

ネットワークトポロジーの最適化とスキーマ解析

Optimization of Network Topology and Schema Analysis

小林 直樹^{*1}

Kobayashi Naoki

白山 晋^{*2}

Shirayama Susumu

^{*1}東京大学大学院工学系研究科

School of Engineering, The University of Tokyo

^{*2}東京大学人工物工学研究センター

RACE, the University of Tokyo

Synchronization in networks of agents has been studied related to the effect of the network structure. In the field of complex network research, most of those studies have focused on the most cost-effective synchronization system. In this paper, we propose a new approach to find a network topology that is the most suitable for the synchronization using schema analysis.

1. はじめに

人間関係のネットワーク、Web ページ間のリンク関係を表すネットワーク、生態系に現れるネットワーク等、現実中存在するネットワークの構造が複雑ネットワーク分析によって明らかにされつつある。また、それらの構造がネットワーク上で起こる様々な現象に影響を与えていることが示されている。

近年、複雑ネットワークの分野で行われる研究には、この構造と現象の関係を探るようなものが多い。ほとんどの場合、ネットワーク構造を特徴づける統計的指標とそれに対応する現象を調べ上げ、二つの関係を探るというアプローチがとられている。平均次数、平均頂点間距離、クラスタリング係数と感染症の伝播や情報伝播との関係、次数分散と協調の進化との関係などである。本稿で扱う同期現象に関しても、同期し易さを示す容易度 θ は、 $\theta \sim \lambda_{max}/\lambda_2$ というように、ラプラシアン行列の第二最小固有値 λ_2 と最大固有値 λ_{max} によって特徴づけられる [Pecora 98]。また、Bernaldo らは、同期容易度を次数相関によって特徴付けている [Bernardo 06]。

しかしながら、実ネットワークの構造は様々な特徴が複雑に絡み合っており、一つの指標で構造を代表させることはできない。したがって、一つの指標のみに着目すると相反する知見が得られることもある。例えば、西川ら [Nishikawa 03] は、次数分布がべき乗則に従うネットワークにおいて、べき乗指数と同期容易度 θ との関係を調べ、べき乗指数が小さくなると、同期容易度が下がり同期しにくくなることを示した。これは、ハブで情報が滞留してしまうため、べき乗指数が小さいようなネットワークでは、同期現象が起こりにくくなると考察している。しかし、ハブの影響により同期が容易に達成されるとする研究 [Moreno 04] もあり、これらの知見は相反する。

このため、構造と現象の関係を調べる場合、いくつかの指標を固定し、変化させる指標を限定するという方法が提案されている。代表的な方法に、所与の次数分布に対して、クラスタリング係数や次数相関を変化させるというものがある。

一方、ネットワークに対し何かしらの作用を加えたときに構造から発現する現象を、ネットワークの機能からのものと考え^{*1}、機能を調べるという方向性がある。所与の機能を備えたネットワークをデザインする、あるいは所与の機能を備えるよ

うにネットワークの構造を変化させるという方向性でもある。この中に、ある目的関数に対して最適なネットワーク構造を探る研究がある。花田らは、多目的遺伝的アルゴリズムを用いて目的関数に応じた最適ネットワークを作成する手法を提案している [花田 07]。また、山本らは、同期容易度を目的関数として遺伝的アルゴリズムによって最適ネットワークを設計している [山本 09]。しかし、それらの研究は、目的に応じた最適ネットワークを作成する手法を提案するにとどまり、生じたネットワークの分析や適用可能性に対する検討までは行っていない。特に、適用可能性という点では、現状のネットワークに対する変更という手段が望ましいが、そのような最適化は行われていない。また、実ネットワークの多くがべき乗則に従う次数分布をもち、それがネットワークの特性に影響を与えているという指摘から、次数分布が改めて注目されている。このため、次数分布の制御可能な最適化が望まれている。しかしながら、既存研究において得られた最適ネットワークの次数分布はべき乗則に従うものよりもガウス分布に近いものが多く、その制御は難しい。

本研究では、次数分布を変えず、目的関数を最適化するネットワークのトポロジーを求める方法を提案する。また、最適化の過程をスキーマ解析によって分析することで、ネットワークの構造と機能の関係を明らかにする。なお、本稿では、目的関数をコンセンサス問題の収束速度、およびクラスタリング係数や次数相関などの統計的指標とする。

2. 提案手法

蛍の同時発光現象、人間の拍手のリズムが同期していく現象、人間社会での意見の合意形成メカニズムなどは、複数エージェントの協調制御問題として、メカニズムの解明とともに、最適ネットワークの設計において用いられるものである。その一つがコンセンサス問題である。本稿においては、ネットワークトポロジーの最適化に対して、ノード上のエージェント間のコンセンサス問題を扱う。

2.1 モデル

ノード数 n 、エッジ数 m とし、ノード $V = \{1, 2, \dots, n\}$ 、エッジ $E \subseteq V \times V$ で構成される無向グラフ $G = (V, E)$ を考える。ノード i 上のエージェントを i とし、 t ステップにおけるエージェント i の状態量を $x_i(t)$ とする。状態量を次式の推

連絡先: 小林直樹, 東京大学大学院システム創成学専攻修士 2 年, naoki_kobayashi@race.u-tokyo.ac.jp

^{*1} 複雑ネットワークの分野では、現象を機能そのものとして扱う場合もある

移則によって変化させる．

$$dx_i/dt = \sum_{j \in N_i} (x_j(t) - x_i(t)) \quad (1)$$

ここで、 N_i は、 i の隣接ノードの集合である．
はじめに、式 (1) を離散化する．

$$x_i(t + \Delta t) = x_i(t) + \Delta t \sum_{j \in N_i} (x_j(t) - x_i(t)) \quad (2)$$

ここで、 Δt は時間刻み幅である．初期状態から式 (2) に従って時間発展させ、すべてのエージェントの状態量が一致したとき、すなわち、

$$x_1 = x_2 = \dots = x_n \quad (3)$$

のときに収束したとみなす．Olfati Saber ら [Olfati-Saber 07] は、実数の初期状態量、かつ $\Delta t < 1/k_{max}$ であれば、式 (2) が収束することを示している．ここで、 k_{max} は最大次数である．ただし、数値誤差等のため、数値シミュレーションでは、式 (3) を満たすことは難しい．そこで、本稿では、 $t = 0$ における i の状態量 $x_i(0)$ を i とし、時間発展させ、 $\max_i |x_i| - \min_j |x_j| < 0.01$ のとき、収束したことにする．目的関数を収束時間とする場合、これを最小化するネットワークのトポロジーを求める．

2.2 最適化手法

2.2.1 次数分布の固定と初期ネットワーク作成

本研究では次数分布を固定して最適化を行う．その利点は、次数分布を制約条件とすることで、例えば、べき乗則に従ったままでの最適化の手段を示すことができることや、次数分布が固定されることにより、分析の際に次数分布の影響が排除できることである．

本稿では、初期ネットワークをコンフィグレーションモデルによって作成する．はじめに、ノード数 n 、エッジ数 m を与える．次に、次数分布をガウス分布、あるいはべき乗分布とし、各ノードの次数を算出し、ノードに次数分のスタブ (stub, 半分のエッジ) を与える (図 1)．すべてのスタブをランダムにリンクする．

2.2.2 コード化

スキーマ解析のために、ネットワークの構造を隣接行列で表現し、それをコード化する．具体的には、山本ら [山本 09] と同様に、隣接行列を図 2 のようにビット列に対応させることでコード化する．

2.2.3 最適化手法

本稿での最適化手法は、遺伝的アルゴリズム (GA) に基づくものである．ただし、現時点では、突然変異に相当する操作のみで、山登り法的に最適化を行っている．目的関数は、コン

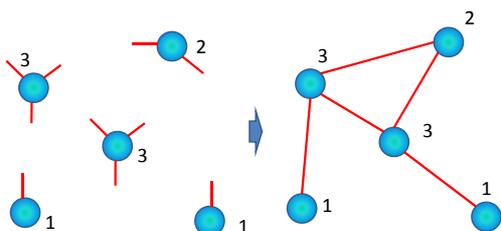


図 1: 初期ネットワーク作成

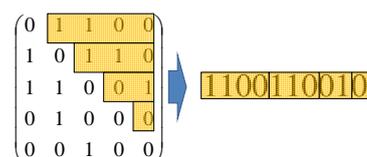


図 2: ネットワークのコード化

センサ問題の収束速度とネットワークの統計的指標のいずれかである．ただし、収束速度とネットワークの統計的指標の関連性を調べるために、目的関数に関わらず、最適化の中で双方を求める．以下に手順を示す．なお、目的関数が統計的指標の場合、最適化の対象となるものを予め決めておく．ここでは、 s で示す．

Step0 初期ネットワークに対してコンセンサス問題を解く．ノード上のエージェントに状態量を与え、初期状態から式 (1) に従って時間発展させ、収束時間を求める．この収束時間を τ_0 とする．また、クラスタリング係数などの統計的指標を求め、最適化の対象となる統計的指標を s_0 とする．

Step1 ノードの共有がない異なる 2 つのエッジをランダムに選び、図 3 のようにつなぎ換えを行い、一時的なネットワークを形成する．つなぎ換えによって多重辺が生じる場合は、エッジを選び直す．

Step2 Step1 で形成されたネットワーク上でコンセンサス問題を解き、収束時間を求める．また、ネットワークの統計的指標を求める．この際、ネットワークが連結であるかの判定も行い、非連結であれば、収束時間 $\tau_d (\gg \tau_0)$ 、および統計的指標 s_d を与えるものとする． s_d は、最小化の場合は、 $s_d \gg s_0$ 、最大化の場合は、 $s_d \ll s_0$ とする．また、コード化された隣接行列、収束時間、統計的指標を履歴情報として蓄える．

Step3 目的関数が収束時間のとき、および統計的指標で最小化のとき、Step2 で求めた目的関数の値が減少する場合、一時的なネットワークを新たなものとして採用する．そうでない場合はネットワークは更新しない．統計的指標で最大化のとき、Step2 で求めた目的関数の値が増加する場合、一時的なネットワークを新たなものとして採用する．そうでない場合はネットワークは更新しない．

Step4 Step1 から Step3 を繰り返す．この際、Step3 においてネットワークが更新されない回数を数え、所与の回数を超えた場合に目的関数の値が最適値になったものとみなし、終了する．

Step2 は突然変異に相当する．また、次数は変わらないので、次数分布を変化させないという制約条件は満たされる．

2.3 スキーマ解析

前項で示した最適化の過程を GA としてみると、隣接行列のコードが個体に、収束時間や統計的指標が適応度に相当する．また、ネットワークが表現型となる．現時点では、一つの初期個体に対して突然変異のみを用いた進化の方法で最適化を進めているが、その過程において、適応度に対応づけられたコードが蓄えられている．そこで、スキーマ解析により個体の特徴を調べる．この際、ビット列の部分的な集合を評価するた

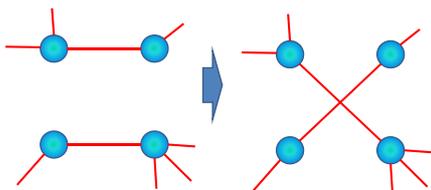


図 3: つなぎ換えの概念図

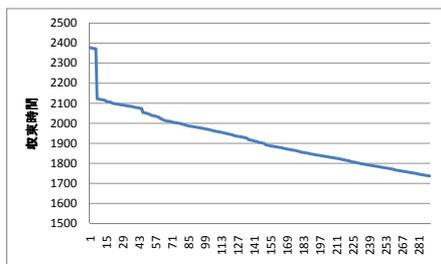


図 4: 収束時間の変化

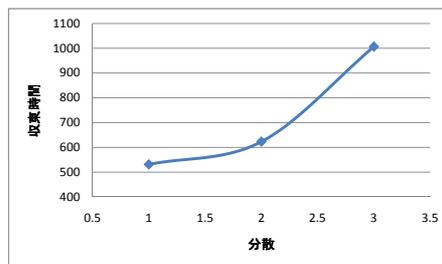


図 5: ガウス分布の分散と収束時間の関係

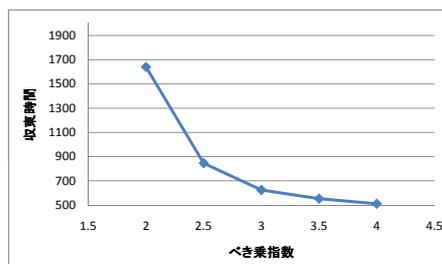


図 6: べき乗指数と収束時間の関係

めに、特徴的な部分集合以外を * で表す。例えば、1000, 1010, 1100, 1110 という個体はスキーマ $1 * * 0$ によって定義される。

スキーマ解析の結果から、適応度の高い、あるいは低い個体の表現型であるネットワークの構造を明らかにする。本稿では、先述の Step3 において、収束時間が改善された（減少した）もの、あるいは統計的指標が小さくなった（大きくなった）ものに対し、変化したノードやエッジの特徴をスキーマ解析を利用して調べる。

3. 実験と考察

本稿では、ノード数を $n = 1000$, エッジ数を $m = 6000$ として数値実験を行う。また、初期の次数分布をガウス分布とべき乗分布とする。ガウス分布の場合、分散 σ^2 が $\sigma^2 = 1, 2, 3$ の 3 種類、べき乗分布の場合、べき乗指数 γ が $\gamma = 2, 2.5, 3, 3.5, 4$ の 5 種類のネットワークを用いる。提案手法では、次数分布が変化しないことに留意されたい。なお、コンセンサス問題を解く際の Δt は、 $1/100$ としている。また、最適化手法 Step4 の所与の回数を 1000 回とした。

はじめに収束時間の最適化を行い、その過程を調べる。次にクラスタリング係数と収束時間、次数相関と収束時間の関係を調べ、最適化されたネットワークのトポロジーについて考察する。最後に、スキーマ解析の一例を示す。

3.1 収束時間の最適化

初期ネットワークの次数分布を $\gamma = 2$ のべき乗分布にした場合の最適化過程を図 4 に示す。縦軸が収束時間、横軸が最適化のステップ数である。281 ステップ目で、ネットワークの非更新数が 1000 回を超えたため、このステップで最適化されたものとみなした。

次に、最適化された収束時間と次数分布の関係を調べる。次数分布がガウス分布の場合における最適化された収束時間と分散の関係を示す（図 5）。分散が小さくなれば、最適な収束時間が短くなることからわかる。図 6 に、べき乗分布の場合の最適化された収束時間とべき乗指数の関係を示す。図からべき乗指数が大きくなれば最適な収束時間が短くなるからわかる。

3.2 クラスタリング係数と収束時間の関係

最適化過程でのネットワークトポロジーの変化をクラスタリング係数によって推測する。初期に与えた次数分布は変化しない。

はじめに、次数分布が $\gamma = 3$ のべき乗分布のネットワークに対して、クラスタリング係数 C を大きくするように最適化する。ここでは、 $C = 0.457$ という高いクラスタリング係数のネットワークが生成された。このネットワークを初期ネットワークとする。

次に、初期ネットワークに対して、収束時間を目的関数にした最適化を行う。最適化の過程におけるトポロジーの変化をクラスタリング係数と収束時間の関係によって調べる。その結果を図 7 に示す。図中の赤線が収束時間を、青線がクラスタリング係数を示す。また、初期ネットワークに対して、クラスタリング係数を目的関数とし、減少するように最適化を行う。図 8 がその結果である。図 7 と図 8 からは、収束時間の減少とクラスタリング係数の減少には正の相関があることがわかる。また、次数分布がべき乗分布の場合、高いクラスタ性をもつネットワークの場合、収束時間が大きいことがわかる。一方、クラスタリング係数はステップ数に対して、ほぼ線形に変化するが（図 8 では変化率は減少するが傾向は同じである）、収束時間はあるステップ数以降、緩やかな減少に転じる。このため、収束時間に対して最適化されるネットワークトポロジーをクラスタリング係数の変化だけで説明することは難しい。

3.3 次数相関と収束時間の関係

最適化過程でのネットワークトポロジーの変化を次数相関によって推測する。

はじめに、次数分布が $\gamma = 3$ のべき乗分布のネットワークに対して、次数相関 r を大きくするように最適化する。ここでは、 $r = 0.28$ というネットワークが生成された。このネットワークを初期ネットワークとする。

次に、初期ネットワークに対して、収束時間を目的関数にした最適化を行う。最適化の過程におけるトポロジーの変化を次

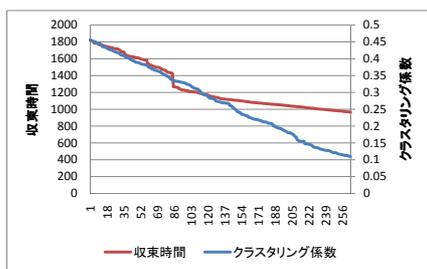


図 7: 収束時間最適化とクラスタリング係数

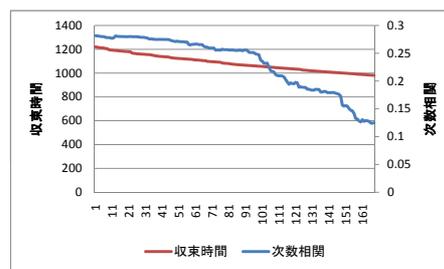


図 9: 収束時間最適化と次数相関

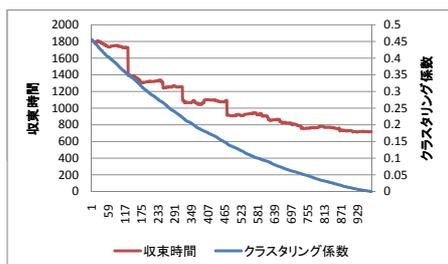


図 8: クラスタリング係数最適化と収束時間

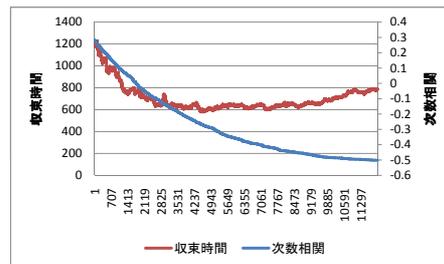


図 10: 次数相関最適化と収束時間

数相関と収束時間の関係によって調べる．その結果を図 9 に示す．図中の赤線が収束時間を，青線が次数相関を示す．また，初期ネットワークに対して，次数相関を目的関数とし，減少するように最適化を行う．図 10 がその結果である． $r = 0.28$ から減少し， $r = -0.50$ となる．図 9 から，収束時間の最適化の過程でトポロジーが変化し，次数相関が減少することがわかる．また，収束時間はほぼ線形に減少するのに対し，次数相関は，100 ステップの近傍で急減することもわかる．一方，次数相関を最適化する場合，図 10 からわかるように，収束時間に大域的な極値が存在する．次数相関の急減と収束時間の極値の存在に関する考察は講演時に示す．

3.4 統計的指標の変化に注目したスキーマ解析

図 4 などの収束時間の最適化において，収束時間が大きく減少するステップがあることがわかる．そのようなステップでのコードの変化に注目する．本稿では，3.2 項で示した収束時間最適化の過程において，クラスタリング係数の変化に着目したスキーマ解析を行う．

はじめに，1 ステップあたり一定以上の収束時間の減少が起きた場合の前後のコードの変化を記録する．提案手法では，4 つのノードが影響するのでそれらを記録すればよい．次に，その際のクラスタリング係数の変化を記録する．その後，統計処理によって変化の平均などを求め，平均以上にクラスタリング係数が変化したときのつなぎ換えに関与したノードを見つける．それらのノードに関連する部分的なコードを抽出し，スキーマ解析によってパターンを求める．そのパターンを分析することで，どのようなトポロジーの変化が収束時間の減少に関わるかを調べる．

3.2 項で示した最適化において，収束時間が 5 以上減少した数は，22 回であった．クラスタリング係数の減少の平均は，0.164 である．この平均よりも大きく変化した場合についてスキーマ解析を行う．結果については講演時に示す．

4. 結語

次数分布を変えず，目的関数を最適化するようにネットワークのトポロジーを変更する方法を構築した．また，最適化の過程において，ネットワーク構造をマクロ的に表す統計的指標と機能から生じる現象の関係を分析する方法を示した．さらに，スキーマ解析によって，局所的なリンク構造と現象の関係を調べるという方法を提案した．

提案手法をコンセンサス問題に適用し，クラスタリング係数，および次数相関と収束時間との関係性を明らかにした．

参考文献

- [Olfati-Saber 07] R. Olfati-Saber, J.A. Fax and R.M. Murray: Proc. of the IEEE, Vol.95, No.1, pp.215-233 (2007)
- [Pecora 98] L.M. Pecora and T.L. Carroll: Phys. Rev. Lett. 80, pp.2109-2112 (1998)
- [Nishikawa 03] T. Nishikawa, A.E. Motter, Y.-C. Lai and F.C. Hoppensteadt: Phys. Rev. Lett. 91, 014101 (2003)
- [Bernardo 06] M. di Bernardo, M. F. Garofalo, and F. Sorrentino: Int. J. Bifurcat. Chaos 17, pp.3499-3506 (2007)
- [Moreno 04] Y. Moreno and A.F. Pacheco: Europhys. Lett., 68 (4), pp.603-609 (2004)
- [花田 07] 花田良子, 佐藤史隆, 廣安知之, 三木光範, 鈴木泰博: コンピュータソフトウェア, Vol.24, No. 1, pp. 91-100 (2007)
- [山本 09] 山本聡彦, 生天目章, 佐藤浩: Proc. of JAWS2009, pp.170-177 (2009)