On-Line Cognitive Radio Learning: from Theory to Practice

Christophe MOY
Centrale Supélec
Rennes Campus
IETR – UMR CNRS 6164

University of Gran Canaria, Las Palmas, Spain
25/05/2016

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

Current and future steps

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

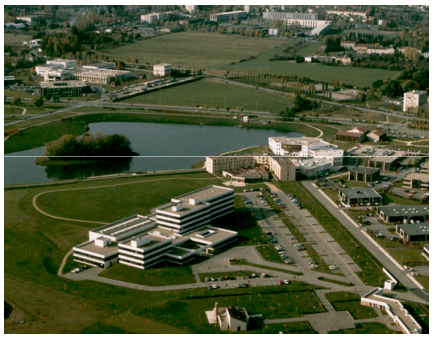
Current and future steps

Centrale Supélec - CS

January 2015: merging of

- Supélec – since 1894 (3 campus)
- Ecole Centrale Paris since 1826

Rennes Campus (west)



Metz Campus (east)

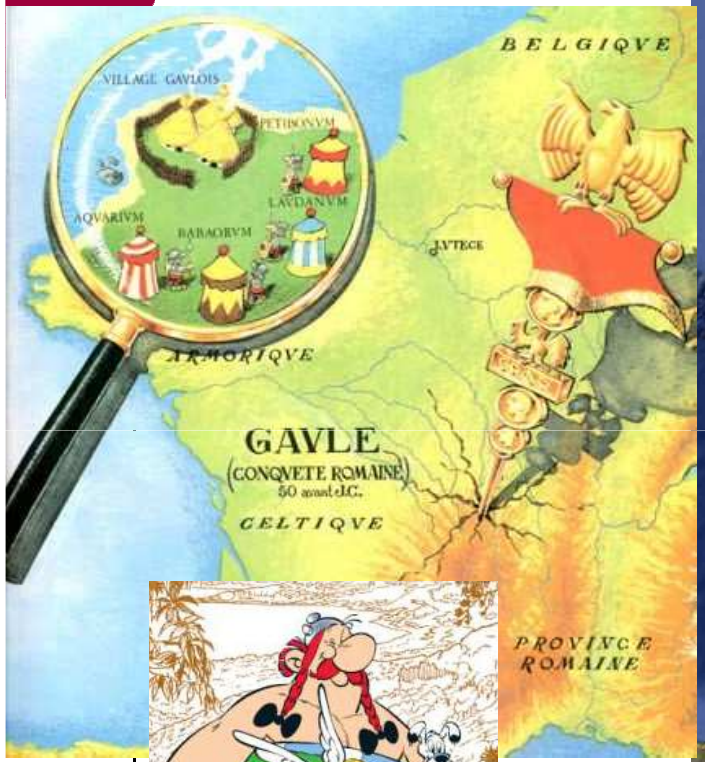
SCEE team is located in Rennes campus

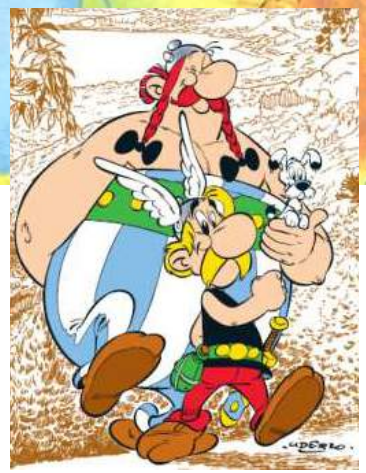
Campus de Châtenay Malabry

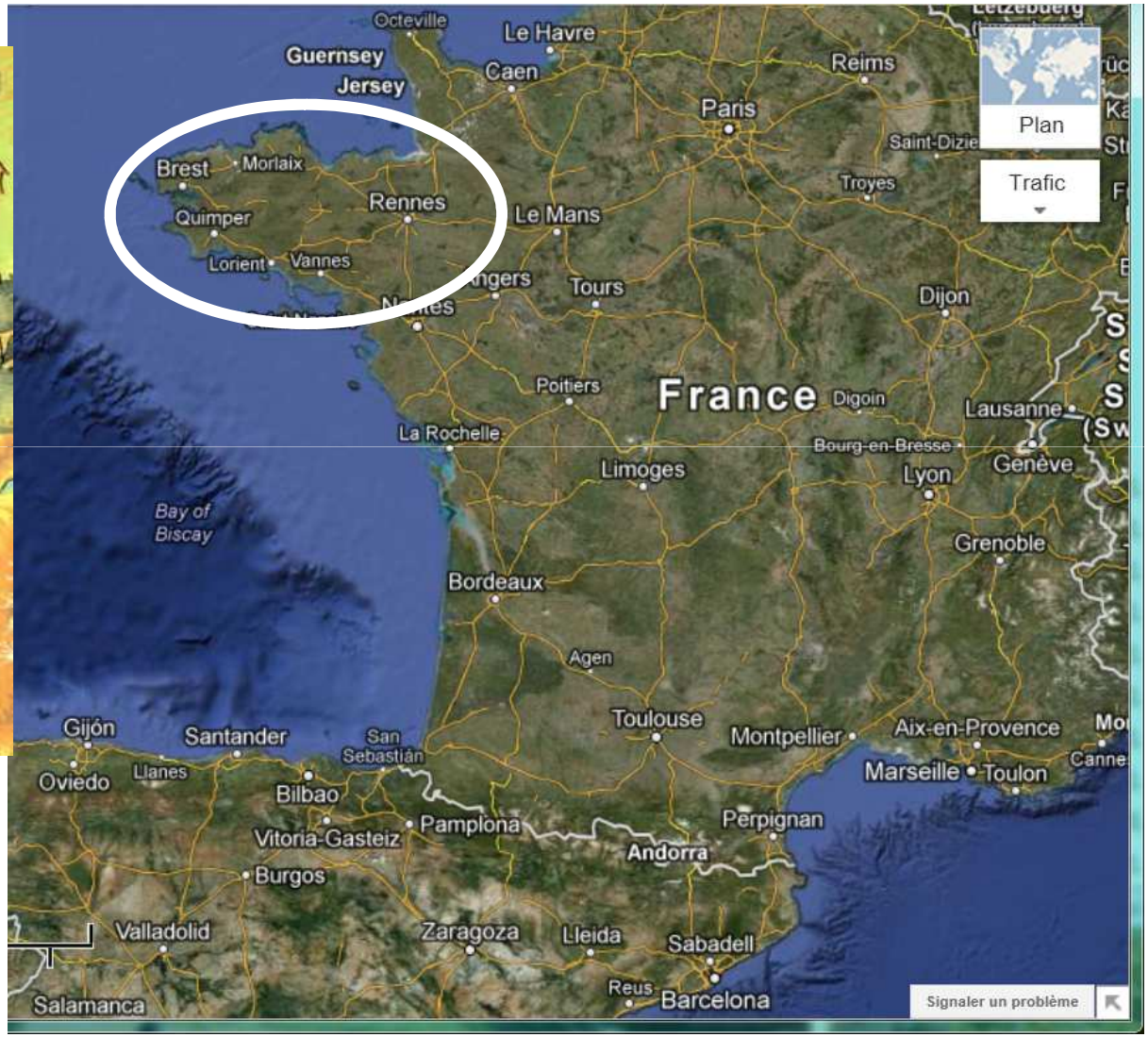
➔

Gif sur Yvette Campus (near Paris)
Future Paris Saclay Campus

Rennes y Bretagna







IETR – Institute of Electronics and Telecommunications of Rennes

IETR members

- CNRS
- Université de Rennes 1
- INSA de Rennes
- Centrale Supélec
- Université de Nantes

IETR Departements

- Antennas et microwave devices
- Communications ← SCEE team
- Image and Automatics
- Propagation, Localisation and Teledetection
- Microelectronics and Microsensors

SCEE faculty

SCEE : Signal, Communications et Electronique Embarquée – permanents :


Prof. Jacques Palicot (**responsable d'équipe**)


Prof. Yves Louet


Prof. Christophe Moy


Prof. Carlos Faouzi Bader


Prof. Amor Nafkha


Prof. Pierre Leray


Prof. Jacques Weiss,


Prof. Gilles Tourneur


Prof. Pascal COTRET

SCEE in years 2014-2015:

- Publications in Journals: ~15
  major conferences: ~70
- 12 PhD students + 3 post-docs
- National and International Projects involvement ≈ 13
- Attracted funding ≈ 1M€


Prof. Honggang Zhang

- **Invited International Expert**
- Zheijang University (Hangzhou,China)
- International Chair Professor, UeB - CominLabs
- Period: 2012-2013
- Main focus: ''Green and Cognitive Radio''

SCEE team research areas

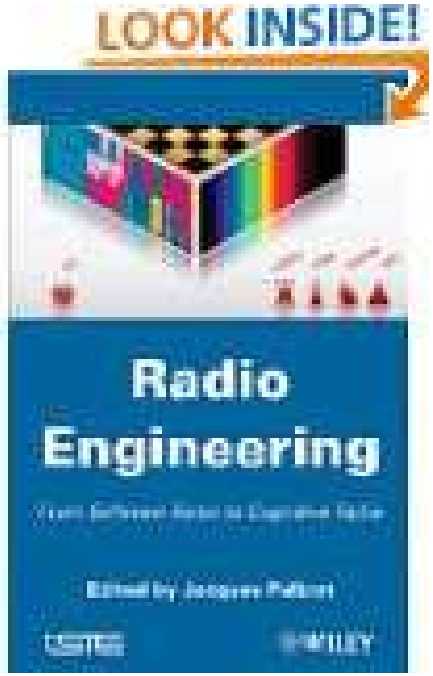
Software Radio / Cognitive Radio / Green Radio

Main applications contexst

- **Digital communications and signal processing (5 PhD)**
  - Non-linearities and PAPR, Blind MIMO demodulation, synchronisation,…
  - Equalization and estimation (blind, semi-blind,...)
  - Green communications and Cognitive Radio
  - Flexible multicarrier waveforms for future wireless networks

- **SDR/SDN and Cognitive Radio architectures (4 PhD)**
  - Reconfiguration and cognitive management for heterogeneous architecture (HDReM, HDCRAM) and systems
  - Learning for cognitive radio equipments for OSA and green criteria
  - SDR/SDN design (operator approach, graph optimization, MDA)
  - Power-efficient communications and electronics

- **Sensors for Cognitive Radio (3 PhD)**
  - Sensorial radio bubble
  - Multilayers sensors
    - » Spectrum holes (white spectrum) detection
    - » Blind Standard Recognition Sensor (BSRS)



Cognitive Systems and 5G Dynamic Future Wireless Networks, Sensor network, Internet of Things, Smart Grid

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

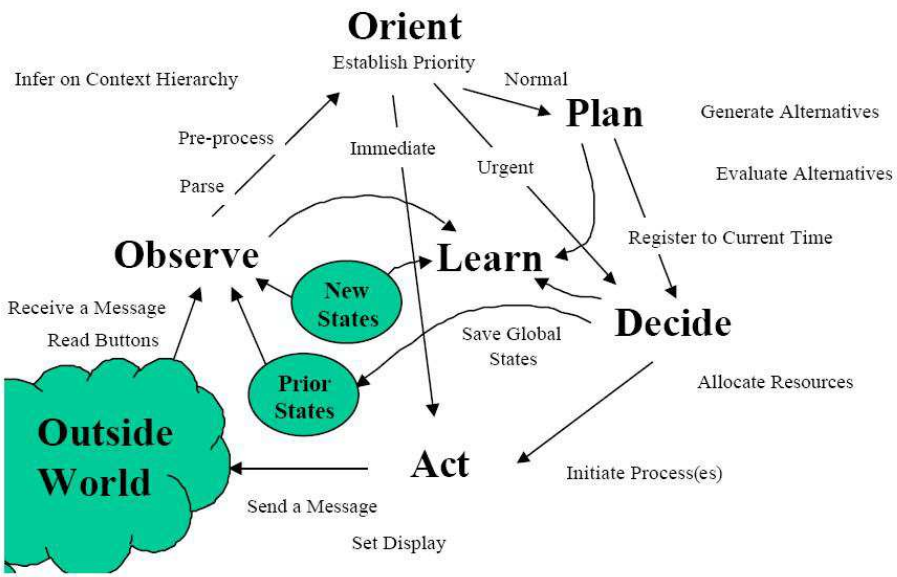
UCB for HF communications

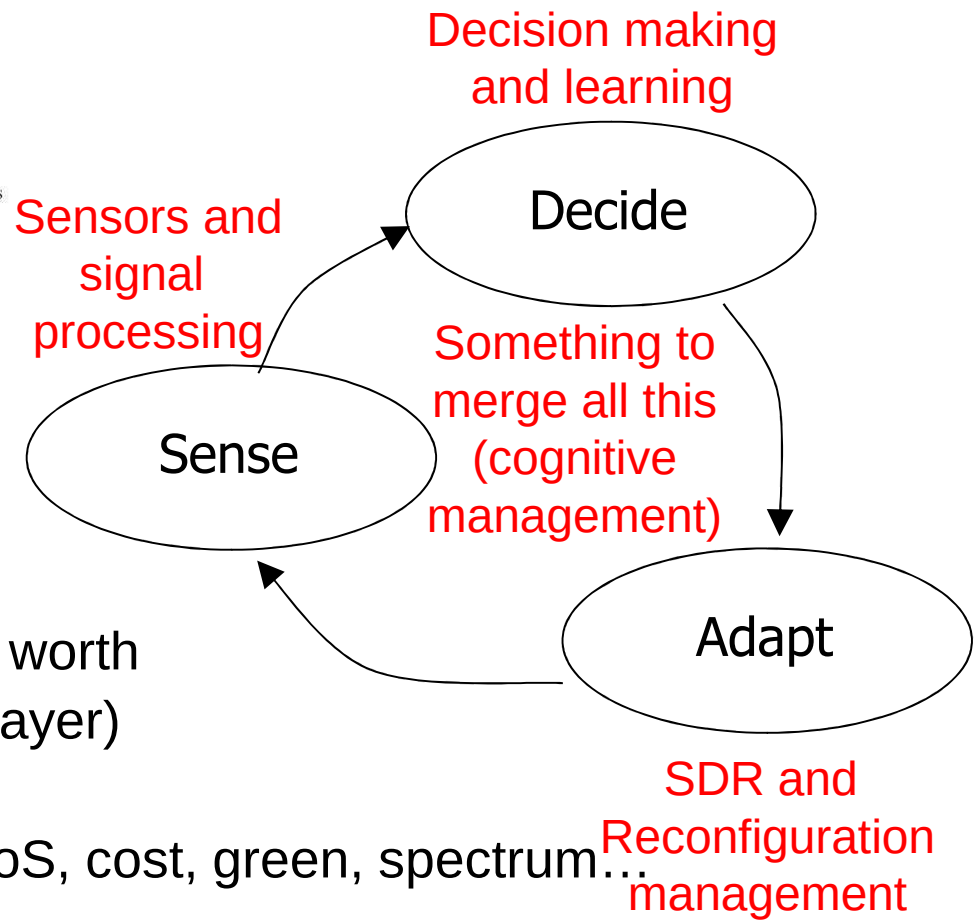
Current and future steps

Cognitive cycle

[1] Joe MITOLA, Cognitive radio: An integrated agent architecture for software defined radio. PhD Thesis, Royal Inst of Technology (KTH) (2000)

[2] Loïg GODARD, Christophe MOY, Jacques PALICOT, "From a Configuration Management to a Cognitive Radio Management of SDR Systems", CrownCom'06, 8-10 June 2006, Mykonos, Greece

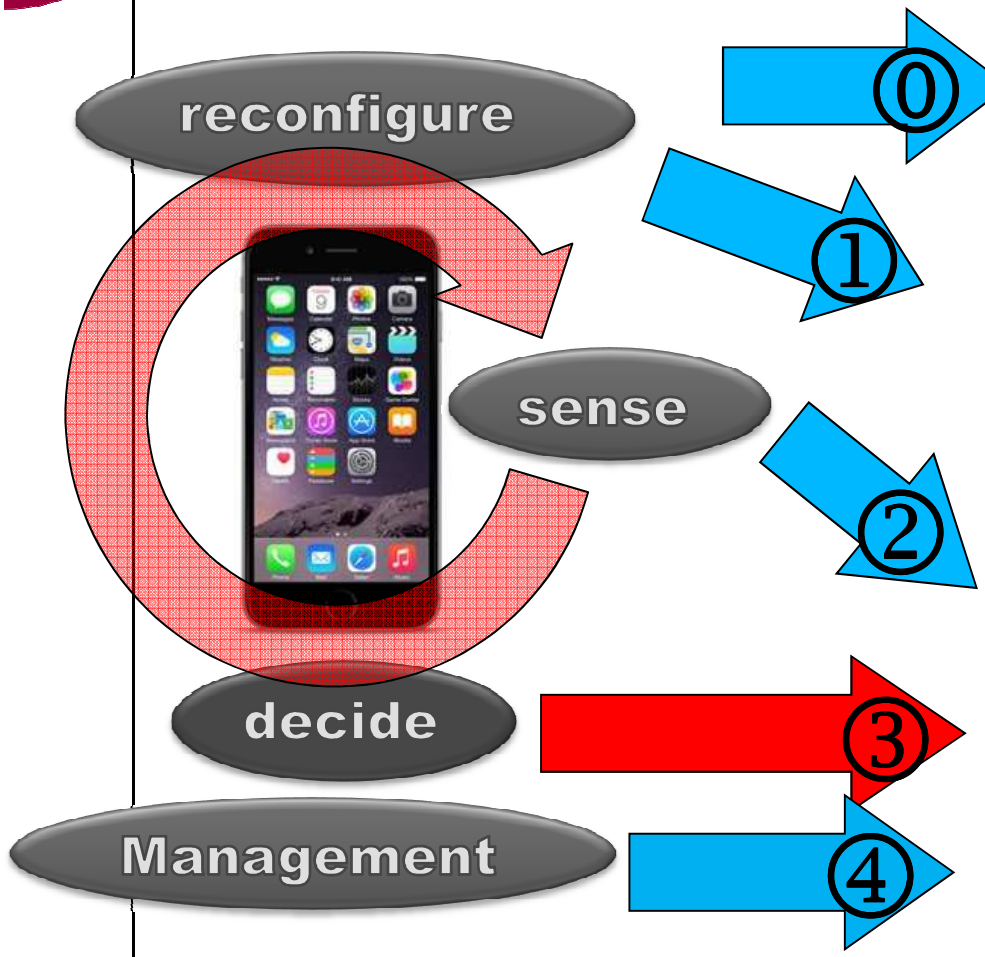




- any captured information is worth
- at any layer (not only PHY layer)
- not only spectrum-oriented
- any criteria can be used: QoS, cost, green, spectrum…

Cognitive radio equipment

A cognitive radio equipment is made of:



- New and legacy waveforms
  - radio for PHY layer
  - also any other layer processing
- Flexible platform and processing
  - multi-processing
  - heterogeneous (DSP, FPGA…)
  - adaptive signal processing
- Sensing signal processing
- Learning & decision making processing
- Management architecture for reconfiguration and cognition

③ Decision making

Context of high uncertainty – no a priori knowledge

- re-inforcement learning (on-line trials success/fail)
- MAB bandit approach for decision making & learning – Upper Confidence Bound UCB
- very simple implementation

➔ first theoretical studies in year 2008

UCB for dynamic configuration adaptation (DCA)

[3] Wassim JOUINI, Damien ERNST, Christophe MOY, Jacques PALICOT, "Multi-Armed Bandit Based Policies for Cognitive Radio's Decision Making Issues", Signal Circuits and Systems Conference, Jerba, Tunisia, 6-8, Nov. 09

UCB for opportunistic spectrum access (OSA)

[4] Wassim JOUINI, Damien ERNST, Christophe MOY, Jacques PALICOT, "Upper confidence bound based decision making strategies and dynamic spectrum access," International Communication Conference, ICC'10, Cape Town, South Africa, 26-29 May 2010

Cognitive radio for green radio communications

Green cognitive radio

- CO_2 emission mitigation
- health – human exposure decrease
- pollution – spectrum savings

[5] Jacques PALICOT, "Cognitive Radio: An Enabling Technology for the Green Radio Communications Concept," IWCMC'09, Leipzig, Germany June 2009

VLSI design and power consumption for CR digital front-end

[6] Navin MICHAEL, Christophe MOY, Prasad VINOD, Jacques PALICOT, “Area-Power Trade-offs for Flexible Filtering in Green Radios”, *Journal of Communications and Networks*, Vol. 12, 2, 30 April 2010

Smart Grid

[7] Jacques PALICOT, Christophe MOY, Benoit RESIMONT, Rémi BONNEFOI, "Application of Hierarchical and Distributed Cognitive Architecture Management for the Smart Grid", *Elsevier Ad-Hoc Networks, 2015*

Internet of Things

[8] Vincent SAVAUX, Yves LOUET, Christophe MOY, Apostolos KOUNTOURIS " Sub-sampling of channels with time and frequency sparsity access ", *submitted to IEEE World Forum on IoT, Reston, USA Dec. 2016*

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

Current and future steps

Cognitive Radio and Dynamic Spectrum Access

Cognitive Radio is often *reduced* to Dynamic Spectrum Access (DSA)

Answer to solve spectrum scarcity

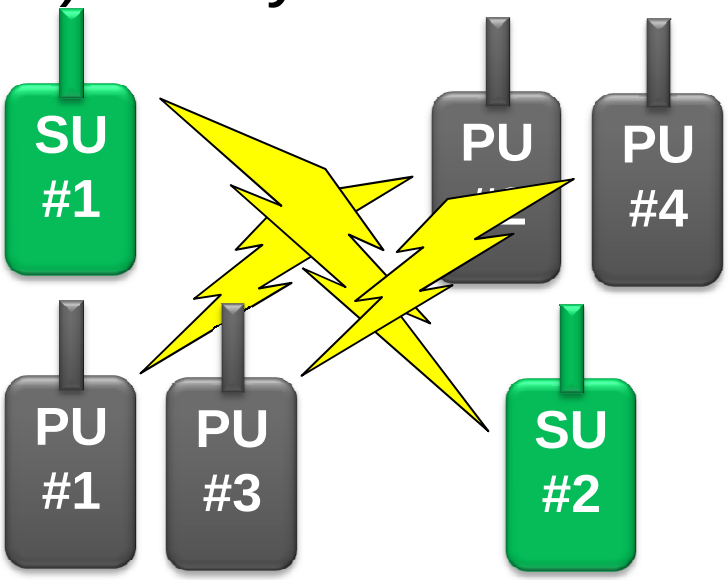
DSA:

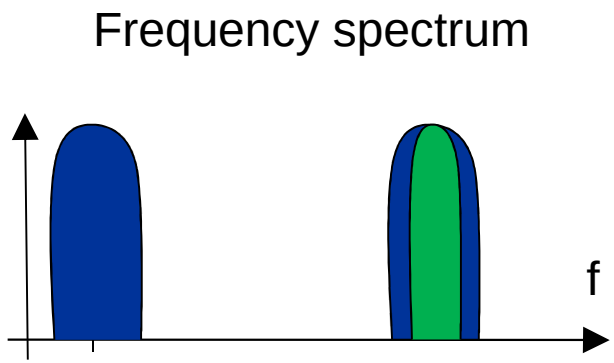
- Regulation: conservative approaches (static)
  - TV White Spaces
  - Licence Sharing Access (LSA)
    - ETSI RRS : LSA in 2.3-2.4 GHz forLTE
    - US SAS in 3.5 GHz band (3-tier)
- Research: dynamic approaches
  - Opportunistic Spectrum Access (OSA)

Opportunistic Spectrum Access (OSA)

Recover dynamicaly (on-the-fly) spectrum opportunities for partially used spectrum

Secondary Users (SUs) are allowed to use in time and space spectrum let vacant by Primary Users (PUs) if they do not interfere with PUs ➔ « LBT »





Added-complexity due to cognition

Cognitive radio equipement on N channels

Either have a N times larger RF to

- sense all N channels in parallel (« LBT »)
- know which channel(s) is (are) vacant at each iteration

➔complexity overhead for RF & digital processing

Or keep RF bandwidth of 1 channel as legacy

- sense 1 channel (light overhead) at a time (iteration)
- need to re-build overall channels knowledge
- to predict which channel(s) to be vacant at next iteration

Re-inforcement learning

A CR equipment (terminal or system) should be able to learn

- from scratch (no a priori knowledge)
- channels unused by PUs in licenced bands

Reinforcement learning: try and regret/reward

- First solution
  - learn then exploit
  - but loose (time or wages) when learning
- Second solution
  - mix learning and exploitation
  - example of multi-armed bandit MAB (machine learning domain)

MAB model for OSA issue

Las Vegas: which armed bandit to select for the best reward (in terms of money)?

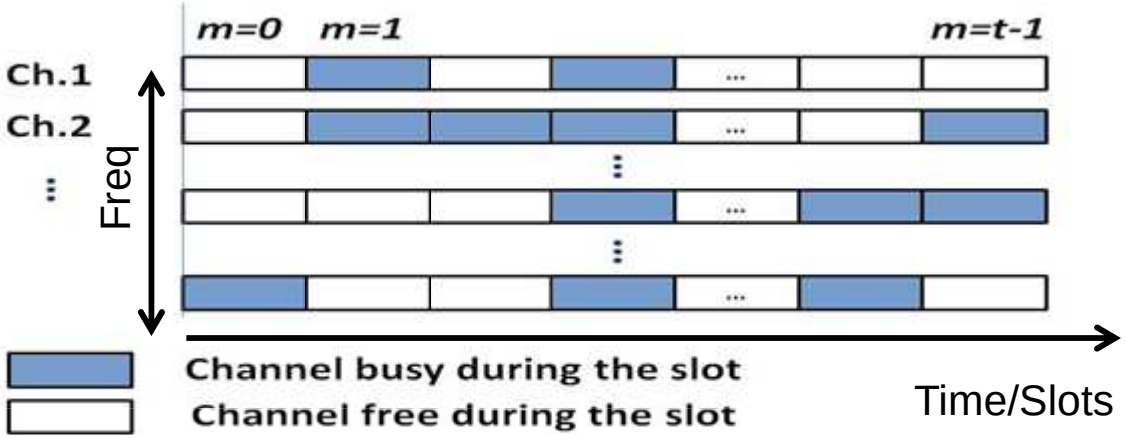
Multi-Armed Bandit (MAB) issue

- K possible choices $k \in [1,K]$ (machines)
- Each choice has a mean performance μ_k
- At each time t, the gambler plays a machine and obtains a reward r_t
- Objective: not to find μ_k , but maximise the cumulative reward (his gain) or minimise regret

[9] Lai, T.L. and Robbins, H. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6:4-22, 1985

Iterative learning process

PUs channels occupancy model



Iterative process for SU, slot by slot

Selects the channel with the « highest vacancy probability »

Sensing	Transmission ?	Reward
Detects presence or absence of a PU in this channel	Communication if possible (when there is no PU, e.g. channel is free)	Performance evaluation

Selection of channel that maximizes reward

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

Current and future steps

Naive approach

Let's consider a <u>uniform</u> test *{1-vacant/0-occupied}* of each channel at each iteration during an exploration phase (round robin)

We obtain an approximative value of the empirical mean of vacancy $\overline{X_k}$ for all channels

The probability to choose a sub-optimal channel after the exploration phase is not null

- either because exploration not long enough
- bad scenari may exist statistically

➔ linear regret in t ("regret" means "compared to the best")

UCB: a solution to MAB problems

Bernouilli *{1-vacant/0-occupied}* empirical mean on each channel

What if $\overline{X}_k \neq \mu_k$?

$$\overline{X}_{k,T_k(t)} = \frac{\sum_{m=0}^{t-1} r_m \mathbf{1}_{\{a_m=k\}}}{T_k(t)}$$

$$\mathbf{1}_{\{a_m=k\}} = \begin{cases} 1 & \text{if } a_m = k \\ 0 & \text{if } a_m \neq k \end{cases}$$

A bias is added in function of the number of activations of each channel

Upper Confidence Bound (UCB_1) index

- $$B_{k,t,T_k(t)} = \overline{X}_{k,T_k(t)} + A_{k,t,T_k(t)}$$
- for each arm k, in function of time t

  $\overline{X}$: empirical mean of the reward

  A: bias (avoid forgetting good solutions)

  T_k: number of times config. k has been chosen

Simplicity: just a sum!
(for each arm)

Regret criteria of convergence

Regret is a off-line efficiency measure of algos (just an evaluation criteria, not used on-line)

Regret is what has been lost compared to the choice of the best case at each t

$$R_t = t.\mu^* - \sum_{m=0}^{t-1} r_m$$

μ^*: reward of the best case

$$E[R_t] = \sum_{k=1}^{K} (\mu^* - \mu_k).E[T_k(t)]$$

Cumulative regret

- under certain conditions, it can be proven that asymptotically UCB converges to the best solution (depending on the policy)

$$\lim_{t\to\infty} \frac{E[R_t]}{t} = 0$$

$$\lim_{t\to\infty} \frac{\sum_{m=0}^{t-1} r_m}{t} = \mu^*$$

Bias functions

Literature proposes several bias functions

- UCB_1 policy

$$A_{k,t,T_k(t)} = \sqrt{\frac{\alpha.\ln(t)}{T_k(t)}}$$

- Also exist UCB_V policy

$$A_{k,t,T_k(t)} = \sqrt{\frac{2.\xi.V_k(t).\ln(t)}{T_k(t)}} + \frac{3.c.\xi.\ln(t)}{T_k(t)}$$

[10] P. Auer, N. Cesa-Bianchi, P. Fischer, "Finite time analysis of multiarmed bandit problems", Machine learning, 47(2/3):235–256, 2002

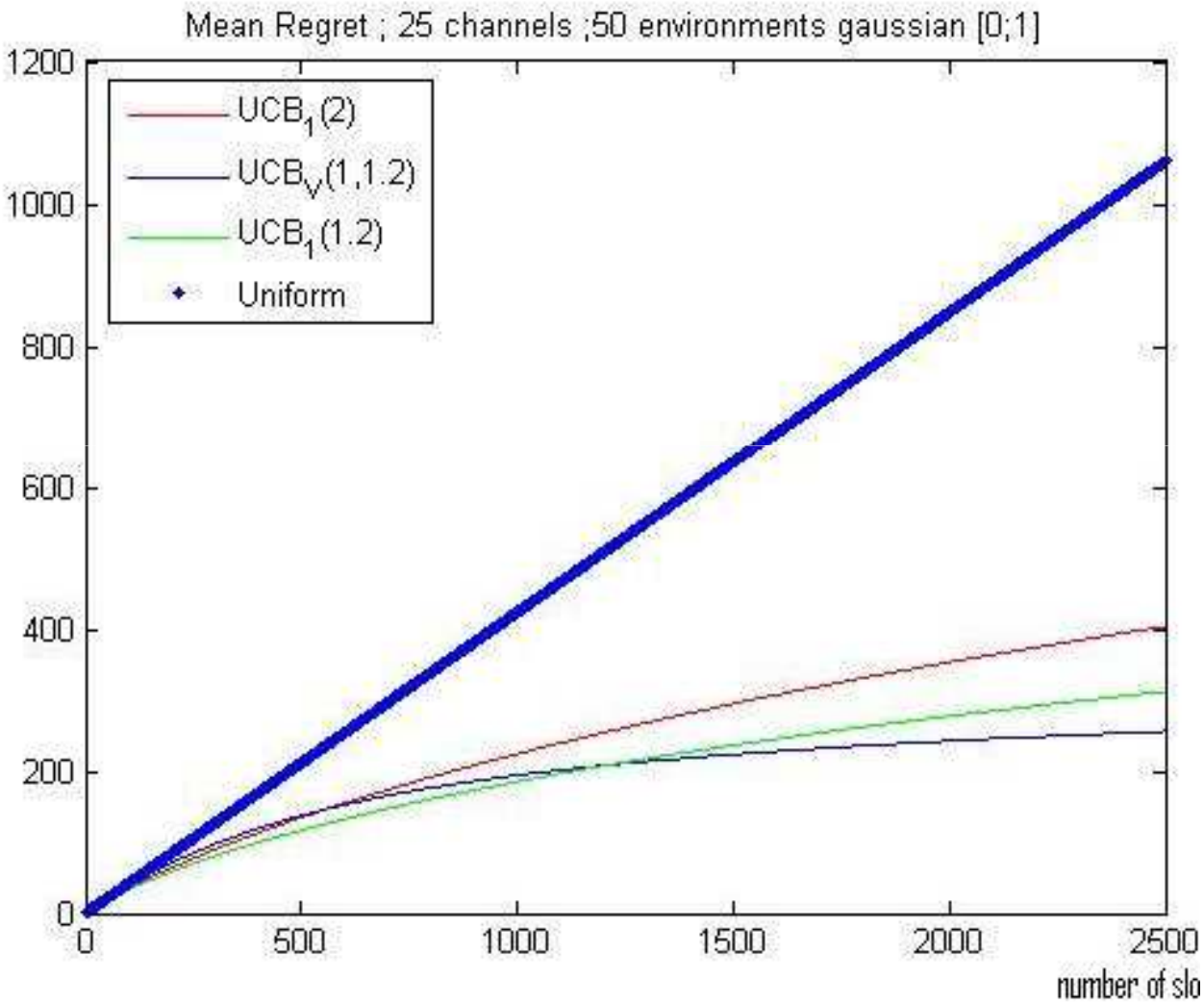
[11] J.Y. Audibert, R. Munos, C. Szepesvari, "Tuning bandit algorithms in stochastic environments", International Conference on Algorithmic Learning Theory, 2007.

Evaluation criteria for machine learning: regret and best choice %

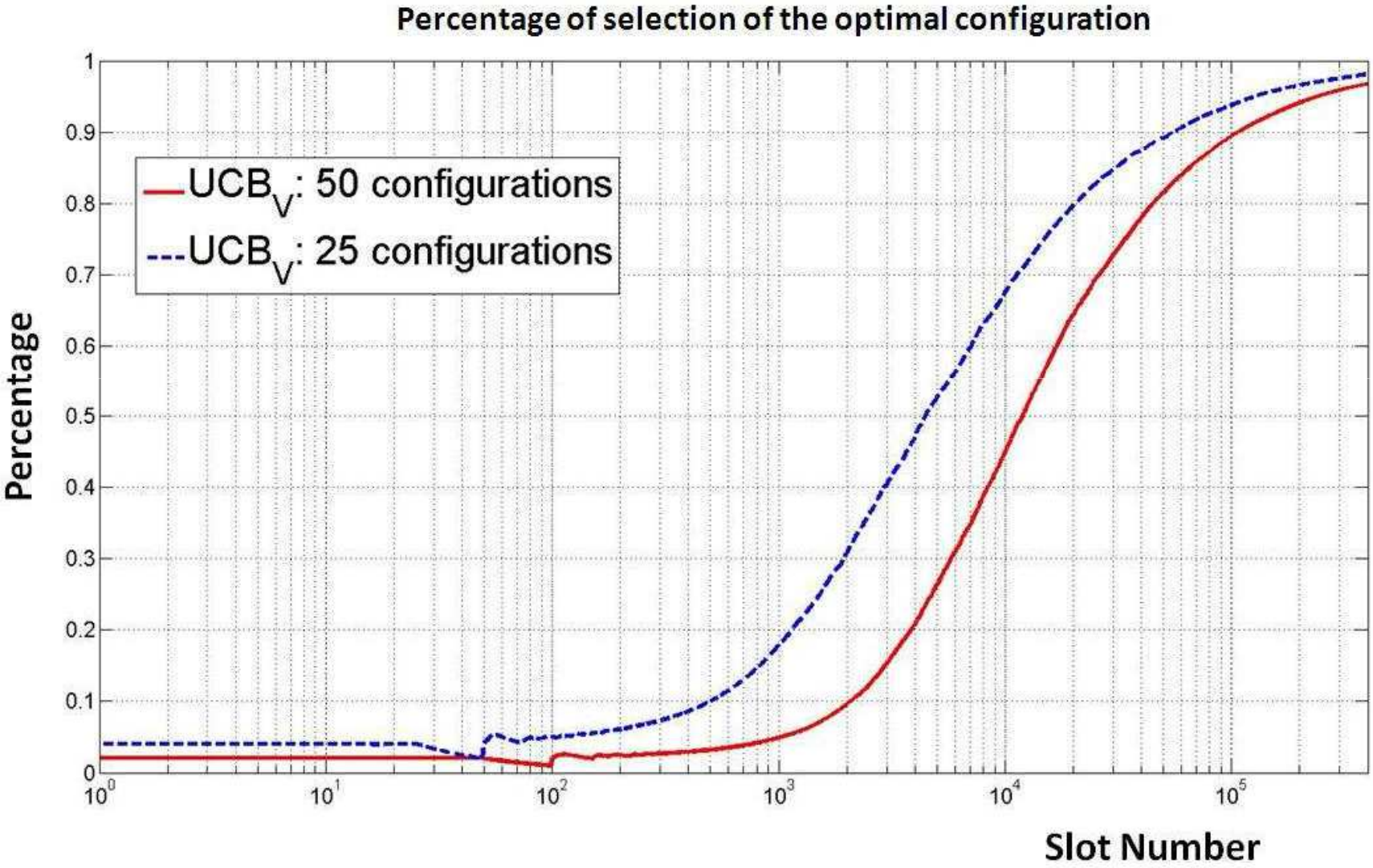
Regret is a machine learning view to evaluate how much it works

Percentage of best solution selection over time is another

BUT nothing to do with telecoms (you can not directly convert it in bits/s)



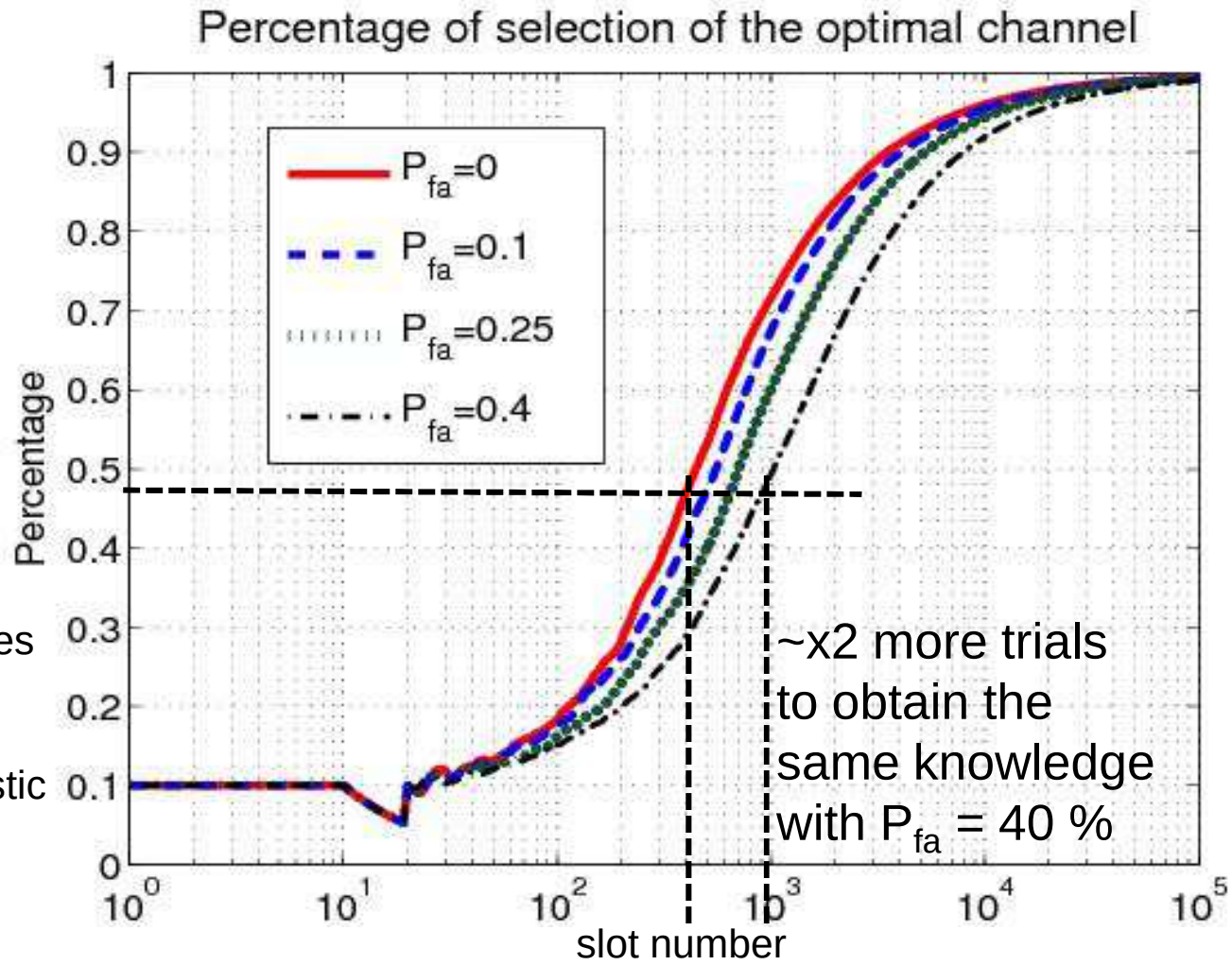
Percentage of selection of the optimal choice – perfect sensing



UCB robustness to learning errors
Percentage of choice of best channel

10 channels

for several false alarm probabilites

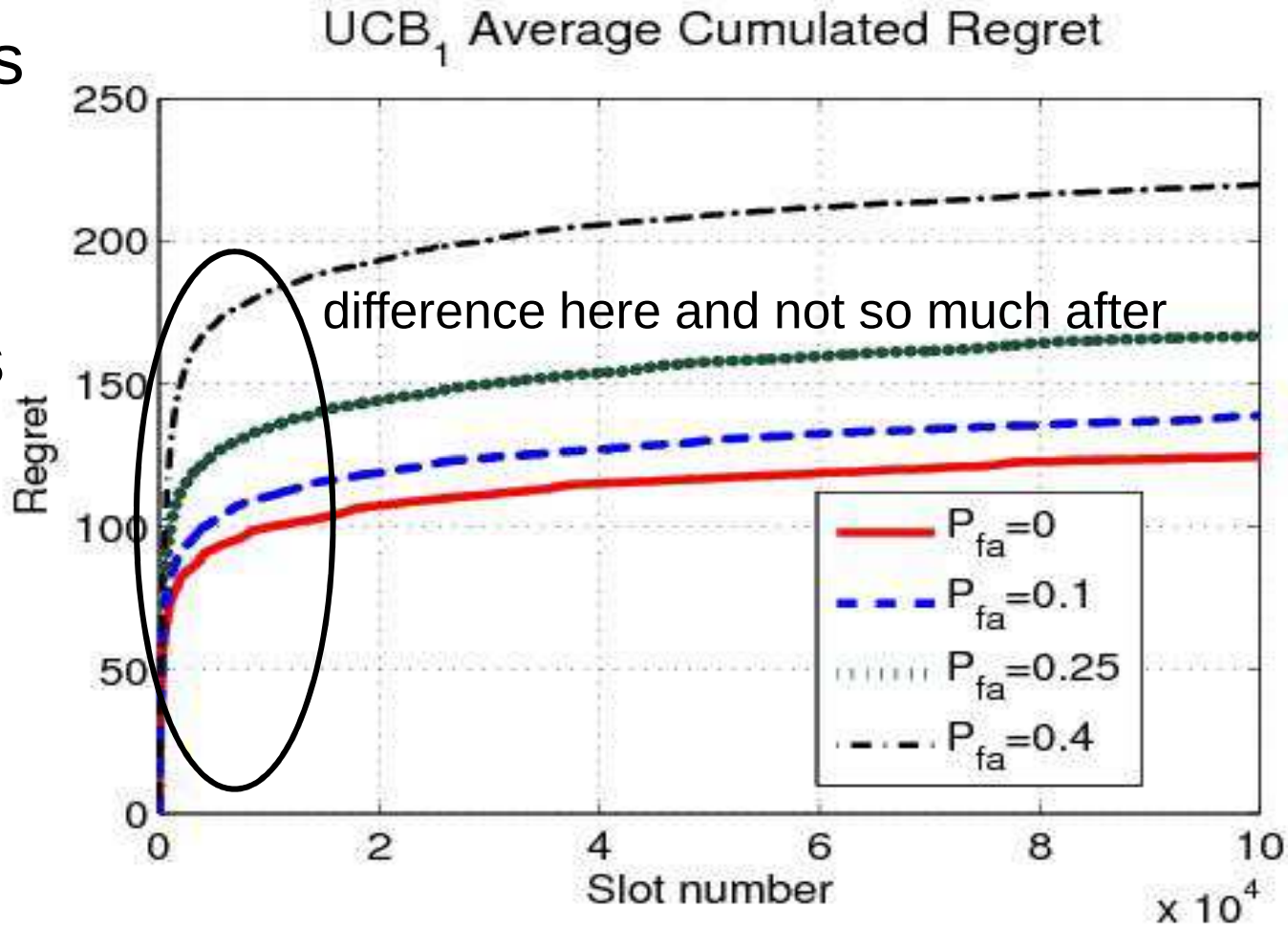


[12] Wassim JOUINI, Christophe MOY, Jacques PALICOT, "Upper Confidence Bound Algorithm for Opportunistic Spectrum Access with Sensing Errors", CrownCom'11, 1-3 June 2011, Osaka, Japan

UCB robustness to learning errors
Regret analysis

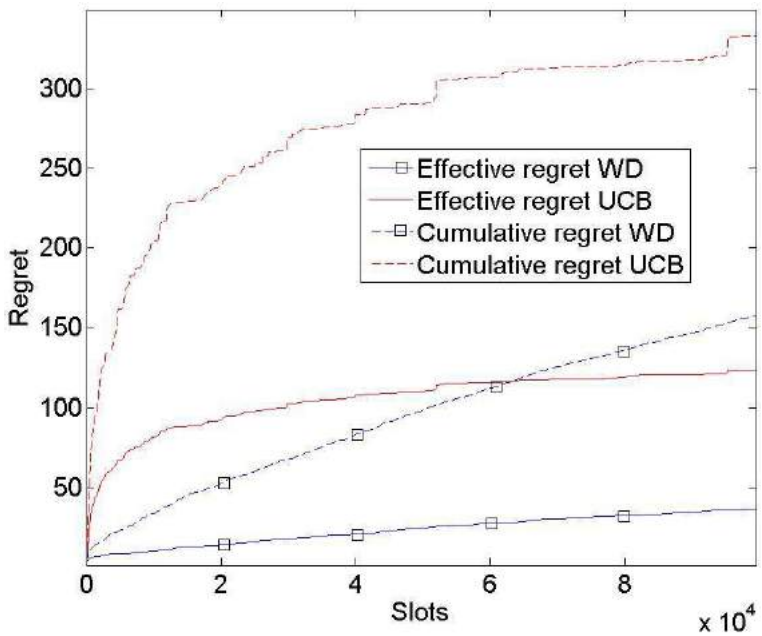
10 channels

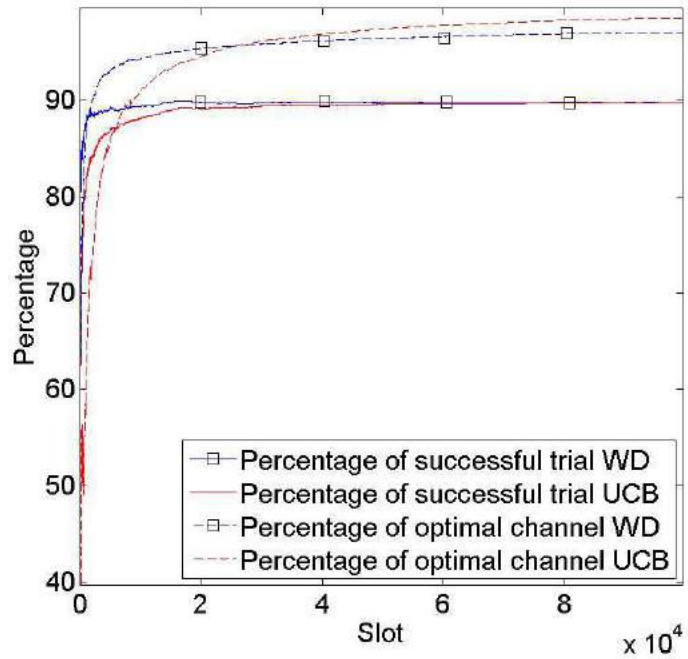
for several false alarm probabilites



New criteria for learning evaluation

Regret and percentage of best solution selection are pure Machine Learning points of view





[13] Clément ROBERT, Christophe MOY, Cheng-Xiang WANG, "Reinforcement Learning Approaches and Evaluation Criteria for Opportunistic Spectrum Access", *IEEE International Conference on Communications, ICC'14, Sydney, Australia, 10-14 June 2014*

Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

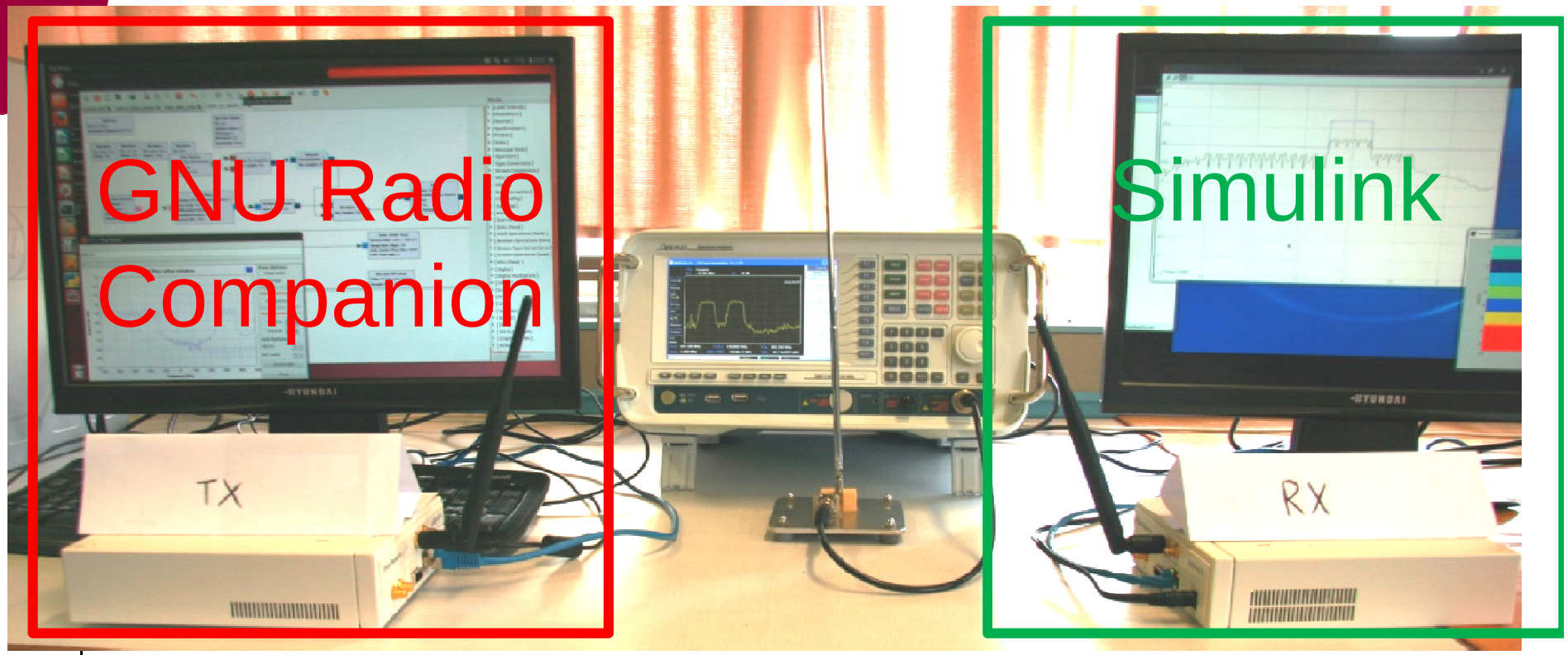
Current and future steps

Real implementation – does it confirm theoretical promises?

What if UCB is implemented on real signals?

- convergence speed
  - is it too long or acceptable with radio usage?
- implementation complexity
  - is it so simple compared to other cognitive functions?
- convergence guarantee
  - are there other real factors that make UCB diverge in real world?
- primary network stationarity
  - what happens when primary users change their behavior?
- game theory
  - are multiple users competing
- centralized/decentrlized
  - should we use resources to exchange knowledge

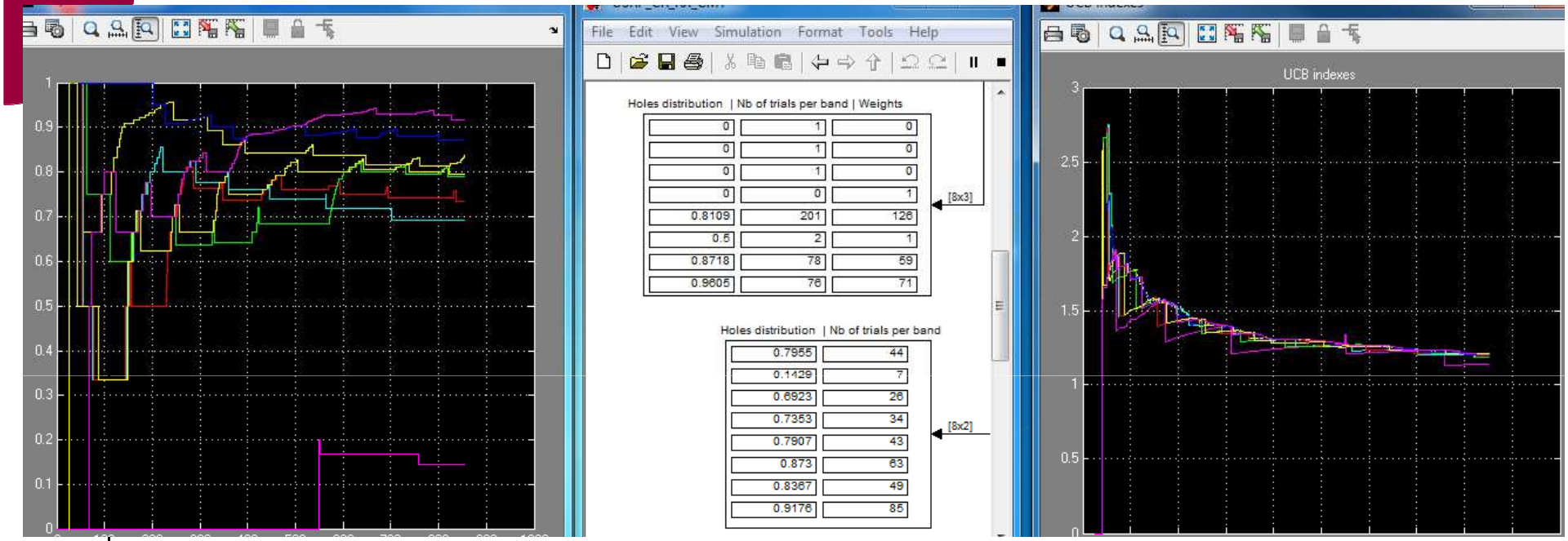
Proof-of-Concept Demonstrator



LEFT: primary network transmission (TX)
RIGHT: one secondary user learning algorithm, implementing an energy detector as a sensor (RX).

[14] Christophe MOY, "Reinforcement Learning Real Experiments for Opportunistic Spectrum Access" , *Karlsruhe Workshop on Software Radios, WSR'14, Germany, March 2014*

Proof-of-Concept
Results display



Evolution during first 350 iterations of an experiment

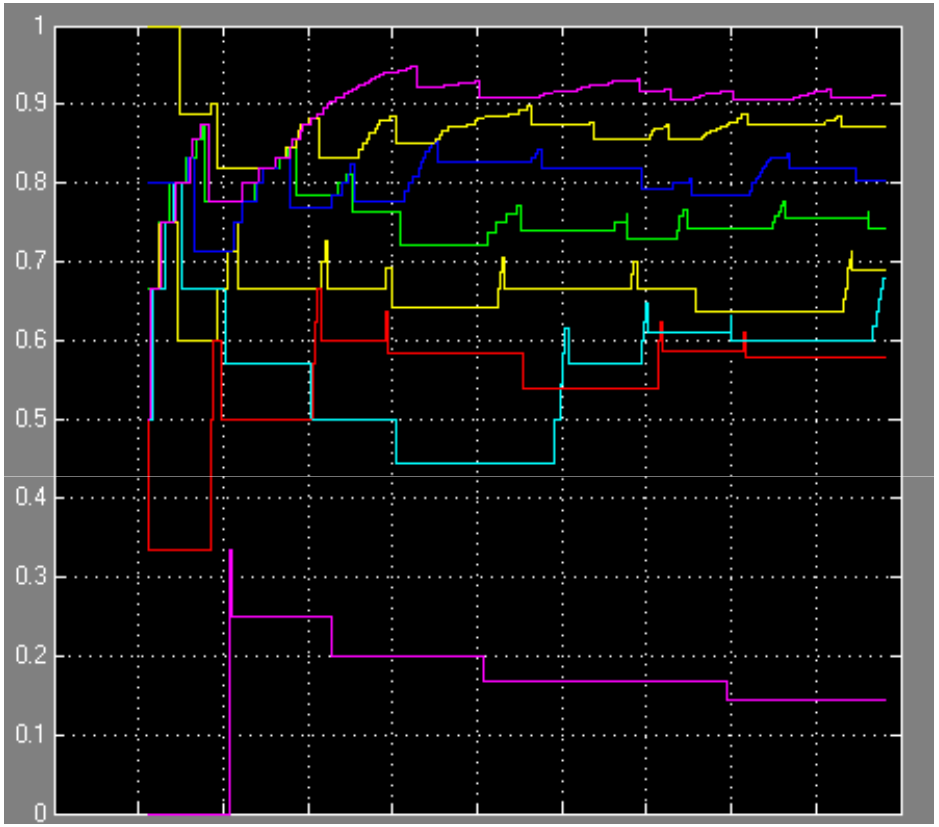
Left hand side: empirical average vacancy rate $\overline{X}_{k,T_k(t)}$ of the 8 channels derived from UCB.

- Right hand side: UCB indexes $B_{k,T_k(t)}$ evolution.
- Bottom middle table: UCB results
- Top middle table: comparison with another algo (WD) results

Proof-of-Concept
Empirical probability of vacancy

Derived by UCB algorithm during 350 iterations:

- #8 violet 0.91 0.9
- #7 yellow 0.87 0.8
- #6 blue 0.80 0.7
- #5 green 0.74 0.6
- #4 red 0.58 0.5
- #3 lht bl. 0.68 0.4
- #2 purple 0.14 0.3
- #1 light yell. 0.80 0.5



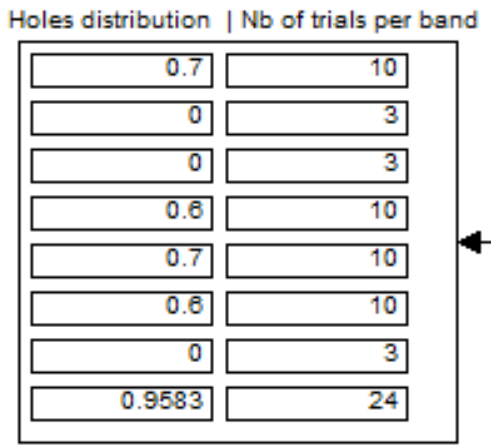
- **Very fast and good estimation and convergence**
- **Can be obtained in ms**

Proof-of-Concept
UCB evolution after 73 iterations

After 73 iterations

- uniform search would have tried 10 times each
- UCB really favors best channel (not at infinity)

	emp.prob	nb trials	target prob
channel#1	0.70	10	0.5
channel#2	0	3	0.3
channel#3	0	3	0.4
channel#4	0.6	10	0.5
channel#5	0.70	10	0.6
channel#6	0.60	10	0.7
channel#7	0	3	0.8
channel#8	**0.96**	**24**	**0.9**



Proof-of-Concept
UCB evolution after 1500 iterations

After 1500 iterations

- best 3 channels selected 80 % of time

Holes distribution	Nb of trials per band
0.7273	110
0.3158	19
0.4483	29
0.5581	43
0.7956	181
0.8444	315
0.8442	308
0.8949	628

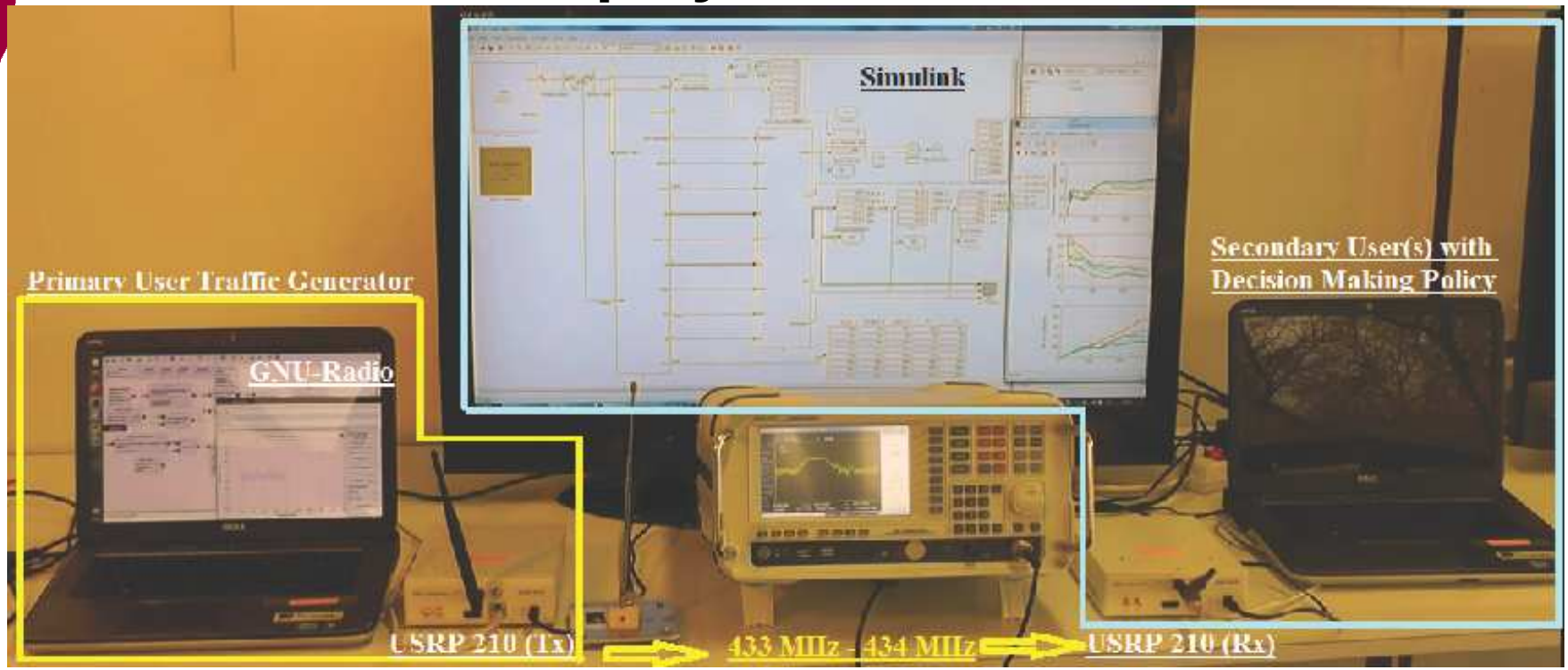
Proof-of-Concept
UCB evolution after 7000 iterations

After 7000 iterations

- best channel selected more than 50%
- as 90% probability of vacancy for this channel
- for the two best channels: up to 75%

Holes distribution	Nb of trials per band
0.7032	219
0.2692	26
0.5352	71
0.573	89
0.7864	557
0.8022	718
0.8414	1532
0.8718	3783

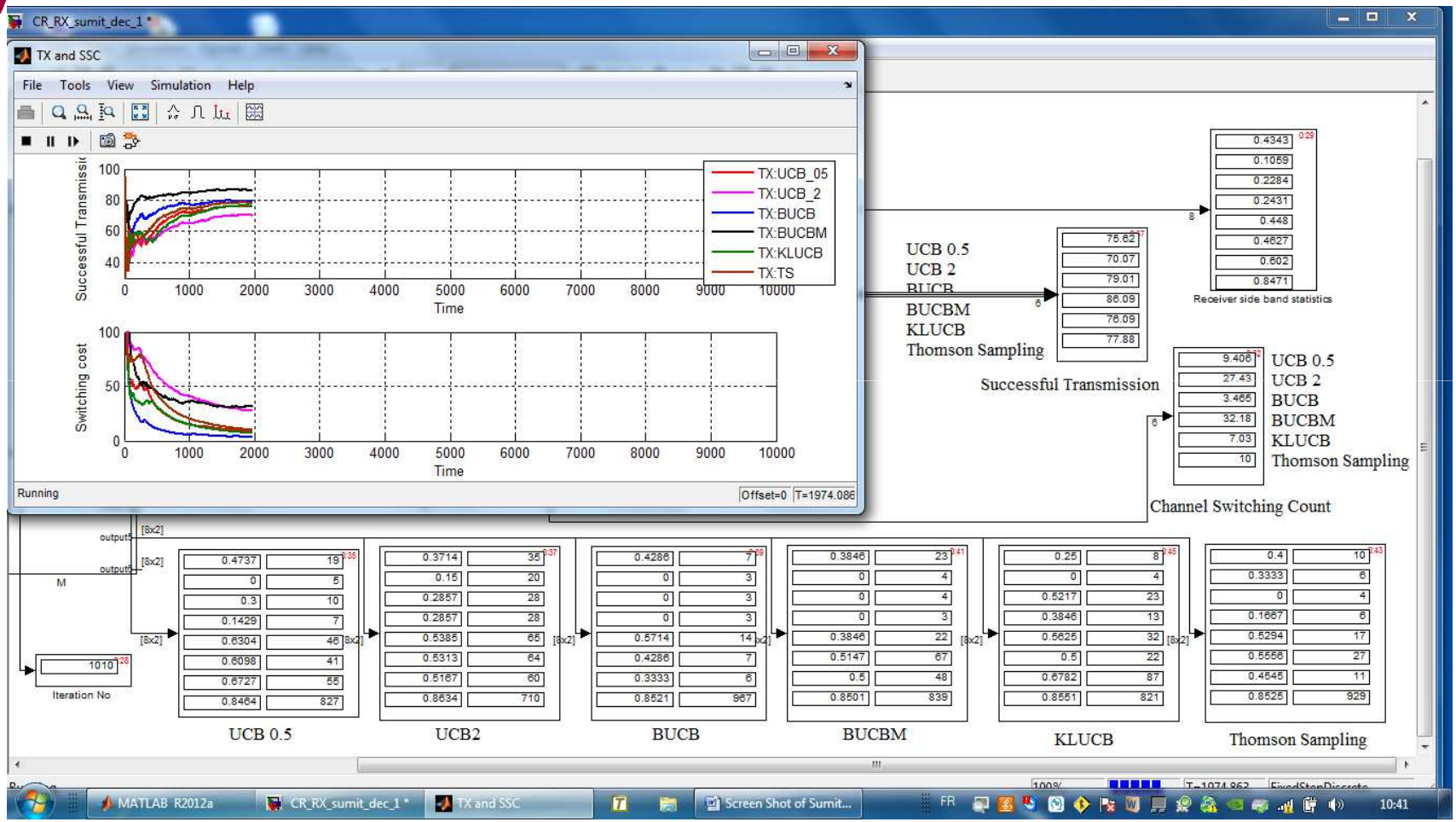
New demo – new algos and multi-player demo



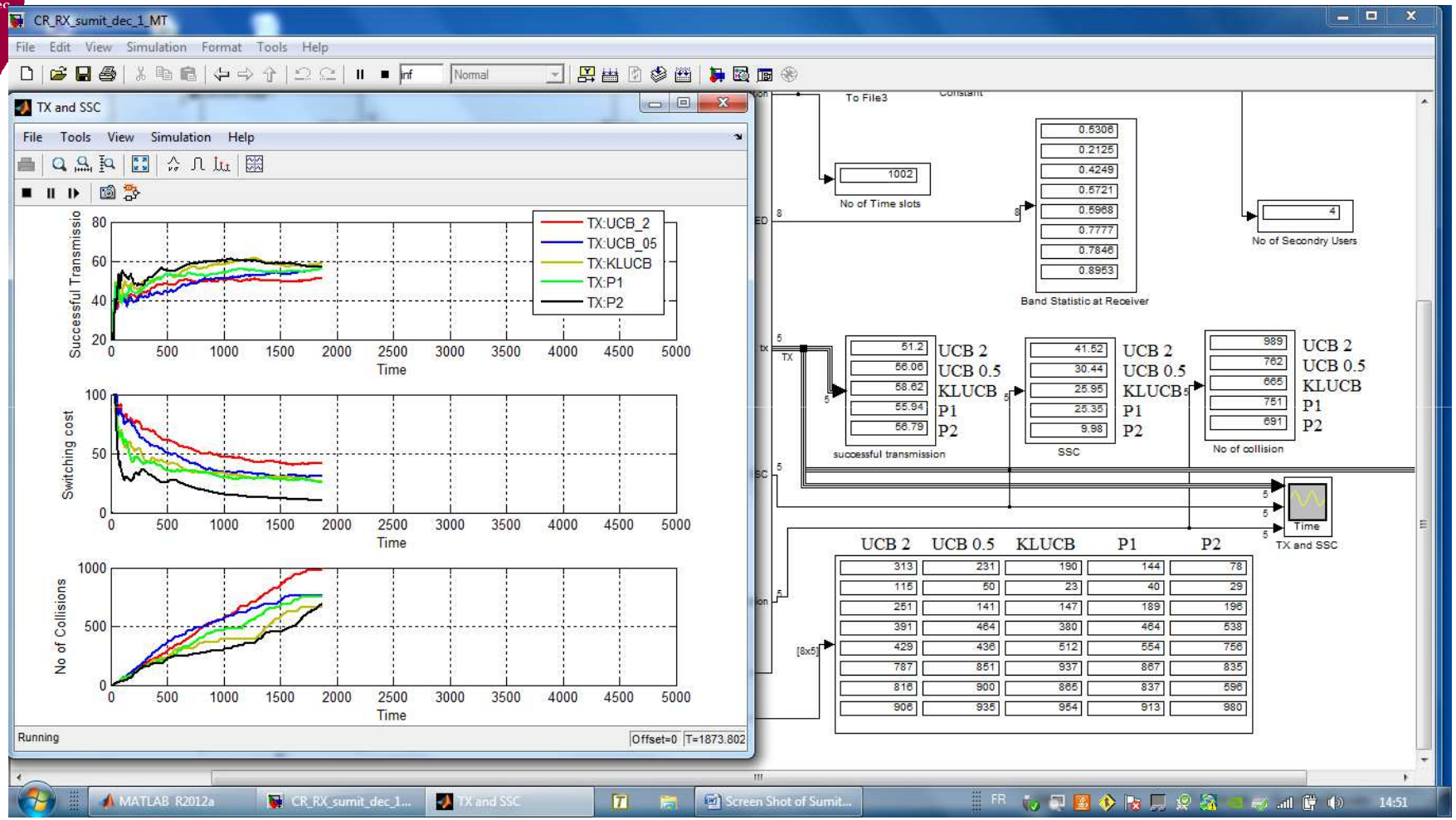
[15] Sumit DARAK, Navikkumar MODI, Amor NAFKHA, Christophe MOY, "Spectrum Utilization and Reconfiguration Cost Comparison of Various Decision Making Policies for Opportunistic Spectrum Access Using Real Radio Signals" , *CrownCom'16, 30 May 2016*

[16] Sumit DARAK, Amor NAFKHA, Christophe MOY, Jacques PALICOT, "Is Bayesian Multi-armed Bandit Algorithm Superior?: Proof-of-Concept for Opportunistic Spectrum Access in Decentralized Networks" *, CrownCom'16, 30 May, Grenoble, France*

Other algorithms



Multi-player games



Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

Current and future steps

Cooperation with IDeTIC

University of Las Palmas, Gran Canaria

Often complain: yes but there is no channel which is perfectly i.i.d. or following a Markovian process

What about facing this approach to a real channel?

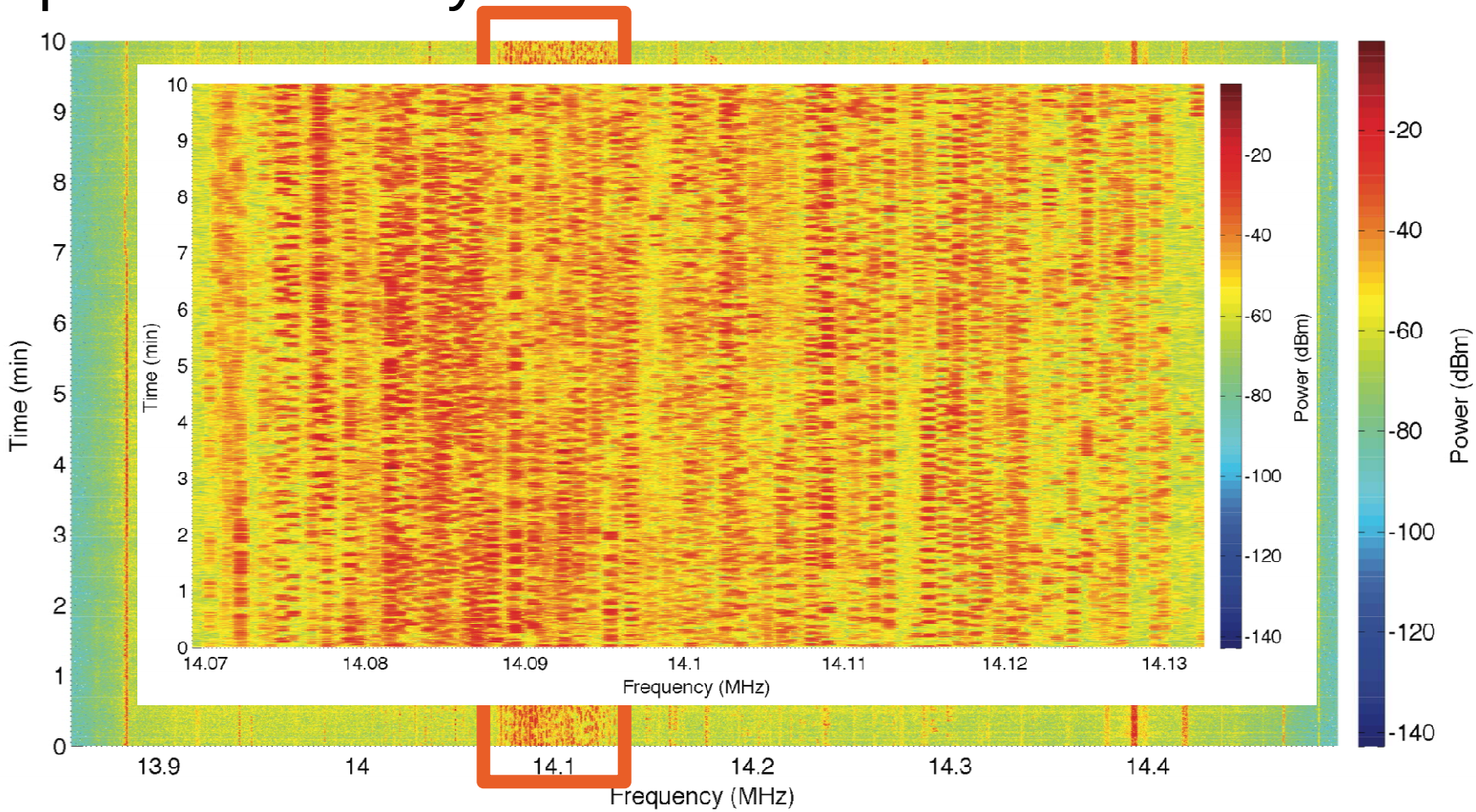
➔ opportunity on radio-amateur HF channel

[17] Laura Melián-Gutiérrez, Navikkumar Modi, Christophe Moy, Iván Pérez-Álvarez, Faouzi Bader, Santiago Zazo, "DSA with Reinforcement Learning in the HF Band“, URSI AT-RASC, Gran Canaria, Spain, May 2015

[18] Laura MELIAN-GUTIERREZ, Navikkumar MODI, Christophe MOY, Faouzi BADER, Ivan PEREZ-ALVAREZ, Santiago ZAZO , "Hybrid UCB-HMM: A Machine Learning Strategy for Cognitive Radio in HF Band" **,** *IEEE Transactions on Cognitive Communications and Networking, Year 2016, Volume: PP, Issue 99*

HFSA_IDeTIC_F1_V01 Database

Spectrum activity in the 14 MHz amateur band.

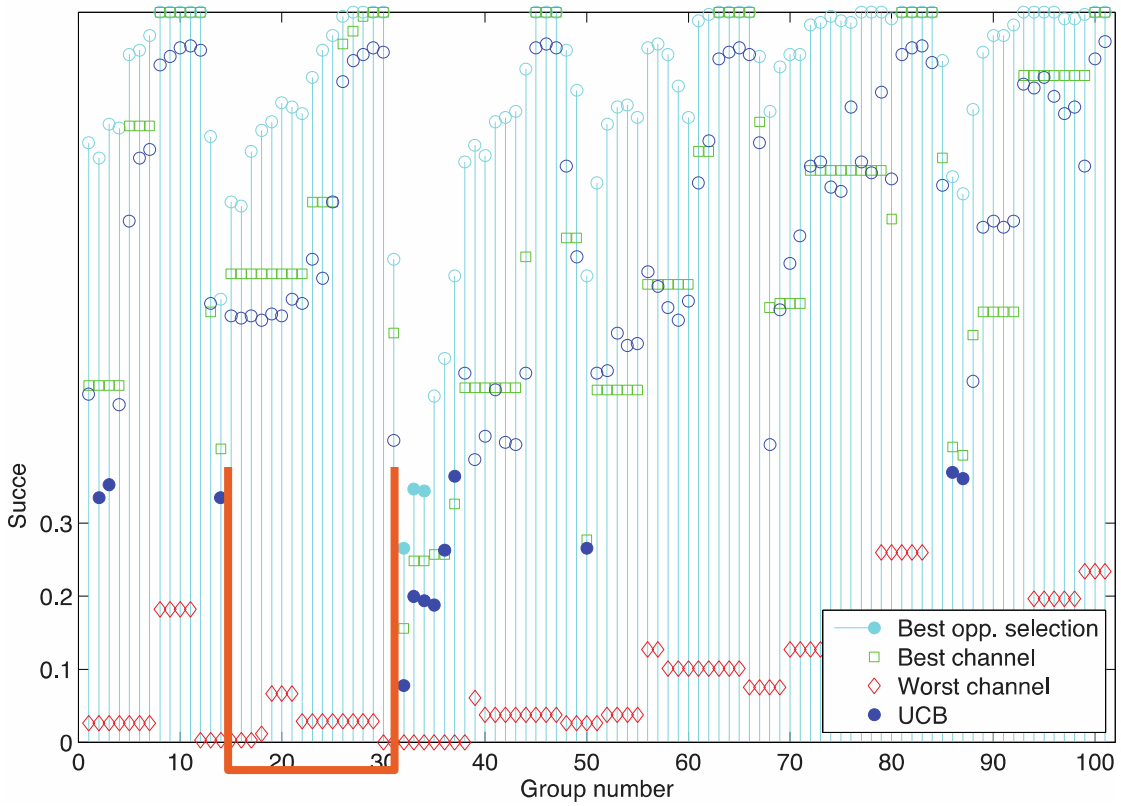


UCB on HF

Tries to exploit 1 over 8 channels

Compare with

- worst channel
- best channel
- genie aided



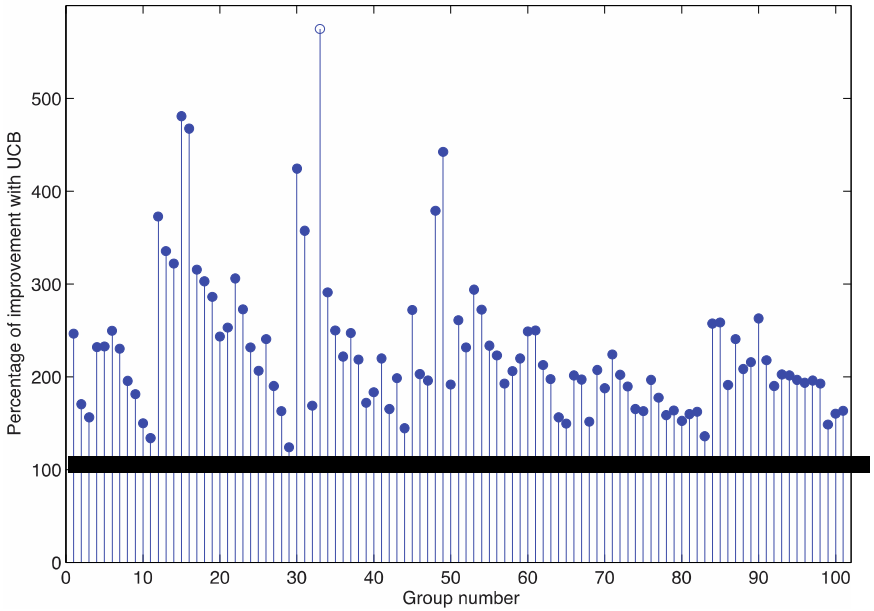
Example

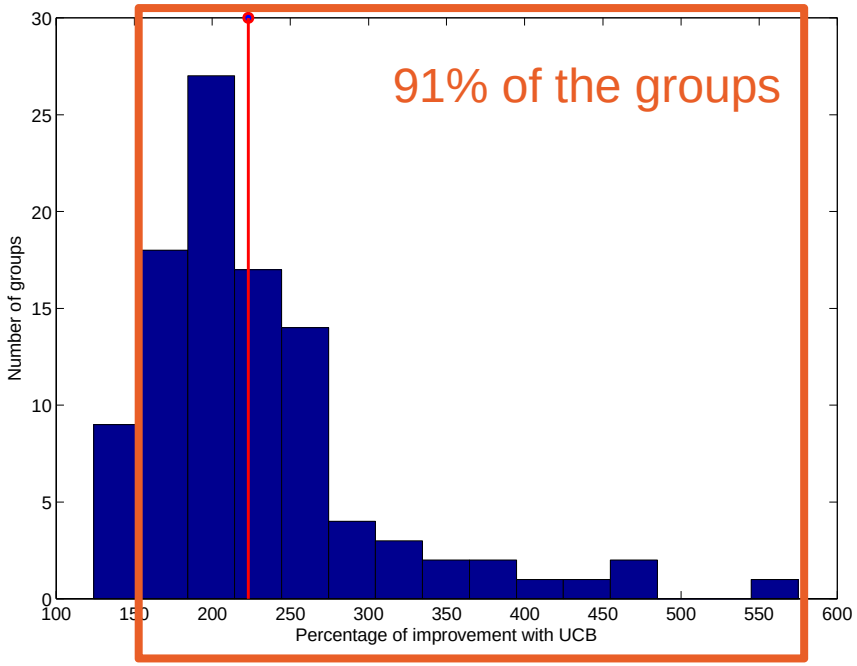
- there are not always vacant channels (genie aided<100%)
- however UCB finds 60% to 95% of opportunities, nearly as much as if he would always use the best channel

Comparison with a (stupid) pure random access

- Always more than twice better pure random policy
- Mean of 225% improvement
- More than 90% of times over 150% of improvement

Percentage of improvement for 100x8 HF channels





Outline of the presentation

Centrale Supelec and SCEE research team

Cognitive Radio Equipments and Systems

On-Line Learning for Cognitive Radio

UCB – Upper Confidence Bound

UCB demo for OSA

UCB for HF communications

Current and future steps

Current and future steps

New UCB taking into account quality of the band in addition to vacancy

- PhD of Navikkumar MODI
- Green quality criteria

➔ patent and curretnly under IEEE Trans review (50 pp)

➔ ON/OFF base station (stochastic geometry)

- Learning potential of a given scenario

RF Energy Harvesting (RFEH)

- collaboration with Prof. Sumit DARAK (IIIT Dehli, India)

Confront on-line learning to other real channels

Thank you for your attention

christophe.moy@centralesupelec.fr

Web page: http://www.rennes.supelec.fr/ren/perso/cmoy/Welcome.php

GoogleScholar:http://scholar.google.com/citations?hl=fr&user=qKv7crcAAAAJ&sortby=pubdate&view_op=list_works&pagesize=100