進化ゲームダイナミクスによるネットワーク通行権取引制度のインプリメンテーション*
Evolutionary Game Dynamics of Multi-Agent Systems for Implementing Tradable Network Permits *

菊地志郎**・赤松隆***・和田健太郎****
By Shiro KIKUCHI**・Takashi AKAMATSU***・Kentaro WADA

１．はじめに

交通混雑緩和のために適切な TDM (Transportation Demand Management) 方策を実施する際には，一般に，道路管理者が利用者の選好情報（eg. 支払意思額）を把握している必要がある．しかし，このような私的選好情報を正確に把握することは，実際には容易ではない．そして，誤った私的選好情報に基づく施策は，必然的に社会的損失を生む．つまり，道路管理者と利用者間での情報の非対称性が問題となる．

この情報の非対称性の問題を解消する施策として，ネットワーク通行権取引制度が赤松ら[1),2)]によって提案された．この制度は，以下のような 2つの仕組みから構成される制度である： (*a*) ネットワーク上の各リンクに対して，そのリンクを予め指定された時刻に通行できる権利（ボトルネック通行権）を道路管理者が発行し，(*b*) その時刻別の通行権を自由に売買できる市場を創設する．この制度の下では，道路管理者は各リンクの交通容量のみを把握していればよく，情報の非対称性の問題が解消される． さらに，この制度導入後の均衡状態は，社会的交通費用を最小化する状態を実現できる．

しかし，ネットワーク通行権取引制度に関するこれまでの研究では，その具体的なインプリメンテーションは考えられていない : (*i*) 均衡状態への到達プロセスが非明示的，(*ii*) ネットワーク利用者による煩雑な通行権取引が必要．そこで，(*i*) に関して，均衡状態への到達プロセスを表現できる枠組みが必要となる．また，この制度をインプリメントするためには，(*ii*) に関してネットワーク利用者の情報処理負荷を軽減する必要がある．

このような要請に応える方法の一つとして，本研究は，multi-agent system を提案する．これは，ネットワーク利用者に要請される煩雑な情報処理を車両に搭載した agent software に代行させるものである．より具体的には，各 agent software は，予め与えられた利用者の選好情報をもとに，経路・出発時刻等を選択し，そのために必要となる通行権の取得（通行権市場での取引）を自動的に実行する．このような機能を持つagent softwareは，最低限，以下の要件が求められるどろう : (*i*) 各 agent が局所的情報のみで経路・到着時刻を選択できる（行動の自律性），(*ii*) agent の戦略決定に必要な計算量が少ない（行動ルールの簡潔性），(*iii*) 個々の agent 行動の集積結果として決定される交通流・通行権価格の evolutionary dynamics が均衡状態に必ず収束し，その収束が速い（agent 群行動ダイナミクスの安定性）．ここで，通行権取引制度下の均衡状態では，社会的交通費用が最小化されることに注意しよう．つまり，上記性質を満たした，multi-agent system を設計できれば，自律・分散的な agent 行動の結果として，社会的最適状態（交通流・通行権価格）を実現できる．

そこで，本研究では，ネットワーク通行権取引制度下で，均衡状態への収束が期待できるagent 行動ルールを設計し，その特性を明らかにすることを目的とする．より具体的には，agent のミクロな行動ルールを提示し，それに対応するマクロな交通量および通行権価格の day-to-day ダイナミクスを導出する．さらに，そのダイナミクスの安定性を解析し，交通量および通行権価格の期待値が社会的最適状態に収束することを示す．

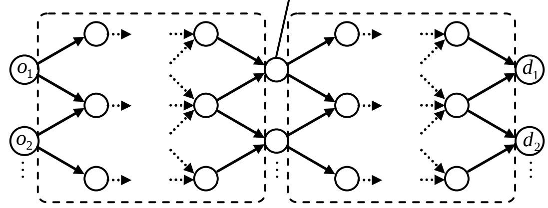

図-1　本研究で対象とするネットワークの概念図

２．モデルの設定

（１）状況設定

本研究では，高速道路ネットワーク（上流側）と一般道ネットワーク（下流側）から成るネットワーク*G*（図-1参照）を解析対象とする．高速道路ネットワークと一般道ネットワークは，オフランプ・ノードで分離される．ここで，高速道路ネットワークでは，渋滞 (queuing congestion) は発生するが，交通量の増加に伴う

---

*キーワーズ : 交通流，交通情報，TDM，ITS

**正会員，工修，国土交通省 中国地方整備局
岡山国道事務所　工務課
（〒700-8539　岡山市富町2丁目19-12 TEL 086-214-2220）
***正会員，工博，東北大学大学院　情報科学研究科
（〒980-8579　仙台市青葉区荒巻字青葉06
TEL:022-795-7507，FAX:022-795-7505）
****学生員，東北大学大学院　情報科学研究科

旅行時間の遅れである混雑 (flow congestion) は発生しないと仮定する．一方，一般道ネットワークでは，混雑は発生するが，渋滞は発生しないと仮定する．

ネットワークG上のノード集合をN，有向リンクの集合をLと表記する．ノード集合N はその部分集合として，トリップが発生する起点の集合Oおよび終了するノードDを含む．また，各ノードは整数の連番で区別され，各リンクはその上流ノードが i，下流ノードが j である場合に (i, j) と表される．さらに，トリップの起点・終点ペア(ODペア)を (o, d) と表し，その集合をWとする．ODペア(o, d) の経路 $r(o, d)$ の集合を$R(o, d)$ とする．

高速道路ネットワーク上の各リンク(i, j) は，単位時間当たりの交通容量が μ^{ij} の point queue モデルで表されると仮定する．また，高速道路ネットワークでは flow congestion は発生しないため，リンク旅行時間c^{ij} は所与の定数である．一方，一般ネットワークでの旅行時間は，単位時間当たりの交通量 y の関数として$c^{ij}(y)$ と表せると仮定する．ただし，$\partial c^{ij}(y)/\partial y \geq 0$ である．

本研究では，within-dayでの時刻別の交通流パターンと，その交通流パターンのday-to-dayダイナミクスを考える．そのために，時間の流れを日付$t \in T$ および t 内での時刻$s \in S$ の2つの変数で表現する．ここで，t, s は共に離散的な変数とし，時間の流れに沿って整数の連番で区別する．ネットワークの利用者は，起点から終点へ1日 1回のトリップを行う．ここで，ODペア(o, d) 別の交通量Q^{od} は所与の定数とする．

（２）導入する制度

ネットワーク通行権取引制度では，queuing congestion を解消することはできるが，flow congestionを解消することは困難である．そこで，本研究では一般道ネットワークでの flow congestion を解消するために，ネットワーク通行権取引制度に加えて進化ゲーム理論に基づく混雑料金制度を導入する．

a）ネットワーク通行権取引制度

道路管理者は，高速道路ネットワークでのqueuing congestion を解消するため，全ての時刻で高速道路ネットワーク上の全てのリンクに対してボトルネック通行権を発行する．ここで，単位時間当たりの通行権発行枚数は，ボトルネック容量μ^{ij} 以下とする．この条件下では，各リンクの流入率が常にボトルネック容量以下となり，渋滞は原理的に発生しない．

通行権市場では，全リンクの発行時間帯内の全時刻の通行権が取引されているものとする．また，市場は独占・寡占等の生じない完全競争的市場であり，時刻別通行権の超過需要量に応じて価格が調整されると仮定する．

b) 混雑料金制度

道路管理者は，一般道ネットワーク上での flow congestion を解消するため，一般道ネットワーク上の全てのリンクに対して混雑料金を賦課する．ここで，道路管理者は交通量は観測できるが，ネットワーク利用者の需要関数情報を知ることはできない．そこで，進化ゲーム理論に基づいて，実現した交通量に対して myopic に混雑料金を調整する[3]．この具体的な混雑料金の調整方法は，4 章で詳しく述べる．

（３）**Multi-Agent System**

本研究では，各ネットワーク利用者の通行権取引を，利用者ごとの車両に搭載された agent software が代行する状況を考える．つまり，各 agent は，利用者に代わって一般化交通費用を最小化する（と予想される）終点到着時刻・経路を選択する．そのために，以下の状況を想定する : (*i*) 全ての利用者は，車両に agent software を搭載しているものとする．そして，(*ii*)各agent software は，何らかのルールに基づいて終点到着時刻・経路を選択する．さらに，(*iii*)各agent software が市場で通行権取引を行い，その結果に基づいて利用者がトリップを行う．

３．達成すべき社会的最適配分状態

本論文では，agent system により達成すべき状態を，利用者が費やす一般化交通費用の総和（社会的交通費用）が最小化される状態（社会的最適状態）と考える :

[P-1] $$\min_{(\mathbf{q},\mathbf{y})\in\Omega} \gamma \sum_{s\in S}\sum_{ij\in L} y_s^{ij} c^{ij}(y_s^{ij}) + \sum_{od\in W}\sum_{s\in S} q_s^{od} d_s \tag{3.1}$$

ここで，γ は時間価値係数，y_s^{ij} はリンク(i, j)の時刻 s での流入交通量である．また，q_s^{od} は終点到着時刻 s の OD 交通量，d_s は時刻sに終点に到着した場合のスケジュール費用である．また，許容領域Ω は，フローが物理的に満たすべき以下の制約条件を満たす領域である : (*i*) 利用者数の保存則，(*ii*) 各ノードでのフロー保存則，(*iii*) リンクの容量制約，(*iv*) 非負条件.

４．制度下での価格ダイナミクス

（１）通行権価格

通行権価格 p_s^{ij} は，本来は組合せオークションなどにより，agentの行動と市場のミクロな条件から決定されるものである．しかし，組合せオークションの枠組みを直接的に適用しようとすると，市場は大規模な中央集権的システムとなる．このようなシステムは管理者に大きな負担を強いるとともに，自律分散的システムに比べて安定的な運用が難しい．このオークション・システムを分権化するのは今後の課題である．そこで，本研究では，市場のミクロな条件までは記述せず，通行権価格はday t での超過需要に基づいてマクロに記述するに留める．

リンク(i, j) の時刻sの通行権の需要量をD_s^{ij} と表す．このとき，供給量は常に一定値μ^{ij} であるため，通行権の t での超過需要は$(D_s^{ij}(t) - \mu^{ij})$ と表せる．このとき，

通行権価格 $p_s^{ij}(t)$ (≥ 0) は超過需要に応じて次のように調整されると仮定する．

$$p_s^{ij}(t+1) = \max.[(D_s^{ij}(t+1) - \mu^{ij}) \cdot K + p_s^{ij}(t),\ 0] \qquad (4.1)$$

ここで，Kは正の定数である．

（2） 混雑料金

道路管理者はネットワーク利用者の需要関数情報を知らないため，最適な混雑料金を各dayで賦課することはできない．そこで，tでの交通量$\mathbf{y}(t)$ に対して次式で混雑料金をmyopicに更新するものとする．

$$\lambda_s^{ij}(t) = \gamma\, y_s^{ij}(t) \cdot \frac{\partial c^{ij}(y_s^{ij}(t))}{\partial y_s^{ij}} \qquad (4.2)$$

上式は，社会的最適状態$\mathbf{y}^*$ ではflow congestion を解消する混雑料金と一致するような関数となっている．

５． Agentの行動ルール

（１）Agentの行動の枠組み

本研究では，各agentのミクロな行動ルールを進化・学習ゲーム理論[4)]の枠組みで考える．進化・学習ゲーム理論は，動学化されたゲーム理論であり，繰返しゲームを行うことを想定する．進化・学習ゲーム理論では，従来のゲーム理論でプレイヤーに仮定されていた先読みの合理性を必要とせず，temporalに実現する結果に対して近視眼的に戦略を選択するプレイヤーを仮定する．

本研究では，各 agent をゲームにおけるプレイヤーとし，各 agent が day t ごとに繰返しゲームを行うものとする．Day t で agent が選択する経路・終点到着時刻の組合せ$\{r,s\}$を t におけるα の戦略と定義する．また，α がとり得る戦略の集合を戦略空間と定義する．このとき，ODペア(o,d)の agent の戦略空間は$\{R(o,d),S\}$である．つまり，同一ODペアを持つ agent の戦略空間は等しい．また，本研究では，それぞれの戦略を選択する agent 数に応じて利得（一般化交通費用）が決定される．

（２）行動ルールの設定

本研究では，day t で実現した交通流パターンのみを参照して day t +1 の戦略を決定する perturbed best response (PBR) モデルを考える．

PBR モデルの枠組みでは，各 agent はday t での戦略$\{r,s\}$ごとの利得を参照する．ここで，ODペア(o,d)の agent の戦略$r(o,d)$に関する利得は，経路$r(o,d)$に含まれるリンクのコストの和として定義される：

$$U^{r(o,d)}(t) \equiv -\sum_{ij \in L} (p^{ij}(t) + \lambda^{ij}(t) + \gamma\, c^{ij}(t) + d^{ij}) \cdot \delta_{ij,r(o,d)} \qquad (5.1)$$

ただし，$\delta_{ij,r(o,d)}$ はリンク(i,j)が経路 $r(o,d)$に含まれれば 1，そうでなければ 0 をとる Kronecker の delta である．各エージェントは，式 (5.1) で定義される利得と確率的な誤差項の和を最大化する戦略を選択する．

本研究では，agent の戦略$\{r,s\}$の選択行動は次のような 2 段階の選択になっていると考える．すなわち，agent α は下位選択として，終点到着時刻 s を与件として，PBR により経路 r を選択する：

$$-\pi_s^{\alpha}(t) = \max_{r \in R(o,d)}.\ U_s^{r(o,d)}(t) + \xi_{r(o,d)}^{\alpha}(t) \qquad (5.2)$$

$$r_s^{\alpha}(t+1) = \arg.\max_{r \in R(o,d)}.\ U_s^{r(o,d)}(t) + \xi_{r(o,d)}^{\alpha}(t) \qquad (5.3)$$

次に，上位選択として，下位選択での結果$\pi_s^{\alpha}(t)$をもとに，終点到着時刻sを選択する：

$$s^{\alpha}(t+1) = \arg.\max_{s \in S}. -\pi_s^{\alpha}(t) - d_s + \varepsilon_s^{\alpha}(t) \qquad (5.4)$$

ただし，$\xi_{r(o,d)}^{\alpha}, \varepsilon_s^{\alpha}$ は確率的な誤差項である．つまり，agent は各 day t で混合戦略をとる．

さらに，$\xi_{r(o,d)}^{\alpha}, \varepsilon_s^{\alpha}$ が独立な Gumbel 分布に従うと仮定する．このとき，ランダム効用理論より， OD ペア(o,d)を持つ agent α の上位／下位選択での経路 r，終点到着時刻sの選択確率$l_s^{r(o,d)}, h_s$ は次のような nested logit モデルで表せる：

$$l_s^{r(o,d)}(t+1) = \exp[\theta U^{r(o,d)}(t)] / \sum_{\tilde{r} \in R(o,d)} \exp[\theta U^{\tilde{r}(o,d)}(t)] \qquad (5.5)$$

$$h_s(t+1) = \exp[-\eta(\pi_s^{od}(t) + d_s)] / \sum_{\tilde{s} \in S} \exp[-\eta(\pi_{\tilde{s}}^{od}(t) + d_{\tilde{s}})] \qquad (5.6)$$

ただし，$\pi_s^{od}(t)$ は下位選択に対応する期待最大効用である．また，θ, η はともに logit パラメータである．これらの選択確率$l_s^{r(o,d)}, h_s$ を用いて，agent α の戦略$\{r,s\}$の選択確率$b_s^{r(o,d)}$ は次のように求められる．

$$b_s^{r(o,d)}(t) = l_s^{r(o,d)}(t) \cdot h_s(t) \qquad (5.7)$$

つまり，同一のODペア(o,d)を持つエージェントは，選択確率$\mathbf{b}$ が等しくなる．

なお，本節で定めた行動ルールに従って各 agent が戦略を選択すると，あるリンク(i,j)の時刻sの通行権の需要量が供給量μ^{ij} を上回る（ i.e. 売切れる ）という場合が発生する．この場合には，μ^{ij} を上回った agent は通行権の売切れたリンクを含まない戦略空間に対して，再度 perturbed best response により戦略を決定するものとする．これを有限回繰り返すことにより，最終的にリンクの容量制約を満たすような agent の戦略$r^{\alpha}(t)$が定まる．

ただし，day t+1での通行権はday t でのリンク旅行時間を参照して購入されるため，通行権時刻は，実際のリンク流入時刻と必ずしも一致しない．つまり，想定していた時刻よりも早く／遅くリンクに流入するという状況が生じる．本研究では，このような早着／遅着の影響は少なく，queuing congestionは発生しない（eg. 道路管理者が通行権の発行数を実際の容量よりも低めに設定している）と仮定する．実際， 道路管理者は，利用者の確率的な到着を考慮して発行枚数を決定すると考えられるため，このような仮定は現実性を損なうものではない．

６．ダイナミクスの収束性

（１）期待値のダイナミクス

到着時刻別の経路交通量 $f_s^{r(o,d)}(t)$ と通行権価格 $p_s^{ij}(t)$ のダイナミクスは確率的である．そこで，以下では $\mathbf{f}(t),\mathbf{p}(t)$ の期待変化率を考えることにより，ダイナミクスの収束性を議論する．

経路交通量および通行権価格を，まとめて $\mathbf{X}(t)\equiv[\mathbf{f}(t)\ \ \mathbf{p}(t)]^T$ と書く．ここで，$\mathbf{X}(t)$ がある状態 $\mathbf{x}=[\mathbf{f}\ \ \mathbf{p}]^T$ であるとすると，利得 $\mathbf{U}$ および各戦略の選択確率 $\mathbf{b}$ は day t での状態 $\mathbf{x}$ のみの関数として $\mathbf{U}(t)=\mathbf{U}(\mathbf{x})$，$\mathbf{b}(t)=\mathbf{b}(\mathbf{x})$ と表せる．このとき，経路交通量 $\mathbf{f}(t)$ の期待変化率は，状態 $\mathbf{x}$ のみの関数として以下のように表せる．

$$E\,[\mathbf{f}(t+1)-\mathbf{f}(t)\,|\,\mathbf{X}(t)=\mathbf{x}]=\hat{\mathbf{F}}(\mathbf{x})-\mathbf{f} \tag{6.1}$$

ただし，$\hat{F}_s^{r(o,d)}(\mathbf{x})$ は状態 $\mathbf{x}$ に対して戦略 $\{r,s\}$ を採用する agent 数の期待値である．つまり，期待変化率は，day t での状態 $\mathbf{f}$ から $\hat{\mathbf{F}}$ への変化そのものである．

５章（２）で定めたルールに従って通行権が配分されるとき，戦略 $\{r, s\}$ を採用する agent 数の期待値 $\hat{F}_s^{r(o,d)}(\mathbf{x})$ は，次の最適化問題の解と一致する（証明は菊地等[5] 参照）．

[P-2]

$$\hat{\mathbf{F}}(\mathbf{x})=\arg\max_{\hat{\mathbf{f}}\geq\mathbf{0},\hat{\mathbf{f}}\in\Theta}.\mathbf{U}^T(\mathbf{x})\hat{\mathbf{f}}-\hat{\mathbf{q}}^T(\hat{\mathbf{f}})\mathbf{v}_{\mathbf{L}}(\hat{\mathbf{f}})-\mathbf{Q}^T\mathbf{v}_{\mathbf{H}}(\hat{\mathbf{f}}) \tag{6.2}$$

ただし，Θ は(i) 利用者数の保存則，(ii) リンクの容量制約，(iii) 非負条件を満たす経路交通量の許容領域，$\mathbf{v}_{\mathbf{L}},\mathbf{v}_{\mathbf{H}}$ はそれぞれ下位／上位選択に対応したエントロピー項である．

一方，式 (4.1) から，通行権価格の期待変化率も，状態 $\mathbf{x}$ のみの関数として次のように表せる．

$$E[\mathbf{p}(t+1)-\mathbf{p}(t)\,|\,\mathbf{X}(t)=\mathbf{x}]=\max.[(\hat{\mathbf{D}}(\mathbf{x})-\boldsymbol{\mu})\cdot K,-\mathbf{p}] \tag{6.3}$$

ただし，$\hat{D}^{ij}(\mathbf{x})$ は t での交通量・通行権価格から計算される，(i,j)の通行権需要の期待値である．

以上から，$\mathbf{f}(t),\mathbf{p}(t)$ の期待変化率が式 (6.1),(6.3) で表された．さらに，$\mathbf{X}(t)$の期待変化率は連続時間では次のようなマルコフ過程に従うことが導くことができる．

$$E[\dot{\mathbf{X}}(t)\,|\,\mathbf{X}(t)=\mathbf{x}]=\begin{bmatrix}\kappa_f\cdot\{\hat{\mathbf{F}}(\mathbf{x})-\mathbf{f}\}\\ \kappa_p\cdot\mathbf{H}(\mathbf{x})\end{bmatrix} \tag{6.4}$$

ただし，κ_f,κ_p は正の定数，H^{ij} は次のような関数：

$$H_s^{ij}(\mathbf{x})=\max[\hat{D}_s^{ij}(\mathbf{x})-\mu^{ij},0] \tag{6.5}$$

このとき，$\mathbf{X}(t)$のダイナミクスの休止点について，次の命題が成立する（証明は菊地等[5] 参照）．

命題１：式(6.4)で表されるマルコフ過程について，$E[\dot{\mathbf{X}}(t)\,|\,\mathbf{X}(t)=\mathbf{x}]=\mathbf{0}$ となるような点 $\mathbf{x}^*$ は，$\theta\to\infty$ の極限では，問題 [P-1] で表される社会的最適状態に一致する．

（２）収束証明

$\mathbf{X}(t)$の収束性を議論するために，連続・微分可能な次のような関数$\boldsymbol{\Pi}(\mathbf{x})$ を考える．

$$\begin{aligned}\Pi(\mathbf{x})\equiv&-\mathbf{U}^T(\mathbf{x})\mathbf{f}+\hat{\mathbf{q}}^T(\mathbf{f})(\mathbf{v}_{\mathbf{L}}(\mathbf{f})+\mathbf{w}_{\mathbf{L}}(\mathbf{x}))\\&+\mathbf{Q}^T(\mathbf{v}_{\mathbf{H}}(\mathbf{f})+\mathbf{w}_{\mathbf{H}}(\mathbf{x}))\\&+\frac{1}{2}(\boldsymbol{\mu}-\hat{\mathbf{D}}(\mathbf{x}))^T(\boldsymbol{\mu}-\hat{\mathbf{D}}(\mathbf{x}))\end{aligned} \tag{6.6}$$

ただし，$\mathbf{w}_{\mathbf{L}}(\mathbf{x}),\mathbf{w}_{\mathbf{H}}(\mathbf{x})$ は下位／上位選択に対応した期待最大効用である．式(6.6)で定義される$\boldsymbol{\Pi}(\mathbf{x})$ に定数を加えた関数は，式(6.4)で表される期待変化に対する Lyapunov 関数となる．よって，確率近似理論により，期待変化率に関して次の命題が成立する．

命題２：5章で与えた agent の行動ルールの下では，経路交通量$\mathbf{f}(t)$，通行権価格$\mathbf{p}(t)$の期待変化率のダイナミクスは大域的に収束する．

以上の命題 1, 2 から，本研究で設定した agent の行動ルールから導かれるフローパターンは大域的に収束し，その均衡状態は社会的最適状態と一致することが示された．

７．おわりに

本研究では，混雑の存在するネットワークを対象として，ネットワーク通行権取引制度を自律分散的にインプリメントするような multi-agent systemを提案した．より具体的には，agentのミクロな行動ルール群を設定し，それに対応した交通量，通行権価格のマクロなダイナミクスを導いた．さらに，その方程式系を解析し，交通量・通行権価格の期待値のダイナミクスが社会的最適状態に収束することを示した．

参考文献

1) 赤松隆：一般ネットワークにおけるボトルネック通行権取引制度, *土木学会論文集D*, Vol.63, No.3, pp.287-301, 2007.
2) 赤松隆, 佐藤慎太郎, Nguyen Xuan Long：時間帯別ボトルネック通行権取引制度に関する研究, *土木学会論文集D*, Vol.62, No.4, pp.605-620, 2006.
3) Sandholm, W.H.: Evolutionary Implementation and Congestion Pricing, *Review of Economic Studies*, pp.667-689, 2002.
4) Vega-Redondo, F.: *Economics and the Theory of Games*, Cambridge university press, 2003.
5) 菊地志郎，赤松隆：進化ゲーム理論に基づいたネットワーク通行権取引制度の自律分散的インプリメンテーション，*土木計画学研究・論文集*，（刊行予定），2008.