A dynamic approach for load balancing

Dominique Barth, Olivier Bournez, Octave Boussaton, Johanne Cohen

A Dynamic Approach for Load Balancing

Dominique Barth
Université de Versailles.
Laboratoire PRiSM,
45, avenue des Etats-Unis
78000 Versailles, FRANCE
Dominique.Barth
@prism.uvsq.fr

Olivier Bournez
Ecole Polytechnique.
Laboratoire LIX,
91128 Palaiseau Cedex,
FRANCE
Olivier.Bournez
@polytechnique.fr

Octave Boussaton
LORIA,
615 Rue du Jardin Botanique,
54602 Villers-Lès-Nancy,
FRANCE
Octave.Boussaton
@loria.fr

Johanne Cohen
CNRS.
Université de Versailles,
45, avenue des Etats-Unis
78000 Versailles, FRANCE
Johanne.Cohen
@prism.uvsq.fr

ABSTRACT

We study how to reach a Nash equilibrium in a load balancing scenario where each task is managed by a selfish agent and attempts to migrate to a machine which will minimize its cost. The cost of a machine is a function of the load on it. The load on a machine is the sum of the weights of the jobs running on it. We prove that Nash equilibria can be learned on that games with incomplete information, using some Lyapunov techniques.

1. INTRODUCTION

We consider in this paper a now classical settings for selfish load balancing, where some dynamical aspects are introduced. We consider tasks or jobs (players) whose objective is to minimize their own cost. Such games are sometimes called the *KP model* [21, 28]. We consider a particular learning scenario for tasks in this game that we prove to converge to Nash equilibria of the game. Our learning algorithm, already considered in [26] for general games, is proved to be convergent on these allocation games, through some Lyapunov function.

The settings is the following [21, 28]: weighted tasks or jobs (players) shall be assigned to a set of machines having different speeds so that tasks try to minimize their own cost. Their cost under an assignment A depends on the load of that machine. The load on a machine under A is the sum of the weights of the jobs running on it. We consider a system in which jobs are unsplittable so that a player's entire job will go to a machine accordingly to this player's probability vector.

More formally, we consider a set of n independent tasks (players) with weight $w_1, \dots, w_n$ respectively. We assume the weights to be positive integers. These tasks are to be associated with E, a set of m resources (machines). Let $[n] = \{1, \cdots, n\}$ denote the set of tasks and $[m] = \{1, \cdots, m\}$ the set of machines.

An *assignment* $A : [n] \to [m]$ specifies for each task i, the machine on which it runs: each task i selects exactly one machine $A[i]$ (also denoted by a_i by abuse of notation).

The load λ_ℓ of machine ℓ under the assignment A is the the sum of the weights of the tasks running on it: $\lambda_\ell = \sum_{i:a_i=\ell} w_i$. The *cost* of machine ℓ corresponds to its finish time, i.e is a function depending on its load denoted by $C_\ell(\lambda_\ell)$. The cost of task i under the assignment A corresponds to the cost on machine a_i, i.e., its cost is $c_i = C_{a_i}(\lambda_{a_i})$.

We will focus in this paper on linear (affine) cost functions: $C_\ell(\lambda) = \alpha_\ell \lambda + \beta_\ell$, with $\alpha_\ell, \beta_\ell \geq 0$. This is sometimes called the *uniformly related machine* case [28].

We assume from now that all costs are not greater than 1. This is without loss of generality, since dividing all costs by some big enough constant in what follows preserve dynamics.

We assume that tasks (players) migrate selfishly without any centralized control and only have a local view of the system. All of the tasks know how many resources (machines) are available. At each elementary step $t+1$, they know their current cost and the machine they chose at step t. We study how to reach a (mixed) Nash equilibrium in this load balancing scenario. Each agent i selects a machine (at each time step or instance of the repeated game) using a mixed strategy $q_i(t)$ at step t, with $q_{i,\ell}(t)$ denoting the probability for agent i to select machine ℓ at step t.

We are interested in building distributed learning algorithms for tasks so that $(q_i(t))_{i\in[n]}$ converges towards some (possibly mixed) Nash equilibrium of the system.

Notice that such games are known to always admit at least one pure Nash equilibrium: see [15, 9] for a proof. At least one of them is a global optimal.

2. RELATED WORK

The allocation problems that we consider in this paper can also be formally defined as a routing problem between two nodes connected by parallel edges with possibly different speeds. Each agent as an amount of traffic to map to one of the edges such as the load on this edge is as small as possible [21, 28].

Price of anarchy, introduced by [21], comparing cost of Nash equilibria to cost of the optimal (social cost) has been intensively studied on these games: see e.g. [28] for a reference introduction.

Similar games but with more general cost functions have also been considered: see e.g. [28, 4, 7, 23], and all references in [28]. Refer to [8, 12, 22] for surveys about results for selfish load balancing and selfish routing on parallel links.

There are a few works considering dynamic versions of these allocation games, where agents try to learn Nash equilibria, in the spirit of this paper.

First, notice that the proof of existence of a pure Nash equilibria can be turned into a dynamic: players play in turn, and move to machines with a lower load. Such a strategy can be proved to lead to a pure Nash equilibrium. Bounds on the convergence time have been investigated in [9, 10]. Considered algorithms are centralized and require complete information games. Obtained bounds are mostly obtained by bounding the possible variations of suitable potential functions. Since players play in turns, this is often called the *Elementary Step System*. Other results of convergence in this model, have been investigated in [16, 23, 25].

Concerning models that allow concurrent redecisions, we can mention the followings works. In [11], tasks are allowed in parallel to migrate from overloaded to underloaded resources. The process is proved to terminate in expected $O(\log\log n+\log m)$ rounds. The analysis is restricted to the case of unitary weights, and with identical machines. The considered process requires a global knowledge: one must determine whether ones load is above or under average.

A Stochastic version of best-response dynamic avoiding that latter problem has been investigated in [2]. It is proved to terminate in expected $O(\log\log n + m^4)$ rounds for uniform tasks, and uniform machines. This has been extended to weighted tasks and uniform machines in [3]. The expected time of convergence to an ϵ-Nash equilibrium is in $\mathcal{O}(nmW^3\epsilon^{-2})$ where W denotes the maximum weight of any task.

The dynamic considered in this paper has been studied in [26] for general stochastic games. It is proved in [26] that this dynamic is weakly convergent to some function solution of an ordinary differential equation. This ordinary differential equation turns out to be a replicator equation. A sufficient condition for convergence is proved. No Lyapunov function is established for systems similar to the ones considered in this paper there.

Replicator equations have been deeply studied in evolutionary game theory [19, 29]. Evolutionary game theory doesn't restrict to these dynamics, but considers a whole family of dynamics that satisfy folk theorems in the spirit of Theorem 1.

Bounds on the rate of convergence of fictious play dynamics have been established in [17], and in [20] for the best response dynamic. Fictious play has been reproved to be convergent for zero-sum games using numerical analysis methods, or more generally stochastic approximation theory: fictious play can be proved to be a Euler discretization of a certain continuous time process [19].

Evolutionary game theory has been applied to routing problems in the Wardrop traffic model in [14, 13].

A replicator equation for the allocation games considered in this paper has been considered in [1], where a Lyapunov function is established. The dynamics considered in [1] is not the same as us: we have a replicator dynamic where fitnesses are given by true costs, whereas, [1] considers marginal costs.

In [5, 6], the replicator dynamics for particular allocation games have been proved to converge to a pure Nash equilibrium by modifying game costs in order to obtain Lyapunov functions. It has also been proved that only pure Nash equilibria can be learned.

3. GAME THEORETIC SETTINGS

Each one of the n players (tasks) has the same finite set of actions (machines) E, of cardinality m: machines are known by and available to all of the players. An element of E, i.e. of $[m]$, is called a *pure strategy*.

Define functions $d_i : \prod_{j=1}^{n} E \to [0,1], 1 \leq i \leq n$, by:

$$d_i(a_1, a_2, \cdots, a_n) = c_i(j \text{ plays strategy } a_j \in E, 1 \leq j \leq n) \tag{1}$$

where $(a_1, \cdots, a_n)$ is the set of pure strategies played by the player team.

In our case, $d_i(a_1, a_2, \cdots, a_n) = C_{a_i}(\sum_{j:a_j=a_i} w_j)$.

We call it the payoff function, or utility function of player i and the objective of each player is to minimize his payoff.

We assume without loss of generality in this paper that payoffs stay in $[0, 1]$. It this doesn't hold, an affine transformation on payoffs yields this property.

As usual, we want to extend the payoff function to mixed strategies. To do so, let $\mathcal{S}$ denote the simplex of dimension

m: $\mathcal{S}$ is the set of m-dimensional probability vectors:

$$\mathcal{S} = \{q = (q_1, \cdots, q_m) \in [0,1]^m : \sum_{\ell=1}^{m} q_\ell = 1\}.$$

An element of $\mathcal{S}$ is called a *mixed strategy* for a player.

Let $K = \mathcal{S}^n$ be the space of mixed strategies.

Payoff functions d_i defined on pure strategies in equation (1) can be extended to functions $\overline{d}_i$ on the space of mixed strategies K as follows:

$$\begin{array}{rcl} \overline{d}_i(q_1, ..., q_n) & = & E[c_i | 1 \leq j \leq n \text{ plays mixed strategy } q_j] \\ & = & \sum_{j_1, \cdots, j_n} d_i(j_1, \cdots, j_n) \times \prod_{s=1}^{n} q_{s, j_s} \end{array} \quad (2)$$

where $(q_1, \cdots, q_n)$ is the set of mixed strategies played by the player team and $q_{j,\ell}$ denotes the probability for player j to play strategy ℓ.

Definition 1. *The n-tuple of mixed strategies $(\tilde{q}_1, \cdots, \tilde{q}_n)$ is said to be a Nash equilibrium (in mixed strategies), if for each i, $1 \leq i \leq n$, we have: $\overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, \tilde{q}_i, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) \leq \overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, q, \tilde{q}_{i+1}, \cdots, \tilde{q}_n)$, for all $q \in \mathcal{S}$.*

It is well known that every n-person game has at least one Nash equilibrium in mixed strategies [24].

Let K^* be defined by $K^* = (\mathcal{S}^*)^n$, where $\mathcal{S}^* = \{q \in \mathcal{S} |\ q$ is a m-dimensional probability vector with one component unity$\}$. In other words, the set K^* denotes the corners of simplex K.

Clearly, K^* is one to one correspondence with pure strategies, i.e. with $\prod_{j=1}^{n} E$: any pure strategy ℓ corresponds to unit vector e_ℓ, with ℓ^{th} component unity.

Definition 2. *The n-tuple of actions $(\tilde{a}_1, \cdots, \tilde{a}_n)$ (or equivalently the set of strategies $(e_{\tilde{a}_1}, \cdots, e_{\tilde{a}_n})$) is called a Nash equilibrium in pure strategies if for each i, $1 \leq i \leq n$, $d_i(\tilde{a}_1, \cdots, \tilde{a}_{i-1}, \tilde{a}_i, \tilde{a}_{i+1}, \cdots, \tilde{a}_n) \leq d_i(\tilde{a}_1, \cdots, \tilde{a}_{i-1}, a_i, \tilde{a}_{i+1}, \cdots, \tilde{a}_n)$, for all $a_i \in \mathcal{S}^*$.*

Unlike general games, the allocation games considered in this paper always admit Nash equilibria in pure strategies [15, 9].

Now the learning problem for the game can be stated as follows: Assume that the game repeats at time $k = 0, 1, 2, \ldots$. At any instant k, let $q_i[k]$ be the strategy employed by the i^{th} player. Let $a_i[k]$ and $c_i[k]$ be the actual action selected and the payoff received by player i respectively at time k ($k = 0, 1, 2, \ldots$). Find a decentralized learning algorithm for the players, that is, design function T_i, where $q_i[k+1] = T_i(q_i[k], a_i[k], c_i[k])$ such that $q_i[k] \to \tilde{q}_i$ as $k \to +\infty$ where $(\tilde{q}_1, \cdots, \tilde{q}_n)$ is a Nash equilibrium of the game.

4. THE CONSIDERED RANDOMIZED DISTRIBUTED ALGORITHM

We consider the following learning algorithm, already considered in [26].

Definition 3 (Considered Algorithm). *1. At every time step, each task (player) chooses an action according to his current strategy or Action Probability Vector (APV). Thus, the i^{th} player selects machine $\ell = a_i(k)$ at instant k with probability $q_{i,\ell}(k)$.*

2. Each player obtains a payoff based on the set of all actions. The rewards to player i at time k is $c_i(k)$.

3. Each player updates his APV according to the rule:

$$q_i(k+1) = q_i(k) + b \times (1 - c_i(k)) \times (e_{a_i(k)} - q_i(k)), \quad (3)$$

where $0 < b < 1$ is a parameter and $e_{a_i(k)}$ is a unit vector of dimension m with $a_i(k)^{th}$ component unity.

Notice, that componentwise, Equation (3) can be rewritten:

$$q_{i,\ell}(k+1) = \begin{cases} q_{i,\ell}(k) - b(1 - c_i(k)) q_{i,\ell}(k) & \text{if} \quad a_i \neq l \\ q_{i,\ell}(k) + b(1 - c_i(k))(1 - q_{i,\ell}(k)) & \text{if} \quad a_i = l \end{cases} \quad (4)$$

Decisions made by players are completely decentralized, at each time step, player i only needs c_i and q_i, respectively her payoff and strategy, to update his APV.

Let $Q[k] = (q_1(k), \cdots, q_n(k)) \in K$ denote the state of the player team at instant k. Our interest is in the asymptotic behavior of $Q[k]$ and its convergence to a Nash Equilibrium. Clearly, under the learning algorithm specified by (3), $\{Q[k], k \geq 0\}$ is a Markov process.

Observe that this dynamic can also be put in the form

$$Q[k+1] = Q[k] + b \cdot G(Q[k], a[k], c[k]), \quad (5)$$

where $a[k] = (a_1(k), \cdots, a_N(k))$ denotes the actions selected by the player team at k and $c[k] = (c_1(k), \cdots, c_N(k))$ their resulting payoffs, for some function $G(.,.,.)$ representing the updating specified by equation (3), that does not depend on b.

Consider the piecewise-constant interpolation of $Q[k], Q^b(.)$, defined by

$$Q^b(t) = Q[k], t \in [kb, (k+1)b), \quad (6)$$

where b is the parameter used in (3).

$Q^b(.)$ belongs to the space of all functions from $\mathbb{R}$ into K. These functions are right continuous and have left hand limits. Now consider the sequence $\{Q^b(.) : b > 0\}$. We are interested in the limit $Q(.)$ of this sequence as $b \to 0$.

The following is proved in [26], and follows for example from [27, theorem 11.2.3].

Proposition 1. *The sequence of interpolated processes $\{Q^b(.)\}$ converges weakly, as $b \to 0$, to $Q(.)$, which is the (unique) solution of Cauchy problem*

$$\frac{dQ}{dt} = \phi(Q), Q(0) = Q_0 \tag{7}$$

where $Q_0 = Q^b(0) = Q[0]$, and $\phi : K \to K$ is given by

$$\phi(Q) = E[G(Q[k], a[k], c[k])|Q[k] = Q],$$

where G is the function in Equation (5).

Recall that a family of random variable $(Y_t)_{t\in\mathbb{R}}$ weakly converges to a random variable Y, if $E[h(X_t)]$ converges to $E[h(Y)]$ for each bounded and continuous function h. As discussed in [26], weak convergence is admittedly a very weak type of convergence.

The proof of Proposition 1 is based on weak-convergence methods, non-constructive in several aspects (mainly relies on existence of limit sequences in some functional spaces), and does not provide error bounds.

Using (4), we can rewrite $E[G(Q[k], a[k], c[k])]$ in the general case as follows.

$$\begin{aligned}
&E[G(Q[k], a[k], c[k])]_{i,\ell} \\
&= q_{i,\ell}(1 - q_{i,\ell})(1 - E[c_i|Q(k), a_i = \ell]) \\
&\quad - \textstyle\sum_{\ell'\neq\ell} q_{i,\ell'} q_{i,\ell}(1 - E[c_i|Q(k), a_i = \ell']) \\
&= q_{i,\ell}[\textstyle\sum_{\ell'\neq\ell} q_{i,\ell'}(1 - E[c_i|Q(k), a_i = \ell]) \\
&\quad - \textstyle\sum_{\ell'\neq\ell} q_{i,\ell'}(1 - E[c_i|Q(k), a_i = \ell']] \\
&= -q_{i,\ell} \textstyle\sum_{\ell'}(E[c_i|Q(k), a_i = \ell] - q_{i,\ell'} E[c_i|Q(k), a_i = \ell']),
\end{aligned} \tag{8}$$

using the fact that $1 - q_{i,\ell} = \sum_{\ell'\neq\ell} q_{i,\ell'}$.

Let $h_{i,\ell}$ be the expectation of the payoff for i if player i plays pure strategy ℓ, and players $j \neq i$ play (mixed) strategy q_j. Formally,

$$h_{i,\ell}(q_1, \cdots, q_{i-1}, q_{i+1}, \cdots, q_n) = E[c_i|Q(k), a_i = \ell].$$

Let $\overline{h_i}(Q)$ the mean value of $h_{i,\ell}$, in the sense that

$$\overline{h_i}(Q) = \sum_{\ell'} q_{i,\ell'} h_{i,\ell'}(Q).$$

We obtain from (8),

$$E[G(Q[k], a[k], c[k])]_{i,\ell} = -q_{i,\ell}(h_{i,\ell} - \overline{h_i}(Q)).$$

Hence, the dynamics given by Ordinary Differential Equation (7) is componentwise:

$$\frac{dq_{i,\ell}}{dt} = -q_{i,\ell}(h_{i,\ell} - \overline{h_i}(Q)). \tag{9}$$

This is a (multi-population) replicator equation, that is to say a well known and studied dynamics in evolutionary game theory [19, 29]. In that context, $h_{i,\ell}$ is interpreted as a fitness of a given game, and $\overline{h_i}(Q)$ is the mean value of $h_{i,\ell}$ in the above sense.

In particular, solutions are known to satisfy the following theorem (an instance of the so-called evolutionary game theory's folk theorems) [19].

Theorem 1 (see e.g. [19]). *The following are true for the solutions of replicator equation (9):*

- *All Nash equilibria are stationary points.*
- *All strict Nash equilibria are asymptotically stable.*
- *All stable stationary points are Nash equilibria.*

Hence, stable stationary points correspond to Nash equilibria.

However, unstable stationary points can exist. Actually, all corners of simplex K are stationary points, as well as, from the form of (9), more generally any state Q in which all strategies in its support perform equally well. Such a state Q is not a Nash equilibrium as soon as there is an unused strategy (i.e. outside of the support) that performs better.

As limit points of a dynamic correspond to stationary points, from this theorem, we can conclude that the dynamics (9), and hence the learning algorithm when b goes to 0, will have stable limit points that correspond to Nash equilibria.

However, for general games, there is no convergence in the general case.

We will now show that for our allocation games, there is always convergence. It will then follow from previous discussions that our considered learning algorithm converges towards Nash equilibria, i.e. solves the learning problem for allocation games.

First, we specialize the dynamics for allocation games. Here, we have

$$\begin{aligned}
c_i &= C_{a_i}(\lambda_{a_i}) \\
&= \alpha_{a_i} w_i + \alpha_{a_i} \textstyle\sum_{j\neq i: a_j = a_i} w_j + \beta_{a_i} \\
&= \alpha_{a_i} w_i + \alpha_{a_i} \textstyle\sum_{j\neq i} \mathbf{1}_{a_j = a_i} w_j + \beta_{a_i}
\end{aligned}$$

where $\mathbf{1}_{a_j=a_i}$ is 1 whenever $a_j = a_i$, 0 otherwise.

Taking expectations, using $E[\mathbf{1}_{a_j=a_i}] = q_{j,a_i}$, we get

$$E[c_i] = \alpha_{a_i} w_i + \beta_{a_i} + \alpha_{a_i} \sum_{j\neq i} q_{j,a_i} w_j,$$

and

$$h_{i,\ell} = \alpha_\ell w_i + \beta_\ell + \alpha_\ell \sum_{j\neq i} q_{j,\ell} w_j.$$

We claim the following.

Theorem 2 (Extension of Theorem 3.3 from [26]). *Suppose there is a non-negative function*

$$F : K \to \mathbb{R}$$

such that for some constants $w_i > 0$, for all i, ℓ, Q,

$$\frac{\partial F}{\partial q_{i,\ell}}(Q) = w_i \times h_{i,\ell}(Q). \tag{10}$$

Then, for any initial condition, the algorithm will converge to a stationary point.

PROOF. We claim that $F(.)$ is a Liapunov function of the dynamic, i.e. that F is monotone along trajectories.

Indeed,

$$\begin{aligned}
&\frac{dF(Q(t))}{dt} \\
&= \sum_{i,\ell} \frac{\partial F}{\partial q_{i,\ell}} \frac{dq_{i,\ell}}{dt} \\
&= -\sum_{i,\ell} \frac{\partial F}{\partial q_{i,\ell}}(Q) q_{i,\ell} \sum_{\ell'} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)] \\
&= -\sum_{i,\ell} w_i h_{i,\ell}(Q) q_{i,\ell} \sum_{\ell'} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)] \\
&= -\sum_i w_i \sum_\ell \sum_{\ell'} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&= -\sum_i w_i \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)]^2 \\
&\leq 0
\end{aligned} \tag{11}$$

In above formula, we used the following fact:

LEMMA 1.

$$\begin{aligned}
&\sum_\ell \sum_{\ell'} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&= \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)]^2
\end{aligned}$$

PROOF. Indeed, we have

$$\begin{aligned}
&\sum_\ell \sum_{\ell'} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&= \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&\quad + \sum_\ell \sum_{\ell' < \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&= \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&\quad + \sum_{\ell'} \sum_{\ell' < \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)]
\end{aligned}$$

(by permuting indices)

$$\begin{aligned}
&= \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q)^2 - h_{i,\ell}(Q) h_{i,\ell'}(Q)] \\
&\quad + \sum_\ell \sum_{\ell < \ell'} q_{i,\ell'} q_{i,\ell}[h_{i,\ell'}(Q)^2 - h_{i,\ell'}(Q) h_{i,\ell}(Q)]
\end{aligned}$$

(by changing notation of indices)

$$= \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)]^2$$

□

This is clear that a replicator equation dynamic preserves simplex K, and hence that trajectories stay in compact set K.

From Liapunov Stability theorem ([18] page 194), asymptotically all trajectories will be in the set $K' = \{Q^* \in K : \frac{dF(Q^*)}{dt} = 0\}$.

Now, from (11), we know that $\frac{dF(Q^*)}{dt} = 0$ implies

$$q_{i,\ell} q_{i,\ell'}[h_{i,\ell}(Q) - h_{i,\ell'}(Q)] = 0$$

for all i, ℓ, ℓ', hence that Q^* is a stationary point of the dynamic.

A solution of Dynamic (9) that starts from some stationary points (i.e. for example a corner of K) will clearly stay invariant. Such a starting point can be unstable.

There may also exist some unstable solutions of Dynamic (9): Consider for example a dynamic that stays on some face of K where some well-performing strategy is never used. However, such trajectories are extremly unstable: We may expect the underlying stochastic learning algorithm to leave such face almost-surely.

In other words, unless a trajectory is started from some unstable corner of K, we may expect its limit points to be stable stationary points.

From Theorem 1, stable stationary points are precisely Nash equilibria. Hence, if the stochastic algorithm is not started from a corner, if a function satisfying the hypotheses of Theorem 2 can be found, the stochastic algorithm will converge to a Nash equilibrium.

We claim that such a function always exists for our allocation games. Notice that our dynamic is actually quite different from the one considered in [1], for which another potential function has been established: we have a replicator dynamic where fitnesses are given by true costs, whereas, for some reasons, [1] considers marginal costs. Hence, constructions from [1] cannot be used directly.

PROPOSITION 2. *For allocation games considered in this paper, the following function F satisfies the hypotheses of previous theorem:*

$$F(Q) = \sum_{\ell=1}^{m} \left[\beta_\ell \sum_{j=1}^{n} q_{j,\ell} w_j + \frac{\alpha_\ell}{2} (\sum_{j=1}^{n} q_{j,\ell} w_j)^2 + \alpha_\ell \sum_{j=1}^{n} q_{j,\ell} w_j^2 (1 - \frac{q_{j,\ell}}{2}) \right]$$

PROOF. We have

$$\begin{aligned}
\frac{\partial F}{\partial q_{i,\ell'}}(Q) &= \beta_{\ell'} w_i + \alpha_{\ell'} (\textstyle\sum_{j=1}^{n} q_{j,\ell'} w_j) w_i + \alpha_{\ell'} w_i^2 - \alpha_{\ell'} q_{i,\ell'} w_i^2 \\
&= w_i[\beta_{\ell'} + \alpha_{\ell'} (\textstyle\sum_{j=1}^{n} q_{j,\ell'} w_j) + \alpha_{\ell'} w_i - \alpha_{\ell'} q_{i,\ell'} w_i] \\
&= w_i[\alpha_{\ell'} w_i + \beta_{\ell'} + \alpha_{\ell'} (\textstyle\sum_{j \neq i} q_{j,\ell'} w_j)] \\
&= w_i \times h_{i,\ell}(Q),
\end{aligned}$$

as requested. □

Notice that the hypothesis of affine cost functions is crucial here.

PROPOSITION 3. *Suppose for example that cost functions were quadratic :*

$$C_\ell(\lambda) = \alpha_\ell \lambda^2 + \beta_\ell \lambda_\ell + \gamma_\ell,$$

with $\alpha_\ell, \beta_\ell, \gamma_\ell \geq 0$, $\alpha_\ell \neq 0$.

A function F of class $\mathcal{C}^2$ that satisfies (10) *for all i, ℓ, Q, and general choice of weights $(w_i)_i$ can not exist.*

PROOF. By Schwartz theorem, we must have

$$\frac{\partial}{\partial q_{i',\ell'}}(\frac{\partial F}{\partial q_{i,\ell}}) = \frac{\partial}{\partial q_{i,\ell}}(\frac{\partial F}{\partial q_{i',\ell'}}),$$

and hence

$$W_i \frac{\partial h_{i,\ell}}{\partial q_{i',\ell'}} = W_{i'} \frac{\partial h_{i',\ell'}}{\partial q_{i,\ell}},$$

for all i, i', ℓ, ℓ', for some constants $W_i, W_{i'}$. This is easy to see that this doesn't hold for general choice of Q and weights $(w_i)_i$ in that case. □

Coming back to our model (affine costs), we obtain.

THEOREM 3. *For allocation games considered in this paper, for any initial condition in $K - K^*$, the considered learning algorithm converges to a (mixed) Nash equilibrium.*

We now discuss the time of convergence.

From the dynamics(9), this is clear that if $q_{i,\ell}$ is 0 (respectively 1) at time 0, it will stay null (1). Hence, if player i starts with some strategy $q_i(0)$ at time 0, we know that at any time t, its strategy $q_i(t)$ will have a support included in the support of $q_i(0)$:

$$q_{i,\ell}(0) = 0 \Rightarrow q_{i,\ell}(t) = 0, \forall t \geq 0,$$

or equivalently,

$$\forall t \geq 0, q_{i,\ell}(t) > 0 \Rightarrow q_{i,\ell}(0) > 0.$$

This motivates the following definitions:

DEFINITION 4. *Given a strategy $q \in \mathcal{S}$, we write $\mathcal{S}(q) \subset \mathcal{S}$ for the strategies q' whose support is included in the support of q: i.e. such that*

$$q'_{i,\ell} > 0 \Rightarrow q_{i,\ell} > 0.$$

DEFINITION 5. *The n-tuple of mixed strategies $(\tilde{q}_1, \cdots, \tilde{q}_n)$ is said to be a relative ϵ-Nash equilibrium, if for each i, $1 \leq i \leq n$, we have: $\overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, \tilde{q}_i, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) \leq \overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, q, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) + \epsilon$, forall $q \in \mathcal{S}(\tilde{q}_i)$.*

In other words, a n-tuple of mixed strategies $(\tilde{q}_1, \cdots, \tilde{q}_n)$ is not a relative ϵ-Nash equilibrium, if there exists some i, $1 \leq i \leq n$, and some $q \in \mathcal{S}(\tilde{q}_i)$ with

$$\begin{aligned} &\overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, q, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) \\ &< \overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, \tilde{q}_i, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) - \epsilon. \end{aligned}$$

Clearly, if there exists such an i, and $q \in \mathcal{S}(\tilde{q}_i)$, there must exist some pure strategies ℓ, and ℓ' in $\mathcal{S}(\tilde{q}_i)$ with

$$\begin{aligned} &\overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, e_\ell, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) \\ &< \overline{d}_i(\tilde{q}_1, \cdots, \tilde{q}_{i-1}, e_{\ell'}, \tilde{q}_{i+1}, \cdots, \tilde{q}_n) - \epsilon, \end{aligned}$$

that is to say, such that

$$h_{i,\ell} < h_{i,\ell'} - \epsilon,$$

and hence

$$[h_{i,\ell}(Q) - h_{i,\ell'}(Q)]^2 \geq \epsilon^2.$$

Since we know from (11) that

$$\frac{dF(Q(t))}{dt} = -\sum_i w_i \sum_\ell \sum_{\ell' > \ell} q_{i,\ell} q_{i,\ell'} [h_{i,\ell}(Q) - h_{i,\ell'}(Q)]^2,$$

which means that concerning a point $Q(t)$ that is not a relative ϵ-Nash equilibrium, we know that

$$\frac{dF(Q(t))}{dt} \quad \leq \quad - w_i q_{i,\ell} q_{i,\ell'} \epsilon^2$$

for some i, ℓ, ℓ'.

THEOREM 4. *Let $Q(0)$ be any initial condition in $K - K^*$. Let*

$$T \leq \frac{F(Q(0))}{(\min_i w_i)\epsilon^2 \delta^*}.$$

Assume

$$\delta_i(Q) = \min_{\ell, \ell' : q_{i,\ell} q_{i,\ell'} \neq 0} q_{i,\ell} q_{i,\ell'}$$

stay bounded by below for all i by some positive δ^*, on time $t \in [0, T]$.

Then a relative ϵ-Nash equilibrium is reached by the dynamic (9) at a time less than T.

PROOF. We have

$$\frac{dF(Q(t))}{dt} \leq -(\min_i w_i)\epsilon^2 \delta^*,$$

whenever $Q(t)$ is not a relative ϵ-Nash equilibrium. This implies that

$$F(Q(t)) \leq F(Q(0)) - (\min_i w_i)\epsilon^2 \delta^* t,$$

if $Q(t)$ is not a relative ϵ-Nash equilibrium from time 0 to t.

Since F is non-negative, we get that a relative ϵ-Nash equilibrium must be reached at a time T with

$$T \leq \frac{F(Q(0))}{(\min_i w_i)\epsilon^2 \delta^*},$$

which yields the theorem. □

5. CONCLUSION

In this paper we considered a classical settings [21] for selfish load balancing, but extended with some dynamical aspects.

We considered a learning algorithm proposed by [26]. We proved that this learning algorithm learns mixed Nash equilibria of the game, extending several results of [26].

To do so, we proved that the learning algorithm is asymptotically equivalent to an ordinary differential equation, which turns out to be a replicator equation. Using a folk theorem from evolutionary game theory, one knows that, if the dynamics converges, it will be towards some Nash equilibria. We proved using a Liapunov function argument that the dynamics converges in our considered settings. We also showed that it is actually possible to provide a general estimation that one could call an epsilon convergence time for any suitable starting condition.

6. REFERENCES

[1] E. Altman, Y. Hayel, and H. Kameda. Evolutionary dynamics and potential games in non-cooperative routing. In *Wireless Networks: Communication, Cooperation and Competition (WNC3 2007)*, 2007.

[2] Petra Berenbrink, Tom Friedetzky, Leslie Ann Goldberg, Paul Goldberg, Zengjian Hu, and Russell Martin. Distributed selfish load balancing. In *SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 354–363, New York, NY, USA, 2006. ACM.

[3] Petra Berenbrink and Oliver Schulte. Evolutionary equilibrium in bayesian routing games: Specialization and niche formation. In *ESA*, pages 29–40, 2007.

[4] I. Caragiannis, M. Flammini, C. Kaklamanis, P. Kanellopoulos, and L. Moscardelli. Tight bounds for selfish and greedy load balancing. *Proceedings of the 33rd International Colloquium on Automata, Languages, and Programming (ICALP'06)*, pages 311–322.

[5] Pierre Coucheney, Corinne Touati, and Bruno Gaujal. Fair and efficient user-network association algorithm for multi-technology wireless networks. In *Proc. of the 28th conference on Computer Communications miniconference (INFOCOM)*, 2009.

[6] Pierre Coucheney, Corinne Touati, and Bruno Gaujal. Selection of efficient pure strategies in allocation games. In *Proc. of the International Conference on Game Theory for Networks*, 2009.

[7] A. Czumaj, P. Krysta, and B. Vöcking. Selfish traffic allocation for server farms. In *ACM Symposium on Theory of Computing (STOC)*. ACM Press New York, NY, USA, 2002.

[8] Artur Czumaj. *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*, chapter Selfish Routing on the Internet. CRC Press, 2004.

[9] E. Even-Dar, A. Kesselman, and Y. Mansour. Convergence time to Nash equilibria. *30th International Conference on Automata, Languages and Programming (ICALP)*, pages 502–513, 2003.

[10] Eyal Even-Dar, Alexander Kesselman, and Yishay Mansour. Convergence time to Nash equilibrium in load balancing. *ACM Transactions on Algorithms*, 3(3), 2007.

[11] Eyal Even-Dar and Yishay Mansour. Fast convergence of selfish rerouting. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 772–781, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[12] Feldmann, Gairing, Lucking, Monien, and Rode. Selfish routing in non-cooperative networks: A survey. In *MFCS: Symposium on Mathematical Foundations of Computer Science*, 2003.

[13] S. Fischer, H. Räcke, and B. Vöcking. Fast convergence to Wardrop equilibria by adaptive sampling methods. *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 653–662, 2006.

[14] S. Fischer and B. Vocking. On the Evolution of Selfish Routing. *Algorithms–ESA 2004: 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, 2004.

[15] D. Fotakis, E.K. Spyros Kontogiannis, M. Mavronicolas, and P. Spirakis. The Structure and Complexity of Nash Equilibria for a Selfish Routing Game. *Automata, Languages and Programming: 29th International Colloquium, ICALP 2002, Málaga, Spain, July 8-13, 2002: Proceedings*, 2002.

[16] Paul W. Goldberg. Bounds for the convergence rate of randomized local search in a multiplayer load-balancing game. In *PODC '04: Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*, pages 131–140, New York, NY, USA, 2004. ACM.

[17] C. Harris. On the Rate of Convergence of Continuous-Time Fictitious Play. *Games and Economic Behavior*, 22(2):238–259, 1998.

[18] Morris W. Hirsch, Stephen Smale, and Robert Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier Academic Press, 2003.

[19] J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 4:479–519, 2003.

[20] J. Hofbauer and S. Sorin. Best response dynamics for continuous zero-sum games. *Discrete and Continuous Dynamical Systems-Series B*, 6(1), 2006.

[21] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Symposium on Theoretical Computer Science (STACS'99)*, pages 404–413, Trier, Germany, 4–6March 1999.

[22] Elias Koutsoupias. Selfish task allocation. *EATCS Bulletin*, 81, 2003.

[23] L. Libman and A. Orda. Atomic Resource Sharing in Noncooperative Networks. *Telecommunication Systems*, 17(4):385–409, 2001.

[24] John F. Nash. Equilibrium points in n-person games. *Proc. of the National Academy of Sciences*, 36:48–49, 1950.

[25] A. Orda, R. Rom, and N. Shimkin. Competitive routing in multiuser communication networks. *IEEE/ACM Transactions on Networking (TON)*, 1(5):510–521, 1993.

[26] M.A.L. Thathachar P.S. Sastry, V.V. Phansalkar. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. *IEEE transactions on system, man, and cybernetics*, 24(5), 1994.

[27] D.W. Stroock and SRS Varadhan. *Multidimensional Diffusion Processes*. Springer, 1979.

[28] Berthold Vöcking. *Algorithmic Game Theory*, chapter Selfish Load Balancing. Cambridge University Press, 2007.

[29] Jörgen W. Weibull. *Evolutionary Game Theory*. The MIT Press, 1995.