

## 複雑システムにおけるネットワーク中心性が予測精度に与える影響

鈴木 智也<sup>†</sup> 池田 真一<sup>†</sup>

自然界に存在する複雑システムを分析するために、それを簡潔に模擬した Coupled Map Lattice(CML) や Global Coupled Map(GCM) などのモデルが広く研究されている。本研究では、近年注目されている複雑ネットワーク科学の知見を取り入れるべく、CML を D.Watts らが提案した WS モデルを用いて構成した。これにより、従来の CML のネットワーク構造はレギュラーネットワークが主流であったが、スモールワールドネットワークやランダムネットワークにおいても、複雑システム全体の振舞いや力学的構造を分析できる。

さらに、ネットワークを構成する各要素は、ネットワーク上の位置に応じて、システム全体から受ける影響の仕方が異なる。本研究では、ネットワーク上の位置として、ネットワーク中心性の概念に注目し、各要素の振舞いや力学的構造の特徴を分析した。また、株価変動の予測など、複雑システムを構成する要素の将来変動を予測する場合、システム全体からの影響の受け方に応じて、予測難易度や最適な予測モデルが異なる可能性がある。本研究では、一般の複雑システムにおいて、ネットワーク中心性が予測難易度に影響を与えることを示すために、先に述べたネットワーク構造を可変できる改良 CML を用いて検証した。

### The Influence of Network Centralities on Prediction Accuracy of Complex Network Systems

TOMOYA SUZUKI<sup>†</sup> and SHINICHI IKEDA<sup>†</sup>

To analyze complex systems in real world, the coupled map lattice (CML) and the global coupled map (GCM), which model real systems basically, are often used. In the present paper, we propose the modified CML whose network structure is composed by the Watts-Strogatz (WS) model to discuss complex systems from the viewpoint of science of complex networks. Although the original CML are composed by the regular network, the modified CML can analyze behaviors and dynamical structures of complex systems composed by not only the regular network but also the small-world network and the random network.

In addition, each node composing a network interacts with the whole of system according to the location of node in network. In the present paper, we consider the location as Network Centralities: the degree centrality, the betweenness centrality and the closeness centrality, and then we discuss dynamical structure and predictability of each node. Moreover, in the case of predicting future behaviors, especially in stock markets, it is possible that the predictability of each node is affected by network centralities. In order to confirm the universality of the possibility in general complex systems, we analyzed the modified CML with changing its network structure as a model including the essence of complex systems.

#### 1. はじめに

自然界の実システムは、一般に各要素がネットワークを構成しており、互いに相互作用して時間発展している。さらに、要素に働くダイナミクスに応じて、系全体の振る舞いは単純ではなくなり、要素還元論的なアプローチで理解できる範疇を超える。このようないわゆる複雑系を解明する方法として、構成論的なアプローチが有力であり、計算機の中で仮想的に複雑系を模擬するモデルを構成することによって、現象の本質的理解や新現象の予言を目指す。この構成論的なアプローチで良く用いられるモデルとして、金子らが提案した Coupled Map Lattice(CML) や Global Coupled Map(GCM) がある<sup>1)</sup>。これらは要素に働くダイナミクスとして、主にロジスティック写像を用いており、極めてシンプルな非線形ダイナミクスでシステムを構成したにも関わらず、実に多様な複雑な振舞いを再現することができる。さらに、ダイナミクスのシンプルさゆえ、より複雑な実システムに内在する原理的な性質に着眼していると言える。

従来の CML においては、隣の要素のみと相互作用を行うといった規則的なネットワーク構造が主流である。しかし、近年の複雑ネットワーク科学が示すように、実システムにはスモールワールドネットワーク<sup>2)</sup> やスケールフリーネットワーク<sup>3)</sup> といった構造が多く存在することが明らかとなった。モデルのシンプルさも重要であるが、これらの構造を無視する

ことは有益ではないと考え、本研究では、様々なネットワーク構造を生成できる Watts-Strogatz(WS) モデル<sup>2)</sup> を用いることで、CML を改良し、ネットワーク構造に応じた各要素の振舞いやリアプノフ指数の変化を調べる。

次に、複雑ネットワーク科学の知見として、各要素のネットワーク上の位置に応じてシステム全体から受ける(あるいは与える)影響の仕方が異なるので、ネットワーク中心性に関する次数中心性、媒介中心性や近接中心性などの概念が重要視されている<sup>4)</sup>。外部から受ける影響が異なれば、要素毎の振舞いの仕方やリアプノフ指数などに影響を及ぼし、将来変動の予測問題においても、要素毎で予測難易度が異なると考えられる。ここで複雑システムを株式市場に例えれば、各企業のネットワーク上の位置付け(例えば、関連会社の有無)に応じて、株価変動の予測難易度や最適な予測モデルが異なる可能性がある。そこで本研究では、株式市場に限らず一般の複雑システムにおいて、各要素の中心性が予測難易度に影響を与えるのかを検証する。その際、先に述べた改良 CML を用いて、複雑システムの原理的レベルから検証する。

#### 2. Coupled Map Lattice(CML) の改良

##### 2.1 従来モデルの改良

従来の CML は、隣の要素のみと相互作用を行うといった規則的なネットワーク構造が主流であり、以下の式によって記述される<sup>1)</sup>。

$$X_i(t+1) = (1-\epsilon)f(X_i(t)) + \frac{\epsilon}{2}(f(X_{i+1}(t)) + f(X_{i-1}(t))) \quad (1)$$

ここで、 $i$  はノード番号、 $t$  は時間である。結合係数  $\epsilon$  により

<sup>†</sup> 同志社大学工学部情報システムデザイン学科  
Department of Information System Design, Faculty of Engineering, Doshisha University

要素間の相互作用の強さを表し、各要素は自分自身の出力値と結合している要素の出力値を混合することにより時間発展していく。さらに、各要素に働くダイナミクスとして、主に以下のロジスティック写像が用いられる。

$$f(X(t)) = 1 - aX^2(t)$$

ロジスティック写像は、カオスの典型的モデルとして調べ尽くされており、さらに、最もシンプルに非線形システムを表現できるため、非線形性以外の特殊な効果によって解析結果が左右されることがなく、システムに内在する原理的な特徴に着目することができる。さらに、コントロールパラメータ  $a$  によって非線形性の強さを調節でき、 $2.0 \geq a \geq 1.4011$  の時においてカオス解を示すことが知られている。

本研究では、CMLの相互作用をWSモデル<sup>2)</sup>で表現することにより、ネットワーク構造を可変できる改良CMLを構築する。式(1)は次式のように変更される。

$$X_i(t+1) = (1-\epsilon)f(X_i(t)) + \frac{\epsilon}{k_i} \sum_{j \in G_i(p)}^{k_i} f(X_j(t))$$

$G_i(p)$  はノード  $i$  と結合する要素  $j$  の集合を表し、WSモデルにおける再結合確率  $p$  に応じたネットワーク構造を反映している。さらに、ノードに働くダイナミクスは従来と同様にロジスティック写像を用い、 $a = 2.0$  とした。

### 2.2 ネットワーク構造に応じた中心性の変化

次に、実際に計算機を用いて  $n = 100$  の要素による改良CMLを構成し、ネットワーク中心性に関する性質を調べる。なお、WSモデルでのネットワーク作成課程において乱数を用いているので、 $p$  の値が同じでも、生成されるネットワークは異なる。よって実験の信頼性を向上させるために、本研究の全ての実験では、同一の  $p$  によるネットワークを20個作成して、その試行平均値をもって解析結果を示している。

図1では、ネットワーク構造の違いに応じた、クラスター係数  $C_i$ 、次数中心性  $K_i$ 、近接中心性  $D_i$ 、媒介中心性  $B_i$  を示している。 $p$  が大きくなるにつれてランダムネットワークに近づくので、クラスター係数  $C_i$  は低下し、次数中心性  $K_i$  の標準偏差は拡大し、近接中心性  $D_i$  は低下するが、媒介中心性  $B_i$  に関してはスモールワールド構造の時に最も低い値を示した。

さらに図2にて、ネットワーク構造の違いによる各中心性の相関関係を示す。例えば、各ノードの  $C_i$  と  $J_i$  ( $i = 1, \dots, n$ ) などの相関係数を算出することで各中心性間の関連の強さ評価した。各中心性ともネットワーク上の影響力の強さを表すものであるが、各相関係数の値は低いので、異なる概念であることが分かる。ただし、ネットワーク構造に応じては、相関係数が大きくなる場合があり、特に、スモールワールドネットワークではクラスター係数  $C_i$  と近接中心性  $D_i$  の関係性が、ランダムネットワークでは次数中心性  $K_i$  と近接中心性  $D_i$  の関係性が高い。

### 2.3 改良CMLの振舞いの分析

次に、 $t = 1 \sim 2,000$  まで時間発展させることにより各要素の振舞い  $X_i(t)$  を生成し、同期度やリアプノフ指数の観点から改良CMLのマクロ的性質を分析する。同期度の計算法として、2つの要素の振舞い、例えば  $X_i$  と  $X_j$  の相関係数を  $\Phi_{ij}$  とすると、システム全体の同期度  $\Phi$  は平均化することにより

$$\Phi = \frac{1}{nP_2} \sum_{i \neq j} \Phi_{ij}$$

として評価した。解析結果を図3に示す。モデルパラメータ  $p$  及び  $\epsilon$  の変化に伴い、改良CMLの構造が変化し、各要素の同期度が変化することが分かる。なお、たとえ互いの相関で

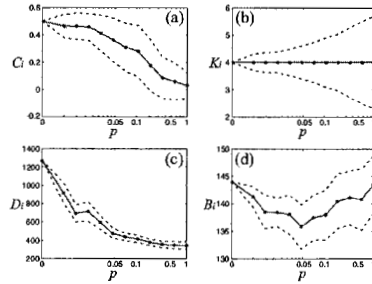


図1 改良CMLにおけるネットワーク構造の違いによる (a) クラスター係数  $C_i$ 、(b) 次数中心性  $K_i$ 、(c) 近接中心性  $D_i$ 、(d) 媒介中心性  $B_i$  の変化。実線は各ノードの平均値を表し、破線は標準偏差を表している。

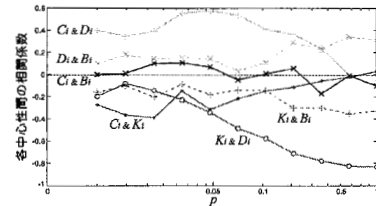


図2 改良CMLにおけるネットワーク構造の違いによる各中心性の相関関係。

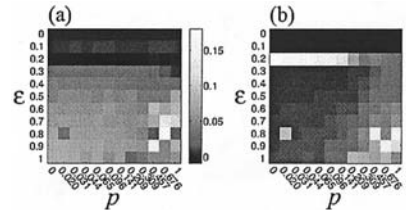


図3 改良CMLにおける各要素の同期度。図(a)は  $\Phi$  を表し、図(b)は  $|\Phi|$  を表している。

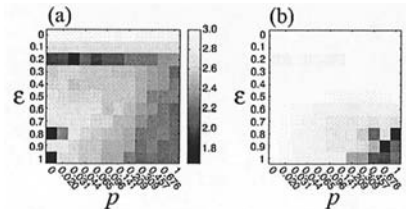


図4 改良CMLにおける (a) 最大リアプノフ指数  $\lambda_M$  と (b) リアプノフ次元  $\lambda_D$ 。

も互いに影響を及ぼし合っているという観点により、図3(b)では絶対値をとることにより

$$|\Phi| = \frac{1}{nP_2} \sum_{i \neq j} |\Phi_{ij}|$$

として同期度を算出している。結合強度  $\epsilon$  が大きくなるほど、互いの影響力が強化されるので同期度は上昇する。さらに、図1(c)で示されるようにランダムネットワークになるほど近接中心性  $D_i$  が低下し、つまりネットワーク全体の距離が短くなるので互いの影響力が強化され、同期度は上昇している。図3(b)に示される  $(p, \epsilon) = (0, 0.2)$  での強い同期は、従来のCMLにおいて観測されるものであるが、 $p$  が異なるネットワーク構造になるにつれ、同期は弱まっていく。

次に、モデルパラメータ  $p$  及び  $\epsilon$  に応じた、改良CMLの力学的構造の変化を調べるために、リアプノフスペクトラム<sup>5)</sup>を算出する。CMLなどの多次元力学系の時間発展は、多次元

状態空間内の軌道として表現されるが、リアプノフスペクトラムは、その軌道の各次元における不安定性を定量化したものである。この値が大きいくほど軌道の不安定性が増し、将来変動の予測難易度に影響を与える。言い換えると、状態空間内においてごく接近した軌道が離れていく度合いを表す量であり、将来変動の予測手法である局所線形近似法のパフォーマンスに直接的な影響を与える。

改良 CML のリアプノフスペクトラム  $\lambda_i$  は、次式のように  $n$  次元状態空間内の軌道の不安定性として算出される。

$$V(t) = \{X_1(t), X_2(t), \dots, X_i(t), \dots, X_n(t)\} \quad (2)$$
算出された  $\lambda_i$  ( $i = 1, \dots, n$ ) において、正の値を示す次元においては軌道が不安定であることを表し、負の値を示す次元においては軌道が安定であることを表している。また、リアプノフスペクトラム  $\lambda_i$  の最大値を最大リアプノフ指数  $\lambda_M$  と呼び、さらに次式により正のリアプノフスペクトラムを用いてリアプノフ次元  $\lambda_D$ <sup>6)</sup> が算出される。これらによって、軌道の不安定性を評価することが一般的である<sup>5)</sup>。

$$\lambda_D = \beta + \frac{\sum_{i=1}^j \lambda_i}{|\lambda_i|}$$

ここで、 $\beta$  は降順に並べ替えられた  $\lambda_i$  の和が正であるような最大の整数である。

図 4 において改良 CML における最大リアプノフ指数  $\lambda_M$  およびリアプノフ次元  $\lambda_D$  の様子を示す。これらの結果は図 3 の同期度の結果と強い関係性があることがわかる。特に同期度が高い場合において、 $\lambda_M$  および  $\lambda_D$  が低下しており、軌道は比較的安定と言える。軌道の安定性は予測難易度に関連するものであるため、この状況における予測難易度は低いと予想される。このようにモデルパラメータ  $p$  と  $\epsilon$  に応じた改良 CML の将来変動の予測難易度を調べるために、式 (2) の軌道  $V(t)$  に対して局所線形近似法<sup>7)</sup>を用いてシステムの将来変動を予測する。

局所線形近似法<sup>7)</sup>では、軌道上の現在点  $V(t)$  の近傍点  $V(t_m)$  を過去の軌道から  $M$  個 ( $m = 1, 2, \dots, M$ ) 見つけ出し、その近傍点の 1 ステップ将来値を平均化することで、予測値  $\hat{V}(t+1)$  を得る。つまり、以下のように求められる。

$$\begin{aligned} \hat{V}(t+1) &= \frac{1}{M} \sum_{m=1}^M V(t_m+1) \\ &= \{\bar{X}_1(t+1), \bar{X}_2(t+1), \dots, \bar{X}_n(t+1)\} \quad (3) \end{aligned}$$

この手法により、全要素の将来値  $\bar{X}_i(t+1)$  ( $i = 1, \dots, n$ ) を得ることができる。

シミュレーションにおいては、近傍数の個数  $M$  を状態空間の次元  $+1$  とし (つまり  $M = 101$ )、後半の  $t = 1001 \sim 2000$  の  $X_i(t)$  を予測対象とした。さらに、 $X_i(T)$  ( $T \geq 1,001$ ) を予測する場合、 $t = 1 \sim T-1$  の時系列データを予測のためのデータベースとした。各ノードの予測精度を評価するために、正解値  $X_i(t)$  と予測値  $\bar{X}_i(t)$  ( $t = 1001 \sim 2000$ ) の相関係数を算出し、その予測精度を  $\gamma_i$  と表す。図 6(a) において、改良 CML における予測精度の平均値  $\langle \gamma_i \rangle_i$  を示す。結合強度  $\epsilon$  が小さくなるにつれ、予測精度の低下を確認できる。これは、影響力の小さい要素まで予測モデルに取り込むので、全体の予測精度が低下したと考えられる。また、図 3 の同期傾向が強まり、図 4 の軌道不安定性が弱まるほど予測精度が向上することが分かる。このように改良 CML のモデルパラメータに応じて、予測難易度が変化することが示された。

### 3. ネットワーク中心性が最大リアプノフ指数や予測精度に与える影響

前章では、改良 CML のマクロ的性質を分析するため、同

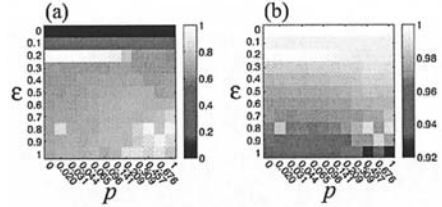


図 5 改良 CML における予測精度の平均値。図 (a) は全要素を用いて予測した場合の  $\langle \gamma_i \rangle_i$  を、図 (b) は単体の要素毎で予測した場合の  $\langle \Gamma_i \rangle_i$  を示している。

期度、リアプノフ指数、予測精度についてシステム全体の様子を分析してきたが、本章では、ノードの個性に応じた特性を調べるため、ミクロ的視点より同種の性質を評価する。

さらに、通常における複雑システムの解析や、予測を行う状況を鑑みれば、式 (2) のようにシステムに内在する全変数を用いる場合は稀である。一般的には、システムに内在する変数を特定することが難しく、また観測にも限界があるので全変数の振舞いを取得するのは困難である。そこで通常では、解析対象とする 1 変量の時系列データから、ターゲットの埋め込み定理を用いて<sup>8)9)</sup>、多次元システムの軌道を再構成することが行われる。つまり、式 (2) の軌道は、

$$V_i(t) = \{X_i(t), X_i(t+\tau), X_i(t+2\tau), \dots, X_i(t+(d-1)\tau)\} \quad (4)$$

のように要素毎の軌道に書き換えられる。ここで、 $d$  は埋め込み次元、 $\tau$  は遅れ時間を表す。この式によって、1 変量の時系列データ  $X_i(t)$  を多次元ベクトル  $V_i(t)$  に変換する。本研究では、最適な埋め込み次元  $d^*$  を決定するために誤り近傍法 (False-nearest neighbor 法)<sup>10)</sup> を用い、最適な遅れ時間  $\tau^*$  は自己相関関数がはじめて 0 となる時間とした<sup>5)</sup>。  $d^*$  や  $\tau^*$  は、ネットワークパラメータの  $p$  や  $\epsilon$  によらず、常に  $d^* = 4$  及び  $\tau^* = 1$  となった。

式 (4) により要素毎に再構成された軌道  $V_i(t)$  を用いることで、2.3 章と同様に最大リアプノフ指数や、局所線形近似法の適用による予測精度を要素毎に求めることが出来る。つまり、式 (3) の予測式は、以下に書き換えられる。

$$\hat{V}_i(t+1) = \frac{1}{M} \sum_{m=1}^M V_i(t_m+1) = \{\bar{X}_i(t+1), \dots, \dots\}$$

これにより、前章と同様に後半の  $t = 1001 \sim 2000$  の  $X_i(t)$  を予測対象とし、予測値  $\bar{X}_i(t)$  ( $t = 1001 \sim 2000$  および  $i = 1 \sim 100$ ) を求める。さらに、正解値との相関係数を計算することによって、ノード毎の予測精度  $\Gamma_i$  を得る。

まず図 5(b) に、前章と同様に予測精度の平均値  $\langle \Gamma_i \rangle_i$  を示す。予測モデルが異なるので  $\langle \gamma_i \rangle_i$  の結果と異なり、特に結合強度  $\epsilon$  が小さい時に他の要素からの影響が少ないので、1 変数による予測の有用性が顕著に現れている。

この予測精度の変化を軌道不安定性からの裏付けを行う場合、先と予測モデルが異なるので単純に図 4 と比較できない。よって次に、各ノードに注目するといったいわばミクロ的な視点より、ネットワーク中心性が最大リアプノフ指数や予測精度に与える影響力を調べるために、クラスター係数  $C_i$ 、次数中心性  $K_i$ 、近接中心性  $D_i$ 、媒介中心性  $B_i$  と、最大リアプノフ指数  $\Lambda_{M_i}$  や予測精度  $\Gamma_i$  の相関係数を計算した。ここで、 $i = 1, 2, \dots, 100$  である。その結果を、図 6, 7 に示す。

図 6(a) では、クラスター係数と最大リアプノフ指数の強い正の相関関係を示し、それに伴い図 7(a) では、クラスター係数と予測精度の強い負の相関関係を示している。つまり、高いクラスター係数を持つ要素は、軌道不安定性が増加するた



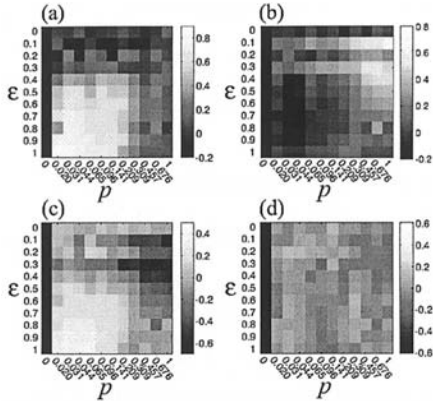


図6 ネットワーク中心性と最大リアプノフ指数  $\Lambda_{M_i}$  の相関関係。図 (a) は  $C_i$  と  $\Lambda_{M_i}$ ，図 (b) は  $K_i$  と  $\Lambda_{M_i}$ ，図 (c) は  $D_i$  と  $\Lambda_{M_i}$ ，図 (d) は  $B_i$  と  $\Lambda_{M_i}$  の相関係数を示している。

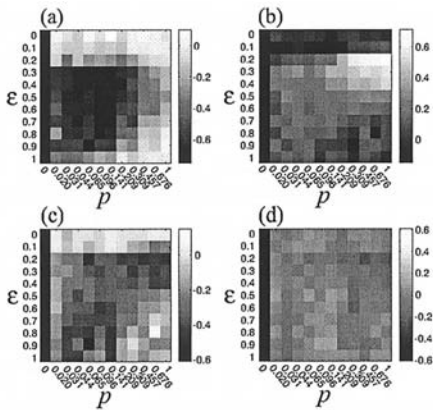


図7 ネットワーク中心性と予測精度  $\Gamma_i$  の相関関係。図 (a) は  $C_i$  と  $\Gamma_i$ ，図 (b) は  $K_i$  と  $\Gamma_i$ ，図 (c) は  $D_i$  と  $\Gamma_i$ ，図 (d) は  $B_i$  と  $\Gamma_i$  の相関係数を示している。

め予測難易度が高いと言える。また、図 6(c) および図 7(c) からも近接中心性に関して同様のことが言える。これは、図 2 で示したように、クラスター係数  $C_i$  と近接中心性  $D_i$  に正の相関関係があることからも伺える。さらに、図 2 より、近接中心性  $D_i$  と次数中心性  $K_i$  にも負の相関関係が伺えるので、図 6(b) および図 7(b) の次数中心性に関する結果は、図 6(c) および図 7(c) と反対の傾向を示している。しかし、媒介中心性に関しては図 6(d) および図 7(d) に示すように、最大リアプノフ指数や予測精度に影響を与えないことが分かる。このように、媒介中心性以外のノードの中心性は、軌道不安定性に作用し、予測難易度に影響を与えることが分かる。

#### 4. まとめ

本研究において、WS モデルを用いて CML を構成し、ネットワーク構造を変えられるように改良した。これにより、ネットワーク構造に応じた要素間の同期度、リアプノフ指数、予測難易度を分析できるようになった。また、従来の CML では各要素の結合パターンは全て同じであったが、改良 CML においては、各要素はネットワーク中心性という個性を持っている。この個性に応じたリアプノフ指数や予測難易度との関係性を調べるのが可能になった。

本研究により得られた知見は、以下の通りである。2.3 章

における実験において、改良 CML の振舞いについてマクロ的な視点より解析を行った結果、

- 改良 CML のモデルパラメータである結合強度  $\epsilon$  やネットワーク構造  $p$  に応じて同期度やリアプノフ指数が変化し、それに伴い予測難易度も変化する。例えば、結合強度が強く、ランダムネットワークに近づくほど、同期傾向が強まりリアプノフ指数は低下する。その場合の予測精度は高い。
- 一方、結合強度が弱いほど、全変数による埋め込みは関係の薄い要素まで予測モデルに取り込むことになるので、予測精度は低下する。
- しかし、1 変数による埋め込みを行えば、関係の薄い要素を予測モデルに含めないで、予測性能は比較的優れている。

さらに、3 章における実験において、改良 CML 上のノードの個性に応じた特性を調べるため、ミクロ的視点よりリアプノフ指数および予測難易度に関して分析した結果、

- ノードの中心性に応じて最大リアプノフ指数が変化し、それに伴い予測難易度も変化する。
- 上記の関連性の強さは、改良 CML のモデルパラメータによって変化する。
- ただし、媒介中心性においては最大リアプノフ指数や予測精度との関連性は見られない。

通常の複雑システムの解析や予測では、状態空間内の軌道を再構成するために式 (4) を用いるので、3 章で得られた知見は特に重要である。

しかし、これらの知見は CML のような基礎的モデルにより得られたものである。実世界に存在するより複雑なシステムにおいて、同種の傾向を見いだすことは簡単ではないかもしれない。しかし、基礎的モデルゆえ、一般の複雑システムに原理的に内在されている性質と捉えることができる。例えば、株価市場を予測する場合、本研究の知見によれば、各企業の関連会社の多さ（ネットワーク中心性に対応）に応じて、株価変動の予測難易度が異なる可能性を示唆している。もしこのような性質が原理的に内在されているならば、予測難易度が低い企業を特定し、予測対象として絞り込みができる。今後、株式市場に限らず様々な実システムにおいてこの可能性を検討し、本研究により得られた知見が有効活用できる事例を見つけていきたい。さらに、ネットワーク中心性が予測精度に作用するメカニズムを検討することも今後の課題である。

なお、本研究の一部は、文科省科研費若手研究（スタートアップ）(No.18800061) のご支援により行われました。

#### 参考文献

- 1) 金子邦彦，“複雑系のカオス的シナリオ，” 朝倉書店，(1997)。
- 2) D. Watts and S. Strogatz, Nature **393**, 440, (1998)。
- 3) A. Barabasi and R. Albert, Science **286**, 509, (1999)。
- 4) 増田直紀，“複雑ネットワークの基礎，” 数理科学 **518**, 36, (2006)。
- 5) 池口徹ら，“カオス時系列解析の基礎と応用，” 合原一幸編，産業図書，(2000)。
- 6) J. Kaplan and J. Yorke, Lecture Notes in Math. **730**, 204, (1979)。
- 7) E. Lorenz, J. of the Atmos. Sci. **26**, 636, (1969)。
- 8) F. Takens, Lecture Notes in Math. **898**, 366, (1981)。
- 9) T. Sauer, et. al., J. of Stat. Phys. **65**(3/4), 579, (1991)。
- 10) H. Abarbanel and M. Kennel, Phys. Rev. E **47**(5), 3057, (1993)。