Imitation-based Spectrum Access Policy for Cognitive Radio Networks

Invited paper

Stefano Iellamo* Lin Chen† Marceau Coupechoux* Athanasios V. Vasilakos‡

* Telecom ParisTech
LTCI CNRS 5141
46 Rue Barrault
Paris 75013, France
{iellamo,coupecho}@enst.fr

†Laboratoire de Recherche en Informatique (LRI)
University of Paris-Sud
91405 Orsay, France
chen@lri.fr

‡Department of Computer and Telecommunications Engineering
Univ. Western Macedonia, Greece
vasilako@ath.forthnet.gr

***Abstract*—In this paper, we tackle the problem of opportunistic spectrum access in cognitive radio networks. We consider a large number of unlicensed Secondary Users (SU) accessing a number of frequency channels partially occupied by licensed Primary Users (PU). Each channel is characterized by an unknown availability probability. We apply evolutionary game theory to model the spectrum access problem and propose imitation-based spectrum access policies based on the proportional imitation rule (PIR) and double imitation (DI) rule. We show that both policies converge exponentially in time to the Nash Equilibrium which is also the system optimum. The proposed spectrum access policies are evaluated by simulations which demonstrate their convergence to a stable equilibrium state which is also the system optimum.**

I. Introduction

Cognitive radio [1], with its capability to flexibly configure its transmission parameters, has emerged in recent years as a promising paradigm to enable more efficient spectrum utilization. Spectrum access models in cognitive radio networks can be classified into three categories, namely exclusive use (or operator sharing), commons and shared use of primary licensed spectrum [2]. In the last model, unlicensed secondary users (SUs) are allowed to access the spectrum of licensed primary users (PUs) in an opportunistic way. In this case, a well-designed spectrum access mechanism is crucial to achieve efficient spectrum usage.

In this paper, we focus on the generic model of cognitive networks consisting of several frequency channels, each characterized by a channel availability probability determined by the activity of PUs on the channel. In such models, from the individual SU's perspective, a challenging problem is to compete (or coordinate) with other SUs in order to opportunistically access the unused spectrum of PUs to maximize its own payoff (e.g., throughput); at the system level, a crucial research issue is to design efficient spectrum access protocols achieving optimal spectrum usage.

We tackle the problem of spectrum access in cognitive radio networks from an evolutionary game theoretic angle. More specifically, we develop a generic spectrum access policy based on imitation, a natural behavior rule widely observed in human society. The proposed spectrum access policy can converge to the stable system equilibrium in a distributed fashion relying only on local interactions among SUs. More specifically, we study the interaction among SUs under the proportional imitation rule and the more advanced adjusted proportional imitation rule with double sampling. Under both imitation rules, each SU strives to improve its individual payoff by imitating other SUs with higher payoff. Compared with the replicator dynamic, the most explored dynamic in evolutionary game theory and its application in wireless networking field, which mimics the effect of natural selection, imitation dynamic captures the spreading of successful strategies through imitation rather than inheritance, which is more adapted in games played by human societies (cf. [3] and references therein). Our work presented in this paper consists of the first step towards systematically applying imitation dynamic to address the spectrum access problem in cognitive radio networks and designing distributed imitation protocols that lead to an efficient and stable system equilibrium.

The rest of the paper is structured as follows. Section II presents the system model followed by the formulation of the spectrum access game. Section III describes the proposed imitation-based spectrum access policy. In Section IV, extensive simulations are performed to evaluate the performance of the proposed policy. Section V concludes the paper.

II. System Model and Spectrum Access Game Formulation

In this section, we present the system model of our work, followed by the game formulation of the spectrum access problem, which serves as the basis of subsequent analysis .

A. System Model

We consider a primary network consisting of a set $\mathcal{C}$ of C frequency channels, each with bandwidth B[1]. The users in the primary network are operated in a synchronous time-slotted fashion. A set $\mathcal{N}$ of N SUs tries to opportunistically access the channels when they are left free by PUs. Let $X_{i,k}$ be the

[1]The heterogeneous case with different channel capacities is left for future work.

random variable (RV) equal to 1 when slot k of channel i is free for SU transmission and 0 otherwise[2]. We assume that the process $\{X_{i,k}\}$ is stationary and independent for each i and k. We assume that at each time slot, channel i is free with probability μ_i, i.e., $\mathbb{E}[X_{i,k}] = \mu_i$. The channel availability probabilities $\mu \triangleq \{\mu_i\}$ are *a priori* not known by SUs. We assume perfect sensing at the SUs, i.e., any transmission of a PU on a channel is perfectly sensed by SUs sensing that channel and thus no collision occurs between PUs and SUs. An important direction of our future work is to address the case of imperfect sensing.

In our work, each SU j is a rational decision maker, striking to maximize the throughput it can achieve, denoted as T_j, which can be expressed as a function of μ_i and n_i, where i is the channel which j chooses, n_i is the number of SUs on channel i. More formally, the expected value of T_j can be written as:

$$\mathbb{E}[T_j] = f(\mu_i, n_i).$$

In order to perform a closed-form analysis, we focus on the scenario where the channel capacity is evenly shared among all SUs on the channel when it is free, i.e.,

$$\mathbb{E}[T_j] = f(\mu_i, n_i) = B\mu_i/n_i,$$

which corresponds to the case of the TDMA-based MAC layer.

B. Spectrum Access Game Formulation

To study the interactions among autonomous SUs and to derive distributed channel access algorithms, we formulate the channel selection problem as a spectrum access game where the players are the SUs. Each player j stays on a channel i to opportunistically exploit the unused spectrum of the PUs to maximize its expected throughput. The game is defined formally as follows:

Definition 1. *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 to maximize its normalized utility function U_j defined as $U_j = \mathbb{E}[T_j]/B = \mu_i/n_i$.*

Using the related theory on congestion games, we can characterize the Nash equilibrium (NE) in the spectrum access game G in the case of a countably infinite players set of cardinality N.

Theorem 1. *In the asymptotic case where N is large, G admits a unique NE. At the NE, there are $x_i^* N$ SUs staying with channel i, where $x_i^* = \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l}$.*

Proof: Given the form of SUs' utility function, it follows from [4] that the spectrum access game is a congestion game. Moreover, in the asymptotic case approximating the game G by a game with a continuous set of users, denote $\mathbf{x} \triangleq \{x_i, i \in \mathcal{C}\}$, we can write the potential function of the congestion game as follows:

$$P(\mathbf{x}) \triangleq \sum_{i\in\mathcal{C}} \int_{\epsilon_0}^{x_i N} \frac{\mu_i}{t} dt,$$

where $\epsilon_0 > 0$ is a small constant introduced to avoid the non-integral point of μ_i/t at 0. We can verify that for a SU j staying on channel i, it holds that :

$$\frac{\partial P(\mathbf{x})}{\partial x_i} = \mathbb{E}[U_j(\mu_i, x_i N)].$$

To derive the NE of G, we seek the maximum of the potential function $P(x)$. To this end, we develop $P(x)$ as

$$P(\mathbf{x}) = \sum_{i\in\mathcal{C}} \frac{\mu_i}{N}(\log x_i - \log \epsilon_0).$$

To find the maximum of $P(x)$, we solve the following optimization problem

$$\max_{\mathbf{x}} \; P(\mathbf{x}) \quad s.t. \quad \sum_{i\in\mathcal{C}} x_i = 1 \text{ and } x_i > 0, \forall i \in \mathcal{C},$$

which has a unique solution because the KKT conditions are necessary and sufficient ($P(\mathbf{x})$ is concave and the constraint is linear). After some straightforward algebraic operations, we can find the maximum $\mathbf{x}^* \triangleq \{x_i^*\}$ as follows:

$$x_i^* = \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l} \quad \forall i \in \mathcal{C}.$$

The maximum $\mathbf{x}^*$ is also the unique NE of G. ■

It can be noted that the above equilibrium derived in Theorem 1 is also a Wardrop equilibrium [5], of which we observe two desirable properties: (1) the NE is optimal from the system perspective as the total throughput of the network achieves its optimum at the NE; (2) the NE ensures that the spectrum resource is shared fairly among SUs. In the sequel analysis, we develop efficient spectrum access algorithms based on local imitation to converge to the unique NE in a distributed fashion without the *a priori* knowledge on μ.

III. Imitation-based spectrum access policy

In this section we develop the imitation-based policy for spectrum access in cognitive networks and study the resulting system dynamics as well as the equilibrium state. As a widely observed behavior rule, imitation [6] captures the behavior of a rational player that mimics the action(s) of other players with higher payoff in order to improve its own payoff. The induced imitation dynamic models the spreading of successful strategies under imitation. In the sequel analysis, we firstly rely on the proportional imitation rule (PIR) and develop the spectrum access policy based on PIR. We then extend our efforts to a more advanced imitation strategy called adjusted proportional imitation rule based on double sampling or double imitation (DI) rule that yields better system performance. We conclude the section by presenting the integrated imitation-based spectrum access policy .

A. Spectrum Access Based on Proportional Imitation

This subsection presents the spectrum access policy based on the proportional imitation rule [6]. As detailed in Algorithm 1, the core idea of the proposed spectrum access mechanism is as follows: at each iteration, each SU uniformly randomly selects another SU; if the payoff of the selected SU is higher than its own payoff, the SU imitates the strategy of the selected SU at the next iteration with a probability proportional to the payoff difference.

[2]Throughout this paper, we use i to refer to the channel index, k to refer to the time-slot index and j the index of the SUs.

Algorithm 1 PIR-based spectrum access policy: executed at each SU j for each iteration

Initialization: set the imitation rate σ and the imitation threshold ϵ_U
Randomly select a SU j'
if $U_j < U_{j'} - \epsilon_U$ **then**
  Switch to the channel of where j' stays with probability $p = \sigma(U_{j'} - U_j)$
end if

It is shown in [3] that the proportional imitation rule generates a population dynamic described by the following set of differential equations:

$$\dot{x}_i = \sigma x_i (V_i - \overline{V}) \quad i \in \mathcal{C}, \tag{1}$$

where V_i denotes the expected payoff of the SUs on channel i, $\overline{V} \triangleq \sum_{i\in\mathcal{C}} V_i$ denotes the expected payoff of all SUs in the network and σ is the imitation rate. One way of setting σ is to set $\sigma = 1/(\omega - \alpha)$, where ω and α are two exogenous parameters such that $U_j \in [\alpha, \omega]$ for each SU j. Injecting $V_i = \mu_i/(x_i N)$ into the differential equations, (1) becomes:

$$\frac{\dot{x}_i}{\sigma} = \frac{\mu_i}{N} - x_i \sum_{l\in\mathcal{C}} \frac{\mu_l}{N}.$$

This equation can be easily solved as:

$$x_i(t) = K e^{-(\sum_{l\in\mathcal{C}} \frac{\mu_l}{N})\sigma t} + \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l}, \tag{2}$$

where $K = x_i(0) - \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l}$.

As the major result of this subsection, the following theorem states the convergence of Algorithm 1 to the unique NE of the spectrum access game derived in Section II-B. The proof, of which the detail is omitted here, follows from (2) and Theorem 1.

Theorem 2. *The proposed spectrum access policy based on proportional imitation rule generates a dynamic that converges exponentially in time to the NE of the spectrum access game G.*

B. Spectrum Access Based on Double Imitation

In this subsection, we turn to a more advanced imitation rule, the proportional imitation rule with double sampling [7]. Under this imitation rule, each SU randomly samples two SUs and imitates them with a certain probability determined by the utility difference. The spectrum access policy based on the double imitation is detailed in Algorithm 2, in which each SUs randomly samples two other SUs j_1 and j_2 (without loss of generality, assume that j_1 and j_2 operate on channel i_1 and i_2 respectively, with corresponding utilities $U_{j_1} \le U_{j_2}$) and updates the probabilities of switching to channels i_1 and i_2, denoted as p_{j_1} and p_{j_2} respectively.

The double imitation rule generates an aggregate monotone dynamic [7], [8], which is defined as follows:

$$\dot{x}_i = \frac{x_i}{\omega - \alpha}\left[1 + \frac{\omega - \overline{V}}{\omega - \alpha}\right](V_i - \overline{V}) \quad \forall i \in \mathcal{C}$$

Injecting $V_i = \mu_i/(x_i N)$ into the differential equations, we have:

$$\dot{x}_i = \frac{\overline{V}}{\omega - \alpha}\left(1 + \frac{\omega - \overline{V}}{\omega - \alpha}\right) - \frac{\overline{V}}{\omega - \alpha}\left(1 + \frac{\omega - \overline{V}}{\omega - \alpha}\right) x_i,$$

Algorithm 2 DISAP: executed at each SU j settling on channel i for each iteration

Initialization: Let i be the channel on which SU j is settling.
Let U_j be the payoff of SU j on channel i.
Set the parameters ω, α and $\sigma = 1/(\omega - \alpha)$.
Define $[A]^+ \triangleq \max\{0, A\}$ and $Q(U) \triangleq 2 - \frac{U-\alpha}{\omega-\alpha}$.
Randomly sample two SUs j_1 and j_2.
Let i_1 (resp. i_2) be the channel on which SU j_1 (resp. j_2) is settling.
Let U_{j_1} (resp. U_{j_2}) be the payoff of SU j_1 (resp. j_2) on channel i_1 (resp. i_2).
Suppose without loss of generality that $U_{j_1} \le U_{j_2}$.
if $|\{i, i_1, i_2\}| = 1$, i.e., $i = i_1 = i_2$ **then**
  Stay on the same channel.
else if $|\{i, i_1, i_2\}| = 2$ **then**
  if $i = i_1$, $i \ne i_2$ and $U_j \le U_{j_2}$ **then**
    $p_{j_2} = \frac{\sigma}{2} Q(U_j)(U_{j_2} - U_j)$.
    Switch to channel i_2 w.p. p_{j_2} and stay on the same channel w.p. $1 - p_{j_2}$.
  else if $i_1 = i_2$, $i \ne i_1$ and $U_j \le U_{j_1} = U_{j_2}$ **then**
    $p_{j_1} = \frac{\sigma}{2}(Q(U_{j_1}) + Q(U_j))(U_{j_1} - U_j)$.
    Switch to channel i_1 w.p. p_{j_1} and stay on the same channel w.p. $1 - p_{j_1}$.
  end if
else if $|\{i, i_1, i_2\}| = 3$ **then**
  if $U_j \le U_{j_1} \le U_{j_2}$ **then**
    $p_{j_1} = \frac{\sigma}{2}\left[Q(U_j)(U_{j_1} - U_{j_2}) + Q(U_{j_2})(U_{j_1} - U_j)\right]^+$.
    $p_{j_2} = \frac{\sigma}{2}\left[Q(U_{j_1})(U_{j_2} - U_j) + Q(U_{j_2})(U_{j_1} - U_j)\right] - p_{j_1}$.
    Switch to channel i_1 w.p. p_{j_1}, to channel i_2 w.p. p_{j_2} and stay on the same channel i w.p. $1 - p_{j_1} - p_{j_2}$.
  else if $U_{j_1} \le U_j \le U_{j_2}$ **then**
    $p_{j_2} = \frac{\sigma}{2}\left[Q(U_{j_1})(U_{j_2} - U_j) + Q(U_{j_2})(U_{j_1} - U_j)\right]^+$.
    Switch to channel i_2 w.p. p_{j_2} and stay on the same channel w.p. $1 - p_{j_2}$.
  end if
else
  Stay on the same channel.
end if

whose solution is

$$x_i(t) = K e^{-\frac{\overline{V}}{\omega-\alpha}\left(1 + \frac{\omega - \overline{V}}{\omega-\alpha}\right)t} + \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l}, \tag{3}$$

where $\overline{V} = \sum_{l\in\mathcal{C}} \mu_l / N$ and $K = x_i(0) - \frac{\mu_i}{\sum_{l\in\mathcal{C}} \mu_l}$.

The following theorem stating the major result in this subsection follows immediately.

Theorem 3. *The proposed spectrum access policy based on double imitation genrates a dynamic that converges exponentially in time to the NE of the spectrum access game G.*

Algorithm 3 Integrated imitation-based spectrum access policy (ISAP): executed as each SU

Initialization: Set ϵ_t
At each iteration t
  With probability $1 - \epsilon_t$ perform imitation (PIR or DI)
  With probability ϵ_t switch to a random channel

C. Imitation with Exploration

In this subsection, we integrate the exploration into the proposed spectrum access policy which allows each SU to exploit all possible channels. Integrating the exploration can avoid the case where some channels have no SUs operating on them, making the imitation fail to work. Algorithm 3 details the integrated imitation-based spectrum access policy. At each iteration t, each SU exploits a new channel with a small probability ϵ_t and performs imitation with probability $1-\epsilon_t$.

IV. Performance Evaluation

In this section we conduct extensive simulations to evaluate the performance of the proposed imitation-based spectrum access policies (ISAP).

A. Simulation Setting

We simulate a cognitive radio network of $N = 50$ SU and $C = 3$ channels, whose mean availabilities (assumed constant) are expressed by the vector $\mu = [0.3, 0.5, 0.8]$. Both the proportional imitation and the double imitation rules applied in Algorithm 3 are investigated. In all simulations we set the horizon (simulation time) equal to $3{\cdot}10^4$ time slots and the learning guide variable ϵ equal to $\max\{\epsilon_{min}, 1-erf(\frac{b\cdot t}{horizon})\}$, where erf is the error function, b is a constant dependent on the horizon, and ϵ_{min} is a very small value (e.g. 10^{-4}) used in simulations to avoid dead-lock situations.

B. Learning and Convergence

Fig. 1 and Fig. 2 display the results of ISAP using PIR and DI, respectively. In particular the convergence trends in terms of number of SUs per channel and average utility per user per channel are depicted. Note that the small variation of the trajectories in both figures from the converged curve is due to the probabilistic nature of the SUs' strategy and has only very limited impact on the system as a whole. The noise during the early phase is caused by the fact that SUs gradually learn the channel availabilities. We notice that in both cases convergence is rapidly achieved after the learning phase and all SUs' payoff stabilizes at a converged value, which shows that the proposed spectrum access policies ensure fairness among SUs. Consequently, the channels with higher availabilities are chosen by more individuals. This can easily be verified (Fig. 1 for PIR and Fig. 2 for DI) that after convergence the major part of population settles permanently in channel 3, i.e. the channel with highest availability.

C. Convergence Speed

We now focus on convergence speed. Unfortunately, it is difficult to observe it on simulation results because of the noise introduced by the learning process. We thereby concentrate upon the replicators/aggregate monotone dynamics equations (2) and (3). Fig. 4 shows the number of SUs on channel 3 as a function of time. Here it's evident how aggregate monotone dynamic (DI) outperforms replicators dynamic (PIR) in convergence speed. By observing the 100^{th} time slot on the x-axis for instance, one can notice that aggregate monotone dynamics have come to a complete convergence for both the considered values of ω, while replicators are still embroiled in the convergence phase.

D. Switching Cost

We now turn to the analysis of the switching cost, i.e., the global number of channel switches. Due to the drastic cost of changing frequencies in current wireless devices in terms of delay, packet loss and protocol overhead, an efficient channel access policy should avoid frequently channel switching, unless necessarily. Fig. 5 shows the switching cost as a function of time for PIR and DI-based spectrum access policies and for different values of ω (α is set to 0). We observe that PIR outperforms DI in terms of switching cost for a given ω.

Moreover, we report that in both replicator and aggregate monotone dynamics the underlying dynamics of PIR and DI-based spectrum access policies, the switching cost depends on ω. More specifically, increasing ω reduces the trend of the switching cost on the long term (in the stability region). On the other hand, increasing ω also decreases the convergence speed as illustrated in Fig. 4. This indeed impacts the trajectories generated by the system of differential equations describing the dynamics, but has no effect on its final outcome (for a reasonable large t). Despite this trend, DI still remains faster than PIR for a fixed ω. Hence, ω should be carefully tuned to strike a balance between switching cost and convergence rate.

V. Conclusion and Further Work

In this paper, we have analyzed the problem of opportunistic spectrum access from the evolutionary game theoretic angle (a tool widely applied in biology and economy but rarely investigated systematically in the wireless networking field), more specifically, using imitation-based rules. We proposed two imitation-based spectrum access policies based on the proportional imitation rule (PIR) and the adjusted proportional imitation rule with double sampling or double imitation (DI), both of which are proved to converge exponentially to the Nash Equilibrium which is also the system optimum.

Following the first step of studying imitation-based spectrum access policies presented in this paper, we plan to explore the more practical scenario where SUs are limited to imitate only cognitive radios transmitting on the same channel. At the current stage, we have developed an adapted algorithm whose convergence is demonstrated by simulations. Our research plan is to derive the dynamic of the adapted algorithm, to show analytically its convergence.

References

[1] S. Haykin. Cognitive Radio: Brain-Empowered Wireless Communications. *IEEE J. on Selected Areas in Communications*, 23(2):201–220, 2005.

[2] M. Buddhikot. Understanding Dynamic Spectrum Access: Models, Taxonomy and Challenges. In *Proc. IEEE DySPAN*, Apr. 2007.

[3] J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *Bull. Amer. Math. Soc.*, 2003.

[4] I. Milchtaich. Congestion Games with Player-Specific Payoff Functions. *Games and Economic Behavior*, 13, 1996.

[5] J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(36), 1952.

[6] K. H. Schlag. Why Imitate, and if so, How ? Discussion paper, University of Bonn, Feb. 1996.

[7] K. H. Schlag. Which One Should I Imitate ? *Elsevier J. of Mathematical Economics*, 31, 1999.

[8] L. Sammuelson and J. Zhang. Evolutionary Stability in Asymmetric Games. *J. of Economic Theory*, 57, 1992.

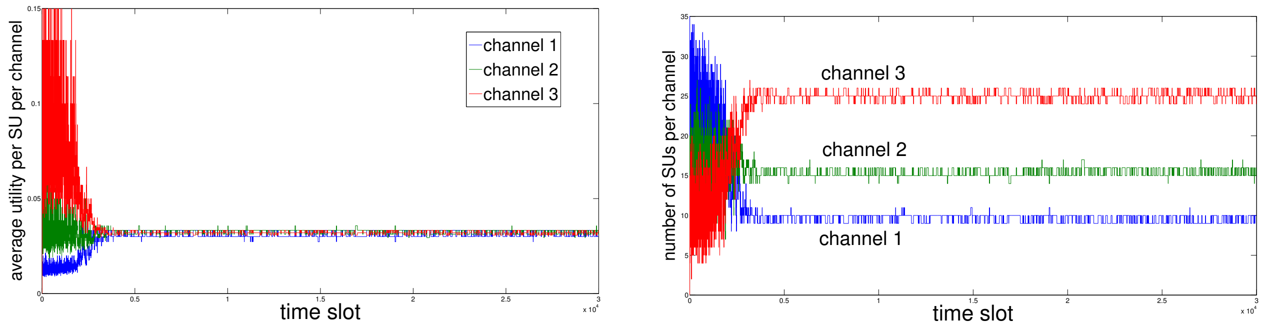

Fig. 1. PIR-ISAP: average per SU utility for different channels (left) and SUs distribution (right) as a function of time

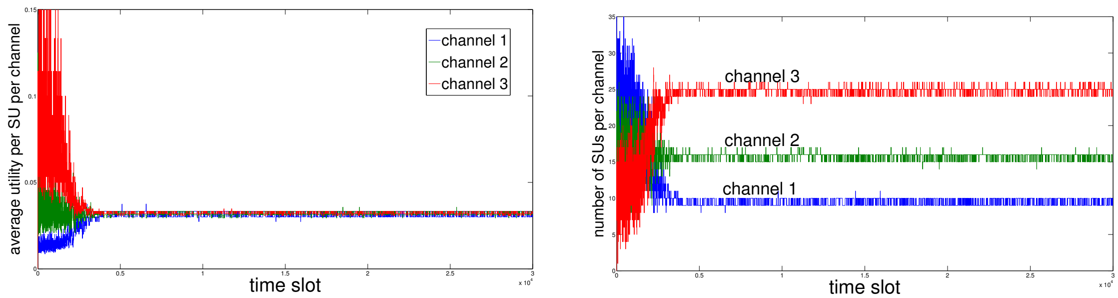

Fig. 2. DI-ISAP: average per SU utility for different channels (left) and SUs distribution (right) as a function of time

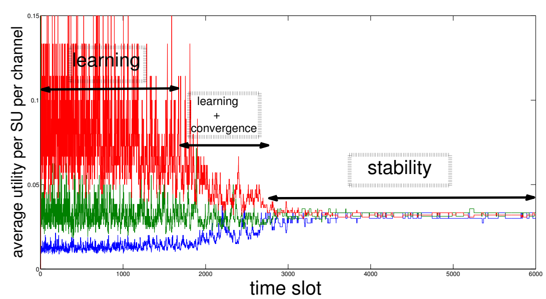

Fig. 3. Three main phases of DI-ISAP over $7 \cdot 10^3$ slots

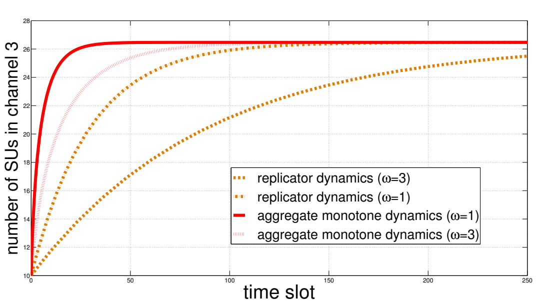

Fig. 4. Convergence delay comparison

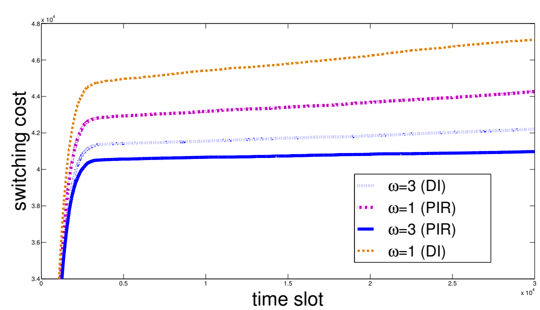

Fig. 5. Switching cost comparison