

自己利得の大きさによる戦略コピー耐性を Pairwise-Fermi に組み込んだ 囚人ジレンマゲームのネットワーク互恵

志垣圭蔵¹, 小窪聡¹, 谷本潤², 萩島理²

¹九州大学大学院 総合理工学府 環境エネルギー工学専攻

²九州大学大学院 総合理工学研究院 エネルギー環境共生工学部門

概要

2×2 ゲームのネットワーク互恵モデルでは隣人と利得差もしくは隣人利得で戦略適応を行うが、現実の意志決定を考えると自身利得の絶対値が社会平均より大きいほど隣人戦略をコピーするインセンティブが失われると推量される。本研究では、この機構を従来の Pairwise-Fermi に組み込んだ戦略適応モデルを構築して、数値シミュレーションを行った。その結果、通常の Pairwise-Fermi に比べて大きな協調 enhance 効果が確認された。

A Revised Pairwise-Fermi process considering copy-resistance caused by the amount of own payoff enhances cooperation in Prisoner's Dilemma Games on networks.

Keizo Shigaki¹, Satoshi Kokubo¹, Jun Tanimoto¹, Aya Hagishima¹

¹Interdisciplinary Graduate School of Engineering Sciences, Kyushu University

Abstract

We propose a revised Pairwise-Fermi process considering “copy-resistance” for copying strategy from a neighbor, which implies a focal agent with affluent payoff than social average might be negative to copy a neighbor's strategy even if her payoff is less than the neighbors. Simulation results reveal this mode revised for strategy adaptation significantly enhances cooperation for Prisoner's Dilemma Games on time-constant networks.

1 諸言

人間を含む生物圏で普遍的に観察される個体群の協調的行動の自己組織化は、囚人ジレンマ (Prisoner's Dilemma, PD) ゲームを基本モデルとする進化ゲーム理論を適用したアプローチにより多くの研究が行われている。Nowak らは協

調創発を可能にする 5 つの互恵メカニズムについて論じている[1]。就中、ネットワーク上で隣人とだけゲームを行い、隣人から戦略コピーをすることで well-mixed な状況から社会粘性を大きくするメカニズムにより協調創発を促すネットワーク互恵に関する研究報告が大部を占めると云っても過言でない。これは、エージェント

の大きな知性（メモリ）を前提としない、結晶構造や相転移との相似性から統計物理学との関連性がある等の理由により様々な分野の関心を喚起しているからだろう。また、ここで言う社会粘性とは特定のエージェントとのみ繋がっていることや、繋がりの数、すなわち次数の制限などに当たる。最近のネットワーク互惠に関する研究としては、一層の協調 enhance 効果を現出させるには、ネットワーク上のゲームに加えて、更にどんな付加機構を入れる必要があるかとの点に話題が集まっている。例えば、Wang & Perc [2], Perc & Wang [3], Tanimoto et al.[4]は戦略適応法としてオーソドックスな Pairwise-Fermi による確率プロセスを行う際、Pairwise 比較する隣人をランダム選択に代わって利得に応じて確率戦略する仕組みを入れると大きな enhance 効果が顕れると報じている。また、Chen et al. [5], Szolnoki et al. [6]は、Fermi 関数で決まるコピーイベントの生起確率をコピー主体の focal エージェントの learning activity level により圧縮すると enhance 効果が生じることを見出した。これらは、いずれも Tanimoto の説明によれば[7], 協調クラスターを形成しているエージェントたちの戦略適応速度をゲーム進行速度に比して遅くしてやることで、彼らの戦略適応、すなわち裏切り戦略へ転じる可能性を縮減している機構であると考えられる。

本研究では、既往の流れを踏んで、ネットワーク互惠に付加する仕組みとして、戦略コピーの耐性と云うアイデアを考究する。自他の利得差に応じて Fermi 確率で隣人から戦略をコピーする際に、もし自己の利得が社会平均に対して十分に高ければ、喩え、選択した隣人の利得が focal より高くともコピーしようとのインセンティブは減じると思われる。「十分にうまくいっている者は敢えて自戦略を変えようとのインセンティブは持たない」との考えは、人間の意志決定を観るとき十分な妥当性がある[8]。本稿では、数値実験結果に基づき、上記を考慮したモデルが更なるネットワーク互惠効果をもたらすことを報告する。

2 モデル

2.1. 2×2 ゲーム

エージェントの戦略を2つの離散値とし、夫々、協調 (Cooperation, C), 裏切り (Defection, D) とする。また 2×2 ゲームの利得構造を P

(Punishment, 自他の手組は D-D, 以下同様), R (Reward, C-C), S (Saint, C-D), T (Temptation, D-C) で表す。Tanimoto & Sagara[9]に倣って、Chicken 型ジレンマ ($T-R$ の大きさ), SH 型ジレンマ ($P-S$ の大きさ) を夫々 D_g, D_r で表す。 $R=1, P=0$ で固定すると、ゲーム構造は、

$$G = \begin{matrix} & C & D \\ \begin{matrix} C & D \end{matrix} & \begin{matrix} C & D \end{matrix} \\ \begin{matrix} C & D \end{matrix} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} = \begin{matrix} C & D \\ D & D \end{matrix} \begin{pmatrix} 1 & -D_r \\ 1+D_g & 0 \end{pmatrix} \end{matrix} \quad (1).$$

本稿では、 $D_g \in [0,1], D_r \in [0,1]$ の PD ゲームのクラスを考察対象とする。

2.2. ネットワーク

全エージェント数 $N=4900$, 平均次数 $\langle k \rangle=4$, ネットワーク構造は格子グラフ (Lattice), BA アルゴリズム[10]による Scale-Free (SF) グラフを用いた。エージェントはこれらのネットワーク上で進化ゲームを繰り返す。

2.3. 戦略適応とその更新方法

エージェント i は全隣人とゲームしてその合計利得 π_i を引数とする。そして Pairwise-Fermi プロセスで相手戦略をコピーするか否かを決定する。Pairwise-Fermi プロセスとは自分の隣人のうち一人をランダムに選択し、確率的にその相手の戦略をコピーするというものである。以下がその式である。

$$P_{i \leftarrow j} = \frac{1}{1 + \exp \left[\frac{(1-a)(\pi_i - \pi_j) + a(\pi_i - \langle \pi \rangle)}{\kappa} \right]} \quad (2)$$

ここで $\langle \pi \rangle$ は全エージェントの平均利得、すなわち社会平均利得、温度係数 κ は 0.1 とした。自他利得差の影響と自己利得と社会平均との差の影響を考慮する重み a は、モデルパラメータである。 $a=0$ のとき通常の Pairwise-Fermi と一致する (以下、このケースをデフォルトとする)。更新方法は全エージェントが一斉に戦略更新するシンクロ更新を用いた。

2.4. 実験方法

あるジレンマ強さ D_g, D_r のゲーム構造において、アンサンブル平均 50 回を採って、解析対象とする。各エピソードでは初期に協調率 0.5 でランダムに C, D エージェントを配置する。各試行は、戦略値と利得の平均の摂動が十分小さくなり擬似均衡と見なせるまで続ける。摂動が大きく均衡に達さなかった場合は、上限である

10000 時間ステップの最終 100 ステップの平均データを算出する。

3 結果

3.1. 格子グラフ

図 1 は $\langle k \rangle = 4$ の Lattice で PD 全域の協調率を示している。デフォルト ($a=0$) と比較して、パラメータ a の値を上げていくと、協調域が広がっていくことが確認できる。特に $a=0.6$ で最も広い協調域となった。併存平衡から徐々にジレンマを大きくして、協調エージェントが全く生き残れなくなるジレンマ強さをもって臨界ジレンマ強さと定義すると、 $a=0.6$ 前後で臨界ジレンマ強さが最大になる resonance 効果を示していると言えよう。 a が過大になると却って協調 enhance 効果は低下する。

図 2 は図 1 の 45deg 線上の値を a で比較した結果である。 $D_g - D_r$ 平面の 45deg 線上のゲームは Donor & Recipient Game (DRG ; $D_g = D_r$ なる PD のサブクラス) である。

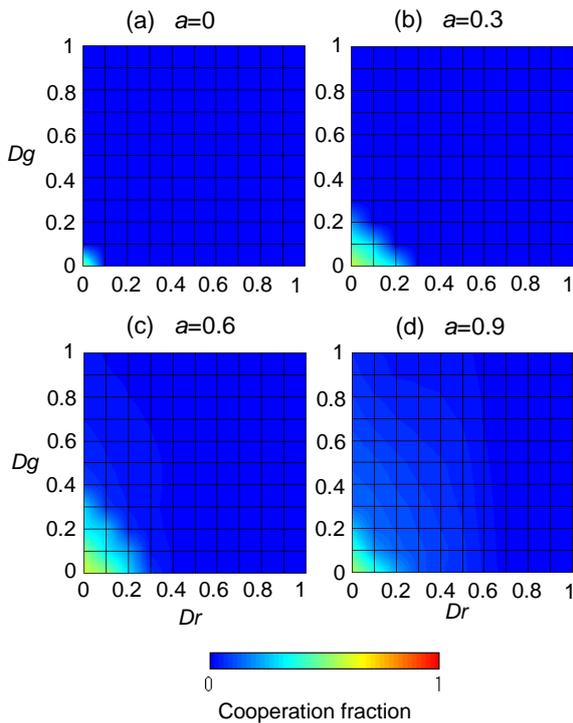


Figure 1: Averaged cooperation fraction. Games are played on lattice networks with average degree $\langle k \rangle = 4$. The dilemma class is PD, where $D_g \in [0, 1]$, $D_r \in [0, 1]$.

これらから、Lattice, $\langle k \rangle = 4$ のデフォルト ($a=0$)

は微弱なジレンマに対しても敏感で、すぐさま裏切り社会に落ちてしまうことが分かる。それに対し、提案モデルにより自己利得の大きさによりコピー耐性を考慮した場合は、明らかに大きなネットワーク互惠効果がある。比較的弱いジレンマ ($r=0.01 \sim 0.15$) では $a=0.6 \sim 0.7$ でより大きな協調率となり、比較的強いジレンマ ($r=0.15 \sim$) ではそれより大きなパラメータ a で粘りを見せている。

この協調 enhance 効果のメカニズムとしては、次のことが考えられる。ジレンマが強い領域で裏切りエージェントの数が多くなっていくと裏切り合って(1)式中の P を取り合うイベントが増えるので、結果として社会の平均利得が減少していく。すると、戦略変更の際に、自分の利得と比較する社会平均利得の値が小さいためエージェントは戦略変更をしにくくなる。よって、その時点で生き残っていた C クラスタが、その後も生き残り易くなる。

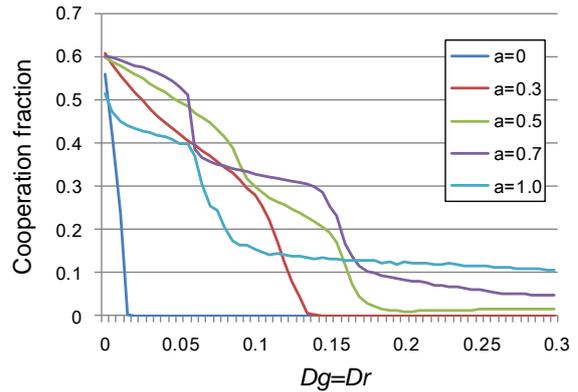


Figure 2: Averaged cooperation fraction. Games are played on lattice networks with average degree $\langle k \rangle = 4$. The dilemma class is DRG in PD, where $D_g = D_r$.

3.2. BA-Scale Free グラフ

続いて BA-Scale Free グラフでの結果を示す。前節同様、図 3 に $\langle k \rangle = 4$, エージェント数 $N=4900$, PD 全領域の結果を、図 4 に DRG 領域の結果を示す。パラメータ a を上げていくにつれ、弱いジレンマ領域での協調率が落ちていくものの、強いジレンマ領域では併存平衡が維持されている。つまり、全員が裏切りに吸引される臨界値である臨界ジレンマ強さは a 大ほど大きい。

これらの理由は以下が考えられる。提案モデルでは自利得を社会利得と比較するため、社会

利得がある臨界点よりも低くなると、エージェントの戦略適応速度が遅くなる。これにより社会利得より大きな利得を得ているエージェントの適応速度が遅くなるため、C クラスタが生き残る確率が高くなる。しかしながらそれと同時に隣人にCがいるDエージェントも同様に戦略適応速度が遅くなる。初期にハブエージェントにCかDかが割り当てられるかはランダムであるため、ハブを中心としてCクラスタとDクラスタが複数形成されるだろう。よって、この初期のCハブクラスタとDハブクラスタがともに適応速度が遅くなって頑強になることが、ジレンマ強弱によって異なる結果に繋がる。つまり、ジレンマが弱ければ、本来（デフォルトであれば）Cクラスタから広まる協調が系全域に至るのが、頑強なDクラスタが部分的に存在するため、そこまで高い協調率には至り得ない。逆に、ジレンマが強ければ、本来、Dクラスタに侵略されて全域で裏切りとなるところが、部分的に存在する頑強なCクラスタの存在により、併存平衡が続く。

4 結語

ネットワーク上の進化ゲームにおける戦略適応方法 Pairwise-Fermi に自己利得の大きさによる戦略コピー耐性を考慮したモデルを提示し、大きな協調 enhance 効果が顕れることをシミュレーションにより示した。これモデルは、自己利得が高ければ、隣人からコピーして戦略を変更しようとのインセンティブは低下する、との人間の意志決定を模擬した枠組みであると思われる。

参考文献

- [1] Nowak, M.A., *Science* 314 (2006) 1560–1563.
- [2] Wang, Z., Perc, M., *PRE* 82 (2010) 021115.
- [3] Perc, M., Wang, Z., *PLoS one* 5(12) (2010) e15117
- [4] Tanimoto, J., Nakata, M., Hagishima, A., Ikegaya, N., *Physica A* (2011).
- [5] Chen, X., Fu, F., Wang, L., *Modern Physics C* 19 (9), (2008) 1377-1387.
- [6] Szolnoki, A., Szabo, G., *Europhys. Lett.* 77 (2007) 30004.
- [7] Tanimoto, J., *Sociobiology* 58 (2), (2011) 315-325.
- [8] Zhang, H., Small, M., Yang, H., Wang, B., *Physica A* 389 (2010) 4734-4739.

- [9] Tanimoto, J., Sagara, H., *BioSystems* 90(1) (2007) 105-114.
- [10] Santos, F.C., Pacheco, J.M., *Phys. Rev. Lett.* 95 (2005) 098104.

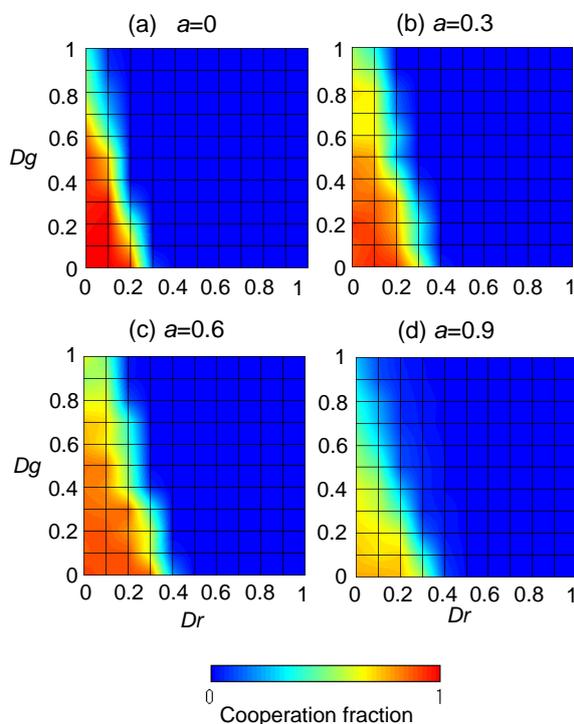


Figure 3: Averaged cooperation fraction. Games are played on Scale-Free networks by B-A model with average degree $\langle k \rangle = 4$. The dilemma class is PD, where $D_g \in [0, 1]$, $D_r \in [0, 1]$.

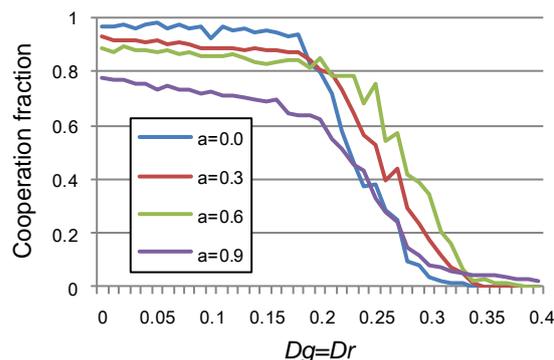


Figure 4: Averaged cooperation fraction. Games are played on Scale-Free networks by B-A model with average degree $\langle k \rangle = 4$. The dilemma class is DRG in PD, where $D_g = D_r$.