

Social Network Rebuilder:
金融不安伝播ネットワーク推定ツール

市川昊平・竹村敏彦・村上雅俊・峰滝和典・前田太陽



文部科学大臣認定 共同利用・共同研究拠点
関西大学ソシオネットワーク戦略研究機構
関西大学ソシオネットワーク戦略研究センター
(文部科学省私立大学学術フロンティア推進拠点)

Research Center of Socionetwork Strategies,
“Academic Frontier” Project for Private Universities, 2003-2009
Supported by Ministry of Education, Culture, Sports, Science and Technology
The Research Institute for Socionetwork Strategies,
Joint Usage / Research Center, MEXT, Japan
Kansai University
Suita, Osaka, 564-8680 Japan
URL: <http://www.rcss.kansai-u.ac.jp>
<http://www.kansai-u.ac.jp/riss/index.html>
e-mail: rcss@ml.kandai.jp
tel: 06-6368-1228
fax. 06-6330-3304

Social Network Rebuilder: 金融不安伝播ネットワーク推定ツール

市川昊平^{*}・竹村敏彦[†]・村上雅俊[‡]・峰滝和典[§]・前田太陽[¶]

概要

本研究の目的は、アンケート調査などの手段で収集されたデータから、人々が構成する社会ネットワークの推定手法を確立することである。社会現象をミクロな視点から分析するうえで、社会ネットワークに関する研究は近年盛んに行われている。社会現象の多くはミクロな視点で分析すると、社会を構成する人々の社会ネットワーク上における相互作用によって形成されていることが分かる。情報・噂の伝播やパンデミックなどがその典型的な例である。しかしながら、今日までに、実際の社会ネットワークを観測されたデータなどに基づいて分析・シミュレートする手法はまだ確立されていない。本研究はこの問題に対し、メタヒューリスティックな解決手法の一つである遺伝的アルゴリズム (GA) を用い、社会ネットワークを間接的に観測したデータから、実際の社会ネットワークを再構築する手法の確立を目指す。その結果として、本稿では分析対象となる社会ネットワークの次数分布を用いて、その社会ネットワークの特徴をある程度反映したネットワークを再構築可能であることを示す。

キーワード: 社会ネットワーク, 遺伝的アルゴリズム, 社会シミュレーション

* 大阪大学情報基盤推進本部助教

関西大学ソシオネットワーク戦略研究機構研究員

関西大学ソシオネットワーク戦略研究センター研究員

Email: ichikawa@cmc.osaka-u.ac.jp

† 関西大学ソシオネットワーク戦略研究機構助教

関西大学ソシオネットワーク戦略研究センター研究員

Email: takemura@rcss.kansai-u.ac.jp

‡ 関西大学ソシオネットワーク戦略研究機構助教

関西大学ソシオネットワーク戦略研究センター研究員

Email: murakami@rcss.kansai-u.ac.jp

§ 関西大学ソシオネットワーク戦略研究機構統計分析主幹

関西大学ソシオネットワーク戦略研究センター研究員

Email: minetaki@rcss.kansai-u.ac.jp

¶ 関西大学ソシオネットワーク戦略研究機構ポストドクトラルフェロー

Email: maeda@pglab.kansai-u.ac.jp

Social Network Rebuilder: A Tool to Estimate Social Network Toward Bank Run Simulation

KOHEI ICHIKAWA*, TOSHIHIKO TAKEMURA†, MASATOSHI MURAKAMI‡,
KAZUNORI MINETAKI§, TAIYO MAEDA¶

Abstract

The objective of this paper is to establish a new method to estimate real social networks based on observed data which is collected by questionnaire surveys. The studies on social networks have been increasing today to analyze social phenomena from a micro viewpoint. Most social phenomena can be explained by interactions among people with micro viewpoint. Spread of rumor and pandemic are the typical examples of that. However, methods which allows us to analyze real social networks based on observed data have not been developed well. This study tries to exploit genetic algorithm (GA) method, which is one of meta-heuristics methods, for this problem and establish a method to rebuild a social network based on the data which is observed indirectly from real social networks. This paper introduces that our proposed method allows us to rebuild a social network to some extent from degree distributions of a target real social network.

KEYWORD: Social Network, Genetic Algorithm, Social Simulation

* Assistant professor, Office for Information Infrastructure Design, Osaka University
Researcher, RISS, Kansai University
Research fellow, RCSS, Kansai University
Email: ichikawa@cmc.osaka-u.ac.jp

† Assistant professor, RISS, Kansai University
Research fellow, RCSS, Kansai University
Email: takemura@rcss.kansai-u.ac.jp

‡ Assistant professor, RISS, Kansai University
Research fellow, RCSS, Kansai University
Email: murakami@rcss.kansai-u.ac.jp

§ Senior Researcher for Statistical Analysis, RISS, Kansai University
Research fellow, RCSS, Kansai University
Email: minetaki@rcss.kansai-u.ac.jp

¶ Postdoctoral fellow, RISS, Kansai University
Email: maeda@pglab.kansai-u.ac.jp

1 はじめに

社会ネットワークとは、友人関係や親類、コミュニケーション、金銭的やり取り関係など、社会的な関係構造を表したネットワークである [1]。社会を構成する人々の間で伝播する情報はこの社会ネットワークを基盤に伝播していくことから、多くの社会現象をミクロな視点から分析するには社会ネットワークの構造を詳細に分析することが非常に重要となる。例えば、金融不安に関する情報伝播や年金未納行動の伝播など、人々の間で交換される噂や情報が社会現象を形成するようなモデルにおいては、社会ネットワークの分析は重要である。

このような人々の間で伝播される情報に基づいて発生する社会現象をミクロな視点で分析・シミュレートする方法としてマルチエージェントシミュレーション (MAS) [2] が例として挙げられる。MAS において社会ネットワークを扱う際には、ネットワークのノードに当たる部分をエージェントとして設計し、ネットワークのエッジは各エージェントが情報を収集する際の情報チャンネルとして定義される。そして、各エージェントは情報チャンネルから得られる情報と自身の属性や過去の経験・情報から何らかの意思決定をし、行動を起こすものとして設計される。われわれが本研究の延長上に見据えている金融不安の伝播シミュレーションもこのような概念に基づいている。

社会科学分野において社会ネットワークに関する研究に注目が集まりつつある一方で、現実の社会ネットワークをシミュレーションの対象として取り扱うのは極めて困難であるのが現状である。その一番の原因は社会ネットワーク全体を直接的に観測することが不可能である点にある。そのため、現在はある限定されたコミュニティの内部においてのみ社会ネットワークを観測した結果、または制御した環境において得られた実験結果から実際の社会ネットワークの振る舞いを推測したり、一般的な社会ネットワークを統計的に分析し、そこで見いだされた社会ネットワークが従う普遍的なルールに基づきシミュレーション対象となる社会ネットワークをモデル化したりといった手法にとどまっている。例えば、BA モデル [3, 4] が定義するスケールフリーネットワークや WS モデル [5] が定義するスモールワールドネットワークは、一般的な社会ネットワークにおいては、次数分布がべき法則に従ったり、平均最短距離がある特定の値以下になることを主張し、そのような特性を持つネットワークを仮想的に構築するアルゴリズムを提供している。しかし、これらのモデルはある特定の社会ネットワークに対して、べき法則の係数や平均最短距離の最大値などの具体的な値を提示するものではなく、これらのモデルで仮想的に構築されたネットワークは経験的な仮定に基づいて成り立っている。したがって、これらのモデルを用いたシミュレーションはそのシミュレーションが対象とする現実の社会ネットワークを必ずしも反映しているとは言い難いモデルの上での議論にとどまっているのが現状である。

本研究は、このような現状を鑑み、分析・シミュレーションの対象となる社会ネットワークを構成する人々に対し Web ベースのアンケート調査などを実施し、アンケート調査で得られたデー

タに基づいて、対象となる社会ネットワークを推定する手法の確立を目的とする。本研究で行う Web ベースのアンケート調査も社会ネットワーク全体を直接的に観測する手法ではない。本研究は、社会ネットワークを構成する一部の人々から間接的に観測されたデータを基に対象となる社会ネットワークの特性を抽出し、その特性を再現する社会ネットワークを仮想的にコンピュータ上で再構築する手法を開発することを目指している。

本論文の構成は以下の通りである。2 節ではまず本研究が取り組む社会ネットワークの推定問題に関して定義する。3 節では、2 節いで定義された問題の解決手法として本稿が提案する手法に関して解説する。4 節では提案手法の評価を行うため、実験的に生成した既知の社会ネットワークに関して、本手法がどの程度有効であるか評価し、さらに 5 節では実際にわれわれが実施した Web ベースのアンケート調査結果に適用した例を示す。最後に 6 節にて、まとめと今後の課題に関して述べる。

2 社会ネットワークの推定問題

本節では、本研究で取り組む社会ネットワークの推定問題を定義する。われわれは構築した社会ネットワーク上における金融不安の伝播シミュレーションを実施することを本研究の延長上に考えている。そのため、本研究ではソシオネットワーク戦略研究機構が収集・管理している共同利用・共同研究データベースの一つである、金融行動データベース II (Web アンケート：2009 年 3 月実施、個票数：1677) [6] を用いている。以下の問題定義では、この Web アンケート調査を適宜例示し、説明する。

1 節で述べた通り、社会ネットワーク全体を直接的に観測する手段は存在しない。したがって、観測可能な一部のデータから元の社会ネットワークを推定する手段が必要とされている。推定の流れは図 1 に示す通りとなる。まず、1) 推定対象となる未知の社会ネットワークに対し、アンケート調査等による観測を行い、2) 観測されたデータからネットワークの特性情報、例えば次数分布

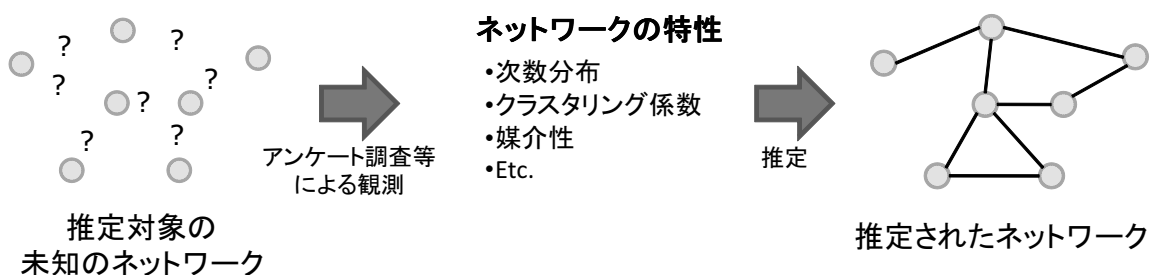


図 1 社会ネットワークの推定

やクラスタリング係数、媒介性などを抽出し、3) 抽出されたネットワーク特性を基に元のネットワークの形を推定するという流れになる。

ここで、次数分布とはあるノードから他の複数ノードへの接続の数（次数）の分布である。一般的な社会ネットワークでは次数分布は知人の数の分布として定義されているものであり、べき法則に従うものと仮定されている。一方で、われわれの研究対象は金融不安情報の伝播であるため、ここでいう次数分布とは金融不安情報を流す可能性のある知人の数の分布となる。したがって、金融行動データベース II の Web アンケート調査では「あなたは金融機関の破たんなどの噂を聞いた場合、それを何人の知人に伝えますか？」というような設問を設置している。このようなある特定の問題領域に基づいたネットワーク特性は一般的な社会ネットワークの仮定を前提として BA モデルや WS モデル等を用いてモデル化できるとは限らないため、本研究が取り組むような観測データからのネットワークの推定手法の確立が必要である。

次数分布はある程度直接的な設問で観測可能な情報であるが、クラスタリング係数や媒介性などはアンケート回答者も把握していない情報であるため、直接的に観測することが困難な情報である。例えば、クラスタリング係数はある人物 A を中心にその知人 B, 知人 C 同士も知人関係であるというような三角形の関係がどの程度存在するかを表す指標であるが、アンケート回答者に自身の知人全てに関して、このような三角形の関係を全て数え上げるように指示しても正確な回答が得られる可能性は低い。そのため、金融行動データベース II の Web アンケート調査では、回答者の知人の中で交友関係が広く一番知人の数が多いであろうという考えられる方に関して、その方の知人と自身の知人で共通の知人は何人いるかというような、ある程度限定した範囲の情報を観測するにとどめている。このような現状から、現実の社会ネットワークを対象とした推定においては、不完全な情報を基に推定を実施する必要がある。

現実の社会ネットワークの推定は、このような間接的に観測された不完全な情報からの推定問題となるため、解析的に推定できるような問題ではない。したがって、解の候補となるネットワーク構成を一つ一つ調べ、そのネットワークの特性が観測された情報と一致する解を探索する方法が必要となる。しかしながら、考えられる解の候補全てに関して、このようなチェックをすることは現実的ではないことは明白である。例えば、エージェント数（ノード数）が k の社会ネットワークのモデルを作成する場合、考えられるネットワーク構成の候補数 N は以下の式で表される。

$$N = 2^C, \quad C = \frac{k(k-1)}{2} \quad (1)$$

ここで、 C はノード数 k のネットワークにおいて考えられる全てのエッジ数である。 N はその全てのエッジにおいて、接続があるか、ないかの 2 通りを考えたネットワーク構成の数であるので、式 (1) のように求められる。したがって、わずか $k = 10$ の場合であっても、 N は約 3.5×10^{13} に達することが分かる。

以上のことから、社会ネットワークの推定問題は NP 困難な問題に属することが分かる。した

がって、観測されたネットワークの特性情報に完全に一致する最適解を求めることは現実的な時間の範囲では不可能である。そのため、求める解を完全な最適解ではなく準最適解に緩和し、何らかのヒューリスティックな解法を用いて解くことが有効であると考えられる。

3 遺伝的アルゴリズムを用いた社会ネットワークの推定手法

本研究では、2 節で説明した推定問題を解くにあたり、メタヒューリスティックな手法の一つである遺伝的アルゴリズム (GA) [7] を用いて社会ネットワークの推定を試みるプログラム、Social Network Rebuilder (SN_Rebuilder) を開発した。SN_Rebuilder は 2 節で述べたアンケート調査の情報を基に、コンピュータ上で仮想的に元の社会ネットワークを再構築するプログラムである。現在の実装ではアンケート調査結果のうち、次数分布のみを用いて推定する。本節は SN_Rebuilder が実装する GA を用いた社会ネットワーク推定手法に関して、その詳細について述べる。

GA は、解の候補をコンピュータで取り扱える何らかの形にエンコードし、それを遺伝子と見立て、それら解の候補に結びついた遺伝子の交叉、突然変異、自然選択を通して生物の進化の過程を模し、最適な解を探索する手法である。GA は多くの最適化問題に対して汎用的に適用できる手法であるが、その対象問題領域に応じて適切に解のエンコード、交叉、突然変異、自然選択操作と自然選択の際の解の選択基準に利用する目的関数を定義する必要がある。図 2 に GA を用いて社会ネットワークを推定する手法の概要を示す。図に示す通り、GA では最初にランダムに生成した解候補を目的関数によって評価し、その評価結果に応じて次の世代に残す解候補 (遺伝子) を決定し、交叉・突然変異を通して新たな世代の解を作成するという行為を繰り返すことによって、最適解を

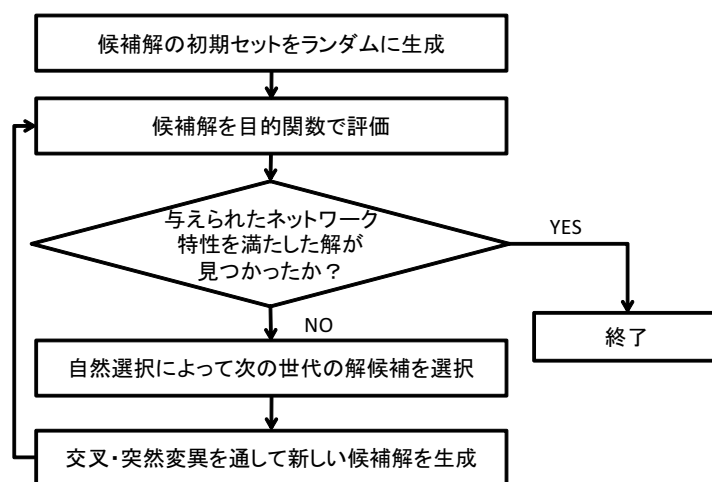


図 2 遺伝的アルゴリズムの概要

探索する。本節の残りの部分は、社会ネットワーク推定問題に対し GA を適用するにあたって、われわれが考案した目的関数の定義、解のエンコード手法、交叉、突然変異、自然選択操作について述べる。

3.1 目的関数の定義

GA における目的関数は、候補解がどれだけ最適解に近いかを評価するための指標として使われる。この目的関数を最大化または最小化させる解を見つけるのが GA での目的となる。本研究ではアンケート調査によって観測された次数分布を用いて社会ネットワークを再構築することを目的とするため、目的関数は候補解であるネットワークの次数分布が、観測された次数分布にどれだけ一致しているかを評価する関数として定義する。実際には、構築するネットワークモデルに応じて、ネットワーク全体のノード数は変化するため、次数分布そのものではなく、次数の確率密度分布を用いて評価する。以下に、提案した目的関数 F の定義を示す。

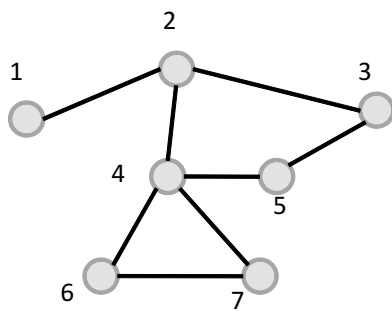
$$F = e^{-\varepsilon} \quad (2)$$

$$\varepsilon = \sum_{j \in J} \left(D(j) e^{kD(j)} \right), \quad D(j) = |PDF_{true}(j) - PDF_{cand}(j)| \quad (3)$$

ここで、 ε は候補解の次数分布が観測された次数分布とどれだけ差があるかを示す指標であり、完全に一致する場合は 0 となり、次数分布の差が大きいほど大きな値を示す。したがって、式 (2) で表される目的関数は候補解と観測された次数分布が一致する場合は、最大値 1 となり、差が大きくなるにつれて小さな値を取るようになる。 ε は式 (3) で表される。ここで、 $PDF_{true}(d)$ 関数は観測された社会ネットワークの次数の確率密度分布関数である。 $PDF_{cand}(d)$ 関数は候補解の次数の確率密度分布関数である。 J は取りうる次数のセットである。 k は定数である（本研究では $k = 10$ とした）。式 (3) は取りうる次数全てに関して、確率密度の差を求め、それを指数関数で評価した値を合計している。指数関数で評価することによって、分布において差異が大きい箇所があると、それがより大きく評価されて、目的関数 F の値に影響するように定義している。このように定義することによって、確率密度分布全体において差が少ない候補解の方をより高く評価し、分布の特定の箇所において差が大きいような候補解はより低く評価される。これにより、分布関数の形状をよりよく反映した候補解の方がより高く評価されるように工夫している。GA では、この目的関数 F を最大化するような解が探索される。

3.2 解のエンコード手法

GA では候補解をコンピュータで扱える数値や文字の配列として表現しなければならない。これをエンコードという。エンコード手法の選択は後述する交叉や突然変異の手順にも影響するため、対象とする問題領域に応じて適切に定義する必要がある。



	1	2	3	4	5	6	7
1		1					
2	1		1	1			
3		1			1		
4		1			1	1	1
5			1	1			
6				1			1
7				1		1	

図3 ネットワークの接続行列によるエンコードの例

本研究では、ネットワークの接続行列を用いて、候補解である社会ネットワークを数値列化する。図3にエンコードの例を示す。接続行列は行列の行・列それぞれがネットワーク上のノードに対応しており、ネットワーク上において接続がある場合は該当要素が1となる行列である。例えば図3の例では、ノード2はノード1, 3, 4と接続があるため、接続行列の2行目には1, 3, 4列目に1が入っている（同様に2列目には1, 3, 4行目に1が入っている）。

ネットワーク接続行列はネットワークを数学やコンピュータ上で取り扱う際に単純で分かりやすく、また各種の行列演算によって容易に扱えることから、広く利用されている。しかしながら、GAでそのまま用いる際には問題がある。GAでは、ある一つの候補解は一意に一つの数値列にエンコードされるべきであるが、接続行列はノード番号の振り方によって、その行列表現が変化する。そのため、同じネットワーク構造を示しているのにも関わらず、エンコードされた接続行列の表現が全く変わってしまう。GAでは候補解の中で評価の高い解を親とし、交叉させ、新しい候補解を作成するという操作が最適解探索の上での核心部分である。したがって、評価の高い候補解同士は似た接続行列として表現される必要がある。

本研究では、このような候補解の表現上の比較を正しく評価可能とするため、接続行列の正規化方法を提案する。つまり、ある一定のルールにしたがってノードの番号付けを行うことによって、ノード番号の振り方によって接続行列が変化しないエンコードを行う。今回開発したSN_Rebuilderは次数分布に基づいてネットワークを推定するプログラムであるため、次数に基づいて各ノードを順序付けし、その順序付けに従ってノード番号を振る。具体的には、以下の手順でノードの順序付けを行う。1) 各ノードを次数の大きさに降順に並べる。2) 同じ次数のノードがある場合は、その該当ノードの接続先ノードの次数の大きさに降順に並べる。このように二段階に順序付けすることによって、同じ次数を持つノードを適切に順序付けできる。

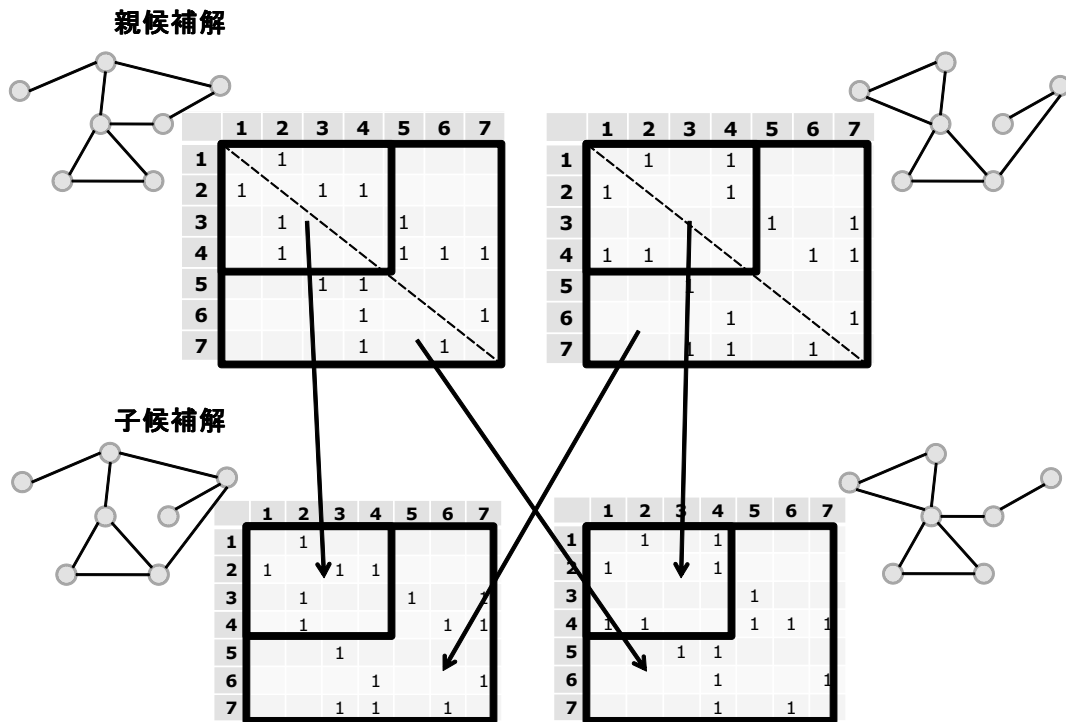


図 4 交叉の例

3.3 交叉操作

GA における交叉は現在の候補解を親とし、親の特徴を引き継ぎつつも、新たな候補解を生成する操作である。本研究では具体的には図 4 のように、2 つの候補解を親とし、接続行列を小四角形と L 字型部分に切り分け、それらの部分を交換した接続行列を子候補解として生成する手法をとる。切り分ける箇所は交叉ごとにランダムに選択する。接続行列は対称行列のため、整合性を保つためには、切り分ける際もこのように対称となるように小四角形と L 字型に分けなければならない。

図 4 では左上に小四角形、右下に L 字型ができるように切り分けているが、実際には左上に L 字型、右下に小四角形を作る切り分け方もあり得る。したがって、実際の実装では 2 つの親候補解から切り分け方の異なる 4 つの子候補解を生成し、その中で目的関数による評価が高い 2 つを子候補解として次世代に残すようにしている。接続行列は前節に述べた正規化のため、次数の大きいノード順に上から並んでいるため、前者の切り分け方では次数が大きいノード同士の接続が子候補解によりよく継承され、後者の切り分け方では次数が小さいノード同士の接続が子候補解によりよ

く継承される傾向にある。以上の仕組みにより、親候補解の特徴をそれぞれ継承した子候補解が生成される。

3.4 突然変異操作

突然変異は GA の候補解探索が局所最適解に陥ってしまい、真に最適な解への探索が中断されてしまうことを避けるため、解候補の値をランダムに変化させる操作である。本研究では図 5 に示す通り、ランダムに選択されたノード間を接続したり、切断したりすることによって実現する。ただし、接続行列が対称行列であることには注意を払って、図のように対称な 2 箇所を同時に更新する必要がある。

3.5 自然選択操作

本研究では選択手法に、GA に一般的に利用されているルーレット選択を用いる。ルーレット選択は目的関数による評価の大きさにより、次世代に残す候補解の選択の確率を決定する選択手法である。例えば、候補解 i の選択確率は以下のように定義される。

$$P_i = \frac{F_i}{\sum_{k \in N} F_k} \quad (4)$$

ここで、 F_i は候補解 i の目的関数による評価である。 N は候補解全体を表すセットである。つまり、候補解全体の評価の合計値で、候補解 i が占める割合に応じて選択確率が決定される。このような仕組みにより、評価値が低い候補解も一定の確率で選択される可能性を有している。GA では真に最適な解を探索するためには局所解に陥ってしまうことを避け、候補解の多様性を維持することが重要であることが一般的に言われている。したがって、評価値が高い候補解のみを残すのでは

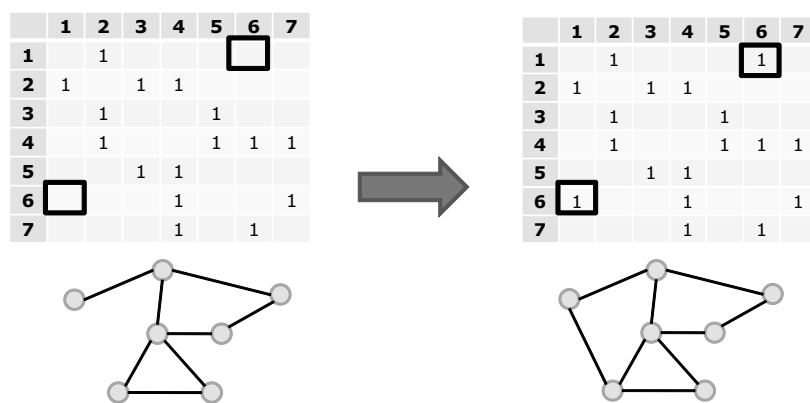


図 5 突然変異の例

なく、評価値が低い候補解であっても一定の割合で次世代に残すこういった仕組みが重要となる。

4 提案手法の評価

本研究の目的は図 1 に示した通り、未知の社会ネットワークを観測されたデータから推定し、ネットワークを再構築することであるが、本節では提案した手法の有効性を評価するため、まずは構造が既知のネットワークに対し提案手法による推定がどの程度有効であるか評価する。そのために、まず、擬似的にネットワークを作成し、そのネットワークの次数分布を観測する。そして、その観測された次数分布を基にネットワークの推定を行い、ネットワークの再現性をみる。図 6 は評価目的のために擬似的に生成したネットワークである。ノード数は 20 個である。黒丸がノードを表し、黒丸を結んでいる線がエッジである。

図 7 に GA によるネットワーク推定過程における次数分布の推移を示す。第 1 世代としているのが初期にランダムに生成した候補解と、擬似ネットワークの比較である。第 1 世代としているグラフ内の左側の山型のグラフが擬似ネットワークから観測された次数分布である。右側のグラフが GA による推定過程のネットワークの次数分布である。この左側の次数分布に一致するようなネットワークを探索するのが推定問題の目的である。図 7 には GA による推定過程の第 10 世代、20 世代と最終的に擬似ネットワークの次数分布と一致した第 22 世代の次数分布の比較を抜粋している。この結果から、GA の推定世代を重ねるに従って、ネットワーク推定による次数分布が擬似ネット

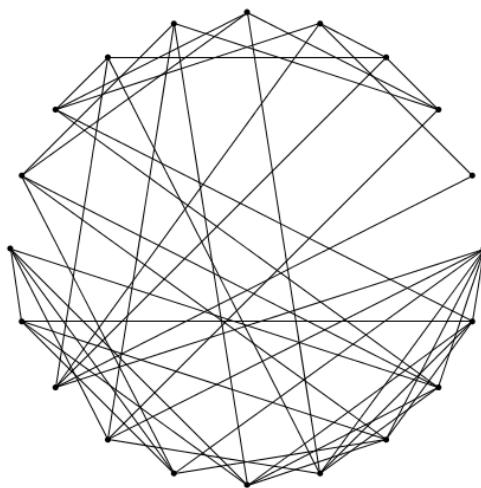


図 6 擬似ネットワーク

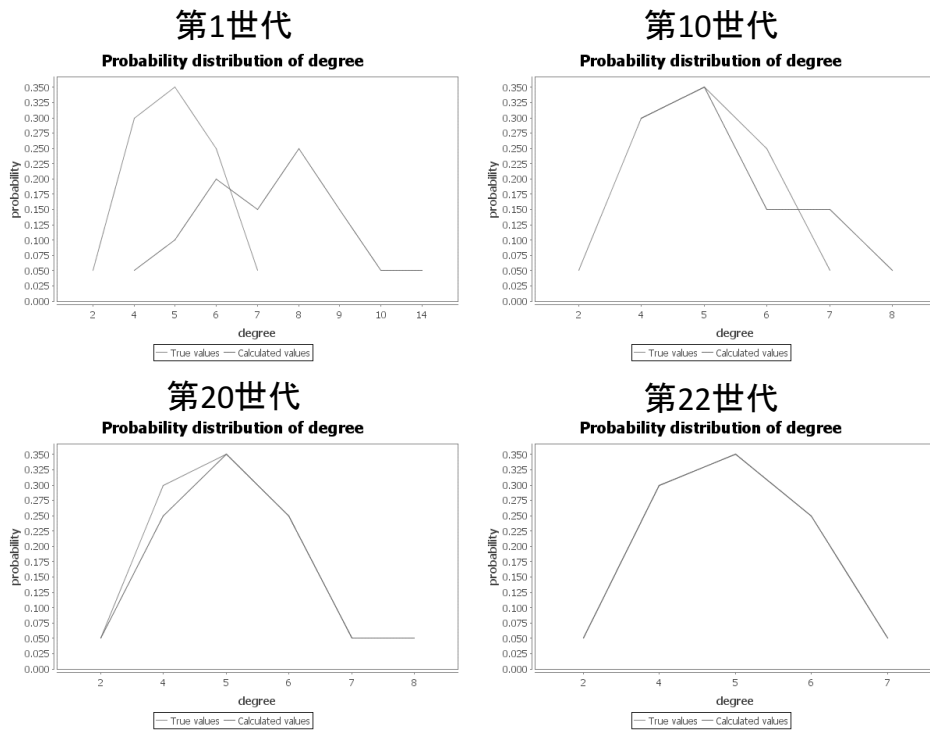


図 7 擬似ネットワークの推定過程における次数分布の推移

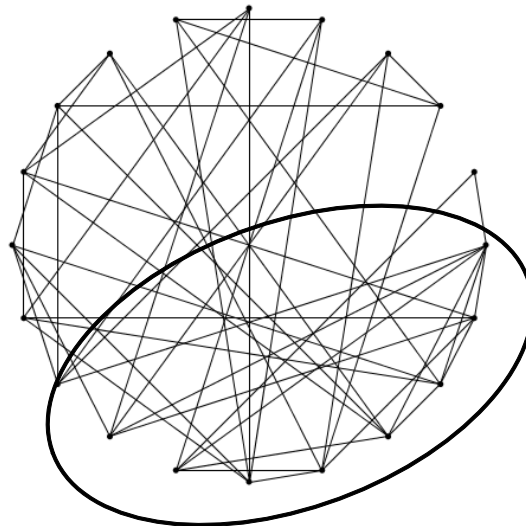


図 8 擬似ネットワークの推定結果

ワークの次数分布に近づいていき、最後には完全に一致する様子が分かる。

図 8 は最終的に推定されたネットワークを示す。現在の SN_Rebuilder では推定に次数分布しか使用しないため、推定精度はまだそれほど高くない。例え次数分布が完全に一致していたとしても、そのようなネットワークの候補は複数存在することが考えられるからである。ただし、現状の推定結果でも、図 8 内の丸で囲んだ次数が大きいノード部分に関しては、図 6 に示す元の擬似ネットワークと共通しているエッジが比較的多いことが確認できる。

5 実際のアンケート調査結果への適用例

本節では、実際に金融行動データベース II の Web アンケート調査結果に対して、提案手法を適用してネットワーク推定した結果を報告する。金融行動データベース II は本調査として 1677 人分のデータが含まれているが、事前調査として 4500 人に対する次数のみの調査も含んでいる。本節で報告する推定結果は、この 4500 人に対する次数分布調査の結果を利用した結果である。

本実験では、構築する社会ネットワークのノード数を 100 とした。GA による推定結果は第 848 世代までの計算でほぼ収束した。図 9 はアンケート調査による次数分布と GA の推定による次数分布を比較したものである。図より次数分布は観測値と推定値がほぼ一致していることが分かる。図 10 は推定結果のネットワークである。前節までの議論により、この推定結果のネットワークはまだあまり精度は高くないことが予想される。ただし、次数が大きいノードに関しては比較的实际のネットワークを反映しているのではないかと予想される。

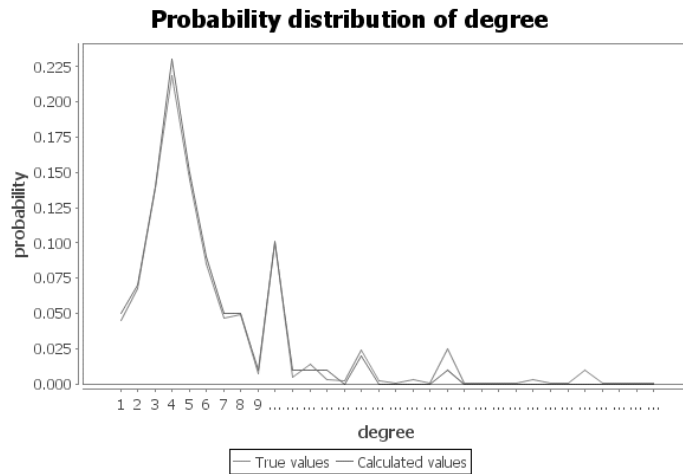


図 9 アンケート調査による次数分布と GA の推定による次数分布の比較

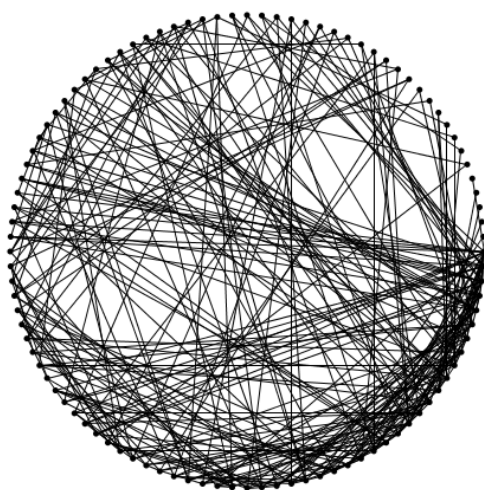


図 10 ネットワークの推定結果

6 まとめと今後の課題

本研究では、ミクロな視点から社会現象を分析する際に必須となる社会ネットワークの分析に関して、実際に観測されたデータからの社会ネットワークの再構築を試みた。ネットワーク推定問題は NP 困難な問題に属することから、本研究ではヒューリスティックな手法として遺伝的アルゴリズム (GA) を応用してネットワーク推定を行うツール, SN_Rebuilder を提案・開発した。その結果、現在の実装は次数分布のみを用いて推定するが、次数が大きいノードに関しては比較的正しく推定できることが分かった。

現在の実装はまだ試作段階のため、以下のような課題が今後の課題としてあると考えられる。まず、現在の実装は次数分布のみを用いて推定するため、推定結果は必ずしも正しくない。今後、クラスタリング係数や媒介性など他のネットワーク特性を考慮した推定手法を確立することによって、精度向上を図る必要がある。また、候補解のエンコード手法に接続行列を利用しているため、構築しようとする社会ネットワークのノード数が大きくなると、ノード数の 2 乗に比例してメモリを消費する。したがって、使用できるメモリの制約を受けて、あまり大きなモデルは推定できない。GA では最適解の探索のために候補解の多様性を維持するために多数の候補解を生成する必要があるが、より少ない候補解で多様性を維持可能な Minimal Generation Gap[8, 9] などの世代交代手法の適用などにより、メモリ消費を抑えられることが考えられる。

謝辞

本研究は文部省科学研究費補助金（課題番号：19653027）および、文部科学省私立大学学術研究高度化推進事業（学術フロンティア推進事業）、文部科学省人文学及び社会科学における共同研究拠点の整備の推進事業による助成を受けて行った研究成果である。

参考文献

- [1] L. C. Freeman: “The Development of Social Network Analysis: A Study in the Sociology of Science”, Empirical Press (2004).
- [2] A. Drogoul and J. Ferber: “Multi-agent simulation as a tool for modeling societies: Application to social differentiation in ant colonies”, Lecture Notes in Computer Science, **830**, pp. 2–23 (1994).
- [3] A. L. Barabasi and R. Albert: “Emergence of scaling in random networks”, Science, **286**, 5439, pp. 509–512 (1999).
- [4] S. N. Dorogovtsev, J. F. F. Mendes and A. N. Samukhin: “Structure of growing networks with preferential linking”, Phys. Rev. Lett., **85**, 21, pp. 4633–4636 (2000).
- [5] D. J. Watts and S. H. Strogatz: “Collective dynamics of ‘small-world’ networks”, Nature, **393**, 6684, pp. 440–442 (1998).
- [6] <http://www.kansai-u.ac.jp/riss/shareduse/database.html> (2009).
- [7] J. H. Holland: “Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence”, The MIT Press (1992).
- [8] H. Sato, I. Ono and S. Kobayashi: “A new generation alternation model of genetic algorithms and its assessment”, Journal of Japanese Society for Artificial Intelligence, **12**, 5, pp. 734–744 (1997).
- [9] M. Yamamura, H. Satoh and S. Kobayashi: “An analysis on generation alternation models by using the minimal deceptive problems”, Journal of Japanese Society for Artificial Intelligence, **13**, 5, pp. 746–756 (1998).