を、コード化し表現する。その上で、交叉と突然変異を行わせる。このあと、親個体と子個体の中で最適な個体を選択し、残りは淘汰する。これを繰り返すことで最適化したネットワークを得る。

世代交代法として、MGGと呼ばれるモデルを用いる[4]。ただし、選択方法としては、最良の2個体を選択する。

まず、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。ノード数を固定し、リンクを張り替えて、新しいネットワークを生成する。初期世代のネットワークとして、ある存在確率でリンクを生成したものをを用いる。本研究では、ネットワークが密度 p_0 を持つポアソン分布になるように生成する。ここで、 p_0 は $7/n C_2$ 、個体数は10、交叉率は0.5、突然変異率は $2/n C_2$ である。

ネットワークの隣接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、隣接行列の上に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。

上記コード化により、隣接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。その後、ある確率で隣接行列の一部を反転させる。

上記の交叉や、突然変異では、独立したネットワークが生じてしまう可能性がある。ここで、それを修正するための操作として、新たに生成されたネットワークが分離しているネットワーク、つまりノード間の距離が 0 になったら、評価値を高めて、それを選べないようにする。

(3)式で与えたネットワークの適応度を最適することで所望のネットワークを探索する。このアルゴリズムの概念図は、図 1 に示す。

4. 実験結果

4.1 初期ネットワークの生成

最初に生成したネットワークを、図 2 に示す。全ノードは約 700 個(100 ノードの場合)、そして、各ノード当たりの平均リンク数は、7 である。ポアソン分布に従って発生させたことから、初期ネットワークはランダムネットワークになっている。このようなランダムネットワークを 10 個生成し、それらに遺伝アルゴリズムを適用して、(3)式の適応関数が最適になるように進化させる。

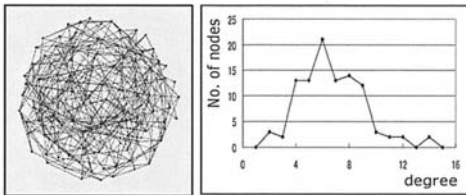


図2. 初期ネットワーク構造と次数分布

4.2 生成されたネットワーク(その 1)

(1) 混雑指数の最小化($\omega=1$)

(3)式の適応関数において、 ω を 1 とおいたときのシミュレーション結果は、完全連結グラフになった(図3)。この場合、平均リンク数は 98、リンク密度は 0.99 である。完全連結グラフのように、各ノードが他のすべてのノードとつながることで、トラヒックの混雑は生じないことになる。

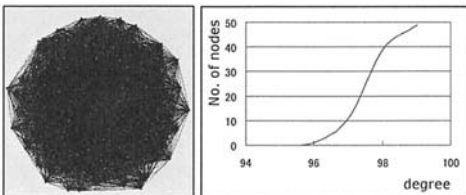


図3. 混雑指数の最小化により生成されたネットワーク構造と次数分布($\rho=0.3$)

(2) リンク密度の最小化($\omega=0$)

(3)式において、 ω を 1 においた結果、木構造のネットワークが得られた(図4)。これは、リンク数が最も少ないネットワークである。

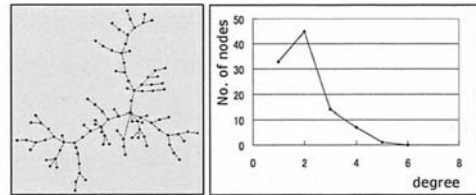


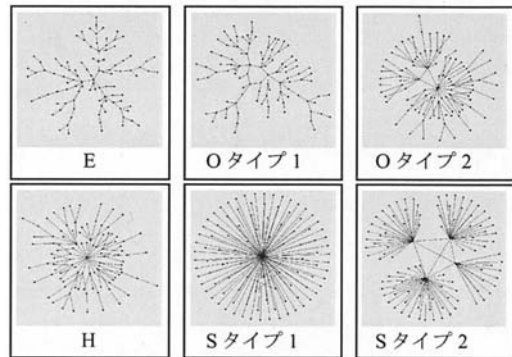
図4. リンク密度の最小化により生成されたネットワーク構造と次数分布($\rho=0.3$)

以上の2つの特異点でのシミュレーション結果より、3節で述べた進化的アルゴリズムが有効に機能することを確認められた。

4.3 生成されたネットワーク(その 2)

次に、(3)式の重み変数 $\omega(0 < \omega < 1)$ とトラヒック発生率 $\rho(0 < \rho < 1)$ の値をさまざまな値に変えたとき、どのようなネットワークが進化するかを実験した。その結果を、図 5 に示す。3 回ずつシミュレーションした結果を示してある。

実験の結果、パケット生成率 ρ と重み ω が低いときは、指数分布型のネットワークが現われる(図 5 の E)。それらの値が高くなれば、ハブネットワーク(図 5 の H)、スター型ネットワーク(図 5 の S タイプ 1)が順に現われる。このように、パケット生成率 ρ と重み ω を高く(混雑度の重視)するとき、木構造ネットワーク、指数分布ネットワーク、そして、ハブネットワーク、最後にスター型ネットワーク、密集ネットワークの順番で進化する事が分かった。



$\rho \backslash \omega$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	E	E	E	E	O1,E,H	E,H	O2,H	O2,H	H
0.2	E	E,H	O2,H	O2,H	O2,H	O2,H	O2,H	H	S1
0.3	O1,E	E	H	O	H	O2,H	H	S1,S2	S1
0.4	E	H	O2,H	O2,H	H	O2,H	O2,H	S1,S2	S1
0.5	E,H	H	H	O	O2,H	O2,E,H	S1,S2	S1	S1
0.6	O1,E	H	H	O	O	H	S1	S1	S1
0.7	O1,E	O2,H	O2,H	O2,H	H	S1,S2	S1,S2	S1	S1
0.8	O2,E,H	H	H	O2,H	H	S1,S2	S1,S2	S1	S1
0.9	O2,E,H	O2,H	S1	O2,H	S1,S2	S1	S1	S1	S1

図5. 最適ネットワーク

4.4 進化的に生成されたネットワークの分析

スケールフリーネットワークは、パケット生成率が低ければ、リンク制約が低い所で、一方で、パケット生成率が高ければ、リンク制約が高い所で現われる。これは、パケット生成率が高ければ、リンクの制約も同様に高くなければならぬことを意味している。結局、多く使われるネットワークほど、ネットワークリンクを追加することになるので、リンクの制約が最も重要になる。

本シミュレーションで現われた独特なネットワークは、中心部にコア構造があるネットワークである(図5のO)。中間的な値をとる ω 、 ρ の領域で、このネットワークは現われた。このネットワークは、ハブがお互いにつながってコアを形成し、ハブは2段以上の構造になっている。このような特徴は、指数分布型ネットワークと、スケールフリーネットワークの両方で現われた。

次に、パケット生成率を $\rho=0.3$ のとき(重み係数 ω)の違いにより4つのタイプのネットワークが生成される)、ネットワーク構造を詳しく調べた。混雑指数や、リンク密度などの値は、コアと階層構造を持つネットワークの差はない(図6.(1))。だが、最大betweennessを見ると(図6.(2))、コアノードと階層構造を持つネットワークは、他のネットワークよりも低くなっていることが分かる。すなわち、中央のコアの存在によりトラヒックの経路がうまく分散されているため、最大betweennessが低くなっている。このようなコアノードの存在により、トラヒックのボトルネックが解消されることになるという解釈できる。

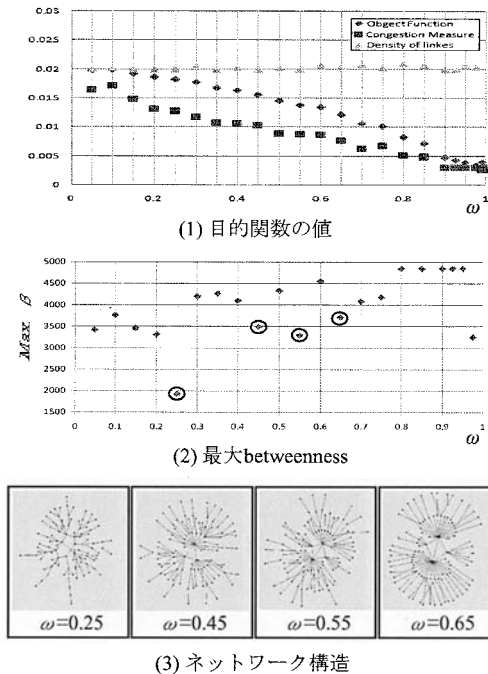


図6. 生成されたネットワークの特性($\rho=0.3$)

また、いくつかのスター型ネットワークがつながってコア構造が現われるネットワークも出現した(図5のSタイプ2)。このようなネットワークの最大betweennessの値も、他のスター型ネットワークの値と差がある。このことから、コア構造を作ることだけでも、トラヒックの混雑を減らすことができることを意味する。

スター型のハブもコアを通じてトラヒックの経路を分散させることで、スター型ネットワークの高いbetweennessは減らすことができる。一方で、スター型ネットワークは、ハブノードが高いbetweennessを持つので、ハブの負荷が高い。そのため、ハブがつながっていても、トラヒックのボトルネックは存在する。結局、コアとなるノードの構造だけでは、トラヒックのボトルネックを解消することはできない。そして、ハブノード間での階層構造が重要になることが分かる。

つまり、多くのリンクを持ついくつかのハブ(スター型ではない)が中央にコア構造を形成することで、トラヒックのボトルネックを無くすことができる。

5. まとめと今後の課題

本研究では、進化的アルゴリズムにリンク密度とトラヒックによる混雑の両方を最小にする最適なネットワークを提案した。

現在の研究は、ノード数が少ないため、進化的に生成したネットワークの構造の特性を詳細にする上であいまいさが残る。したがって、今後はノード数を増やすことが必要である。また、ネットワークシミュレーションを通じて最適化されたネットワークにパケットを流して、実際に適用することで効果の確認をする必要がある。

References

- [1] Àlex Arenas, Antonio Cabrales, Albert Díaz- Guilera, Roger Guimerà and Fernando Vega: Optimal Information Transmission in Organizations: Search and Congestion. (2005)
- [2] Albert, R., and Barabási, A.-L., Statistical mechanics of complex networks, Reviews of Modern Physics 74, pp. 47-97 (2002)
- [3] S. A. Kauffman: The Origins of Order: Self- Organization. Oxford University Press, New York. (1993)
- [4] Sato Hiroshi, Ono Isao, Kobayashi Shigenobu: A New Generation Alternation Model of Genetic Algorithms and Its Assessment, Journal of Japanese Society for Artificial Intelligence, Vol. 12, No. 5, pp. 734-744. (1997)
- [5] S.H. Strogatz: Exploring complex networks. Nature 410, 268-276. (2001)
- [6] Watts D.J. and Strogatz S.H., Collective dynamics of small-world networks, Nature 393, pp. 440-442 (1998)