2F2-01

ネットワーク上での囚人のジレンマゲーム
Prisoner's Dilemma Game on the Network

小野真裕 石塚満
Masahiro ONO Mitsuru ISHIZUKA

東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo

We study an evolutionary dynamics of the Prisoner's Dilemma game, played by agents on the network. It is assumed that a network is changed by each agent which has meta strategy and link change strategy in the evolutionary process. We propose the network formation model and observe a cooperative behavior with a movement.

1. はじめに

囚人のジレンマゲームにおいては，従来から空間構造を扱った研究ががある ([Nowak 92][Lindgren 94] など). また，格子モデルよりも実際のプレーヤー間の関係を適切に表現できるネットワーク上におけるゲームも研究されてきた ([Watts 98][Abramson 01][小野 05] など). しかし，これまでは静的なネットワークを仮定しており，動的なネットワーク変化を考慮していなかった．本稿では，プレーヤーの戦略とネットワーク形成のダイナミクスを調査した．

2. 背景

n 個のノードから構成される無向ネットワークは $2^{n(n-1)/2}$ 通りの可能性がある．しかしながら，実際の世界では n が大きく莫大な数の可能性があるにも関わらず，多くのネットワークで SW(Small World)[Watts 98]，SF(Scale Free)[Barabasi 99] といった性質を示すことが多い．これまで SW,SF の性質を示すネットワークとなるようなネットワーク形成方法はいくつか提出されてきた．しかし，ネットワークを構成するノード戦略の変化は同時に考慮されてこなかった．

一方，ネットワーク形成を戦略とし自分の利得を最大にするようゲーム理論的に考察すると，限定されない合理性を持つプレーヤーにより形成されるネットワークは，ナッシュネットワークと呼ばれるホイールやスター型のネットワークになることが知られている [Bala 00]

プレーヤーの戦略とネットワーク形成の統合的な研究はこれまで十分でない．本稿では，囚人のジレンマゲームにおいて，ノードの戦略変化を包含するネットワーク形成モデルを提案し，戦略とネットワークのダイナミクスを調査する．

3. ネットワーク形成モデル

提案するネットワーク形成モデルを示す．以下の説明においては，ネットワーク構成要素をノードとリンクと呼んでいる．また，ノードのネットワーク知識に関する限定合理性を仮定し，2 ホップ先までのノード情報までを知ることができるとした．

$N = \{1, 2, \dots, n\}$ を n 個の意思決定主体であるノードとする．以下，$i, j \in N$ とする．各ノードは m 本のリンクを所有し，所有するリンクに関して他のノードへの切断・接続の操作を自由に行えることとする．世代 t の時点のネットワーク $g(t)$ は，i, j の全ての組み合わせにおけるリンクの有無を示す要素 $g(t)_{i,j}$ からなる (式 (1)).

$$g(t)_{i,j} = \begin{cases} 1 & j \text{ is linked by } i \\ 0 & \text{absence} \\ -1 & i \text{ is linked by } j \end{cases} \tag{1}$$

値 $1, -1$ の違いはリンクの所有ノードの違いのみをあらわし，リンクを通してなされる相互作用自体には影響しない．i に接続されているリンク数を l_i とする．

ここで，i の隣接ノード集合は $N^1(i, g(t)) = \{k \in N | g(t)_{i,k} \neq 0\}$ である．また i の所有するリンクで結ばれた隣接ノード集合は $N^1_o(i, g(t)) = \{k \in N | g(t)_{i,k} = 1\}$ となる．更に 2 ホップ先のノード集合を考えると $N^2(i, g(t)) = \{k \in N | k \in N^1(j, g(t)), j \in N^1(i, g(t))\}$ となる．

ネットワーク上では，リンクで結ばれているノード同士は，利得行列を $(T, R, P, S) = (5, 3, 1, 0)$ とした 2 人対称ゲームを行うこととし，全てのリンクについてゲームを行う．i が j との対戦の結果に得る利得を $p_{i,j}$ とすると，i の平均利得は $a_i = \Sigma_{j \in N^1(i,g(t))} p_{i,j} / l_i$ となる．

対戦後，全てのノードについてリンク張替え操作を行う．この操作は概念的には物理的な移動に近い．操作は 1:切断，2:接続からなり，ノード i については次のような操作を行う．

1. $N^1_o(i, g(t+1)) = N^1_o(i, g(t)) - N^{del}_i (\in N^1_o(i, g(t)))$
2. $N^1_o(i, g(t+1)) = N^1_o(i, g(t)) + N^{add}_i (\in N^2(i, g(t)))$

つまり，所有リンクで接続されたノードの中から N^{del}_i を選択し切断し，2 ホップ先のノードの中から N^{add}_i を選択し接続する．ここで，隣接ノードの指標を $U(i, j) = \{l_i, l_j, p_{i,j}, p_{j,i}, a_i, a_j\}$，2 ホップ先のノードの指標を $V(i, j, k) = \{l_i, l_j, l_k, p_{i,j}, p_{j,i}, p_{j,k}, p_{k,j}, a_i, a_j, a_k\}$，指標の差分も考慮した指標を $W(U) = \{r - s | r \in U, s \in U \cup \phi\}$，指標の評価関数を $H = \{\mathrm{argmin}, \mathrm{argmax}, \mathrm{random}\}$ と定義する．また，ノードは $h \in H, w \in W$ を遺伝子コード化し保持し，リンク張替え操作は遺伝子に従って行うとする．N^{del}_i は式 (2) に従い選択される．

$$N^{del}_i = H_x \left(W(U(i, x)) | x \in N^1_o(i, g(t)) \right) \tag{2}$$

N^{add}_i は，2 ホップ先のノード集合を一度に評価する式 (3)，または隣接ノード集合評価を 2 度行う式 (4) のいずれかに従い選択される．ただし，式 (4) に従う場合，ノードは $h' \in H, w' \in W$

連絡先: 小野真裕，東京大学大学院情報理工学系研究科，mono@miv.t.u-tokyo.ac.jp

も保持する．

$$N_i^{add} = h_x\left(w(V(i,j,x))|x{\in}N^1(j,g(t)), j{\in}N^1(i,g(t))\right) \quad (3)$$

$$N_i^{add} = h'_x\left(w'(U(j,x))|x{\in}N^1(j,g(t))\right)$$

$$where \;\; j = h_j\left(w(U(i,j))|j{\in}N^1(i,g(t))\right) \quad (4)$$

ただし，N_i^{del}, N_i^{add} に複数の要素が含まれる場合，および $h = \mathrm{random}$ の場合は，集合の要素から一様ランダムに一つ選択することとする．

以上のネットワーク形成操作を，$t = 0$ の時にランダムネス 0 の SW ネットワークから開始し，各世代，全対戦後，全てのノードについて行う．

4. シミュレーション結果

各ノードは，自分と相手の戦略を 1 回記憶する 5bit のメタ戦略と，上述のリンク張替え操作遺伝子を持つ．$n = 400, m = 3$ に固定しシミュレーション評価を行い，ノードはそれぞれ独立にメタ戦略，リンク張替え操作を行う場合を調べた．

まず，リンク張替え操作を行う場合と行わない場合を比較したものを図 1 に示す．両ケースにおいて，一度非協調的な戦略が栄えた後，協調的な戦略が栄えるというシナリオは同じである．しかし，リンクを張り替える場合には，張り替えない場合に比べて収束が早いことがわかる．この過程においては，非強調的な戦略は協調的な戦略を求めてリンク張替えを行い，協調的な戦略は協調的な戦略同士でリンクし安定した集団を作ろうとする．ただし接続の嗜好性から，非協調的な戦略は，全リンクは自分の所有するリンクのみであるのに対し，協調的な戦略は被リンクも存在する．このようにして最終的には被リンクの分だけ利得を稼ぐ協調戦略が栄えることになる．

次に図 2 にその他の指標の推移例を示す．

パス長はほぼ単調減少である．この例においては，一度上昇しているが，これはリンク数の多いノードが死亡してしまったためである．クラスタ係数は一度落ち込んだあと，徐々に上昇し最終的には 1 に近くなっている．このネットワークでは，リンク数が非常に大きい，少数のスーパーノード同士がが互いにリンクしあい，その他のノードはそれらのスーパーノードにリンクする構成となっている．そのためクラスタ係数が 1 に近い値となっている．

ある時点においてノードがそれ以前にリンクしたことがあるノード数の平均値を移動距離と定義すると，この例ではクラスタ係数が 1 に漸近するのと同時に 30 程度に収束した．移動距離が収束する条件は，N^{add} の選択にリンク数を指標とするノードが多いことである．初期においては，当然自ノードの得る利得を高めるような選択方法が生き残る．しかし，その後はリンク数が多いノード＝協調的なノードとなるため，利得とリンク数は指標としてほぼ等価となり，リンク数を指標とするノードも生き残ることが可能となる．今回のシミュレーション条件では，あえてリンク数を指標とすべき特別な理由は存在しないため，進化の過程において最初に多数を占めた指標が支配的になると考えられる．それぞれの指標の特徴として，リンク数の順位は安定傾向があるが，利得は隣接ノードに応じて変動があり比較的不安定である．そのため，利得を指標とするノードが多数を占める場合には，移動距離はほぼ単調に増加する．また，N^{del} に関しては，自分の利得を最も小さくするノードを選択し切断する操作が生き残ることがわかった．

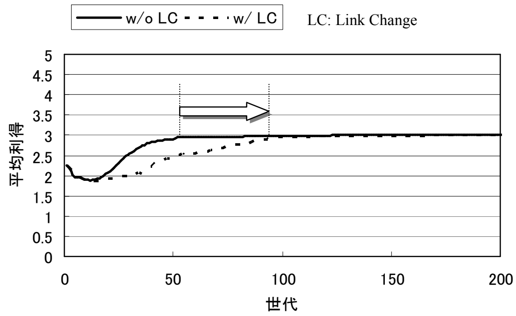

図 1: 平均利得の推移

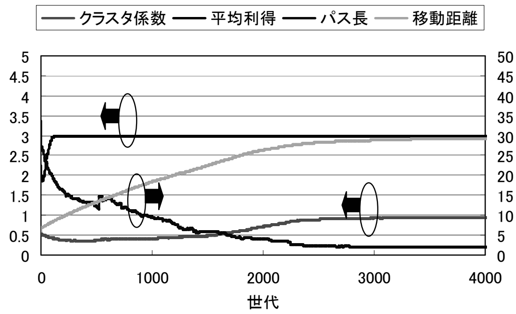

図 2: 進化ダイナミクス例 (noise=0, mutation=0)

5. おわりに

ネットワーク形成過程を取り入れたネットワーク上での囚人のジレンマゲームを提案し，シミュレーション評価を行った．

その結果，移動の概念を取り入れることにより戦略の収束が遅くなることがわかった．また，その際に特殊なネットワークの発生を観察した．

参考文献

[Abramson 01] Abramson, G. and Kuperman, M.: Social games in a social network, *Physical Review E* (2001)

[Bala 00] Bala, V. and Goyal, S.: A Noncooperative Model of Network Formation, *Econometrica* (2000)

[Barabasi 99] Barabasi, A. and Albert, R.: Emergence of scaling in random networks, *Science*, Vol. 286, No. 5439, pp. 509–512 (1999)

[Lindgren 94] Lindgren, K. and Nordahl, M. G.: Evolutionary dynamics of spatial games, *Physica D*, Vol. 75, pp. 292–309 (1994)

[Nowak 92] Nowak, M. and May, R.: Evolutionary games and spatial chaos, *Nature*, Vol. 359, pp. 826–829 (1992)

[Watts 98] Watts, D. J. and Strogatz, S. H.: Collective dynamics of ‘small-world’ networks, *Nature*, Vol. 393, pp. 440–442 (1998)

[小野 05] 小野 真裕, 石塚 満：囚人のジレンマゲームにおけるネットワーク構造の影響, 情報処理学会全国大会 (2005)