An Anonymous Sequential Game Approach for Battery State Dependent Power Control

Eitan Altman, Yezekael Hayel and Piotr Więcek

POPEYE meeting
Avignon, 9th of april, 2009

Informal description

- A large population of mobile terminals competing for a wireless access.
- Each terminal attempts transmission over a sequence of time slots.
- At each attempt, it has to take the decision on the transmission power based on its battery energy state.
- The transmission ends when the battery is empty.
- The aim of each player to maximize his throughput minus the cost of the transmission over whole lifetime of his battery.

Informal description

- A large population of mobile terminals competing for a wireless access.
- Each terminal attempts transmission over a sequence of time slots.
- At each attempt, it has to take the decision on the transmission power based on its battery energy state.
- The transmission ends when the battery is empty.
- The aim of each player to maximize his throughput minus the cost of the transmission over whole lifetime of his battery.

Informal description

- A large population of mobile terminals competing for a wireless access.
- Each terminal attempts transmission over a sequence of time slots.
- At each attempt, it has to take the decision on the transmission power based on its battery energy state.
- The transmission ends when the battery is empty.
- The aim of each player to maximize his throughput minus the cost of the transmission over whole lifetime of his battery.

Informal description

- A large population of mobile terminals competing for a wireless access.
- Each terminal attempts transmission over a sequence of time slots.
- At each attempt, it has to take the decision on the transmission power based on its battery energy state.
- The transmission ends when the battery is empty.
- The aim of each player to maximize his throughput minus the cost of the transmission over whole lifetime of his battery.

Challenges and mathematical tools I

- Problem has 2 dimensions:
  1. Increase in the transmission power has a crucial impact on the achievable throughputs.
  2. Each mobile has to take into consideration the battery state in the decision taking, which influenced by its past choices.

  A mobile that is interested in maximizing the amount of information during its lifetime, has to balance the two.

  $\Longrightarrow$ Dynamic model (stochastic game)
- In a CDMA type cellular system all mobiles transmit simultaneously to a common base station.
  1. Performance of a mobile determined by the distribution of the actions used by all other mobiles.
  2. Mobiles are indistinguishable.

  $\Longrightarrow$ An anonymous game

Challenges and mathematical tools I

- Problem has 2 dimensions:
  1. Increase in the transmission power has a crucial impact on the achievable throughputs.
  2. Each mobile has to take into consideration the battery state in the decision taking, which influenced by its past choices.

  A mobile that is interested in maximizing the amount of information during its lifetime, has to balance the two.

  $\Longrightarrow$ Dynamic model (stochastic game)
- In a CDMA type cellular system all mobiles transmit simultaneously to a common base station.
  1. Performance of a mobile determined by the distribution of the actions used by all other mobiles.
  2. Mobiles are indistinguishable.

  $\Longrightarrow$ An anonymous game

Challenges and mathematical tools I

- Problem has 2 dimensions:
  1. Increase in the transmission power has a crucial impact on the achievable throughputs.
  2. Each mobile has to take into consideration the battery state in the decision taking, which influenced by its past choices.

  A mobile that is interested in maximizing the amount of information during its lifetime, has to balance the two.

  $\Longrightarrow$ Dynamic model (stochastic game)
- In a CDMA type cellular system all mobiles transmit simultaneously to a common base station.
  1. Performance of a mobile determined by the distribution of the actions used by all other mobiles.
  2. Mobiles are indistinguishable.

  $\Longrightarrow$ An anonymous game

Challenges and mathematical tools I

- Problem has 2 dimensions:
  1. Increase in the transmission power has a crucial impact on the achievable throughputs.
  2. Each mobile has to take into consideration the battery state in the decision taking, which influenced by its past choices.

  A mobile that is interested in maximizing the amount of information during its lifetime, has to balance the two.

  $\Longrightarrow$ Dynamic model (stochastic game)
- In a CDMA type cellular system all mobiles transmit simultaneously to a common base station.
  1. Performance of a mobile determined by the distribution of the actions used by all other mobiles.
  2. Mobiles are indistinguishable.

  $\Longrightarrow$ An anonymous game

Challenges and mathematical tools II

The framework that combines the two mentioned above is that of anonymous sequential game

- Introduced by Jovanovic and Rosenthal (1988), used subsequently in some economic models.
- Only for discounted cost.

Our goal: To extend it to the cost criterion used here.

Challenges and mathematical tools II

The framework that combines the two mentioned above is that of anonymous sequential game

- Introduced by Jovanovic and Rosenthal (1988), used subsequently in some economic models.
- Only for discounted cost.

Our goal: To extend it to the cost criterion used here.

Challenges and mathematical tools II

The framework that combines the two mentioned above is that of anonymous sequential game

- Introduced by Jovanovic and Rosenthal (1988), used subsequently in some economic models.
- Only for discounted cost.

Our goal: To extend it to the cost criterion used here.

Model for individual player I

We associate with each player a Markov Decision Process with the following parameters:

- $S = \{0, 1, \dots, N\}$ is a finite set of states.
- $Q = \{q_1, \dots, q_K\}$ is a finite set of actions.
- P is the transition probability law. Namely, $p(s'|s, q)$ is the probability that the next state is s' given that the actual state is s and the action taken is q.

Model for individual player II

- The state s^i of player i = his battery energy state
- The action of a player = energy level at which he transmits
  The set of actions available to a player in state s is

  $$Q_s = \{q_1, \ldots, q_s^+\} \subset Q$$

  For any two $s_1 < s_2$, $q_{s_1}^+ \leq q_{s_2}^+$.
- The transitions are defined for individual states of the players.
  The probability of staying in state s by a player taking action q is $p(q)$.

  $$p(q) = 1 - \alpha q - \gamma,$$

  where α and γ are some fixed positive coefficients.
  At any given time a mobile whose battery is empty may have it recharged with probability p_{0N}

Model for individual player II

- The state s^i of player i = his battery energy state
- The action of a player = energy level at which he transmits
  The set of actions available to a player in state s is

$$Q_s = \{q_1, \ldots, q_s^+\} \subset Q$$

For any two $s_1 < s_2$, $q_{s_1}^+ \leq q_{s_2}^+$.

- The transitions are defined for individual states of the players.
  The probability of staying in state s by a player taking action q is $p(q)$.

$$p(q) = 1 - \alpha q - \gamma,$$

where α and γ are some fixed positive coefficients.
At any given time a mobile whose battery is empty may have it recharged with probability p_{0N}

Model for individual player II

- The state s^i of player i = his battery energy state
- The action of a player = energy level at which he transmits
  The set of actions available to a player in state s is

$$Q_s = \{q_1, \ldots, q_s^+\} \subset Q$$

  For any two $s_1 < s_2$, $q_{s_1}^+ \leq q_{s_2}^+$.
- The transitions are defined for individual states of the players.
  The probability of staying in state s by a player taking action q is $p(q)$.

$$p(q) = 1 - \alpha q - \gamma,$$

  where α and γ are some fixed positive coefficients.
  At any given time a mobile whose battery is empty may have it recharged with probability p_{0N}

Model for individual player II

- The state s^i of player i = his battery energy state
- The action of a player = energy level at which he transmits
  The set of actions available to a player in state s is

$$Q_s = \{q_1, \ldots, q_s^+\} \subset Q$$

  For any two $s_1 < s_2$, $q_{s_1}^+ \leq q_{s_2}^+$.
- The transitions are defined for individual states of the players.
  The probability of staying in state s by a player taking action q is $p(q)$.

$$p(q) = 1 - \alpha q - \gamma,$$

  where α and γ are some fixed positive coefficients.
  At any given time a mobile whose battery is empty may have it recharged with probability p_{0N}

Model of interactions

- The global state of the system at time t is

$$X^t = (X_0^t, \ldots, X_N^t),$$

such that $\sum_{n=0}^{N} X_n^t = 1$
X_s^t = fraction of mobiles with battery state s.

- The reward at time t for a user in state s_t playing q_k^t when the vector of proportions of players using different actions is w^t is

$$R^t(q_k^t, s_t; w^t) = \frac{q_k^t}{\sigma^2 + C\sum_{l=1}^{K} q_l^t w_l^t} - \beta q_k^t,$$

C – interference parameter
σ^2 – noise power
β – energy cost.

Model of interactions

- The global state of the system at time t is

$$X^t = (X_0^t, \dots, X_N^t),$$

such that $\sum_{n=0}^{N} X_n^t = 1$
X_s^t = fraction of mobiles with battery state s.

- The reward at time t for a user in state s_t playing q_k^t when the vector of proportions of players using different actions is w^t is

$$R^t(q_k^t, s_t; w^t) = \frac{q_k^t}{\sigma^2 + C \sum_{l=1}^{K} q_l^t w_l^t} - \beta q_k^t,$$

C – interference parameter
σ^2 – noise power
β – energy cost.

Policies

A stationary policy of a player u is a map from the set of possible individual states S to the set of probability measures over set Q, $P(Q)$.

- We do not consider general (history/time-dependent) policies.
- We assume the game is in a stationary regime
$\Longrightarrow$ State of the system X^t does not change over time.

Policies

A stationary policy of a player u is a map from the set of possible individual states S to the set of probability measures over set Q, $P(Q)$.

- We do not consider general (history/time-dependent) policies.
- We assume the game is in a stationary regime
  $\Longrightarrow$ State of the system X^t does not change over time.

Policies

A stationary policy of a player u is a map from the set of possible individual states S to the set of probability measures over set Q, $P(Q)$.

- We do not consider general (history/time-dependent) policies.
- We assume the game is in a stationary regime
  $\Longrightarrow$ State of the system X^t does not change over time.

Policies

A stationary policy of a player u is a map from the set of possible individual states S to the set of probability measures over set Q, $P(Q)$.

- We do not consider general (history/time-dependent) policies.
- We assume the game is in a stationary regime
  $\Longrightarrow$ State of the system X^t does not change over time.

The equilibrium

The objective of player is to maximize his reward over the whole lifetime of his battery:

$$J(v,u) = E^{p,X^0} \sum_{t=1}^{\tau} R^t(u,v).$$

τ – lifetime of a player (random variable)

A stationary policy u is in an equilibrium if

$$J(u,u) = \max_{v \in \mathcal{U}} J(v,u).$$

The equilibrium

The objective of player is to maximize his reward over the whole lifetime of his battery:

$$J(v,u) = E^{p,X^0} \sum_{t=1}^{\tau} R^t(u,v).$$

τ – lifetime of a player (random variable)
A stationary policy u is in an equilibrium if

$$J(u,u) = \max_{v \in \mathcal{U}} J(v,u).$$

Threshold policies

$$\mathcal{U}_0 = \Bigg\{ u \in \mathcal{U} : \exists s_0 \in S, \exists r \in [0,1],$$

$$u_s = \left\{ \begin{array}{ll} \delta[q_1], & s < s_0 \\ r\delta[q_1] + (1-r)\delta[q_s^+], & s = s_0 \\ \delta[q_s^+], & s > s_0 \end{array} \right\} \Bigg\}.$$

$\delta[q]$ – probability measure concentrated in q.

Equilibrium characterization theorem I

Theorem

The game under consideration always possesses an equilibrium $u \in \mathcal{U}_0$. Moreover:

(i) *This equilibrium is unique in the set $\mathcal{U}_0$.*

(ii) *If*

$$\beta CN > \frac{N\alpha(1-\beta\sigma^2)+N\gamma}{\alpha q_1^+ + \gamma} \quad \text{and} \quad p_{0N} \leq \frac{\alpha(1-\beta\sigma^2)}{\beta CN - \frac{N\alpha(1-\beta\sigma^2)+N\gamma}{\alpha q_1^+ + \gamma}}$$

then $u^+(s) = \delta[q_s^+]$ is the equilibrium.

Equilibrium characterization theorem II

(iii) *If*

$$\beta C q_1 > 1 - \beta\sigma^2 \quad \textit{and} \quad p_{0N} \geq \frac{(\alpha q_1 + \gamma)(1 - \beta\sigma^2)}{N(\beta C q_1 - (1 - \beta\sigma^2))}$$

then $u^-(s) = \delta[q_1]$ *is the equilibrium.*

The proof I

Step 1:

For each policy $u \in \mathcal{U}$ we compute a stationary global state $X(u)$.

Step 2:

We introduce in the set $\mathcal{U}_0$ a linear ordering under which $\mathcal{U}_0$ is homeomorphic with the interval $[1, N]$ (h – the homeomorphism). We show that the global interference

$$A(u) = \sigma^2 + C \sum_{k=1}^{K} w_k(u) q_k$$

is continuous increasing on $\mathcal{U}_0$.

The proof I

Step 1:

For each policy $u \in \mathcal{U}$ we compute a stationary global state $X(u)$.

Step 2:

We introduce in the set $\mathcal{U}_0$ a linear ordering under which $\mathcal{U}_0$ is homeomorphic with the interval $[1, N]$ (h – the homeomorphism). We show that the global interference

$$A(u) = \sigma^2 + C \sum_{k=1}^{K} w_k(u) q_k$$

is continuous increasing on $\mathcal{U}_0$.

The proof II

Step 3:

We show that the equilibrium policy is a solution to the equation

$$A(u) = \frac{1}{\beta},$$

which exists and is unique, if A admits values both greater and smaller than $\frac{1}{\beta}$, since A has an intermediate value property.

Step 4:

We compute when $A(u) > \frac{1}{\beta}$ for every $u \in \mathcal{U}_0$ and when $A(u) < \frac{1}{\beta}$ for every $u \in \mathcal{U}_0$. This gives the conditions for u^- or u^+ to be optimal. $\square$

The proof II

Step 3:
We show that the equilibrium policy is a solution to the equation

$$A(u) = \frac{1}{\beta},$$

which exists and is unique, if A admits values both greater and smaller than $\frac{1}{\beta}$, since A has an intermediate value property.
Step 4:
We compute when $A(u) > \frac{1}{\beta}$ for every $u \in \mathcal{U}_0$ and when $A(u) < \frac{1}{\beta}$ for every $u \in \mathcal{U}_0$. This gives the conditions for u^- or u^+ to be optimal. $\square$

Theorem

The equilibrium $u^* \in \mathcal{U}_0$ in the game under consideration can be computed using bisection applied to function

$$\phi(a) = A(h^{-1}(a)) - \frac{1}{\beta}$$

on the interval $[1, N]$. The approximation of u^* will be given by $h^{-1}(a^*)$, where a^* is the (approximate) zero of function ϕ. If the zero does not exist then either u^+ (when $\phi < 0$) or u^- (when $\phi > 0$) is the equilibrium.

Simple example with three battery states I

Three possible states of the battery:

- empty (E),
- almost empty (A),
- full (F).

Whenever the player is in state F he has two actions possible:

- To transmit at high power h,
- To transmit at low power l.

If he is in state A, he can only transmit at low power l.

Simple example with three battery states II

The strategy u^- is optimal when

$$\beta Cl > 1 - \beta\sigma^2$$

and

$$p_{EF} \geq \frac{(\alpha l + \gamma)(1 - \beta\sigma^2)}{\beta Cl - (1 - \beta\sigma^2)}.$$

The strategy u^+ is optimal when

$$\beta\sigma^2(\alpha(l+h) + 2\gamma) > (l+h)(C\beta\gamma - \alpha) + 2(C\beta\alpha hl - \gamma)$$

or

$$p_{EF} \leq \frac{(\alpha l + \gamma)(\alpha h + \gamma)(1 - \beta\sigma^2)}{(l+h)(C\beta\gamma - \alpha) + 2(\beta\alpha hl - \gamma) - \beta\sigma^2(\alpha(l+h) + 2\gamma)}$$

Simple example with three battery states II

The strategy u^- is optimal when

$$\beta C l > 1 - \beta\sigma^2$$

and

$$p_{EF} \geq \frac{(\alpha l + \gamma)(1 - \beta\sigma^2)}{\beta C l - (1 - \beta\sigma^2)}.$$

The strategy u^+ is optimal when

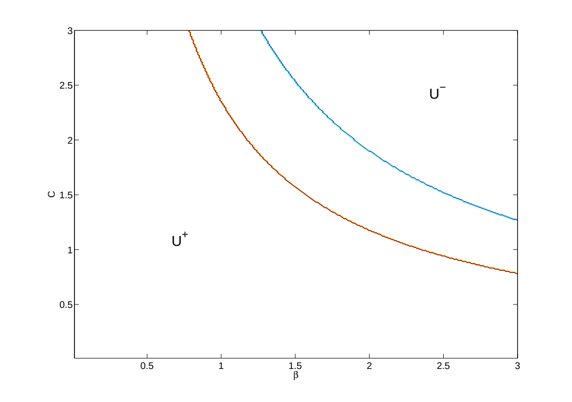
$$\beta\sigma^2(\alpha(l+h) + 2\gamma) > (l+h)(C\beta\gamma - \alpha) + 2(C\beta\alpha h l - \gamma)$$

or

$$p_{EF} \leq \frac{(\alpha l + \gamma)(\alpha h + \gamma)(1 - \beta\sigma^2)}{(l+h)(C\beta\gamma - \alpha) + 2(\beta\alpha h l - \gamma) - \beta\sigma^2(\alpha(l+h) + 2\gamma)}$$

Simple example with three battery states III

The regions where strategies u^+ and u^- are optimal as a function of β and C for $p_{EF} = 0.5$, $\alpha = 0.8$, $\gamma = 0.05$, $h = 1$, $l = 0.5$, $\sigma^2 = 0.00001$.



Simple example with three battery states IV

Equilibrium strategies for the region where neither u^- nor u^+ are optimal:

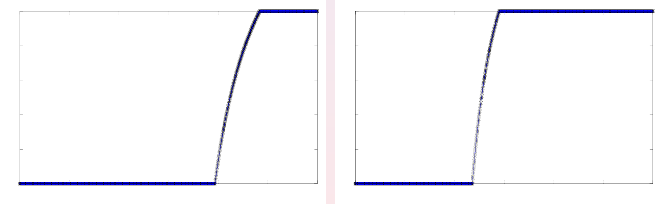
$$u^*(s) = \begin{cases} \delta[l], & s = L \\ r\delta[l] + (1-r)\delta[h], & s = H \end{cases}$$

where

$$\begin{aligned} r &= \frac{Cp_{EF}[2\alpha lh + \gamma(l+h)]}{(l-h)[(1-\beta\sigma^2)\alpha(p_{EF}+\alpha l+\gamma) - Cp_{EF}(2\alpha l+\gamma)]} \\ &- \frac{(1-\beta\sigma^2)[\alpha p_{EF}(l+h) + 2p_{EF}\gamma + (\alpha l+\gamma)(\alpha h+\gamma)]}{(l-h)[(1-\beta\sigma^2)\alpha(p_{EF}+\alpha l+\gamma) - Cp_{EF}(2\alpha l+\gamma)]}. \end{aligned}$$

Simple example with three battery states V

The dependance of r on β for $p_{EF} = 0.5$, $\alpha = 0.8$, $\gamma = 0.05$, $h = 1$, $l = 0.5$, $\sigma^2 = 0.00001$, and for $C = 1.2$ (left) and $C = 2$ (right)



Work in progress

Introducing a distributed learning algorithm

- Checking whether in our model the dynamics converge for variants of known evolutionary dynamics.
- Finding the algorithms that converge (proving their convergence).
- Numerical experiments.

Work in progress

Introducing a distributed learning algorithm

- Checking whether in our model the dynamics converge for variants of known evolutionary dynamics.
- Finding the algorithms that converge (proving their convergence).
- Numerical experiments.

Work in progress

Introducing a distributed learning algorithm

- Checking whether in our model the dynamics converge for variants of known evolutionary dynamics.
- Finding the algorithms that converge (proving their convergence).
- Numerical experiments.

Work in progress

Introducing a distributed learning algorithm

- Checking whether in our model the dynamics converge for variants of known evolutionary dynamics.
- Finding the algorithms that converge (proving their convergence).
- Numerical experiments.

Possible extensions

1. General theory for anonymous sequential games with total and average costs.
2. Extension of the model presented here to include different types of users (differing by some battery characteristics like maximal achievable battery energy or probabilities $p(q)$).

Possible extensions

1. General theory for anonymous sequential games with total and average costs.
2. Extension of the model presented here to include different types of users (differing by some battery characteristics like maximal achievable battery energy or probabilities $p(q)$).

Thank you!