Stochastic Population Games with individual independent states and coupled constraints*

Hamidou Tembine
LIA/CERI
University of Avignon
339, chemin des Meinajaries
Agroparc BP 1228
F-84911 AVIGNON Cedex

Eitan Altman
INRIA, MAESTRO Group
2004 Route des Lucioles
F-06902,Sophia-Antipolis
Cedex

Rachid ElAzouzi
LIA/CERI,
University of Avignon
339, chemin des Meinajaries
Agroparc BP 1228
F-84911 AVIGNON Cedex

Yezekael Hayel
LIA/CERI
University of Avignon
339, chemin des Meinajaries
Agroparc BP 1228
F-84911 AVIGNON Cedex

ABSTRACT

This paper studies non-cooperative population games with several individual states and independent Markov process. Each member of each class of the population has (i) its own state (ii) its actions in each state, (iii) an instantaneous reward which depends on its state and the population's profile, (iv) a time-average (coupled) constraints. We apply this model to battery-dependent power control in wireless networks with several types of renewable energies. We show that the game has an equilibrium in stationary strategies under ergodic assumptions and we present a class of evolutionary game dynamics which converge to stationary equilibria.

Keywords

Population games, Markov decision process, power control, access control.

1. INTRODUCTION

We study in this paper a multiclass stochastic population game model with individual states. We consider several large subpopulations (classes or groups) of players. Each player from each subpopulation is associated with a controlled Markov chain, whose transition probabilities depend only on the action of that player (individual state). Each player interacts with a large number (possibly infinite) of others players. It does not know the states of, and the actions taken by other players. There are payoff (called also fitness, reward, utility) functions (one per subpopulation) that depend on the individual state and actions of all players.

1.1 Contributions

We characterize and establish the existence of stationary equilibrium in the stochastic population game model with individual independent states with time-average constraints under ergodic properties. A probabilistic representation of $\epsilon-$equilibrium for time-average Cesaro payoff is obtained in the general (non-)communicating stochastic population game under feasibility conditions. We apply this model to dynamic renewable energy-state dependent power control and access control in wireless networks.

In the battery model, players (which correspond to users, mobiles etc) have their own battery. The state of each battery is described as a Markov decision process (MDP) [1]. Several types (classes) of batteries and several modes of rechargeable batteries are considered (renewable energy: solar, wind etc). In the battery-dependent power control population game, a non-decreasing function of the signal to interference plus noise ratio (SINR) is used as the instantaneous reward of the user. An equilibrium is explicitly determined in that case and the equilibrium payoff is expressed as function of the stationary distribution associated to that equilibrium. This model offers us a new class of repeated games: *constrained repeated games with individual states and unknown horizon* in a large population. We show that this class of games has a constrained $0-$equilibrium (theorem 6.1.2) under ergodic and Slater conditions respectively on each individual Markov chain and constraints .

1.2 Related Works

Shapley [15] introduced the model of *atomic* stochastic games and proved that every two-player zero-sum discounted stochastic game has a discounted value. Moreover, there are stationary equilibrium profiles. We refer the reader to [21] and the references therein for details and recent results on stochastic two-player games. Fink [8], Takahashi [18] generalized this result for $n-$player *atomic* stochastic games.

*This work was partially supported by the INRIA ARC Program POPEYE.

GameComm 2008 October 20, 2008, Athens, GREECE.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

A single decision-maker stochastic game is called Markov Decision Process (MDP).

The n-player stochastic game model with individual state has been first introduced by Altman et al. 2005 [6], see also [5]. A product state formulation of their model called *product game* has been proposed by Flesh et al.[9] in 2007. They showed that an equilibrium exists under aperiodicity conditions on the transition law between the states. Note that in stochastic population games with periodic Markov decision process (common state for all populations), the existence result does not holds in this case as well as in atomic stochastic games. For two players (atomic) stochastic games, Vieille [20] showed existence of $\epsilon-$equilibria for all $\epsilon > 0$. See for example the modified Big Match and Paris Match games in Sorin[16] where $0-$equibria do not necessarily exists.

1.3 Structure of the paper

The remainder of this paper is organized as follows. In next section we present two simple examples that motivate us to consider with stochastic population games. In Section 3 we present the model of stochastic multi-population game with individual states. We give a class of dynamics which converge to stationary equilibria (when it exists). In Section 4 and 2.2, we apply our model to power control and energy management in wireless networks. In Section 5, we extend our model to the case where each player has *(coupled) constraints* on its strategies.

2. MOTIVATIONS, ILLUSTRATING EXAMPLES

2.1 Battery-State Dependent Power Control as a Stochastic Population Game

Consider a large number of mobiles terminals controlling their transmission power and a distributed base stations. Each mobile has an amount of energy E when its battery is new (typically it is the case if the battery is new or if the battery is completely recharged). Each mobile implements a power control policy where the transmission power is allowed to depend on the energy level (state) of its battery. The available action (reachable base stations and powers) depends on the state of the battery. Given the remaining energy of its battery, the mobile have to choose the optimal power level. One of the important element for each mobile is its instantaneous throughput which can be characterized as a function of the signal to interference plus noise ratio (SINR) at the base station where he transmits. The battery is replaced only when it is completely empty. The cost of new battery cost is C. The new battery has the same energy of E. The mobile have to control both the power consumption as well as the time at which the batteries are changed. At each slot, each mobile is faced to a random number [19] of interacting players which transmit at the same base station. Each battery life-time game corresponds to a stochastic population game with finite horizon (absorbing state of battery when the energy is very small). The aim here is to find jointly the power levels and the base stations such that all users achieve as high payoff as possible, minimum guarantee (e.g. QoS requirement thresholds) but also to control the battery-state.

When batteries are recharged dynamically with different types of alternative energy such as renewable energies (solar, wind etc). The battery transition state becomes irreducible Markov decision process under each policy depending on an exogenous parameter which characterizes the good weather (good weather will correspond to the sun for the solar-power systems and to the wind for wind-powered systems). In this case, the interaction becomes a stochastic population game with infinite horizon and we shall consider time-average reward (discounted or not).

2.2 Energy Management in Distributed Hybrid ALOHA Networks

Consider a distributed Aloha network with large number of mobile terminals. Each mobile can choose both the channels and powers (this is in contrast to standard Aloha model in which users are associated to the closest receivers). Each terminal is faced to a random number of interacting players which transmit at the channel. A terminal attempts transmissions during a finite horizon of times depending on the state of its battery energy. At each slot, each terminal have to take a decision on the transmission power based on the battery state. At each state of the battery, there are a finite power levels. At the lowest state of battery no power is available and the mobile have to replaced the battery by a new or to recharge its battery. A transmission is successful if no other user transmit during the slot or the mobile transmits with a power which is bigger than the power of all others transmitting mobiles at the same receiver. The pairwise interactions case of this problem has been studied by Altman and Hayel in [2] as a stochastic evolutionary game. They have considered three states: *Full, Almost Empty* and *Empty,* and simultaneous interactions with more two users are neglected[1]. Their model can be extend to more than two opponent interactions and also to finitely many states as shown in Section 3. We can also extend to the case where each terminal is faced to a random number [19, 3] of interacting terminals which transmit at the same range and each terminal have to control an arbitrary transition state of its energy.

3. A MULTICLASS MARKOV POPULATION GAME MODEL

Consider the following model of population game denoted by

$$\Gamma = (P, (Y^p)_{p\in P}, (A^p(y))_{p\in P, y\in Y^p}, (Q^p)_{p\in P}, (r^p)_{p\in P}),$$

where

- The population is composed as several subpopulations. Each subpopulation contains a large number of players. P denotes the set of subpopulations (we assume that P is finite).
- Each player of each subpopulation p has its own state Y^p (finite) and Markov transition structures Q^p between the states.
- For every player i from the subpopulation $p \in P$ and every state $y \in Y^p$ of i, $A^p(y)$ is the set of actions available. The action space of the subpopulation p is

[1] Note that this assumption does not holds in dense networks.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

given by $\prod_{y\in Y^p} A^p(y)$. The set of all actions at all states is given by All^p where

$$All^p = \{(y,a),\ y\in Y^p,\ a\in A^p(y)\}.$$

- We denote by $\Delta(Y^p)$ the $(|Y^p|-1)$-dimensional simplex of $\mathbb{R}^{|Y^p|}$ and by $q^p : All^p \rightarrow \Delta(Y^p)$ a transition rule between the states. The transition probability distribution between states is defined by

  $$Q^p_{y,a,y'} := q^p(y'|y,a) = q^p(y'|y_1,a_1,\ldots,y_{t-1},a_{t-1},y,a)$$

  for each $y',y\in Y^p, a\in A^p(y)$.

- For every subpopulation $p\in P$,

  $$r^p : \prod_{p'}\prod_{y\in Y^{p'}} \mathcal{X}^{p'}(y) \rightarrow \mathbb{R}^{\sum_{y\in Y^{p'}}|A^{p'}(y)|}$$

  is the vector of all instantaneous payoff functions of a player from the class p',

  $$\mathcal{X}^{p'}(y) = \Big\{ (x^{p'}(y,b))_{b\in A^{p'}(y)} \mid x^{p'}(y,b)\geq 0,$$

  $$\sum_{y\in Y^{p'}}\sum_{b\in A^{p'}(y)} x^{p'}(y,b) = m^{p'} \Big\}$$

  where $m^{p'}$ is the mass associate to the subpopulation p'. Given a state y and strategy profile x^p, x^{-p}, the payoff obtained by playing the action $a\in A^p(y)$ is $r^p_{y,a}(x)$

- The game is played many times.

3.1 Histories and Strategies

Histories A history h_t at time t is a collection of states and actions $(y_1,a_1,x_1,\ldots,y_{t-1},a_{t-1},x_{t-1},y_t)$. We denote by

$$H^p_t = (All^p\times\mathcal{X})^{t-1}\times Y^p$$

the set of histories of a member of the subpopulation p at time t. At $t=1$, $H^p_1 = Y^p$. Let H^p_∞ be the set of all infinite histories of the subpopulation p endowed with the product $\sigma-$field and $H_\infty = \prod_{p\in P} H^p_\infty$.

Strategies

- **Pure strategy** A pure strategy of a player from subpopulation p at time t is a map $\sigma^p_t : H^p_t \longrightarrow A^p(y_t)$. The collection $\sigma^p = (\sigma^p_t)_{t\geq 1}$ of pure strategy at each time constitutes a pure strategy of the subpopulation p. We denote by Σ^p the set of all pure strategies of subpopulation p, by $\Sigma = \prod_p \Sigma^p$ the set of all pure strategy profiles. Note that the number of pure strategies is infinite.

- **Stationary strategy:** σ is stationary strategy if for each population p and every time t and histories,

  $$h_t = (y_1,a_1,x_1,\ldots,y_{t-1},a_{t-1},x_{t-1},y_t),$$

  $$h'_t = (y'_1,a'_1,x'_1,\ldots,y'_{t-1},a'_{t-1},x_{t-1},y'_t)$$

  such that if $y_t = y'_t$ one has $\sigma_t(h_t) = \sigma_t(h'_t)$ i.e a stationary strategy is a history and time independent strategy which depends on the state only.

LEMMA 3.1.1. *The number of pure stationary strategies is $\prod_{p\in P}\prod_{y\in Y^p}|A^p(y)|$.*

- **Behavioral strategy** A behavioral strategy at time t is a function that assigns each finite history to a mixed action profile of the current state: $\sigma^p_t : H^p_t \longrightarrow \prod_p \Delta(A^p(y_t)),\ p\in P$.

- **Mixed strategy** A mixed strategy profile is a collection of probability distributions on Σ. Using Tychonoff's theorem, the set of all these notions of strategies is compact in the product set histories spaces in the sense of the weak-topology. A general mixed strategy is a probability distribution on the behaviorial strategies set.

For any strategy profile $\sigma = (\sigma^p)_{p\in P}$ and every initial state distribution profile $\mu = (\mu^p)_{p\in P}$, a probability measure $\mathbb{P}_{\sigma,\mu}$ is induced by σ and μ. The stochastic process $(y_t,a_t,x_t)_{t\geq 1}$ is defined on H_∞ in a canonical way, where the random variables y_t,a_t,x_t describe the individual state, the action in this state and the population profile.

3.2 Fitness

We examine the limit average Cesaro-type payoff

$$F^p_\mu(\sigma^p,\sigma^{-p}) = \mathbb{E}_{\sigma,\mu}\left[\liminf_{T\longrightarrow+\infty}\frac{1}{T}\left(\sum_{t=1}^{T} r^p_{y_t,a_t}(x_t)\right)\right]$$

where $\mathbb{E}_{\sigma,\mu}$ denotes the expectation over the probability measure $\mathbb{P}_{\sigma,\mu}$ induced by σ,μ on the set of histories endowed with the product $\sigma-$algebra.

Given a strategy σ and a initial state y, we define the expected time-average payoff. We denote by Π^p the stationary limit average matrix:

$$\Pi^p(\sigma^p) = \lim_{t\longrightarrow+\infty}\frac{1}{t}\sum_{j=1}^{t}(Q^p)^j(\sigma^p).$$

The matrix Π^p is well-defined, commutes with Q^p and satisfies the projection equation: $\Pi^p\times\Pi^p = \Pi^p$.

If F^p is the vector $(F^p_y(x))_{y\in Y^p}$, we have that $F^p(x) = \Pi^p r^p(x)$ for all stationary strategy profile x. Then, $F^p = \Pi^p F^p$. Note that the function $x\longmapsto F^p(x)$ is not necessarily continuous because the limit matrix $\Pi^p(x)$ can be discontinuous on x.

DEFINITION 3.3. *A strategy σ is an $\epsilon-$equilibrium if for all p,*

$$F^p(\sigma)+\epsilon \geq F^p(\sigma'^p,\sigma^p),\ \forall\ \sigma'^p\in\Delta(\Sigma^p).$$

A $0-$equilibrium is called equilibrium.

Remarks

- When each member of each subpopulation has a single state, we obtain a population game model which each local interaction is repeated game.

- If there exists a subpopulation p^* such that $|Y^p| = 1$ for all $p\neq p^*$. We obtain a stochastic population game with single class of controllers which is the subpopulation p. We can adapts the model of Vieille, Rosenberg and Solan[22] on two player zero-sum stochastic game with single controller and incomplete information to stochastic population game with incomplete information.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

3.4 Markov Decision Process Decomposition

For each population p, the state y communicates with state y' if it is true that both y is accessible from y' (there exists an integer k such that $\forall a$, $\mathbb{P}(X_k^p = y' \mid X_0^p = y, a) > 0$ and that y' is accessible from y. A set of states $\underline{Y}^p \subseteq Y^p$ is a communicating class if every pair of states in $\underline{Y}^p$ communicates with each other, and no state in $\underline{Y}^p$ communicates with any state not in $\underline{Y}^p$. It is known that communication in this sense is an equivalence relation (reflexivity,symmetry,transitivity).

Given a strategy σ^p, the associate Markov chain is decomposed in communicating class $J_1^p, \ldots, J_m^p$ which constitute a partition of Y^p. We associate to the MDP $-(p,i)$ the restricted state-actions transition to $J_i^p \subset Y^p$. A Markov chain is said to be irreducible if its state space is a communicating class; this means that, in an irreducible Markov chain, it is possible to get to any state from any state. A Markov decision process (a stochastic game with single decision-maker) is irreducible if for any strategy, the induced transition law is irreducible.

Proposition 3.4.1. *Assume that for each subpopulation p and any stationary strategy σ^p, the state process is an irreducible Markov chain with one ergodic class then*

- *For any strategy σ the frequencies (called also occupation measures) $(f_{\sigma,\mu}^{p,t}(y,a))_{p\in P, t\geq 1}$ where*

$$f_{\mu,\sigma}^{p,t}(y,a) = \frac{1}{t}\sum_{k=1}^{t} P_\sigma(X_k = y, a_k = a | y_1 = \mu)$$

are tight.

- *Denote by BR^p the best response correspondence for a player in class p. If $\sigma^p \in BR^p(\sigma^{-p})$ then $f_\mu^p(\sigma)$ is independent of the initial state distribution μ and the linear programming problem : find $z^p(\sigma) = (z^p(\sigma)(y,a))_{y,a}$ that maximizes*

$$\sum_{y,a} r^p(y,a,\sigma^{-p}) z^p(\sigma)(y,a)$$

subject to

$$\sum_{y,a} [\delta_{y'}(y) - Q^p_{yay'}] z^p(\sigma)(y,a) = 0$$

$$z^p(\sigma)(y,a) \geq 0,\ \forall y \in Y^p,\ a \in A^p(y)$$

$$\sum_{y,a} z^p(\sigma)(y,a) = m^p$$

where $\delta_{y'}$ is the Dirac distribution concentrated in y'.

A proof of the proposition is given together with the result 5.2.1.

Definition 3.5. *The stationary strategy x is an equilibrium if*

$$\forall p,\ \sum_{y\in Y^p}\sum_{a\in A^p(y)} (-z^p(y,a) + x^p(y,a)) F^p_{y,a}(x) \geq 0,$$

for all $z^p \in \mathcal{X}^p$ satisfying

$$\sum_{y\in Y^p, a\in A^p(y)} (\delta_{y'}(y) - Q^p_{yay'}) z^p(y,a) = 0,\ \forall y' \in Y^p.$$

Result 3.5.1. *The stochastic population game with individual independent states has an equilibrium in stationary strategies under ergodic properties of the each class of MDP and continuity of the payoff function $r = (r^p)_p$.*

Proof. For any subpopulation p, the existence of a vector $x = (x^p)_{p\in P}$ of the $(-1+\sum_y |A^p(y)|)-$simplex satisfying he variational inequality :$\forall z \in \mathcal{X}^p$, one has

$$\langle x^p - z^p, r^p(x)\rangle := \sum_{y\in Y^p}\sum_{a\in A^p(y)} (-z^p(y,a) + x^p(y,a)) r^p_{y,a}(x) \geq 0,$$

Multiplying the inequality $\langle x^p - z^p, r^p(x)\rangle \geq 0$ by $\varsigma > 0$, and adding $\langle x^p, z^p - x^p\rangle$ to both sides of the resulting inequality, one obtains

$$\langle z^p - x^p, x^p - [x^p + \varsigma r^p(x)]\rangle \geq 0.$$

Recall that the projection map on the simplex which is convex and closed set is characterized by: $w \in \mathbb{R}^{\sum_{y\in Y^p}|A^p(y)|}$, $z'^p = \Pi_{\mathcal{X}^p} w$ is equivalent to $\langle z'^p - w, z^p - z'^p\rangle \geq 0,\ \forall z^p \in \mathcal{X}^p$. Thus, $x^p = \Pi_{\mathcal{X}^p}(x^p + \varsigma r^p(x))$. According to Brouwer's or Schauder's fixed point theorem, given a map $\Psi :\ \mathcal{X}^p \longrightarrow \mathcal{X}^p$, with Ψ continuous, there is at least one $z^p \in \mathcal{X}^p$, such that $z^p = \Psi(z^p)$. Observe that since the projection $\Pi_{\mathcal{X}^p}$ and $(I + \varsigma r^p)$, are each continuous, $\Pi_{\mathcal{X}^p}(I + \varsigma r^p)$ is also continuous.

It follows from compactness of $\mathcal{X}^p$ and the continuity of $\Pi_{\mathcal{X}^p}(I + \varsigma r^p)$ that a such x^p exists. The result follows from Kakutani's fixed point theorem and Theorem 2.6 (ii) in [4] or the Theorem 1 in [10] by adapting to population game concept. $\square$

3.6 Evolutionary game dynamics in stationary strategies

Let $\nu_b^{p,y,a}(x)$ conditional switch rate from the pure strategy a to the strategy b in state y for a player of class p. The flow of the population is specified in terms of the functions $\nu_b^{p,y,a}(x)$ which determine the rates at which an player who is considering a change in strategies opts to switch to his various alternatives. The function $\nu_b^{p,y,a}(.)$ depends on the strategy of the population but also on the payoffs.

The *inflow* into the action a at state y is

$$\sum_{b\in A^p(y)} x^p(y,b)\nu_a^{p,y,b}$$

and *outflow* from the action a in y is

$$x^p(y,a)\sum_b \nu_b^{p,y,a}$$

where $x^p(y,b)$ represents the fraction of players of the subpopulation p in state y which use the pure action b. We assume that the revision protocols satisfy

$$\nu_b^{p,y,a} > 0 \Longrightarrow b, a \in A^p(y).$$

Let $V_F^{p,y,a}(x)$ be the difference between the inflow and outflow of the action a at state y,

$$\sum_{b\in A^p(y)} x^p(y,b)\nu_a^{p,y,b}(x) - x^p(y,a)\sum_{b\in A^p(y)} \nu_a^{p,y,b}(x).$$

The evolutionary game dynamics is given by

$$\dot{x}^p(y,b)(t) = \frac{d}{dt}x^p(y,b)(t) = V_F^{p,y,b}(x(t)), \qquad (1)$$

$$y \in Y^p,\ b \in A^p(y).$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

The revision protocol ν defined by

$$\nu_a^{p,y,b}(x) = \mu^p \left[\max\left\{0, F^p_{y,a}(x) - F^p_{y,b}(x)\right\}\right]^{\theta^p}, \ \mu^p, \theta^p > 0,$$

if x satisfies

$$\sum_{y \in Y^p, a \in A^p(y)} (\delta_{y'}(y) - Q^p_{yay'}) x^p(y,a) = 0$$

and 0 otherwise, induces an evolutionary game dynamic with the following properties:

- (i) every stationary equilibrium of the game is a stationary point of the dynamic.
- (ii) every stationary point of the dynamic is an equilibrium point of the game.

Note that the well-known replicator dynamics does not satisfies the second point (ii). The parameter μ^p can be interpreted as a probability to have a base station/channel around the range of the players (density distribution of the resources in the space). This parameter have positive effect on the rate/speed of convergence of the dynamics. θ^p is a positive number.

4. BATTERY-STATE DEPENDENT POWER CONTROL WITH DIFFERENT TYPES OF RENEWABLE ENERGY

Power control in wireless networks has became an important research area. Since the technology in the current state cannot provide batteries which have small weight and large energy capacity, the design of tools and algorithms for efficient power control is crucial.

Thanks to the renewable energy techniques, designing autonomous mobile terminal and consumer embedded electronics that exploit the energy coming from the environment is becoming a feasible option. However, the design of such devices requires the careful selection of the components, such as power consumption and the energy storage elements, according to the working environment and the features of the application.

Menache and Altman have studied in [13] a battery-energy dependent power control with finite number of mobiles as a dynamic non-cooperative game with power cost assumption. In this model we consider a stochastic population game approach with dynamic rechargeable battery based on renewable energy. Environmental energy is becoming a feasible alternative for many low-power systems, such as wireless sensor/mesh networks. Nevertheless, environmental energy is an exciting challenge. Because of the limited amount of energy over time, the power provided is unpredictable. Power storage elements, such as rechargeable batteries or supercapacitors, in order to have energy available for later use has been proposed. Alternative energy as solar, wind, or nuclear energy, that can replace or supplement traditional fossil-fuel sources, as oil, and natural gas is needed. We refer the reader to [14] for advantageous to use renewable energy in broadband wireless networks such as Wi-Fi, Wimax or mesh networks.

We consider several class of large number of mobiles terminals controlling their transmission power and a distributed base stations. The mobiles with the same type of renewable energy (wind, solar, hydro) are in the same class or subpopulation. Each mobile of the subpopulation p has an amount of energy E^p when its battery is at the full state. Each mobile implements a power control policy where the transmission power is allowed to depend on the energy level (state) of its battery. The available action (reachable base stations) depends on the state of the battery. Given the remaining energy of its battery, the mobile have to choose the optimal power level. One of the important element for each mobile is its instantaneous throughput which can be characterized as a function of the signal to interference plus noise ratio (SINR) at the base station where he/she transmits. The battery is recharged by different techniques of renewable energy (solar-power, wind-power etc). The mobile have to control both the power consumption as well as the level of its battery and its throughput. Each mobile is faced to a non-cooperative stochastic game with individual states with many others mobiles which transmit at the same base station or at the same range. The goal of a terminal is to find jointly the power levels and the base stations such that the terminal achieves as high payoff as possible, minimum guarantee (e.g. QoS requirement thresholds) but also to control the battery-state.

4.1 Battery-state transition

We consider the energy reserve of the battery type p, $(X^p_t)_{t\geq 1}$ and power level management as a Markov decision process. For each state $y \neq 0$, the action space is $A^p(y)$ with at least two elements, and $A^p(0)$ has at most one element (empty or singleton). Given a stationary policy σ and a strategy of all the populations the change in energy reserves of the battery type p is described by the (first order, time-homogeneous) Markov process (X^p_t) with the transition law q^p. $\forall y \neq 0, n^p, \forall a,$, the probability of transition $q^p(X^p_{t+1} = y'|X^p_t = y, a)$ is expressed as

$$\left\{\begin{array}{cc} 1 - R^p_{\gamma^p,y}(a) - Q^p_{\gamma^p,y}(a) & \text{if } y' = y-1 \\ R^p_{\gamma^p,y}(a) & \text{if } y' = y+1 \\ Q^p_{\gamma^p,y}(a) & \text{if } y' = y \\ 0 & \text{otherwise} \end{array}\right. ,$$

$$q^p(X^p_{t+1} = y'|X^p_t = n^p, a) = \left\{\begin{array}{cc} Q^p_{\gamma^p,n^p}(a) & \text{if } y' = n^p - 1 \\ 1 - Q^p_{\gamma^p,n^p}(a) & \text{if } y' = n^p \\ 0 & \text{otherwise} \end{array}\right. ,$$

and

$$q^p(X^p_{t+1} = y'|X^p_t = 0, 0) = \left\{\begin{array}{cc} \gamma^p & \text{if } y' = 1 \\ (1-\gamma^p) & \text{if } y' = 0 \\ 0 & \text{otherwise} \end{array}\right. ,$$

where $\forall\ y, a,\ \gamma^p \longmapsto R^p_{\gamma^p,y}(a) \in [0,1]$ is an increasing function with $R^p_{0,y}(a) = 0$, $0 \leq R^p_{\gamma^p,y}(a) + Q^p_{\gamma^p,y}(a) \leq 1$. The factor γ^p represents a function of the probability to have a "good weather"(for example, the sun for the solar-power battery, the wind for wind-power battery) and the probability for battery of type p to go from state 0 to state 1. If γ^p is zero, the state 0 is absorbing. For $\gamma^p \neq 0$ is the chain is communicating.

Note that each user controls the transition state of its battery: q^p is independent of the decision of the other mobiles.

4.2 Reward

We focus on utility function based on a simplified version of the signal to noise plus interference ratio (SINR). The battery-state have the property that more energy is available

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

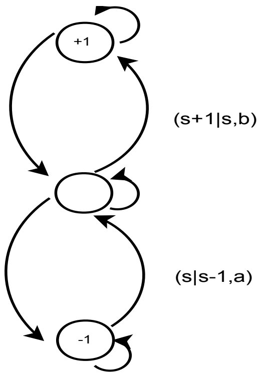

Figure 1: Generic battery state transition rule.

in high state. Hence, that set of powers in $y+1$ contains the set of power available in y. For example,

$$\varnothing \subset A^p(0) = \{pow_0^p\} \subset A^p(1) = \{pow_0^p, pow_1^p\} \subset$$

$$A^p(2) = \{pow_0^p, pow_1^p, pow_2^p\} \subset$$

$$\ldots A^p(n^p) = \{pow_0^p, pow_1^p, \ldots, pow_{n^p}^p\}$$

The signal to noise plus interference ratio of a mobile with the battery type p in state y at the position $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is

$$\mathrm{SINR}_y^p(a, x; \lambda, BS) = \frac{\frac{a g^{pp}}{(\epsilon^2+(\lambda_1-x_0)^2+(\lambda_2-y_0)^2+(\lambda_3-z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \kappa I_{own}(x^p) + \kappa I_{other}(x^{-p})},$$

$$a \in A^p(y), p \in P, y \in Y^p = \{0, 1, 2, \ldots, n^p\}$$

where

$$I_{own}(x^p) = \sum_{y,b} b g^{pp} x^p(y, b) h^{pp}$$

$$I_{other}(x^{-p}) = \sum_{k \neq p} \sum_{y,b} g^{kp} b x^k(y, b) h^{kp},$$

where

$$h^{kp} = \int_{\lambda \in D} \frac{d\mu^{k,BS}}{(\epsilon^2 + (\lambda_1 - x_0)^2 + (\lambda_2 - y_0)^2 + (\lambda_3 - z_0)^2)^{\frac{\alpha}{2}}}$$

$x^j(y, a)$ is the fraction of the sub-population j in state y with the power level a, N_0 is the power of the thermal background noise, $\mu^{p,BS}$ is the distribution of mobiles (in the 3-dimensional space) with the battery type p around the base station BS, $D \subseteq \mathbb{R}^3$ is the domain (geographical placement of base stations and mobiles) and α is the path-loss and κ is the inverse of the processing gain of the system, it weights the effect of interferences, depending on the orthogonality between codes used during simultaneous transmissions. The coefficient κ is equal to 1 in a narrow band system, and is smaller than 1 in a broadband system that uses CDMA. The instantaneous expected reward $r_{y,a}^p(x)$ of an user in state y is expressed as

$$\int_{\lambda \in D} f\left(\mathrm{SINR}_y^p(a, x, \lambda, BS)\right) d\mu^p(\lambda)$$

where f is a non-decreasing function with $f(0) = 0$. ϵ is a positive parameter (to eliminate of continuity problem at zero) and the g^{ij} are positive gain parameters. The $3-$dimensional vector (x_0, y_0, z_0) describes the position of the base station BS in $\mathbb{R}^3$

Computing the interference term in presence of continuum of users

In order to compute explicitly the SINR term, we first need the following lemma:

LEMMA 4.2.1.

$$\nu \geq 1,\ b_\nu = \int_0^{+\infty} \frac{1}{(1+x^2)^\nu} dx = \begin{cases} \frac{\pi}{2} & \text{if } \nu = 1 \\ \frac{\sqrt{\pi}}{2} \frac{\Gamma(\nu - \frac{1}{2})}{\Gamma(\nu)} & \text{if } \nu > 1. \end{cases}$$

where Γ *is the Euler function* $\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1}\ dt$.

PROOF. For $\nu = n$ a positive integer, the polynomial $(1+z^2)^n$ has two zeros $z = \pm i$ each zero with the order n. Consider the circuit $C_{R'} = [-R', R'] \cup \{R' e^{i\theta},\ 0 \leq \theta \leq \pi\}$. Since the complex function $z \in \mathbb{C} \longrightarrow \frac{1}{(1+z^2)^n}$ has no zero on the circuit $C_{R'}$, Using residue's theorem of complex analysis, we obtain the following result: $\int_{x \geq 0} \frac{1}{(1+x^2)^n} dx = \pi i Res(\xi(z), i)$. The residue of ξ around $z = i$, a pole of order n, can be found by the formula:

$$\mathrm{Res}(\xi, i) = \frac{1}{(n-1)!} \lim_{z \longrightarrow i} (\frac{d}{dz})^{n-1} \left[(z-i)^n \xi(z)\right]$$

Thus, $b_n = \frac{\sqrt{\pi}}{2} \frac{\Gamma(n - \frac{1}{2})}{\Gamma(n)}$. We then use the extension of the Euler function Γ on the positive real axis. $\square$

From the lemma 4.2.1, we derive immediately that, $n \geq 2$, $\int_0^{+\infty} \frac{x^2}{(1+x^2)^n} dx = b_{n-1} - b_n$

PROPOSITION 4.2.2. $h^{jp} = \frac{4\pi^2}{\epsilon^{2\alpha-3}} (b_{\frac{\alpha}{2}-1} - b_{\frac{\alpha}{2}})$

PROOF. Using spherical coordinates from cartesian coordinates by the transformation

$$\begin{cases} \lambda_1 = r \sin\theta \cos\phi \\ \lambda_2 = r \sin\theta \sin\phi \\ \lambda_3 = r \cos\theta \end{cases}$$

and the volume element $r^2 dr \sin\theta d\theta d\phi$, one has,

$$h^{jp} = 4\pi^2 \int_0^{+\infty} \frac{r^2}{(\epsilon^2 + r^2)^{\frac{\alpha}{2}}} dr = \frac{4\pi^2}{\epsilon^{2\alpha-3}} \int_0^{+\infty} \frac{r^2}{(1+r^2)^{\frac{\alpha}{2}}}\ dr$$

i.e $h^{jp} = \frac{4\pi^2}{\epsilon^{2\alpha-3}} (b_{\frac{\alpha}{2}-1} - b_{\frac{\alpha}{2}})$ $\square$

PROPOSITION 4.2.3. *The highest payoff that a mobile with the battery type p can obtain against any strategies of others mobiles in the one-shot power control game is given by*

$$\bar{v}_y^p = \int_D f\left(\frac{\frac{u_y^p g^{pp}}{(\epsilon^2+(\lambda_1-x_0)^2+(\lambda_2-y_0)^2+(\lambda_3-z_0)^2)^{\frac{\alpha}{2}}}}{N_0 + \sum_j \sum_y h^{jp} u_y^j m_y^j g^{jp}} \right) d\mu^p$$

where u_y^p *is the maximum power level available in the battery-type* p *in state* y.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

PROOF. Since the payoff decreases when the others players increases their power levels (in average), the minmax point is obtained when they uses their high powers. The maximum payoff that a mobile with the battery type p can obtain against any strategies of others mobiles is then given by

$$\bar{v}^p_y = \max_{a\in A^p(y)} \int_D f\left(\frac{\frac{ag^{pp}}{(\epsilon^2+(\lambda_1-x_0)^2+(\lambda_2-y_0)^2+(\lambda_3-z_0)^2)^{\frac{\alpha}{2}}}}{N_0+\sum_j\sum_y h^{jp}u^j_y m^j_y g^{jp}}\right) d\mu^p$$

$$= \int_D \max_{a\in A^p(y)} f\left(\frac{\frac{ag^{pp}}{(\epsilon^2+(\lambda_1-x_0)^2+(\lambda_2-y_0)^2+(\lambda_3-z_0)^2)^{\frac{\alpha}{2}}}}{N_0+\sum_j\sum_y h^{jp}u^j_y m^j_y g^{jp}}\right) d\mu^p$$

$$= \int_D f\left(\max_{a\in A^p(y)} \frac{\frac{ag^{pp}}{(\epsilon^2+(\lambda_1-x_0)^2+(\lambda_2-y_0)^2+(\lambda_3-z_0)^2)^{\frac{\alpha}{2}}}}{N_0+\sum_j\sum_y h^{jp}u^j_y m^j_y g^{jp}}\right) d\mu^p.$$

This completes the proof. □

RESULT 4.2.4. *Each mobile with the battery type p can guarantee the payoff*

$$\sum_{y\neq 0} \Pi^p_y \bar{v}^p_y$$

for all $\gamma^p > 0$, where $\Pi^p_y = \lim_{t\longrightarrow\infty} \mathbb{P}(X^p_t = y)$ is the probability to be in state y under the maximum power strategy.

PROOF. Π^p_y is the frequency of visit of the battery state in y. From Proposition 4.2.3, each mobile with the battery type p can obtain at least $\bar{v}^p_y$ against any strategies of others mobiles. Each mobile of subpopulation can then obtain at least $\sum_y \Pi^p_y \bar{v}^p_y$ which is an equilibrium payoff. This completes the proof. □

5. CONSTRAINED STOCHASTIC POPULATION GAMES

In addition to the model described in Section 3, we assume that players have (possibly coupled) average constraints on their actions in any state. The payoff of the subpopulation p is

$$E_{\sigma,\mu}\left(\liminf_{t\longrightarrow+\infty} \frac{1}{t}\sum_{k=1}^{t} r^p_{y_k,a_k}(x_k)\right)$$

with $\sigma^p \in \Delta(\Sigma^p)$ subject to

- Orthogonal constraints:

  $$p\in P,\ P_{\sigma^p}\left(\limsup_{t\longrightarrow+\infty} \frac{1}{t}\sum_{k=1}^{t} D^p(y_k,a_k) \le \beta^p\right) = 1,$$

  where $D^p : All^p \longrightarrow \mathbb{R}$ is an individual cost function (independent of the strategies of the others players), $\beta^p \in \mathbb{R}$ is a given cost threshold.

- Coupled constraints:

  $$p\in P,\ P_{\sigma,\mu}\left(\limsup_{t\longrightarrow+\infty} \frac{1}{t}\sum_{k=1}^{t} C^p_{y_k,a_k}(x_k) \le \alpha^p\right) = 1$$

  where $C^p : All^p \times \mathcal{X} \longrightarrow \mathbb{R}$ is a cost function which depends on the individual state-action but also on the population profile i.e the strategies of the others players (in the same class or not).

A strategy σ is a constrained equilibrium if for all p,

$$F^p(\sigma) \ge F^p(\sigma'^p, \sigma^p),\ \forall\ \sigma'^p \in \Delta(\Sigma^p)$$

and $\sigma^p \in \Lambda(\sigma^{-p})$ where $\Lambda(\sigma^{-p})$ is the set feasible strategies (that satisfy the orthogonal and coupled constraints) given the strategies of the others populations σ^{-p}.

5.1 Communicating case

Throughout this subsection fix an $p \in P$. Consider the evolution of the state and action processes for a player in the subpopulation p. We say that a Markov chain is communicating it is a Markov chain with single class. We assume this property in this subsection. We define the following constrained programming problem:

$$\text{CPP}_p:\ v^p(\sigma^{-p}) = \max \sum_{y,a} r^p(y,a,\sigma^{-p}) z^p(y,a)$$

subject to

$$\sum_{y,a} [\delta_{y'}(y) - Q^p_{yay'}] z^p(y,a) = 0, \quad (2)$$

$$z^p(y,a) \ge 0,\ \forall y,a \text{ and } \sum_a z^p(y,a) = m^p_y,\ \forall y, \quad (3)$$

$$\sum_{y,a} z^p(y,a) = \sum_y m^p_y = m^p, \quad (4)$$

$$\sum_{y,a} D^p(y,a) z^p(y,a) \le \beta^p,\ \sum_{y,a} C^p(y,a,\sigma^{-p}) z^p(y,a) \le \alpha^p \quad (5)$$

The relation between CPP_p and best response correspondence is given by the following: there exists a feasible strategy for the subpopulation p against σ^{-p} if and only if CPP_p has a solution. Moreover, a solution is a best reply to σ^{-p}. If CPP_p is feasible for a given strategy σ^{-p} then for each $\epsilon > 0$ there exists a stationary $p-$feasible strategy x^p such that $F_\mu(x^p, \sigma^{-p})$ is independent of the initial state and

$$F^p_\mu(x^p, \sigma^{-p}) + \epsilon > v^p(\sigma^{-p}).$$

5.2 Non-communicating chain: multichain

Each player of the class p can decompose its transition state into a partition $J^p_i,\ i = 1, 2, \ldots$ such that the restricted state-action game $(P, J^p_i, A^p_i(.), r^p)$ where $A^p_i(y)$ is the set of actions such that if the player start in the state $y \in J^p_i$, the state process will remain in the communicating class J^p_i.

RESULT 5.2.1. *Suppose now that σ is feasible and the probability that the process $X^p_t \in J^p_i$ almost any time (a.a.t) is positive (under σ and the initial distribution). The following constrained programming problem $CPP_{(p,i)}$ is feasible and*

$$\mathbb{P}_{\sigma,\mu}\left(\liminf_{t\longrightarrow+\infty} \frac{1}{t}\sum_{k=1}^{t} r^p_{y_k,a_k}(x_k) \le v^p_i \mid X^p_t \in J^p_i\ \ aat\right) = 1$$

where

$$CPP_{(p,i)}:\ v^p_i(\sigma^{-p}) = \max \sum_{y,a} r^p(y,a,\sigma^{-p}) z^p(y,a)$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

subject to

$$\sum_{y,a}[\delta_{y'}(y)-Q^p_{yay'}]z^p(y,a)=0,\ y'\in J^p_i \quad (6)$$

$$z^p(y,a)\geq 0,\ \forall y\in J^p_i, a\in A^p_i(y) \quad (7)$$

$$\sum_{a\in A^p_i} z^p(y,a)=m^p_y,\ \forall y\in J^p_i \quad (8)$$

$$\sum_{y\in J^p_i, a\in A^p_i(y)} z^p(y,a)=m^p \quad (9)$$

$$\sum_{y\in J^p_i, a\in A^p_i(y)} D^p(y,a)z^p(y,a)\leq \beta^p \quad (10)$$

PROOF. See appendix. □

5.3 Constrained stationary strategies

For each player of the subpopulation $p\in P$, its pure action set in state y is $A^p(y)$ and

$$\mathcal{A}^p_y:\mathcal{X}\to 2^{\Delta(A^p(y))}$$

is its restricted constrained correspondence which restricts the strategies in to the subset $\mathcal{A}^p_y(x^{-p})\subseteq\Delta(A^p(y))$ when the state of the population is y.

5.3.1 Example: coupled/orthogonal constraints

$$\mathcal{A}^p(x^{-p})=\left\{a,\ C^p_j(a,x^{-p})\leq\alpha^p_j,\right.$$

$$\left. D^p_l(a)\leq\beta^p_l\ j=1,\ldots,n^p, l=1,\ldots,n'^p\right\}.$$

RESULT 5.3.2. *Assume that*

- *The chain of each player from each subpopulation is ergodic for any stationary strategies.*
- *Slater conditions: For stationary profile x, each population p has some strategy $\sigma^p\in\{a,\ D^p_l(a)\leq\beta^p_l,\ l=1,\ldots,n'^p\}$ such that*

$$C^p_j(\sigma^p,x^{-p})<\alpha^p_j, j=1,\ldots,n^p$$

then a constrained equilibrium exists in stationary strategies.

PROOF. Under the two above assumptions, we can apply the Theorem 2.1 in [5] in which an optimal stationary strategy is obtained using constrained linear programming. □

6. EXTENSIONS

6.1 Constrained stochastic population games with unknown stopping time

In general, the lifetime of individual or system is not known. We shall integrate this considerations in our interaction model. In this section we develop a general formulation of a local interaction with unknown stopping time. Players does not known the length of the local interaction but have a common probability structure on the stochastic local game. At time t, they assign some probability $\mathbb{P}(T=t)$ to the event $\{T=t\}$ that the local interaction ends in time t.

$$t\geq 1,\ \mathbb{P}(T=t)\geq 0,\ \sum_{t\geq 1}\mathbb{P}(T=t)=1.$$

Fix an anonymous member of some subpopulation p, and a sequence of state-actions σ. A player from the class p will receive

$$F^p_\mu(\sigma)=\mathbb{E}_{\sigma,\mu}\left[\liminf_{t\to+\infty}\frac{\sum_{k=1}^t\mathbb{P}(T=k)(\sum_{j=1}^k\ r^p_{y_j,a_j}(x_j))}{\sum_{j=1}^t j\mathbb{P}(T=j)}\right]$$

under the constraints: $p\in P$,

$$P_\sigma\left(\limsup_{t\to+\infty}\frac{\sum_{k=1}^t\mathbb{P}(T=k)(\sum_{j=1}^k\ C^p_{y_j,a_j}(x_j))}{\sum_{j=1}^t j\mathbb{P}(T=j)}\leq\alpha^p\right)=1,$$

$$P_{\sigma^p}\left(\limsup_{t\to+\infty}\frac{\sum_{k=1}^t\mathbb{P}(T=k)(\sum_{j=1}^k\ D^p(y_j,a_j))}{\sum_{j=1}^t j\mathbb{P}(T=j)}\leq\beta^p\right)=1,$$

RESULT 6.1.1.

$$F^p_\mu(\sigma)=\mathbb{E}_{\sigma,\mu}\liminf_{t\to+\infty}\ F^{p,t}_\mu(\sigma)$$

where $F^{p,t}_\mu(\sigma):=\dfrac{1}{\sum_{j=1}^t j\mathbb{P}(T=j)}\displaystyle\sum_{j=1}^t\left(\sum_{k=j}^t\mathbb{P}(T=k)\right)r^p_{y_j,a_j}(x_j)$

PROOF. We apply Fubini's theorem on finite summation to change the order between k and j in the expression of $F^{p,t}_\mu(\sigma)$ where

$$F^{p,t}_\mu(\sigma)=\frac{\sum_{k=1}^t\mathbb{P}(T=k)(\sum_{j=1}^k\ r^p_{y_j,a_j}(x_j))}{\sum_{j=1}^t j\mathbb{P}(T=j)}.$$

□

Examples: This model generalizes the finite and infinite horizon payoff notions:

- If T is the Dirac measure concentrated on t_* i.e $\mathbb{P}(T=j)=0$ if $j\neq t_*$ and $\mathbb{P}(T=t_*)=1$, we obtained the arithmetic average payoff

$$\mathbb{E}_{\sigma,\mu}\frac{\sum_{j=1}^{t_*}r^p_{y_j,a_j}(x_j)}{t_*}$$

- If T is the geometric distribution $\mathbb{P}(T=t)=(1-\delta)\delta^{t-1}$, then we obtain the average discounted payoff:

$$(1-\delta)\ \mathbb{E}_\sigma\sum_{t=1}^{+\infty}\delta^{t-1}r^p_{y_t,a_t}(x_t)$$

- Note that when the expected horizon of local interaction is finite (for example when the lifetime of the system or of the user is finite - in expectation - but the end of the interaction is not known)[2], the average payoff can be rewritten as

$$F^p_\mu(\sigma)=\frac{\sum_{j=1}^{+\infty}\left(\sum_{k=j}^{+\infty}\mathbb{P}(T=k)\right)\mathbb{E}_{\sigma,\mu}r^p_{y_j,a_j}(x_j)}{\sum_{j=1}^{+\infty}j\mathbb{P}(T=j)} \quad (11)$$

$$=\frac{\sum_{t\geq 1}\mathbb{P}(T\geq t)\mathbb{E}_{\sigma,\mu}r^p_{y_t,a_t}(x_t)}{\mathbb{E}(T)} \quad (12)$$

The following theorem generalizes the Theorem 2.1 in [5] for constrained games and also the Theorem 2.6 (ii) in [4] and the Theorem 1 in [10] for unconstrained product games.

[2]Note that the expected horizon can be finite and $\mathbb{P}(T=t)>0$. It is the case for $\mathbb{P}(T=t)=\delta^{t-1}(1-\delta)$, $\delta\in(0,1).\mathbb{E}(T)=\frac{1}{1-\delta}$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359

RESULT 6.1.2. *Assume that each subpopulation has a single (aperiodic) ergodic class under each stationary strategy. Then the stochastic population game with individual independent states and unknown lifetime has an equilibrium in stationary strategies. Moreover, the constrained game has an equilibrium under Slater condition.*

For the proof we need tightness properties of the measure generated by the frequencies state-actions under the distribution of the horizon.

PROPOSITION 6.1.3. *Assume that for each subpopulation p and any stationary strategy σ^p, the state process is an irreducible Markov chain with one ergodic class then, for any strategy σ the frequencies state-actions*

$$(f^{p,t}_{\mu,\sigma}(y,a))_{p\in P, t\geq 1}$$

where

$$f^{p,t}_{\sigma,\mu}(y,a) = \frac{\sum_{j=1}^{t}\left(\sum_{k=j}^{t}\mathbb{P}(T=k)\right)P_\sigma(X_j=y, a_j=a|y_1=\mu)}{\sum_{j=1}^{t} j\mathbb{P}(T=j)}$$

are tight.

PROOF. See Appendix. □

The occupation measures in this extended model are characterized by the following convergence result: $\mathbb{P}_{\sigma,\mu}$ almost surely, the random variables that give the frequencies state-action

$$\frac{\sum_{k=1}^{t}P(T=k)\sum_{j=2}^{k}\delta_{y'}(X_j^p) - \sum_{y,a}Q^p_{yay'}\delta_{(y,a)}(X^p_{j-1}, a^p_{j-1})}{\sum_{j=1}^{t} jP(T=j)} \tag{13}$$

goes to zero when t goes to infinity, for all $y'\in Y^p, p\in P$. Hence, when $E(T)=\sum_{k\geq 0}P(T>k)=\sum_k kP(T=k) < +\infty$ then we obtain the equation:

$$\frac{\sum_{k=1}^{+\infty}P(T=k)\sum_{j=2}^{+\infty}\delta_{y'}(X_j^p) - \sum_{y,a}Q^p_{yay'}\delta_{(y,a)}(X^p_{j-1}, a^p_{j-1})}{\sum_{j=2}^{\infty} P(T\geq j)} \tag{14}$$

7. CONCLUSIONS AND PERSPECTIVES

In this paper we have investigated power control interaction based on stochastic modeling of the remaining energy of different types of battery in large networks. We have showed existence of equilibria in the general model of constrained Markovian population games under ergodic assumptions. This model offers a new class of repeated games: *constrained repeated games with individual transition states and unknown horizon.*

APPENDIX

We first need the following lemma:

LEMMA .0.4. *For all strategies σ and all initial states μ,*

$$\sum_i P_{\sigma^p,\mu}(X^p_t\in J^p_i \text{ almost any time}) = 1, \text{ and}$$

$$P_{\sigma^p,\mu}(a^p_t\in A^p_i(y_t) \text{ almost any time}) = 1$$

PROOF THE RESULT 5.2.1. Combining the Proposition 3.4 and the *strong law of large number for martingale differences* [7] one has the following results:

- $\mathbb{P}_{\sigma,\mu}$ almost surely, the random variables that give the frequencies state-action satisfy

$$\lim_{t\longrightarrow+\infty}\frac{1}{t}\sum_{k=2}^{t}\left[\delta_{y'}(X^p_k) - \sum_{y,a}q^p_{yay'}\delta_{(y,a)}(X^p_{k-1}, a^p_{k-1})\right] = 0$$

for all $y'\in Y^p, p\in P$.

- Let Ω be set $(y_1,a_1,x_1,y_2,a_2,x_2,\dots,)$ in H_∞ that satisfy (i) $a_t\in A^p(y_t),\ t\geq T$ for some integer T. (ii)

$$\lim_{t\longrightarrow+\infty}\frac{1}{t}\sum_{k=2}^{t}\left[\delta_{y'}(X^p_k) - \sum_{y,a}q^p_{yay'}\delta_{(y,a)}(X^p_{k-1}, a^p_{k-1})\right] = 0$$

(iii) $\limsup_{t\longrightarrow+\infty}\frac{1}{t}\sum_{k=1}^{t}C^p_{y_k,a_k}(x_k)\leq\alpha^p$,

(iv) $\limsup_{t\longrightarrow+\infty}\frac{1}{t}\sum_{k=1}^{t}D^p(y_k,a_k)\leq\beta^p$,

Due to lemma .0.4 and the fact that σ is feasible, we have that $\mathbb{P}_{\sigma,\mu}(\Omega)=1$. It suffices therefore to show that $\text{CPP}_{(p,i)}$ is feasible and the event

$$\{X^p_t\in J^p_i \text{almost any time}\}\bigcap\Omega$$

is contained in the event $\{\liminf_{t\longrightarrow\infty}\frac{1}{t}\sum_{k=1}^{t}r^p_{y_k,a_k}(x_t)\leq v^p_i(x)\}$. Let define the random variable associate to the frequencies state-actions $f^{p,t}(y,a)=\frac{1}{t}\sum_{k=1}^{t}\delta_{(y,a)}(y_k,a_k)$ for $y\in Y^p, a\in A^p(y)$. We have that any limit point $\{f^p(y,a)\}$ of $f^{p,t}(y,a)_{t\geq 1}$ is a feasible solution of $\text{CPP}_{(p,i)}$ and that

$$\sum_{y\in J^p_i}\sum_{a\in A^p_i(y)} r^p_{y,a}(x)f^p(y,a)\leq v^p_i(x)$$

Thus, $\mathbb{P}_{\sigma,\mu}$ almost surely, one has,

$$F^p_\mu(\sigma) = \sum_i \delta_{\{X_t\in J^p_i \text{aat }\}}\sum_{y\in J^p_i}\sum_{a\in A^p_i(y)}\Pi^p_i(y)f^p_i(y,a)r^p_{y,a}(x^{-p}),$$

$$\sum_{y\in J^p_i}\sum_{a\in A^p_i(y)}\Pi^p_i(y)f^p_i(y,a)C^p_{y,a}(x^{-p})\leq\alpha^p,$$

$$\sum_{y\in J^p_i}\sum_{a\in A^p_i(y)}\Pi^p_i(y)f^p_i(y,a)D^p(y,a)\leq\beta^p$$

where $\Pi^p_i(y)$ is the unique stationary distribution associate to the state-process restricted to J^p_i, and *aat* :=almost any time. □

PROOF OF THE PROPOSITION 6.1.3. Since each subpopulation has a single (aperiodic) ergodic class under stationary strategies, stationary state probabilities exists (depending on the strategy) and $(Q^p(\sigma))^j$ goes to $\Pi^p(\sigma)$. Hence, $f^{p,t}_{\sigma,\mu}(y,a)$ converges weakly to $\Pi^p(\sigma)_{y,a}\sigma_{y,a}$. Using Lemma 18.2 in [1], The claim for general strategies follows from the bounded convergence theorem: the sequence $\{f^{p,t}_{\sigma,\mu}\}$ is bounded (by one) and converges weakly. Hence, for any strategies $\wp\in\Sigma^p$ we have that, $\lim_t\langle\wp, f^{p,t}_{\sigma,\mu}\rangle = \langle\wp, \lim_t f^{p,t}_{\sigma,\mu}\rangle$. This completes the proof. □

Acknowledgment

The authors thank three anonymous reviewers for helpful comments and suggestions which lead to improvement of the paper.

A. REFERENCES

[1] Altman E.:*Constrained Markov Decision Processes*, Chapman and Hall/CRC, 1999.

[2] Altman E. and Hayel Y.: *A Stochastic Evolutionary Game of Energy Management in a Distributed Aloha Network*, In Proc. of INFOCOM 2008.

[3] Altman E., Elazouzi R., Hayel Y. and Tembine H., *Evolutionary Power Control Games in Wireless Networks*, in Proc. 7th IFIP Networking 2008, Singapore, May 5-9 .

[4] Altman E., Hordijk A., Spiekma F. M.:*Contraction conditions for average and $\alpha-$discounted optimality in countable stat Markov games with unbounded rewards* , 1996.

[5] Altman E. and Shwartz A. , *Constrained Markov Games: Nash Equilibria* , Annals of the International Society of Dynamic Games , vol. 5, Birkhauser, V. Gaitsgory, J. Filar and K. Mizukami, editors, pp. 303-323, 2000.

[6] Altman E., Avrachenkov K.,Marquez R. and :*Zero-sum constrained stochastic games with independent state processes*, Mathematical Methods of Operations Research, 62,375-386,2005.

[7] Anastassiou G., *Quantitative Approximations*, Chapman & Hall/CRC, 2000.

[8] Fink A. M. *Equilibrium in a Stochastic $n-$Person Game*, J. Sci. Hiroshima University, 28, 89-93, 1964.

[9] Flesch J., Thuijsman F., Vrieze O J.:*Stochastic games with additive transitions*, European Journal of Operational Research, 2007.

[10] Flesch J., Schoenmakers G.,Vrieze O J.:*Stochastic games on a product space*, Research Memoranda 010, METEOR, Maastricht Research School of Economics of Technology and Organization, 2007.

[11] Dutta P.K.:*A Folk theorem for stocastic games*, Journal of Economic Theory, 66,pp.1-32, 1995.

[12] Hofbauer H. and Sigmund K. *Evolutionary game dynamics*, Bulletins (New Series) OF AMS, Volume 40, Number 4, pp. 479-519.

[13] Menache I. and Altman E. :*Battery-State Dependent Power Control as a Dynamic Game*, In Proc. of WiOpt'08, Berlin, 31 March–04 April 2008.

[14] Niyato D., Hossain E., and Fallahi A.,*Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: Performance analysis and optimization*. IEEE Transactions on Mobile Computing, 6(2):221-236, 2007.

[15] Shapley L.S:*Stochastic games*, Proc. Nat. Acad. Science, 39: pp.1095-1100, 1953.

[16] Sorin S.:*Asymptotic properties of a non-zero sum game*, International Journal of Game Theory, 15, pp. 101-107, 1986.

[17] Solan E., *Stochastic Games*, Thesis, Hebrew University of Jerusalem, 1998.

[18] Takahashi M.:*Equilibrium points for stochastic non-cooperative $n-$person games*, Journal of Science of Hiroshima University, pp. 95-99.

[19] Tembine H., Altman E., El-Azouzi R. and Hayel Y.: *Evolutionary games with random number of interacting players applied to access control*, in Proc. of WiOpt, Berlin, 31 March-04 April 2008.

[20] Vieille N.:*Equilibrium in two-person stochastic games: a reduction* , Israel Journal of Mathematics, 119, pp.93-126.

[21] Vieille N.:*Stochastic games: Recent results*. In Handbook of Game Theory, pages 1833-1850. Elsevier Science, 2002.

[22] Vieille N., Rosenberg D., Solan E.:*Stochastic Games with a Single Controller and Incomplete Information*, 2004,SIAM Journal on Control and Optimization, 43, 86-110.

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2008.4359
http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4359