Self-Imitation for Cognitive Radio Networks

Marceau Coupechoux*, Stefano Iellamo*, Lin Chen+
* TELECOM ParisTech (INFRES/RMS) and CNRS LTCI
+ University Paris XI Orsay (LRI)

AlgoGT

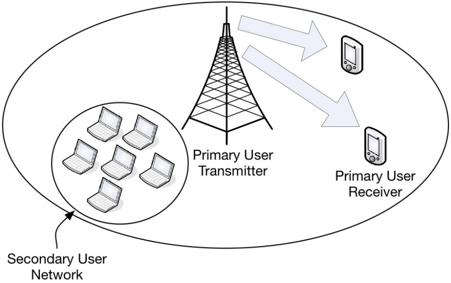
2nd July 2013

Introduction

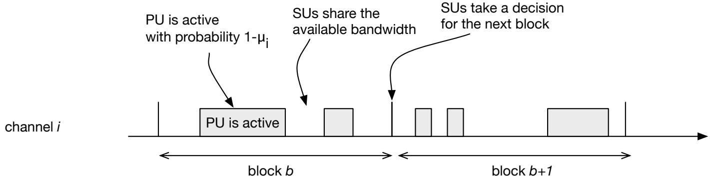
- Opportunistic spectrum access in large-scale cognitive radio networks
- SU access the freq. channels partially occupied by the licensed PU
- Distributed spectrum access policies based on self-imitation
- Convergence analysis based on perturbed Markov chains

System Model I

- A PU is using on the DL a set $\mathcal{C}$ of C freq. channels
- Primary receivers are operated in a synchronous time-slotted fashion
- The secondary network is made of a set $\mathcal{N}$ of N SUs
- We assume perfect sensing



System Model II



- At each time slot, channel i is free with probability μ_i
- Throughput achieved by j on channel i is denoted T_j^i
- Expected throughput $\pi_i(n_i) = E[T_j^i] = B\mu_i S(n_i)$
- S is a function that depends on the MAC protocol and on the number of SUs on channel i, n_i
- We assume $B = 1$, S strictly decreasing and $S(x) \leq 1/x$ for $x > 0$

Spectrum Access Game Formulation

Definition

The spectrum access game G is a 3-tuple $(\mathcal{N}, \mathcal{C}, \{U_j\})$, where $\mathcal{N}$ is the player set, $\mathcal{C}$ is the strategy set of each player. Each player j chooses its strategy $s_j \in \mathcal{C}$ to maximize its payoff function U_j defined as $U_j = \pi_{s_j}(n_{s_j}) = \mathbb{E}[T_j^{s_j}]$.

Lemma

For the spectrum access game G, there exists at least one Nash equilibrium.

Lemma

For N sufficiently large, G admits a unique NE, where all SUs get the same payoff.

Motivation

- Find a distributed strategy for SUs to converge to the NE
- Uniform random imitation of another SU leads to the *replicator dynamics* (see Proportional Imitation Rule in [1, 2])
- Uniform random imitation of **two** SUs leads to the *aggregate monotone dynamics* (see Double Imitation in [1, 2])
- Imitation on the same channel can be approximated by a double replicator dynamics [3]
- We now avoid any information exchange between SUs

LASTra

Algorithm 1 LASTra: executed at each SU j

```
Initialization: set ε(t)
At t = 0 and t = 1, randomly choose a channel to stay and store the payoffs U_j(0) and U_j(1).
while at the end of block t > 1 do
    With probability 1 − ε(t):
    if U_j(t − 1) > U_j(t) then
        Migrate to the channel s_j(t − 1)
    end if
    With probability ε(t): switch to a random channel
end while
```

Perturbed Markov Chain I

- We have a model of evolution with noise:
  - $Z = \left\{ z \triangleq \left(s_j(t), s_j(t-1), \pi_j(n_{s_j}(t)), \pi_j(n_{s_j}(t-1))\right) \right\}$ is the finite state space of the system stochastic process
  - $P = (p_{uv})_{(u,v)\in Z^2}$ is the transition matrix of LASTra without exploration (i.e. $\epsilon(t) = 0 \; \forall t$)
  - $P(\epsilon) = (p_{uv}(\epsilon))_{(u,v)\in Z^2}$ is a family of transitions matrices on Z indexed by $\epsilon \in [0, \bar{\epsilon}]$ associated to LASTra with exploration ϵ
- Properties of $P(\epsilon)$:
  - $P(\epsilon)$ is ergodic for $\epsilon > 0$
  - $P(\epsilon)$ is continuous in ϵ and $P(0) = P$
  - There is a cost function $c : Z^2 \to \mathcal{R}^+ \cup \{\infty\}$ s.t. for any pair of states (u, v), $\lim_{\epsilon\to 0} \frac{p_{uv}(\epsilon)}{\epsilon^{c_{uv}}}$ exists and is strictly positive for $c_{uv} < \infty$ and $p_{uv}(\epsilon) = 0$ if $c_{uv} = \infty$

Perturbed Markov Chain II

- Remarks:
  - ϵ can be interpreted as a small probability that SUs do not follow the rule of the dynamics. When a SU explores, we say that there is a *mutation*
  - The cost c_{uv} is the rate at which $p_{uv}(\epsilon)$ tends to zero as ϵ vanishes
  - c_{uv} can also be seen as the number of mutations needed to go from state u to state v
  - $c_{uv} = 0$ when $p_{uv} \neq 0$ in the unperturbed Markov chain
  - $c_{uv} = \infty$ when the transition $u \to v$ is impossible in the perturbed Markov chain

Perturbed Markov Chain III

Lemma ([4])

There exists a limit distribution $\mu^* = \lim_{\epsilon \to 0} \mu(\epsilon)$

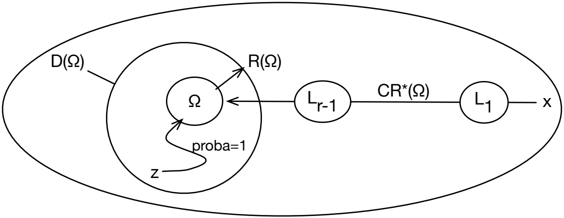
Definition

A state $i \in Z$ is said to be *long-run stochastically stable* iff $\mu_i^* > 0$.

Lemma ([4])

The set of stochastically stable states is included in the recurrent classes (limit sets) of the unperturbed Markov chain (Z, P).

Ellison Radius Coradius Theorem I



- Ω: a union of limit sets of (Z, P)
- $D(\Omega)$: **basin of attraction**, the set of states from which the unperturbed chain converges to Ω w.p.1
- $R(\Omega)$: **radius**, the min cost of any path from Ω out of $D(\Omega)$
- $CR(\Omega)$: **coradius**, maximum cost to Ω
- $CR^*(\Omega)$: **modified coradius**, obtained by substracting from the cost, the radius of intermediate limit sets

Ellison Radius Coradius Theorem II

Theorem (Ellison00, Theorem 2 and Sandholm10, Chap. 12)

Let $(Z, P, P(\epsilon))$ be a model of evolution with noise and suppose that for some set Ω, which is a union of limit sets, $R(\Omega) > CR^*(\Omega)$, then:

- *The long-run stochastically stable set of the model is included in Ω.*
- *For any $y \notin \Omega$, the longest expected wait to reach Ω is $W(y, \Omega, \epsilon) = O(\epsilon^{-CR^*(\Omega)})$ as $\epsilon \to 0$.*

Proof idea

Uses the Markov chain tree theorem and the fact that it is more difficult to escape from Ω than to return to Ω.

LASTra Convergence Analysis I

Lemma

Under LASTra, the limit sets (LS) are absorbing states, called imitation stable states (ISS).

Lemma

Let z^* be an ISS and Ω^* be the union of all ISS at NE. It holds that $R(z^*) = 1 \; \forall z^* \notin \Omega^*$.

Proof idea

For a congestion game $\mathcal{G}$ with player specific decreasing payoff functions, the weak-FIP property holds [5]. Using weak-FIP, we show that a single mutation is enough to leave the basin of attraction of any ISS not in Ω^* and to reach a new ISS.

LASTra Convergence Analysis II

Lemma

Ω^* can be reached from any state $z \notin \Omega^*$ by stepwise mutations.

Proof idea

We show that any z is in the basin of attraction of an ISS. The cost to reach this ISS is thus null. From the weak-FIP property: from any strategy profile, there exists a sequence of single player improvements that terminates at NE after a finite number of steps.

LASTra Convergence Analysis III

Lemma

$CR^*(\Omega^*) = 1$

Proof idea

From any state, there is a path of null cost to reach an ISS and then a path, which is a sequence of ISS. Each ISS has a radius of 1.

Lemma

$R(\Omega^*) > 1$

Proof idea

Comes from the definition of the NE and of LASTra.

LASTra Convergence Analysis IV

Theorem

If all SUs 1) adopt LAstra as self-imitation protocol and 2) adopt a random strategy at each iteration with probability $\epsilon \to 0$, then the system dynamics converges to Ω^*, which is also a pure NE of the game. The expected wait until a state in Ω^* is reached given that the system begins in any state not in Ω^*, is $O(\epsilon^{-1})$ as $\epsilon \to 0$.

Simulation Settings

- We compare our algorithm to Trial and Error (T&E, Pradelski's optimized learning parameters in [Pradelski&Young 2012]) [6] and to the Distributed Learning Algorithm (DLA) [Chen&Huang 2012] [7].
- We consider two networks:
  - **Network 1:** We consider $N = 50$ SUs, $C = 3$ channels characterized by the availability probabilities $\mu = [0.3, 0.5, 0.8]$.
  - **Network 2:** We set $N = 10$, $C = 2$ and $\mu = [0.2, 0.8]$.

Example of Trajectory

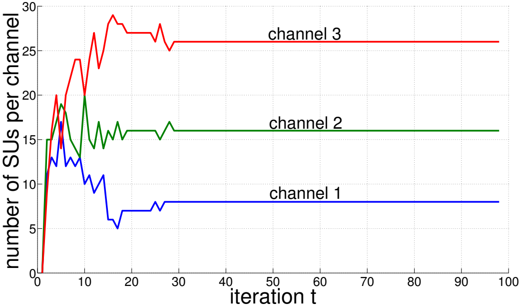


Figure : LASTra trajectory on Network 1.

LASTra vs DLA

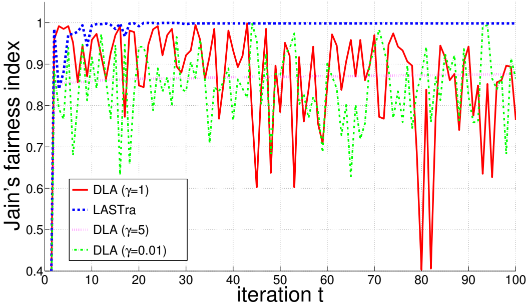


Figure : LASTra and DLA (with temperature γ) fairness index on Network 1.

LASTra vs T&E

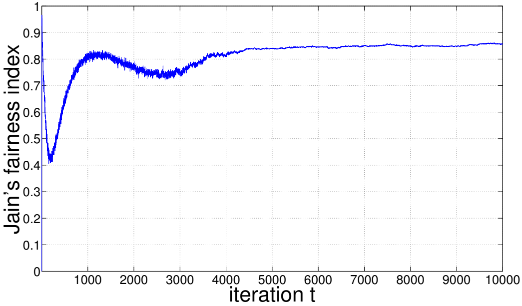


Figure : Trial and Error fairness index on Network 2 (average of 1000 trajectories).

LASTra vs T&E

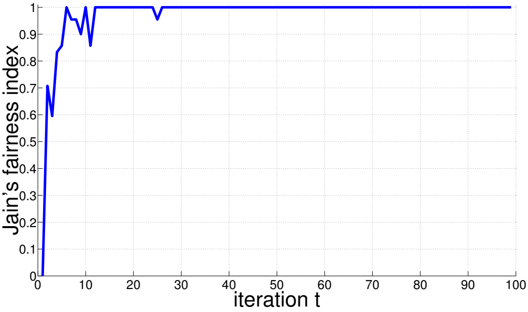


Figure : LASTra fairness index on Network 2.

Conclusion

- We discussed the distributed resource allocation problem in CRNs
- We have proposed a fully distributed scheme without any information exchange between SUs and based on self-imitation
- We have proved convergence using Ellison00 radius-coradius theorem
- We have compared LASTra to T&E [Pradelski&Young 2012] and to DLA [Chen&Huang 2012]

References I

[1] K. H. Schlag.
Why Imitate, and if so, How ? A Boundedly Rational Approach to Multi-Armed Bandits.
Journal of Economic Theory, 78(1):130–156, Jan. 1998.

[2] K. H. Schlag.
Which One Should I Imitate ?
Journal of Mathematical Economics, 31(4):493–522, May 1999.

[3] S. Iellamo, L. Chen, and M. Coupechoux.
Proportional and double imitation rules for spectrum access in cognitive radio networks.
Elsevier Computer Networks, 57(8):1863–1879, June 2013.

References II

[4] G. Ellison.
Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution.
Review of Economic Studies, 67, 2000.

[5] I. Milchtaich.
Congestion Games with Player-Specific Payoff Functions.
Games and Economic Behavior, 13:111–124, 1996.

[6] H. Peyton Young.
Learning by Trial and Error.
Games and Economic Behavior, 65(2):626–643, 2009.

[7] X. Chen and J. Huang.
Spatial spectrum access game: nash equilibria and distributed learning.
In *MobiHoc*, 2012.