

機能クラスに基づく複雑ネットワークのモデル推定

Model estimation for complex networks based on functionality classes

東京大学大学院工学系研究科 ○内田 誠
 東京大学人工物工学研究センター 白山 晋

複雑ネットワークとして表されるシステムについて、ネットワークの構造そのものに基づくのではなく、そこから生じる機能に基づくクラス分類と、そのクラスに基づいてネットワークのモデルを推定する手法を提案する。提案手法を人工市場シミュレーションのためのモデルの推定に適用し、有効性を検証する。

1 はじめに

ネットワークによって表現されるシステムは広範な分野に存在する。そして、システムを構成するネットワークの構造は、それを外部環境としてシステムに引き起こされる動的なダイナミクスと密接に結びつく。そのようなダイナミクスは、ネットワークの機能 (function) として知られている^{1),9),3),7),4)}。

従来のネットワークのクラス分類やモデルの推定は、構造そのものに対して行われているものが主流であったといえる^{3),4)}。そのような構造のクラスやモデルを用いて、それを外部環境としたときにどのような機能が発現するかという問題に取り組んだ先行研究は多い^{9),3),7),4)}。しかしながら、それらはネットワークの構造としてあるクラスを持つネットワークモデルを明示的に仮定した上で、その上で生じる動的なダイナミクスを理論的に、あるいはシミュレーションによって分析するアプローチが主流であったといえる。そのため、現実のネットワークに還元する場合、その知見をそのまま適用できるとは限らないという問題がある。

そこで、本研究ではネットワークの構造そのものに基づくのではなく、そこから生じる機能に基づくクラス分類と、そのクラスに基づいたモデル推定の手法を提案する。提案手法を人工市場シミュレーションのためのモデル推定に適用することで、その有効性を検証する。

2 提案手法

ネットワークの機能 f は、ネットワークのグラフ構造 $G = (V, E)$ と、 G に応じて決定される任意の初期条件 $i(G)$ を入力として、 $f(G, i(G))$ として定義することができる。 $i(G)$ は機能 f に与えられる何らかの初期条件である。たとえば、 f を後に扱う人工市場モデルによって定義される機能モデルとした場合、時刻 $t = 0$ における各エージェント $v \in V$ の状態を $i(G)$ として考えることができる。

機能 f に対して G と $i(G)$ が与えられたとき、出力 o を

$$o = f(G, i(G)), \quad (1)$$

として考える。

提案手法は、機能に対する入力 $i(G)$ と出力 o との関係から、機能の逆関数 f^{-1} によって、以下のように G を求めるという考えに基づく。

$$G = f^{-1}(i(G), o). \quad (2)$$

提案手法を概念図として表したものを Fig. 1 に示す。このように、対象と同じ入出力の関係を与えるネットワーク G を推定するというのが提案手法の核である。

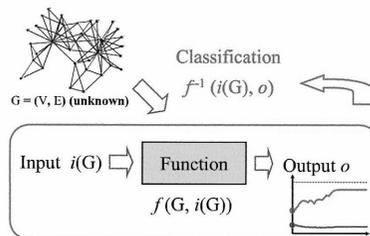


Fig. 1 Framework of network classification based on network functionality.

ところが、機能 f は陽に表現できることは少なく、仮にできたとしても f^{-1} を求めることは難しい。そこで、既知のネットワークモデル g_m をリファレンスとして、機能 f に対する入出力関係をあらかじめ求め、それらをクラスに分類する。そして、未知のネットワークの入出力関係がどのクラスに属するかを調べることで G を推定する。

機能モデル f によって生じる入出力関係のパターンを P とする。既知の g_m によって $P(i(g_m), f(g_m, i(g_m)))$ を求め、それをクラス c_α に分類する。本研究ではこのクラスを機能クラスと呼ぶ。未知のネットワークに対する入出力関係がどの機能クラスに属するかを、適当な距離関数を定めることで以下のように求める。

$$\beta = \arg \min_{\alpha} \|P(i(G), f(G, i(G))) - c_\alpha\|. \quad (3)$$

未知のネットワーク G が機能クラス c_β に属するという結果が得られれば、そのクラスに属する既知のネットワークモデルによって、同等の機能を再現するモデルが推定される。

3 人工市場シミュレーションにおけるモデル推定

提案手法であるモデル推定の方法論を、人工市場シミュレーションを題材として、エージェントの相互作用ネットワークのモデル推定に適用し、有効性を検証する。

具体的には、携帯電話やメール等の利用を念頭においたコミュニケーションサービスの市場を題材とする。二つの事業者がシェアを競争している市場を想定し、利用者のコミュニケーションのネットワーク構造と、そのどちらかを利用する利用者の意思決定を、人工市場のモデルとして考える。そして、市場における2事業者のシェアの変遷を分析する。提案手法によってコミュニケーションネットワークの構造のモデルを推定することを試みる。

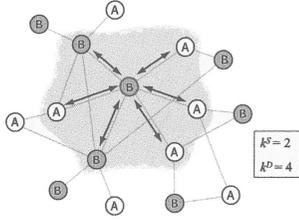


Fig. 2 Schematic representation of communication network.

3.1 人工市場モデル

サービスの利用者をエージェントとするマルチエージェントシステムを考える。エージェントのコミュニケーションの空間構造を Fig. 2 に模式的に示すようなネットワーク構造によって定義する。Fig. 2 では、中央の○で示したエージェントは、事業者 B を利用していて、コミュニケーション相手のエージェントの数が 6 (次数が 6) であり、事業者 A の利用者 4 人、事業者 B の利用者 2 人と恒常的にコミュニケーションを取っていることを表している。

実際の携帯電話利用者の調査では、利用者の次数分布はピークを持たずに高次の部分が厚い、いわゆるヘビータイルな確率分布になることが報告されている^{11),10)}。そこで、そのような空間構造をもつネットワークモデルをリファレンスとして用いる。本稿では、リファレンスとして BA モデル²⁾、KE モデル ($\mu = 0.1$)⁸⁾、CNN モデル¹²⁾を考え、比較のために ER ランダムグラフ⁹⁾、WS モデル¹³⁾ および格子 (Lattice) を加えた 6 種類のネットワークモデルを用いた。

3.2 エージェントの意思決定モデルと効用関数

市場における利用者はエージェントとしてモデル化される。エージェントの効用関数として、簡単のために金銭コストに関する部分のみを考える。利用費用はサービスの利用量およびコミュニケーション相手の人数に依存する費用と、別の事業者に変更する際に発生する乗り換え費用からなるものとし、エージェントは金銭コストを最小化するように事業者を選択するものとする。このとき、単位期間あたり相手 1 人あたりのサービス利用量を x_i とするとき、エージェント i の効用関数 u を次式とする。

$$u(x_i, \alpha^S, \alpha^D, k_i^S, k_i^D, \delta) = -\min \left\{ (\alpha^S k_i^S + \alpha^D k_i^D) x_i, (\alpha^S k_i^D + \alpha^D k_i^S) x_i + \delta \right\}. \quad (4)$$

ここで、添え字 S および D は、それぞれエージェント i が利用するものと同じ事業者、および違う事業者をあらわす。 k_i^S および k_i^D は、エージェント i の周囲で、それぞれ i と同じ、および異なる事業者を利用するエージェントの数である。 α^S および α^D は、それぞれ同じおよび異なる事業者を利用する相手に対する単位量あたりのサービス利用費用を表す。本稿では $\alpha^S < \alpha^D$ なる定数であることを仮定する。 δ は事業者を変更するための乗り換え費用であり、全エージェントについて一律の定数とする。

x_i はある分布 \mathcal{X} に従ってエージェント毎に決定する

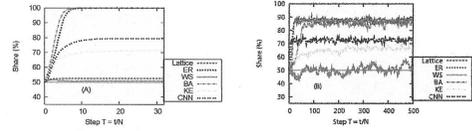


Fig. 3 Temporal evolution of the winner's share. $\langle k \rangle = 10$, $\alpha_S = 20$, $\alpha_D = 25$, $\delta = 500$, (A) $\mathcal{X} = \mathcal{X}_1$ (B) $\mathcal{X} = \mathcal{X}_2$.

ものとし、次の 2 種類のモデルを考える。第一のモデルでは、相手 1 人あたりのサービス利用量は全エージェントについて、平均 \bar{x} に等しいものと仮定する。このとき、エージェント i の総サービス利用量は $(k_i^S + k_i^D) \bar{x}$ となる。第二のモデルでは、各々のエージェントの総サービス利用量 X_{total} が正規分布に従うものと仮定する。すなわち、相手 1 人あたりの利用量 x_i は $x_i = X_{total} / (k_i^S + k_i^D)$ となる。以降、前者をモデル I、後者をモデル II とする。また、モデル I およびモデル II によって決定される x_i を与える分布を \mathcal{X}_1 および \mathcal{X}_2 とする。

これらのモデルに従い、各エージェントは効用関数が最大になるように、現在の事業者を利用し続けるか、異なる事業者に乗り換えるかの意思決定を行う。具体的には、ある時刻 t にエージェントを一つランダムに選択し、式 (4) に従って u を最大化するように、次に利用する事業者を選択する。

3.3 機能モデル表現による定式化

第 2 節で提案したフレームワークと対応づけると、利用者の効用関数および意思決定のモデルが機能モデル f 、利用者のコミュニケーションネットワークが G 、初期状態における 2 事業者の割合 (以降、事業者のシェアと呼ぶ) および分布が初期入力 $i(G)$ 、最終的な 2 事業者のシェアおよび分布が出力 o に対応する。すなわち、エージェントのコミュニケーションを形成するネットワーク構造が市場のモデル、そのネットワークの機能がエージェントの相互作用モデルとなる。また、 x_i の分布 \mathcal{X} 、 α^S 、 α^D および δ は相互作用モデル自身のパラメータであり、これらが機能のパラメータになる。このように、機能モデルに対していくつかの説明変数が陽に表しうるパラメータの形で存在する。そこで、これらのパラメータを \mathbf{p} で表し、式 (1) を以下のように拡張する。

$$o = f^*(G, i(G), \mathbf{p}). \quad (5)$$

本稿の機能モデルでは $\mathbf{p} = (\mathcal{X}, \alpha^S, \alpha^D, \delta)$ である。

4 計算機実験

4.1 準安定状態

初期状態における 2 事業者の割合を r 、終状態において卓越した事業者のシェアを r' で表す。また、二つの事業者の α^S 、 α^D および δ のパラメータはすべて等しく設定する。すなわち、2 事業者のサービスやコストには差異がない状態を考える。

予備事件として、Fig. 3 に $\mathcal{X} = \mathcal{X}_1$ および $\mathcal{X} = \mathcal{X}_2$ の意思決定モデルによるシミュレーションのある 1 回の試行における事業者のシェアの典型的な時間発展を示す。初期状態は 2 事業者が市場全体に 50% ずつのシェアをランダムに持った状態 ($r = 0.5$) とした。エージェント数は $N = 3600$ とした。

Fig. 3 からわかるように、 $\mathcal{X} = \mathcal{X}_1$ 、 \mathcal{X}_2 いずれの意思決定モデルにおいても、初期状態におけるシェアが等しいにもかかわらず、いずれかの事業者が卓越する準安定

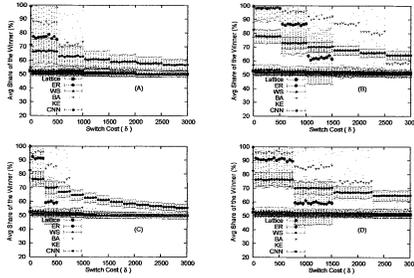


Fig. 4 Functional relationships of δ and r' on network models ($\mathcal{X} = \mathcal{X}_1$, (A) $\bar{X} = 100$, $\langle k \rangle = 8$, (B) $\bar{X} = 100$, $\langle k \rangle = 10$, (C) $\bar{X} = 50$, $\langle k \rangle = 10$, (D) $\bar{X} = 150$, $\langle k \rangle = 10$).

状態に達する。2事業者のパラメータを対称に設定したので、「どちらが」卓越したかという点は意味を持たないが、十分時間が経過した後に卓越している事業者（以降“勝者”と呼ぶ）のシェアが収束する値は試行によらず一定の範囲になることを確認している。

4.2 入出力関係

初期状態における2事業者の割合 r とその分布 I は、3.3項で定式化した機能モデル f^* の入力である。初期状態として2事業者が市場全体に50%ずつのシェアをランダムに持った状態は、 $i(G) = (r, I)$ において、 r を $r = 0.5$ 、 I としてランダムな分布とした場合に相当する。同様に、終状態における勝者のシェア r' は機能 f^* の出力 o として考えることができる。

本節では、 $\mathcal{X}, \alpha^S, \alpha^D$ を固定し、ある初期シェア r とその分布 I に対して、 r' と δ の関係を調べる。そこで、式(5)で定義される機能モデルを改めて \tilde{F} とすれば、

$$r' = \tilde{F}(G, r, I, \delta). \quad (6)$$

と表される関係を調べることになる。

後述する計算機実験では、初期状態におけるシェア r を r_0 に、またその分布 I をランダムな分布に固定している。したがって、 $r-r'-\delta$ 空間において $r = r_0$ 平面での乗り換え費用 δ と勝者のシェア r' の関係を機能クラスとして求めることに相当する。

4.3 ネットワークモデルを用いたシミュレーション

まず、 G としてリファレンスネットワークモデルを用いたシミュレーションを行い、結果について考察する。以降の数値実験では、初期状態におけるシェア r とその分布 I は、それぞれ $r = 0.5$ およびランダムな分布に固定した。また、 $\alpha^S = 20$ 、 $\alpha^D = 25$ とした。

Fig. 4に、 $\mathcal{X} = \mathcal{X}_1$ の意思決定モデルによる、乗り換え費用 δ と勝者のシェアとの関係を示す。図中で、誤差棒は500回の試行における標準偏差を示している。同様に、Fig. 5および $\mathcal{X} = \mathcal{X}_2$ の意思決定モデルによる δ と r' との関係を示す。分布 \mathcal{X}_2 として、各エージェントの総コミュニケーション量 $X_{total} = x_i k_i$ が平均 $\bar{X}(k)$ 、標準偏差 $\bar{X}(k)/3$ に従う正規分布を仮定した。

4.4 機能に基づくクラス分類

ネットワーク G やその他のパラメータの違いによって、勝者のシェア r' と乗り換え費用 δ との関係にいくつかのパターンが確認される。

$\mathcal{X} = \mathcal{X}_1$ による δ と r' の関係は、Fig. 6に示すクラスに分類できる。Fig. 6で、クラス(a)は、すべての δ の

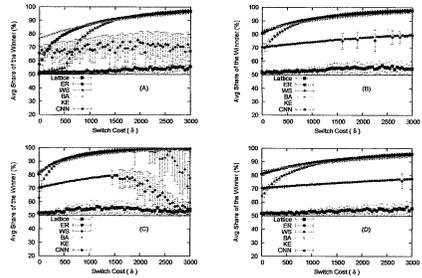


Fig. 5 Functional relationships of δ and r' on network models ($\mathcal{X} = \mathcal{X}_2$, (A) $\bar{X} = 100$, $\langle k \rangle = 8$, (B) $\bar{X} = 100$, $\langle k \rangle = 10$, (C) $\bar{X} = 50$, $\langle k \rangle = 10$, (D) $\bar{X} = 150$, $\langle k \rangle = 10$).

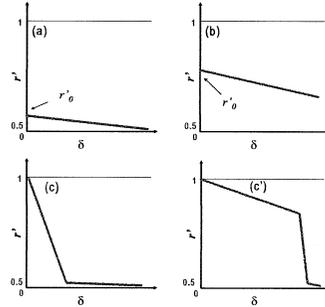


Fig. 6 Typical patterns in the δ - r' relations ($\mathcal{X} = \mathcal{X}_1$).

Table 1 Functional classes of the network models ($\mathcal{X} = \mathcal{X}_1$).

	$\bar{X} = 50$ $\langle k \rangle = 10$	$\bar{X} = 100$ $\langle k \rangle = 10$	$\bar{X} = 100$ $\langle k \rangle = 8$	$\bar{X} = 150$ $\langle k \rangle = 10$
Lattice	(a)	(a)	(a)	(a)
ER	(c)	(c)	(c)	(c)
WS	(a)	(a)	(a)	(a)
BA	(c)	(c')	(c)	(c')
KE	(a)	(a)	(a)	(a)
CNN	(b)	(b)	(b)	(b)

領域において、 r' が $r' \approx 0.5$ に近い値をとるクラスである。クラス(b)は、 r' が $r' > 0.5$ なる値をとるクラスであり、 $\delta \geq 0$ の領域でも $r' \approx 0.5$ には達しない。クラス(c)は、 $\delta = 0$ のときに $r' \approx 1$ となり、 δ の増加と共に $r' \approx 0.5$ に漸近するクラスである。 $r' \approx 0.5$ となる δ の違いにより、(c) および (c') の二つの種類がみられる。

同様に、 $\mathcal{X} = \mathcal{X}_2$ による δ と r' の関係は、Fig. 7に示すクラスに分類できる。Fig. 7で、クラス(a)は実験を行った δ のすべての領域で分散が小さなクラスである。このとき、 r' は δ に対して増加する関数となる。 $\delta \approx 0$ 付近における変曲点の有無により、(a) と (a') の二つの派生クラスに分けることができる。クラス(b)は、すべての δ の領域で分散が大きいクラスである。 r' のばらつきが大きく、明確な関係を持たないパターンとみなすこともできる。誤差棒によって分散が大きいことを示している。平均の r' の大きさにより、(b) と (b') の二つに分類できる。クラス(c)はクラス(a)および(b)の複合であ

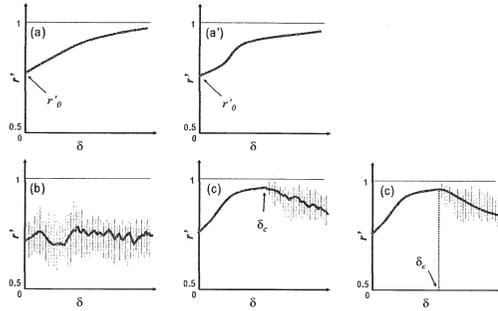


Fig. 7 Typical patterns in the δ - r' relations ($\mathcal{X} = \mathcal{X}_2$).

Table 2 Functional classes of the network models ($\mathcal{X} = \mathcal{X}_2$).

	$\bar{X} = 50$ $\langle k \rangle = 10$	$\bar{X} = 100$ $\langle k \rangle = 10$	$\bar{X} = 100$ $\langle k \rangle = 8$	$\bar{X} = 150$ $\langle k \rangle = 10$
Lattice	(b')	(b')	(b')	(b')
ER	(a)	(a)	(a)	(a)
WS	(c)	(a)	(a')	(a)
BA	(c)	(a)	(a)	(a)
KE	(b)	(b)	(b)	(b)
CNN	(c)	(a)	(b)	(a)

Table 3 Functional classes email correspondence network.

	\bar{X}	10	25	50	100	150	300
$\mathcal{X} = \mathcal{X}_1$		-	-	(b)	(b)	(b)	-
$\mathcal{X} = \mathcal{X}_2$		(c)	(c)	(c)	(a)	(a)	(a)

る。特定の値 δ_c より小さい δ の領域では r' の分散が小さく、 $\delta > \delta_c$ では分散が大きい。 $\delta = \delta_c$ において、 r' は極大値をとる。

リファレンスネットワークモデルによるシミュレーションの結果は、これらのクラスによって Table. 1 および Table. 2 のようにまとめられる。このように、人工市場シミュレーションにおけるエージェントの相互作用と意思決定という機能モデルによって、ネットワークモデルはそれらを特徴づけるいくつかの機能クラスに分類することができる。

4.5 実際のネットワークを用いたシミュレーション

次に、現実のネットワークが属する機能クラスを通じて、実際の市場のネットワークモデルとして再現性の高いモデルについて考察する。

例として、ある組織のEメール送受信のネットワークの実データを用いる。このデータは文献⁶⁾で分析されたものであり、ある期間に少なくとも1往復のEメールのやりとりがあったペアにエッジを定義したものである。1133人(ノード)に5451本のエッジが定義され、平均次数は $\langle k \rangle = 9.62$ である。

Fig. 8に、このネットワークを導入したシミュレーションの結果を示す。これらの結果は、4.4項で分類した機能クラスを用いて Table. 3 のようにまとめられる。

$\mathcal{X} = \mathcal{X}_1$ で比較的大きな δ に対しても r' が比較的緩やかに低下するクラス (b) や、 $\mathcal{X} = \mathcal{X}_2$ で同程度の \bar{X} において近い $\delta = \delta_c$ の値で明確な極大値がみられるクラ

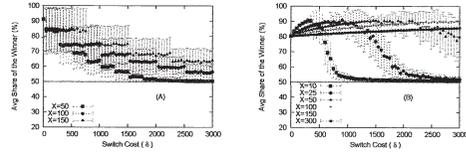


Fig. 8 Functional relationships of δ and r' (on email correspondence network, (A) $\mathcal{X} = \mathcal{X}_1$, (B) $\mathcal{X} = \mathcal{X}_2$).

ス (c) が観察され、CNN モデルと同じ機能クラスに分類される。

一方、Eメールネットワークと同程度のクラスターリング係数 C や平均経路長 L を再現するWSモデルやKEモデルは、実データとは異なるクラスに属するシミュレーション結果を得ている。このように、本稿で提案する機能クラスに基づくモデル推定手法は、ネットワークの統計的な構造からは明らかにならない機能の差異を適切に反映し、実際のシステムと同様の機能を再現するモデルを推定することが可能であるといえる。

5 おわりに

本稿では複雑ネットワークで表されるシステムに対する入出力の関係性をネットワークの機能と定義し、その入出力パターンを機能クラスとして分類する方法論を提案した。また、機能クラスに基づいて現実の複雑ネットワークのモデルを推定する手法を提案した。

提案手法を人工市場シミュレーションのためのエージェントの相互作用ネットワークのモデル推定に適用し、現実のコミュニケーションネットワークと同じダイナミクスを機能として再現するモデルを選び出せることを確認した。このように、提案手法は従来のネットワークの統計的な指標によっては明らかにならない構造の差異と、それに起因する機能の違いを適切に反映するモデル推定を可能にするものである。

参考文献

- 1) A.-L. Barabási. *Linked*. Perseus Publishing, 2002.
- 2) A.-L. Barabási and R. Albert. *Science*, **286**, pp. 509–512, 1999.
- 3) S. Boccaletti, et al.. *Phys. Rep.*, **424**, pp. 175–308, 2006.
- 4) L. d. F. Costa, et al.. arXiv:0711.3199v3, 2008.
- 5) P. Erdős and A. Rényi. *Publicationes Mathematicae*, **6**, pp. 290–297, 1959.
- 6) R. Guimerà, et al.. *Phys. Rev. E*, **68** (065103), 2003.
- 7) 林幸雄. ネットワーク科学の道具箱 -つなかりに隠れた現象をひもとく-. 近代科学社, 2007.
- 8) K. Klemm and V. M. Eguiluz. *Phys. Rev. E*, **65** (057102), 2002.
- 9) M. E. J. Newman. *SIAM Rev.*, **45**, pp. 167–256, 2003.
- 10) J.-P. Onnela, et al.. *PNAS*, **104**, pp. 7332–7336, 2007.
- 11) 佐々木純, 他. 電子情報通信学会ソサイエティ大会講演論文集, **2**, p. 168, 2005.
- 12) A. Vázquez. *Phys. Rev. E*, **67** (056104), 2003.
- 13) D. J. Watts and S. H. Strogatz. *Nature*, **393**, pp. 440–442, 1998.