

2×2 ゲームにおけるネットワークと戦略共進化に基づく ジレンマの解消について

相良博喜*, 谷本 潤**

*九州大学大学院総合理工学府環境エネルギー工学専攻/JST 特別研究員

**九州大学大学院総合理工学府環境共生工学部門

Zimmermann & Eguiluz (2005) のモデルを基に、ネットワークの共進化メカニズムを取り入れた 2 人ジレンマゲームについて考察した。Zimmermann らは、裏切り戦略 (D) をお互いに選択した場合、一定の確率で対戦のリンクを切断し、新しい相手にランダムで接続するというルールを用いて、適応的なネットワークを構築した。本研究では、このモデルを基に様々なゲーム構造に関して検討を行い、R 互恵 (お互いに協調 (C) を選択する互恵関係、囚人のジレンマや Stag Hunt など) が成り立つゲームでは、スケールフリー的なネットワークが生成されることを示した。しかし、このモデルは ST 互恵 (C-D や D-C を交互に取り合う、Leader-Game など) では、非効率であることを示した。

A Study of Dilemma-Solving Effects by the Co-evolution of Both Social Networks and Strategy in a 2×2 Game.

Hiroki Sagara*, Jun Tanimoto **

* JST Fellow, IGSES, Kyushu University

** Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

A 2×2 game model implemented by a co-evolution mechanism of both networks and strategy, inspired by Zimmermann & Eguiluz's work (2005) is established. Network adaptation is the manner in which an existing link between two agents is killed, and how a new one is established to replace it. The strategy is defined as whether an agent offers Cooperation (C) or Defection (D). Both the networks and strategy are synchronously renovated in a simulation time step. A series of numerical experiments, considering various 2×2 game structures, reveals that the proposed co-evolution mechanism can solve dilemmas in several game classes. The effect of solving a dilemma means the mutual-cooperation reciprocity (R reciprocity), which is brought about by emerging several cooperative hub agents who have plenty of links. This effect can be primarily observed in game classes of Prisoner's Dilemma (PD) and Stag Hunt (SH). The co-evolution mechanism, however, seems counterproductive for game classes of Leader and Hero, where the alternating reciprocity (ST reciprocity) is meaningful.

1. 緒言

進化ゲーム理論を適用したジレンマゲーム研究は、情報科学にとどまらず、数理生物学、経済学、その他の社会科学などで広く応用される数理科学上の問題である。よく知られる (2 人 2 戦略ゲーム (2×2 ゲーム)) である囚人のジレンマ (Prisoner's Dilemma, PD) では、2 人のプレーヤ (エージェント) は互いに裏切り (Defect, D) 合うことが自己利得を最大化する上で合理的選択となる (Nash 均衡) が、本来は互いに協調 (Cooperation, C) する方が、より利得が高く、かつ公平な状態 (公平な Pareto 最適) を達成出来る。このように二律背反する状況を数理工学的ジレンマと云う。ジレンマゲーム研究の最大の関心は、抽象化されたゲーム (2×2 ゲームや多人数型ゲームなど) にどのような付加的枠組みを導入すれば、Nash 均衡であるジレンマ状況から脱し、互恵的な状態に達し得るか、にある (このことが解明出来れば、例えば、人間を始め多くの生物種がどのようにして社会的協調行為を進化させてきたのかを推論出来る)。ジレンマを緩解、あるいは解消する付加機構としては、同じ相手との繰り返し対戦 (例えば[1])、対戦拒否戦略 (Walk-away あるいは lonely) の付加 (例えば[2])、タグによる認証 (例えば[3]) などが考えられてきた。その多くが、対戦相手を制限することで相互作用 (ゲーム対戦) の匿名性を軽減し、互恵的に行動することが自己利得の最大化を目指す上で合理的であるようにゲーム構造を (ある意味で) 変容させる機構であると言える。C か D かの単純な戦略しか持たないエージェント間に対戦相手を限定するためのネットワークを仮定する、いわゆる空間型ゲーム (例えば[4] - [9]) も前記と同様の機構と云える。このネットワークには規則的格子状ネットワークから small-world[10]、scale-free ネットワーク[11]まで様々な形態が考

くい)には適用出来ない。

2. モデル

エージェントがリンクを結んでいる相手と互いに 2×2 ゲームを繰り返しながら、自己の利得の最大化を目指して、戦略 (C か D か) 及びネットワークを適応的に進化させる。集団全体のエージェン数を N とする。

2.1 ゲーム構造

自他戦略の組み合わせにより決まる 2×2 ゲームの利得表を Table1 の如く定義する。全ての 2×2 ゲームのゲーム構造は、

$$(2-1) \quad P = x_0 - 0.5 \cdot r_1 \cdot \cos\left(\frac{\pi}{4}\right)$$

$$(2-2) \quad R = x_0 + 0.5 \cdot r_1 \cdot \cos\left(\frac{\pi}{4}\right)$$

$$(2-3) \quad S = x_0 + r_2 \cdot \cos\left(\frac{\pi}{4} + \theta\right)$$

$$(2-4) \quad T = x_0 + r_2 \cdot \sin\left(\frac{\pi}{4} + \theta\right).$$

のようにパラメータ θ [rad] と $r = r_2 / r_1$ で表すことが出来る[17] (x_0 は解可能域がシフトするだけで P, R, S, T の相対関係上意味がないので、実質的には2パラメータとなる)。これら2パラメータで表されるゲーム空間を表すと Figure 1 のようになる。太線で囲われ、ジレンマゲーム名が冠されている領域以外が、ジレンマのない Trivial ゲームである。

Tanimoto ら[17]によると、 2×2 ゲームにおけるジレンマ性の有無は $DL_g = T - R$ (ギャンブル性ジレンマ; Chicken 型ジレンマの大小) と $DL_r = P - S$ (リスク回避型ジレンマ; Stag Hunt (SH) 型ジレンマの大小) の正負で判定することが出来る。PD は Chicken 型, SH 型両方のジレンマのあるゲームクラスである。Hero と Leader は Chicken 型ジレンマがあり、かつ $S + T > 2R$ の条件を満たし、さらに前者では $S > T$ 、後者では $T > S$ なる関係があるクラスである。

2.2 エージェント

各エージェントは自己利得を最大化するために戦略とネットワークをダイナミクスの過程で共進化させる。但し、戦略及びネットワークの更新は全エージェント一斉に (synchronic) 行う。各エージェントはネットワークによりリンクされているエージェントの全てと 2×2 ゲームを行い、その獲得利得総和が各エージェントの利得となる。従って、リンク数の多いエージェントほど高利得を得易い。各エージェントが持つリンク数の平均 (平均次数) を K とする。

2.3 戦略の適応

リンクを張っている対戦相手の中でその時間ステップで最大総利得を上げたエージェントの戦略 (C か D か) を確定的にコピーする Imitation Dynamics による。

2.4 ネットワークの適応

シミュレーション・エピソードの初期に各エージェントは平均次数 K のランダムネットワークで繋がっている。初期の 50 ステップまでは、ネットワークはランダムのまま固定とする。50 ステップ以降、ステップごとに行われるネットワークの適応は、切断と接続の2つのプロセスにより行われる。それぞれ以下の2手法を想定する。切断に際しては、相手に D を出されたときに自ら進んでその相手とのリンクを切らない確率 p_k をパラメータとして定義する。

切断法 1; D-D 間 (利得 P を得た場合) のリンクを確 $1 - p_k^2$ で切断。リンク両端の両者が D を出さない限り切断されない、との設定。

切断法 2 ; D-D 間 (利得 P を得た場合), C-D 間 (利得 S を得た場合) のリンクをそれぞれ $1-p_k^2$, $1-p_k$ で切断。リンク両端の一方のエージェントが D を出せば切断される可能性が発生する。但し、一方が D を出すより両者が D を出す方がそのリンクが切断される可能性が高い、との設定。

接続法 1 ; 切断したリンクの代替えとして、ランダムに選択したエージェントに接続。T 但し、同じリンクが 2 重に存在するのは許さないものとする。

接続法 2 ; 切断したリンクの代替えとして、保有リンク数でルーレット選択したエージェントに接続。B-A アルゴリズム[11]との類似性から、scale-free 的なべき則に従う次数分布が発生することが推測される設定である。但し、同じリンクが 2 重に存在するのは許さないものとする。

3. 数値実験

仮定した数値実験パラメータは、 $K=8$, $N=1000$, $p_k=0.03$, $r_1=1.272$, $x_0=0.55$ である。また、初期の戦略 C の割合は 0.5 とした。検討したゲーム構造は $-\frac{3}{4}\pi \leq \theta \leq \frac{5}{4}\pi$ かつ $0 \leq r \leq 2$ の範囲 (Figure 1) で、以下に示す等値線図はシミュレーション・エピソードごとに得た均衡値 (ダイナミクスが十分安定した定常値) を 5 試行して求めたアンサンブル平均像である。

4. 結果及び考察

Figure 2 に (a)切断法 1 と接続法 1 に基づく (以下, “標準” (Standard Evolutionary Network, SEN)) 適応ネットワークにおける協調率, (b)標準適応ネットワークと解析解との利得差, (c)標準適応ネットワークと固定ランダムネットワークとの利得差, (d)標準適応ネットワークと固定 scale-free との利得差, (e)標準適応ネットワークにおける最大次数, (f)切断法 1 と接続法 2 に基づく適応ネットワークと解析解との利得差, (g)切断法 2 と接続法 1 に基づく適応ネットワークと解析解との利得差, (h)切断法 2 と接続法 2 に基づく適応ネットワークと解析解との利得差を示す。解析解は、 2×2 ゲームの Replicator Dynamics を解析的に求解することで得られ、協調をサポートする機構が一切存在しない場合を意味する。固定ランダムネットワークあるいは固定 scale-free とは、初期に決めたネットワーク上で戦略適応だけを考慮して行ったゲームを意味し、標準適応ネットワークとランダムネットワークとの利得差は、ネットワークを適応させることによる効果を表す (なぜなら適応ネットワークの初期 50 ステップまでは固定ランダムネットワーク上でゲームを行っているから)。なお、利得は 1 ゲーム対戦当たりの値を示す。

Figure 2(a), (b)より、戦略とネットワークを共進化させる機構を適用しても、PD を含む Chicken 型ジレンマが強い領域 (Figure 2(a)中の点線で囲んだ r が大きい領域) ではジレンマは克服されない。しかし、Stag Hunt 型ジレンマが存在する多くのゲーム領域 ((Figure 2(a), (b)中の破線で囲んだ領域) では大略ジレンマが解消されていることがわかる。また、リンクの両端で C と D となる状況 (ST 互恵) の方が互いに協調するよりも高利得となる Leader や Hero の領域では適応型ネットワークは解析解より低利得となっており (Figure 2(b)中の一点鎖線で囲んだ領域)、ST 互恵に対してはこの共進化メカニズムは有意に機能しないことが分かる。

このような適応型ネットワークが有するジレンマ緩解・解消特性の頑強さは、戦略とネットワークの共進化により、ハブエージェントが創発することで説明できる。固定ランダムネットワークに比べて標準適応型ネットワークの平均利得が高い領域 (Figure 2(c)の黒い領域) では、最大次数が大きくなる領域 (巨大ハブエージェントが創発する領域) と一致する (Figure 2(e))。これは Zimmermann ら[14]が指摘するように、scale-free 同様のべき則に従う次数分布となつて、高利得を上げるハブエージェントが C を出し、彼に連なる C を出すエージェント達が階層構造を為すネットワーク上で R 互恵を行うことによって、高利得を上げているからである。Figure 3 に Figure 2(e)中の●プロットで示したゲーム構造 (Figure 1 で $\theta=3\pi/4$, $r=1.8$ なる SH 型ジレンマを有する Anti-Leader) における標準適応ネットワークの次数分布をランダムネットワーク (最大次数 19), scale-free (最大次数 74) とともに

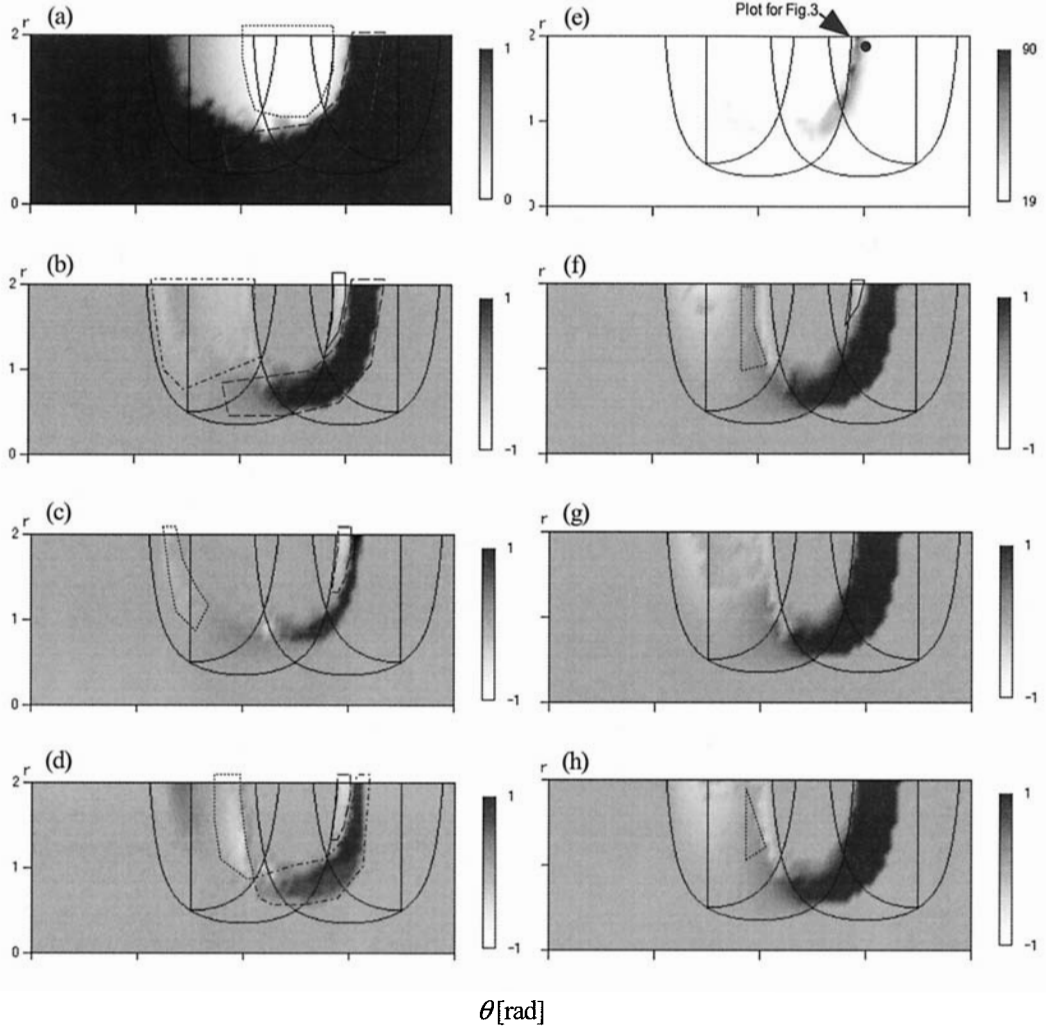


Figure 2. Result of the numerical experiment; (a) cooperation fraction among N in SEN, (b) payoff difference between SEN and the analytical solution, (c) payoff difference between SEN and the fixed random network case, (d) payoff difference between SEN and the fixed scale-free network case, (e) maximum degree in the network in SEN, (f) payoff difference between the evolutionary network based on Severing Method #1 & Connecting Method #2 and the analytical solution, (g) payoff difference between the evolutionary network based on Severing Method #2 & Connecting Method #1 and the analytical solution, (h) payoff difference between the evolutionary network based on Severing Method #2 & Connecting Method #2 and the analytical solution. Each payoff indicates a payoff per single game. Degree distribution of the closed plot in (e) that locates on $\theta=3\pi/4$ and $r=1.8$ is shown in Figure.3.

示す。標準適応ネットワークの最大次数は、平均次数 8 やランダムネットワークの最大次数に比して大きな値であり、scale-free 同様にべき則に従う指数分布が創発していることが分かる。

Figure 2(c)を見ると、一部の領域で標準適応型ネットワークが固定ランダムネットワークに劣る領域が 2 カ所あ

る。

Chicken 型ジレンマのある Hero から Leader にわたる一部のエリア (Figure 2(c)中の点線部) で標準適応型ネットワークの結果が劣るのは、そのゲーム構造によるものである。この領域は $2R < S+T$ なる条件を満たしており、お互いが協調し合う R 互恵よりも異なる戦略を出し合う ST 互恵を取る方が高利得となるゲーム構造を有する。固定ランダムネットワークでは、一定の割合で ST 互恵 (リンクの両端が C と D により構成される) が生起して

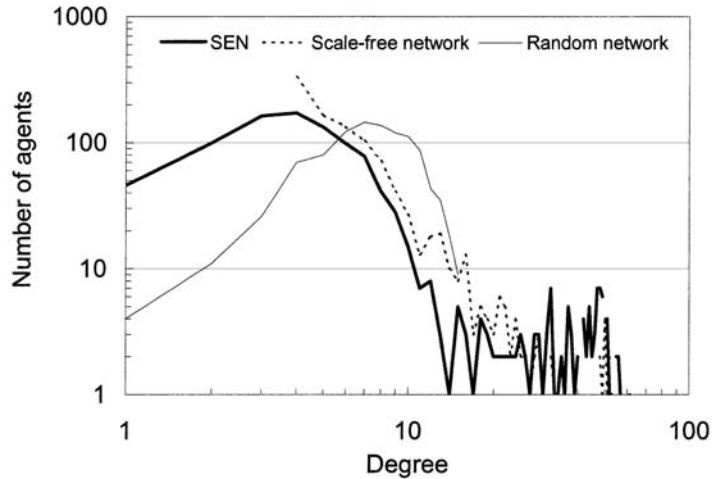


Figure 3. Bold line indicates the degree distribution of the closed plot ($\theta=3\pi/4$ and $r=1.8$ in Figure1) in Figure2 (e), of which game structure is Anti-Leader (see Figure1) having SH-type dilemma. Thin line and dotted line are those of the fixed random network (maximum number of links is 19) and the fixed scale-free network (maximum number of links is 74).

いるが、標準適応型ネットワークでは、適応の過程でそのネットワーク適応ルールの特性により、却って利得の低い R 互恵 (リンクの両端が C と C により構成される) へと進化が誘導されてしまう。

また、PD との境界からやや Anti-Leader 側に入ったエリア (Figure 2(c)中の破線部) において標準適応型ネットワークの利得が低くなっている。この領域では $2P > S+T$ なる条件が満たされている。解析解 (ジレンマ解消メカニズムが一切ない) では、この SH 型ジレンマのあるこのゲームの均衡は初期の C の割合が 0.5 であるなら全エージェントが D を出す状態に吸引される (すなわち全エージェントが P を取る)。このようにゲームの本質は D を出すインセンティブがある構造を有するのに、適応型ネットワークを導入すると、その R 互恵を促す効果が却って徒となって、全ネットワーク中に S と T を取る部分が生じてしまう。このことが社会全体平均利得を押し上げる結果となる (以上 Figure 2(c))。

固定 Scale-Free との利得差 (Figure 2(d)) を見ると、Leader と Anti-Leader の一部 (Figure 2(d)中の点線部及び破線部) で標準適応型ネットワークが低利得となっているが、PD から SH, Anti-Leader へ繋がる領域 (Figure 2(d)中の一点鎖線部) では標準適応型ネットワークが scale-Free を凌駕している。Figure 2(d)中の点線部では、標準適応型ネットワークの最大次数 (Figure 2(e)) は固定 scale-free の 74 に対してごく小さい。大きなハブが進化出来ないことが R 互恵を効率的に行えないことに繋がり、scale-free に対して低利得となったものと思われる。Figure 2(d)中の破線部で低利得となるのは、前記した固定ランダムネットワークに比して低利得となるのと同様の理由による。Hu ら[12]は固定 scale free 上で PD の進化ゲームを実験し、ネットワークのハブを占拠するエージェントが初期に C となるか D となるかでゲームの帰結が協調的か裏切りのかに分岐すると報告している。PD から SH,

Anti-Leader へ繋がる領域で標準適応型ネットワークが scale-free の利得を大きく上回る (Figure 2(d)中の一点鎖線部) のは、標準適応型ネットワークでは戦略とネットワークの共進化の効果によって、ハブエージェントが C を出す協調エージェントへ進化する傾向が一層促され、効率的な R 互恵が創発しているためである。Hu らの論点から言えば、適応型ネットワークは、ネットワーク上へのエージェント配置初期依存性に関して頑強なジレンマ緩解・解消プロトコルと云うことが出来る。

接続法 2 とすることで、Leader のエリアで適応型ネットワークが解析解に対して高利得となっている (Figure 2(f) と Figure 2(h)の点線部; Figure 2(b)と Figure 2(f), Figure 2(g)と Figure 2(h)とを比較)。基本的には適応型ネットワークは ST 互恵が要求される Leader や Hero には有効に機能しないと述べたが、接続法 2 とすることで Leader の一部のゲーム領域では解析解より高利得とすることが出来る。接続法 2 はその B-A アルゴリズムとの類似性から、創発してくるネットワークは scale-free 的な特性をもつ。前述した、このエリアでは適応型ネットワークは大きなハブを成長させられないとの欠点³が、この接続法 2 の特性により克服されている。

また、切断法 2 とすることで、PD との境界からやや Anti-Leader 側に入ったエリアで適応型ネットワークが解析解に比べて低利得になるとの欠点 (これは適応型ネットワークが固定ランダムネットワークに対して低利得になるとした Figure 2(c)中の破線部と一致) が克服されている (Figure 2(b)と Figure 2(f)の細実線部; Figure 2(b)と Figure 2(g), Figure 2(f)と Figure 2(h)とを比較)。これは先に詳述したように、 $2P > S+T$ なるゲーム構造では、切断法 1 による適応型ネットワークを用いると、 R 互恵促進が逆効果となって、全ネットワーク中に S と T を取る部分が生じてしまう点が、リンクの両端で C と D なる組み合わせでも切断してネットワークを適応させることにより解消されているからである。

以上の 2 点のことから、Zimmermann らの提案にはない、切断法 2、接続法 2 を用いることでゲーム構造によってはより高利得を生む互恵への適応を促す共進化メカニズムとすることが可能である。

5. 結論

Zimmermann らのモデルをベースに改良を加えた“戦略とネットワークの共進化モデル”を構築した。様々な構造の 2×2 ゲームにこのジレンマ緩解・解消機構を適用してみたところ、以下の知見を得た。

- (1) このメカニズムは、基本的には R 互恵 (互いに C を出し合う) に対してのみ有効で、Leader や Hero のジレンマに対して求められる ST 互恵 (自手 C, 相手 D を交互に出し合う) には効果がない。
- (2) ただし、“リンク大のエージェントに接続する”との適応ルールを加味すると、外生的にハブエージェントを出現させることとなって、PD に近い Leader の領域では R 互恵をサポートする効果が一層増進され、解析解より高利得を上げ得る。
- (3) また、“リンク両端で D-D に加え C-D の場合でもリンクを切断する”との適応ルールを加味すると、PD から Anti-Leader 側で解析解より低利得となる適応型ネットワークの欠点が克服される。

謝辞

本研究の遂行に際しては当時、九州大学工学部エネルギー科学科 4 年生の伊藤陽平君の労に依るところが大きい。記して、謝意を表する。

参考文献

- [1] Axelrod, R., The Evolution of Cooperation, Basic Books, New York, 1984.
- [2] Michor, F., Nowak, M. A., The good, the bad and the lonely, Nature 419, (2002), 677-679.
- [3] Epstein, J., Zones of cooperation in demographic prisoner's dilemma, Complexity 4 (2), (1998), 36-48.
- [4] Nowak, M., Bonhoeffer, S., May, R. M., Spatial games and the maintenance of cooperation, Proc. National Academy Sciences USA Vol. 91, (1994), 4877-4881.
- [5] Hauert, C., Doebeli, M., spatial structure often inhibits the evolution of cooperation in the snowdrift game, Nature 428, (2004),

643-646.

- [6] Masuda,N., Aihara,K., Spatial prisoner's dilemma optimally played in small-world network, *Physical Letters A* 313, (2003), 55-61.
- [7] Hauert,C., Szabo,G., Game theory and physics, *American Journal of Physics* 73, (2005), 405-414.
- [8] Tomochi,M., Defectors' niches: prisoner's dilemma game on disordered networks, *Social networks* 26, (2004), 309-321.
- [9] Tomassini,M., Luthi,L., Giacobini,M., Hawks and Doves on small-world networks, *Physical Review E* 73, (2006), #016132.
- [10] Watts,D.J., Strogatz,S.H., Collective dynamics of 'small-world' networks, *Nature* 393, (1998), 440-442.
- [11] Albert,R., Barabasi,A.-L., Statistical mechanics of complex networks, *Review of Modern Physics* 74, (2002), 47-97.
- [12] Hu,M.B., Wang,W.X et al., A unified framework for the Pareto law and Matthew effect using scale-free network, *The European Physical Journal B* 53, (2006), 273-277.
- [13] Tang,C.-L., Wang,W.-X., We,X., Wang,B.-H., Effects of average degree on cooperation in networked evolutionary game, *European Physical Journal B* 53, (2006),411-415.
- [14] Zimmermann,M.G., Eguiluz,V.M., Cooperation, social networks, and the emergence of leadership In a prisoner's dilemma with adaptive local interactions, *Physical Review E* 72, (2005), #056118.
- [15] Pacheco,J.M., Traulsen,A., Nowak,M.A., Active linking in evolutionary games, *Journal of Theoretical Biology* 243, (2006), 437-443.
- [16] Pacheco,J.M., Traulsen,A., Nowak,M.A., Coevolution of strategy and structure in complex networks with dynamical linking, *Physical Review Letters* 97, (2006), #258103.
- [17] Tanimoto, J., Sagara, H., Relationship between dilemma occurrence and the existence of a weakly dominant strategy in a two-player symmetric game, *BioSystems*, (2007), *In process*; <http://linkinghub.elsevier.com/retrieve/pii/S0303264706001274>]

著者紹介

谷本 潤. 正会員. 九州大学大学院・教授 (大学院総合理工学研究院, 〒816-8580 春日市春日公園 6-1, tanimoto@cm.kyushu-u.ac.jp). 専門は都市建築環境工学, 人間-環境-社会システム. 空気調和・衛生工学会学会賞, 日本建築学会奨励賞, 日本建築学会学会賞 受賞.