Stationary Equilibrium of Mean Field Games with Congestion-dependent Sojourn Times

Extended Abstract

Costas Courcoubetis
The Chinese University of Hong Kong
Shenzhen, China
costas@cuhk.edu.cn

Antonis Dimakis
Athens University of Economics and Business
Athens, Greece
dimakis@aueb.gr

ABSTRACT

We consider stationary equilibria of mean-field games between agents which follow continuous time semi-Markov decision processes with finite states and actions, when congestion affects their state-sojourn times but not the reward and transition structure. Games of this type arise in situations where selfish agents either traverse or circulate a network of congestible resources, as in routing games and models of driver mobility in ride-hailing platforms.

A variational characterization of equilibria is employed to establish existence and uniqueness of average rewards. In contrast to ordinary routing games, where the price of anarchy can be unbounded, the latter equals 2 when agents never exit.

KEYWORDS

Mean field games; stationary equilibrium; selfish routing

ACM Reference Format:
Costas Courcoubetis and Antonis Dimakis. 2023. Stationary Equilibrium of Mean Field Games with Congestion-dependent Sojourn Times: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

The literature on stationary equilibria in general mean field games (e.g., see [1, 5, 8, 9, 13]) is primarily focused in obtaining existence results and is not very informative about uniqueness and efficiency of the equilibria. To the best of the authors' knowledge no results on the price of anarchy exist in the literature of general mean field games in the stationary case. (Although, some progress has been made in the nonstationary linear case, e.g., see [4, 6, 10][1].) In contrast, when congestion affects only the time evolution (although in a substantially general way) we show that the price of anarchy is 2. Application specific models found in the literature, e.g., for ride-hailing in [2, 3], explore properties beyond existence, but the methods do not seem to generalize and use assumptions pertaining to the specific domain (e.g., competitive pricing, geographical symmetry in [3]).

All the proofs, applications to ride-hailing and selfish routing in closed networks (as opposed to open, studied in [12]) can be found in the full paper version in [7].

[1]We thank the anonymous reviewer for bringing these to our attention

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

2 MODEL OF AN INDIVIDUAL PLAYER

A nonatomic set of players of L types, indexed by $l = 1, \dots, L$, compete for resources. (A player of type l is referred to as an l-player.) All types share the same finite sets of states, S, and actions, A. The state of any l-player evolves according to a continuous time semi-Markov decision process. The actions are chosen at state transitions, and choosing $\alpha \in A$ in $i \in S$ results in the following sequence:

(1) An immediate reward $r^l_{i\alpha}$ is awarded. Let the reward vector be $r^l = (r^l_{i\alpha}, i \in S, \alpha \in A)$.
(2) The next state is chosen independently of the past according to the transition probabilities $p^l_{i\alpha j}, i, j \in S, \alpha \in A$, where $p^l_{i\alpha j} \geq 0, \sum_{j \in S} p^l_{i\alpha j} = 1$. Let $p^l = (p^l_{i\alpha}, i \in S, \alpha \in A)$.
(3) Conditionally on the next state being j, the transition to j occurs after a random time, independently of the past. The mean sojourn time, $\tau^l_{i\alpha}$, of state i when α is chosen, depends on the interaction with the other players, including those of other types, and is defined in Section 3.
(4) Upon arrival to j the player decides the next action and the process continues as above.

3 CONGESTION-DEPENDENT SOJOURN TIMES

Players exhibit a mean field type of interaction where their state sojourn times $\tau^l = (\tau^l_{i\alpha}, i \in S, \alpha \in A)$ depend on the distribution of players on $S \times A$. Let $\mu^l \in \mathcal{M}(S \times A)$ be a measure on $S \times A$, where $\mu^l(\{(i, \alpha)\})$, or simply $\mu^l_{i\alpha}$, is the mass of l-players which choose action α in i. ($\mu^l(S \times A)$ gives the total mass of l-players.) The mean sojourn time is decomposed as

$$\tau^l_{i\alpha} = w^l_{i\alpha}(\mu) + t^l_{i\alpha}(x(\mu)), \tag{1}$$

where $w^l_{i\alpha}(\mu)$ is the time the player waits to collect the resources required for l-players to execute action α in i, and $t^l_{i\alpha}(\cdot)$ is the action execution time which is allowed to depend on the rates $x(\mu) = (x^{l'}_{i'\alpha'}(\mu), i' \in S, \alpha' \in A, l' = 1, \dots, L)$. ($x^l_{i\alpha}(\mu)$ is the rate of l-players entering i and choosing α per unit of time.)

$x(\mu)$ and $w(\mu) = (w^l_{ia}(\mu), i \in S, \alpha \in A, l = 1, \dots, L)$ are defined by Lemma 1 below.

Assumption 1. *There exists convex $G : \mathbb{R}^{L \times S \times A}_+ \longrightarrow \mathbb{R}_+$ with $G(0) = 0$ such that*

$$\frac{\partial G}{\partial x^l_{i\alpha}} = t^l_{i\alpha}, \ \forall i \in S, a \in A, l = 1, \dots, L. \tag{2}$$

For example, $t^l_{i\alpha}(x) = g_{i,\alpha}\left(\sum_{l'} x^{l'}_{i\alpha}\right)$ for nondecreasing $g_{i,\alpha}$, satisfies Assumption 1, since (2) holds with

$$G(x) = \sum_{i,\alpha} \int_0^{\sum_l x^l_{i\alpha}} g_{i,\alpha}(u)du. \tag{3}$$

Lemma 1. *Let Assumption 1 hold, and let $K \in \mathbb{N}$, $a^l_{k,i\alpha} \geq 0$, $b_k > 0$ for all $k = 1, \ldots, K$, $l = 1, \ldots, L$, $i \in S$, $\alpha \in A$. There exists a unique continuous map $\mu \mapsto (x(\mu), w(\mu))$ defined on $\mathcal{M}(S \times A)^L$ such that it satisfies:*

$$\left[w^l_{i\alpha}(\mu) + t^l_{i\alpha}(x(\mu))\right] x^l_{i\alpha}(\mu) = \mu^l_{i\alpha}, \tag{4}$$

$$w^l_{i\alpha}(\mu) = \sum_{k=1}^K a^l_{k,i\alpha} d_k, \tag{5}$$

$$\sum_{l,i,\alpha} a^l_{k,i\alpha} x^l_{i\alpha}(\mu) < b_k \Longrightarrow d_k = 0, \tag{6}$$

$$\sum_{l,i,\alpha} a^l_{k,i\alpha} x^l_{i\alpha}(\mu) \leq b_k. \tag{7}$$

for some $d_k \geq 0$, $k = 1, \ldots, K$.

Equation (4) is Little's identity. Equation (5) is because $w^l_{i\alpha}(\mu)$ is assumed to arise from waiting to obtain a mix of K resources indexed by $k = 1, \ldots, K$. (The k-th resource is referred to as k-resource.) $a^l_{k,i\alpha} \geq 0$ is the amount of k-resource that an l-player requires to execute action α in i, and let $d_k \geq 0$ be the waiting time to obtain a unit of k-resource. Assuming that players wait for one unit of resource at a time yields (5). Equation (6) implies that d_k should be zero if the k-resource constraint (7) is not active, i.e., the k-resource is not exhausted.

We write $\tau^{l,\mu}$ to emphasize the dependence of τ^l on μ.

4 MEAN FIELD GAME

For any initial state i, the ergodic average of rewards of an l-player following a Markovian policy σ is

$$\liminf_T \frac{1}{T} E\left(\sum_{n=1}^{N_T} r^l_{X_n A_n} \middle| X_0 = i\right), \tag{8}$$

where N_T is the number of transitions before time T, X_n is the state visited at the n-th transition, and A_n is the action chosen by σ at that instant. Let $V(r^l, p^l, \tau^{l,\mu})$ be the *optimal average reward* per unit time, i.e., the supremum of (8) over all policies σ, which does not depend on i under the following assumption (see [11]).

Assumption 2 (Weakly communicating model, [11]). *For every $l = 1, \ldots, L$, the transition model p^l is weakly communicating.*

Furthermore, assume that a player always prefers to participate in the game regardless of the strategies of the other players.

Assumption 3 (Participation). *There exists $\mu \in \mathcal{M}(S \times A)^L$ such that the optimal average reward is positive for any player type, i.e., $V(r^l, p^l, \tau^{l,\mu}) > 0$ for every l.*

Next, the equilibrium of the mean field game is defined, in stationarity:

Definition 1. *$\mu_o \in \mathcal{M}(S \times A)^L$ is a (stationary) equilibrium if and only if*

(1) μ^l_o is time-invariant, i.e.,

$$\sum_{j \in S, \alpha \in A} x^l_{j\alpha}(\mu_o)\left(\delta_{ij} - p^l_{j\alpha i}\right) = 0, \quad \text{for each } i \in S, \tag{9}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, and

(2) the optimal average reward equals the aggregate average reward per unit mass, i.e.,

$$V\left(r^l, p^l, \tau^{l,\mu_o}\right) = \frac{\sum_{i,\alpha} r^l_{i\alpha} x^l_{i\alpha}(\mu_o)}{\mu^l_o(S \times A)}, \tag{10}$$

for each type l.

5 MAIN RESULTS

A key result is a characterization of equilibria as the optimal solutions of a convex optimization problem.

Theorem 1. *Under Assumptions 1, 2, 3, $\mu_o \in \mathcal{M}(S \times A)^L$ is an equilibrium if and only if $x(\mu_o) \in \mathbb{R}_+^{L \times S \times A}$ is an optimal solution of:*

$$\max \quad \sum_l \mu^l_o(S \times A) \log\left(\sum_{i,\alpha} r^l_{i\alpha} x^l_{i\alpha}\right) - G(x) \tag{11}$$

$$\text{s.t.} \sum_{l,i,\alpha} a^l_{k,i\alpha} x^l_{i\alpha} \leq b_k, \ , k = 1, \ldots, K, \tag{12}$$

$$\sum_{j,\alpha} x^l_{j\alpha}\left(\delta_{ij} - p^l_{j\alpha i}\right) = 0, \ l = 1, \ldots, L, i \in S, \tag{13}$$

$$\text{over } x = (x^l_{i\alpha}, l = 1, \ldots, L, i \in S, \alpha \in A) \in \mathbb{R}_+^{L \times S \times A},$$

and $w^l_{i\alpha}(\mu_o) = \sum_k a^l_{i\alpha} d_k$ for all l, i, α, where $d_1, \ldots, d_K$ are optimal Lagrange multipliers for the resource constraints (12).

Corollary 1. *Under Assumptions 1, 2, 3, the following hold:*

(a) For any set of player masses $m^l > 0$, $l = 1, \ldots, L$, an equilibrium μ_o exists with $\mu^l_o(S \times A) = m^l$ for every l.

(b) The optimal average reward $V\left(r^l, p^l, \tau^{l,\mu_o}\right)$, $l = 1, \ldots, L$ and $G(x(\mu_o))$ assume the same value for every equilibrium μ_o with the same player masses. That is, their values depend on μ_o only through $\mu^l_o(S \times A)$, $l = 1, \ldots, L$.

5.1 Price of Anarchy

Here we restrict attention to games with a single player type ($L = 1$). (The player type index l is dropped.)

Definition 2. *The optimal aggregate average reward $W(m)$ for player mass m is*

$$W(m) = \sup_{\mu \text{ stationary}, \ \mu(S \times A) \leq m} \sum_{i,\alpha} r^l_{i\alpha} x^l_{i\alpha}(\mu). \tag{14}$$

$W(m)$ and the aggregate average reward at an equilibrium with player mass m, do not coincide in general. The price of anarchy, i.e., the largest possible ratio between the two, is

$$\sup \frac{W(\mu_o(S \times A))}{\sum_{i,\alpha} r^l_{i\alpha} x^l_{i\alpha}(\mu_o)}, \tag{15}$$

where the supremum is taken over all model parameters and corresponding equilibria $\mu_o \in \mathcal{M}(S \times A)$ for which Assumptions 1, 2, 3 are true.

Proposition 1. *The price of anarchy is 2.*

REFERENCES

- [1] Sachin Adlakha, Ramesh Johari, and Gabriel Y Weintraub. 2015. Equilibria of dynamic games with many players: Existence, approximation, and market structure. *Journal of Economic Theory* 156 (2015), 269–316.
- [2] Siddhartha Banerjee, Ramesh Johari, and Carlos Riquelme. 2015. Pricing in Ride-Sharing Platforms: A Queueing-Theoretic Approach. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation* (Portland, Oregon, USA) *(EC '15)*. Association for Computing Machinery, New York, NY, USA, 639. https://doi.org/10.1145/2764468.2764527
- [3] Kostas Bimpikis, Ozan Candogan, and Daniela Saban. 2019. Spatial pricing in ride-sharing networks. *Operations Research* 67, 3 (2019), 744–769.
- [4] Pierre Cardaliaguet and Catherine Rainer. 2019. On the (in) efficiency of MFG equilibria. *SIAM Journal on Control and Optimization* 57, 4 (2019), 2292–2314.
- [5] René Carmona, François Delarue, et al. 2018. *Probabilistic theory of mean field games with applications I-II.* Springer.
- [6] René Carmona, Christy V Graves, and Zongjun Tan. 2019. Price of anarchy for mean field games. *ESAIM: Proceedings and Surveys* 65 (2019), 349–383.
- [7] Antonis Dimakis Costas Courcoubetis. 2023. Stationary Equilibrium of Mean Field Games with Congestion-dependent Sojourn Times. http://stecon.cs.aueb.gr/media/1229/mean_field_game_sojourn_times.pdf
- [8] Hugo A. Hopenhayn. 1992. Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica* 60, 5 (1992), 1127–1150. http://www.jstor.org/stable/2951541
- [9] Boyan Jovanovic and Robert W. Rosenthal. 1988. Anonymous sequential games. *Journal of Mathematical Economics* 17, 1 (1988), 77 – 87. https://doi.org/10.1016/0304-4068(88)90029-8
- [10] Mathieu Lauriere. 2021. Numerical methods for mean field games and mean field type control. *Mean Field Games* 78 (2021), 221.
- [11] Martin L Puterman. 2014. *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons.
- [12] Tim Roughgarden and Éva Tardos. 2002. How Bad is Selfish Routing? *J. ACM* 49, 2 (March 2002), 236–259. https://doi.org/10.1145/506147.506153
- [13] Piotr Wiecek and Eitan Altman. 2015. Stationary Anonymous Sequential Games with Undiscounted Rewards. *Journal of Optimization Theory and Applications* 166, 2 (2015), 686–710. https://doi.org/10.1007/s10957-014-0649-9