Machine learning proof-of-concept for Opportunistic Spectrum Access

Christophe Moy, Rodolphe Legouable

Machine learning proof-of-concept for Opportunistic Spectrum Access

Christophe Moy {christophe.moy@supelec.fr;christophe.moy@b-com.com}
Rodolphe Legouable {Rodolphe.legouable@b-com.com}

1. < Abstract >

The set of 5G requirements show that future radio systems should answer to network capabilities in terms of : capacity, **spectrum** efficiency, **spectrum** agility, low power consumption**,** flexibility for future evolutions, fixed-mobile convergence, integration of 3GPP and non 3GPP RATs, resilience and robustness, cost efficiency, etc.

In parallel, new user experiences will appear such as: homogenous experience over the coverage area, from static to high-speed-trains velocity, from outdoor to deep-indoor and higher typical throughput per user/ application, E2E[i] latency of a few *ms*, connectivity transparency, etc. leading to the need of new enablers for business as: Internet of Things (IoT), V2V[ii], D2D[iii], M2M[iv] communications.

This ecosystem points out that **Opportunistic Spectrum Access** should play an interesting role in 5G network optimization.

2. < Introduction >

The telecommunication community is working on the definition of future 5G that should take place at the Horizon 2020 (H2020). This future system would answer to the increase in mobile data traffic issued from the new services (3D TVHD, gaming, virtual reality, smart cities, e-health, etc.) and new user experiences. Connectivity should be possible at Any Time, Any Where and Any Device (ATAWAD). Thus, 5G should allow a set of requirements such as:

- higher capacity
- higher spectrum efficiency
- spectrum agility
- low power consumption
- Higher network coverage
- flexibility for future evolutions
- fixed-mobile convergence
- integration of 3GPP and non 3GPP RATs
- resilience and robustness
- cost efficiency
- support for high down-to very low bit-rates

Sometimes, some answers to fulfill all these requirements are antagonistic; indeed, increase the system's capacity with higher coverage without increasing the transmit power seems conflicting. So, the system optimization has to be considered in general.

Spectrum is one of the key issues to answer to 5G requirements but spectrum is scarce and expensive.

So, shared spectrum access will be essential to achieve the 5G objectives by considering:

- Licensed and un-licensed bands
- Co-existence between actual legacy and new Radio Access Technologies (RAT) by dynamic use of spectrum.

By introducing spectrum sharing for dynamic spectrum allocation agility in radio resource management, it will allow:

- to use "free" bands for transmission purpose (TV White Space, license Sharing access …) by using geo-localization criteria.
- to limit Interference between systems.

These considerations should contribute to densify the network, increasing its capacity while decreasing the power consumption.

We introduce, in section 3, an opportunistic spectrum access learning method allowing to efficiently exploit unused spectrum and then to optimally use it for data (re)transmission. First the problem statement is given; then the system overview including the platform description is depicted before showing the performance results achieved by the proposed method.

Section 4 introduces some strategic applications viewed by b<>com, mainly, addressing multi-RAT approach where spectrum sharing and management are crucial. In this same section, a brief focus on the b<>com's involvements at ETSI RRS is done. Section 5 concludes this article.

3. < Learning for Opportunistic spectrum access >

3.1.Problem statement

Dynamic spectrum access (DSA) is a new degree of freedom to mitigate spectrum scarcity. There are several ways of tackling this issue [1]. Opportunistic Spectrum Access (OSA) is one of them. It consists in enabling Secondary Users (SUs) to use the spectrum let vacant by licensed Primary Users (PUs). Cognitive radio [2], through its features in terms of permanent adaptability to varying conditions, is foreseen as key technology in order to implement such new schemes in commercial and military spectrum.

Cognitive radio paradigm is all about providing self-adaptation capabilities to the radio equipments and networks so that they can adapt dynamically to real-time conditions [3]. Current radio systems however are designed to support the worst case situation they are supposed to face rarely, which results at almost all instants in a loss in terms of power consumption, battery autonomy, spectrum efficiency and consequently capacity of the global system, etc. In other words, current radio systems are far from optimality, whatever the goal criteria, and cognitive radio is a way to make a further step towards optimality. The facilities a cognitive radio equipment (or a cognitive network) should include in addition to usual radio processing any radio equipment should have, can be summarized as [4]:

- sensing means,
- learning and decision making means,
- adapting means.

The demonstrator presented here focuses on learning and decision making stage through reinforcement learning (RL). RL is a machine learning approach which is based on the "try and evaluate" principle which consists in iteratively trying a set of solutions, evaluating their result and then deriving some quality factor of each trial. The goal is to order solutions, given a quality objective, so that the best one is used at next iteration. In other words, this aims at predicting which solution is giving the best opportunity at next trial.

OSA scenario can be modelled as a MAB (Multi-Armed Bandit) problem. You play one gambling machine at a time in casino and you try to maximize your benefit through playing more and more on the machine which gives you rewards. In the OSA context, gambling machines are radio channels and rewards are transmitted bits. UCB (Upper Confidence Bound) algorithm is one RL algorithm to solve MAB problem [5]. UCB is used here by a secondary user (SU) to learn about channels occupancy in order to derive the best channel to select in an OSA scenario [6]. The considered learning algorithms are able to act in highly unpredictable conditions, e.g. learn from scratch about the spectrum occupancy by primary users (PU), without any *a priori* knowledge.

The advantage of the proposed UCB algorithm for OSA is that in such an approach only one channel is sensed at a time. Then Radio Frequency (RF) front-end and digital processing front-end of the radio do not have to support a larger bandwidth than the bandwidth of one channel which is required for the transmission itself. In other words, there is no need for a wideband RF (and the associated digital processing overhead) to sense all channels of interest in parallel, which represents great savings both at design and operation times. As a consequence, OSA introduction does not require changing the global design of the equipments in this approach, but just add some light digital signal processing as it will be shown in the experiments.

Moreover, UCB learning and decision making are very light in terms of processing and memory demands which is nothing compared to sensing effort. The proposed OSA radio equipment can then be based on a conventional radio with the only addition of sensing function, and a very light overhead for decision making and learning.

3.2.System overview

The proposed demonstration is the first worldwide implementation in real radio conditions of reinforcement learning algorithms for OSA. In the OSA context, learning is derived, at the output of a sensing algorithm detecting the presence of a PU signal (energy detector, cyclo-stationnarity detector, etc.). The learning and decision process aims at:

1. deciding to transmit or not at current iteration,
2. updating learning information,
3. deciding which channel to sense and to choose for transmission at next iteration.

The decision to transmit is done only if the sensor detects the channel vacant at the current iteration. Learning, as well as the decision on which channel to try to transmit at next iteration, are done whatever the detection result.

As the OSA context can be modeled as a MAB issue, the following scheme is implemented: each frequency channel is equivalent to a gambling machine or a bandit arm. If we consider a wide sense stationary context, the figure of merit of a channel is its probability of vacancy, e.g. the probability that a channel is not used by a PU, which is equivalent to the probability for a gambling machine or arm to win a constant amount of money.

The MAB model for OSA implies that time is slotted in iterations. At each iteration, the SU radio system senses a channel. Either the channel is detected vacant and then transmission is done by the SU system on this channel during the rest of the slot. Or the channel is detected occupied and no transmission is done at that iteration. Then, SU system must wait for next iteration to sense another channel. Learning consists in taking into account the past trials' results in order to decide which channel to target at next iteration. The goal is to maximize the probability of success (e.g. trying a vacant channel) in order to maximize transmission opportunities for SU, while learning enough to avoid local minima in the optimization process.

In our experiments, the primary network is made of 8 channels. The probability of vacancy of each channel by the primary users can be set as wanted and has been chosen as follows in the following experimental results of this paper: {0.5;0.3;0.4;0.5;0.6;0.7;0.8;0.9}. This means that the probability of occupancy of channel #1 by PUs is 0.5, the probability of occupancy of channel #2 by PUs is 0.7, the probability of occupancy of channel #3 by PUs is 0.6, and so on until the probability of occupancy of channel #8 which is 0.1 only. So channel #8 has a probability of vacancy of 90% and then is the best channel to offer secondary transmissions opportunities.

The chosen design environment for the primary network radio signal generation is GNU Radio Companion (GRC) and the hardware platform is made of a USRP platform from Ettus Research [7] connected to a laptop running Linux as shown left hand side of Figure 1. The right hand side platform of Figure 1, made of a computer and a USRP platform, represents a secondary user. Only sensing and learning are implemented here, e.g. this platform is only a receiver (RX).

For simplicity purposes, there is one slot per second. This means that the channel occupancy of PUs varies once a second and can be followed by human eye. Nothing technically prevents form accelerating this rate. Algorithms converge in function of the number of trials, so learning algorithms convergence speed is directly a function of this rate. If frames would be 1 ms long, we could directly conclude on a learning speed 1000 times faster than the current experiment.

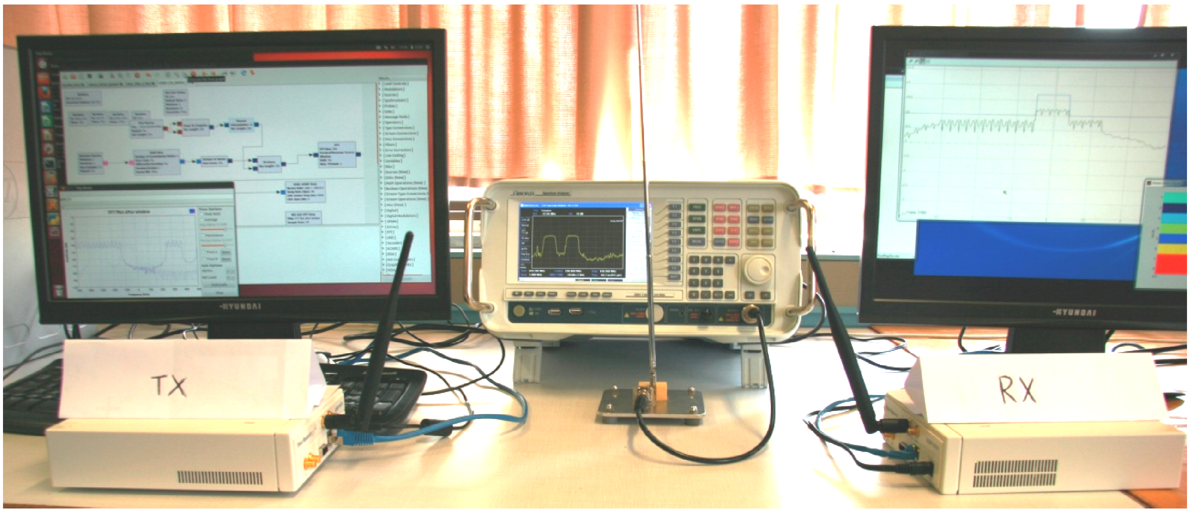

Figure 1 – Experimental testbed for learning in an OSA context. Left hand side (laptop + USRP) is the playing the role of the primary network transmission (TX) with the visualization of the generated traffic on 8 channels. Right hand side (laptop + USRP) is the playing the role of the secondary user learning algorithm, implementing an energy detector as a sensor (RX). A spectrum analyzer inbetween enables to monitor PUs RF signals.

3.3.Performance results

First results concern the ability of UCB to rapidly learn, e.g. converge on the best channels. This kind of reinforcement learning algorithm is mathematically proven to converge at infinite time [6]. The demonstration shows that UCB are very rapidly converging to choose the best channel. Table 1 gives the empirical mean of vacancy derived, for each channel, by UCB algorithm, as learning goes on through one experimental shot, compared to the effective vacancy rate set in the experiments for the primary network channels (column 2).

We can see on table 1, where the 8 channels performance are listed from channel #1 at the top, to channel #8 at the bottom, that after just 72 iterations (a mean of 9 trials per channel), channel #8 has been selected 24 times in column 4, e.g. one third of times. We can see on column 3 that a 95% vacancy rate has been obtained for channel #8 during this period, so the SU was a little bit luckier than what should happen in the long term, as a primary network vacancy rate of 90% has been set for channel #8 (column 2). This illustrates how fast the algorithm learns and that as soon as it starts learning, it starts finding transmission opportunities, while privileging the best channel, compared to a random or uniform strategy. Indeed no exclusive learning period, also called exploration phase, exists. Transmission is done as soon as a channel is found vacant, all along the learning process. This is called exploitation phase. Then learning, or exploration, is improved all along the exploitation process.

Table 1 - Empirical mean of vacancy derived for each channel by UCB as learning goes on through one experimental shot, compared to the effective vacancy rate set in the experiments for the primary channels

		After 72 iterations		After 1633 iterations		After 7000 iterations	
Channel index	Effective vacancy rate	Empirical mean	Nb of trials per band	Empirical mean	Nb of trials per band	Empirical mean	Nb of trials per band
#1	0.50	0.70	10	0.72	110	0.70	219
#2	0.30	0	2	0.32	19	0.27	26
#3	0.40	0	3	0.45	29	0.54	71
#4	0.50	0.60	10	0.56	43	0.56	89
#5	0.60	0.70	10	0.80	181	0.79	557
#6	0.70	0.60	10	0.84	315	0.80	718
#7	0.80	0	3	0.84	308	0.84	1532
#8	0.90	0.96	24	0.90	628	0.87	3783

We can see in columns 5 and 6 of table 1 that after 1633 iterations, best 3 channels have been selected 1251 times so 84% of the time. As already stated in previous section, if slots would have been 1 ms long, this means that after only 1.5 seconds, UCB would have perfectly learnt and would have been using most of the time the best transmission opportunities.

Columns 7 and 8 of table 1 show that after 7000 iterations, UCB has selected the best channel more than half of iterations (3783). As this channel has a 87% probability of vacancy, SU has found many transmission opportunities. If the two best channels are considered, this increases up to ¾ of the attempts. Then, the more the system runs, the better knowledge it acquires and the more the best channel is selected in terms of percentage. Remark that from the machine learning point of view, the selection rate of the best channel is the quality criteria. This is different from the OSA point of view, where the criteria is the percentage of transmission opportunities. Then selecting the best, or the second best, or even the third best channel maybe also considered as a good result, as long as channel is detected vacant [9].

Hence Figure 2 is a snapshot of the Simulink interface tracing the SU learning process. UCB performance, in bottom table, is compared with another algorithm in top table [10]. Tables have the same shape as table 1, e.g. channel #1 is at the top of the tables, and channel #8 at the bottom. Holes distribution column corresponds to the empirical mean column of table 1. The other algorithm has a weight column which is relaed to the number of times the considered channel has been detected vacant minus the number of times it has been detected occupied. This is out of the scope of this paper to describe this algorithm, however we can just see that it does not explore all channels as UCB does, and you can refer to [10] for more information. It concentrates on the a subset of channels it has considered as the best ones, at the very beginning of its learning phase. As a consequence, it may be mistaken. That is what happened here as it has missed channel #8. However it is using channels #7 and #6 which are respectively vacant 80% and 70% of time. However, after 27610 iterations, this algorithm has found 75.2% of successful transmission opportunities (see middle left box of Figure 2, top line), e.g. 75.2% of the time, the channel selected by this algorithm was found vacant, so free to be used for a secondary communication. In the middle left box of Figure 2, bottom line, we can see that UCB has almost found 85% of transmission opportunities, which is very good as only one of the channels has a better average performance than this value. This means that UCB could reach a better performance than the average performance of the second best channel, while getting knowledge about all other channels (at the price of loosing some opportunity). Bottom left box of Figure 2 gives the percentage of time that the best channel has been selected by the 2 algorithms. This is a pure machine learning evaluation criteria. We can see that this makes little sense from the OSA point of view as even if the alternative algorithm has never used the best channel, it has reached more than 75% of transmission opportunities. Note that UCB, for which it is mathematically proven that it will converge, e.g. use most of the time the best channel in these conditions, has a 92% rate of selection of the best channel, far before infinity of iterations.

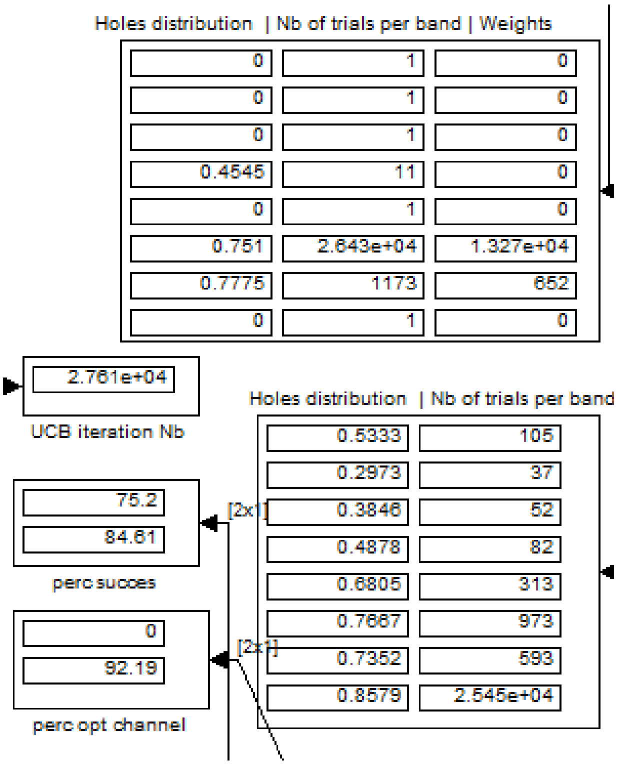

Figure 2 – Learning results on the eight bands after 27610 iterations – top table for WD algorithm and bottom table for UCB algorithm

UCB robustness has been established theoretically and can be verified trough repetitions of this demo when sensing errors occur. Learning still works, even with a very high level of errors (several tens of percent), e.g. UCB still converge to find the best channel, but at the only price of a longer learning phase to obtain the same level of knowledge obtained without sensing errors [8].

Last point concerns the complexity of implementation of UCB. UCB algorithm consists in computing at each slot one index for each channel, each demanding 2 multiplications, 2 divisions, 1 addition, 1 square root, 1 logarithm operations [6]. Adding to this, 2 increments of integers are also necessary whatever the number of channels. In terms of memory, 2 buffers are required per channel as well. In other words, UCB costs nothing in terms of extra computing and memory resources.

3.4.OSA conclusion

OSA is one possible way of implementing DSA. OSA is a futuristic alternative of flexible spectrum use in most scenari. It is probably still too early to use it now in a given number of licensed bands. However, OSA could be used to improve current systems operating in unlicensed bands, such as WiFi for instance.

We show in this demo that implementing OSA is possible. Turning a legacy radio equipment into an OSA equipment only requires to add some sensing part which can detect for sure the presence or not of primary users. Other requirements are already common in radios, as we show that very simple algorithms can be used for learning and decision making. In addition, this algorithm enables to keep a narrow band RF front-end, without the need to sense all bands in parallel, which also would imply an unacceptable level of extra digital processing. Finally last necessary requirement consists in being able to jump from one frequency to another, which is a common feature of any radio equipment. The only remaining challenge indeed concerns the sensing. That is the reason why current dynamic spectrum approaches which are under standardization process are concerned with geo-location databases, such as in ETSI RRS.

4. <Strategic applications>

Spectrum sharing and flexible spectrum usage techniques can be used to optimize spectrum utilization and more importantly to provide opportunities for operators to access additional spectrum typically allocated to licensed or un-licensed radio services. Flexible spectrum usage is more related to spectrum usage optimization within one system while spectrum sharing is for sharing between systems (multi-RAT). Depending on the specific service and user experience, different spectrum usage strategy can be selected individually or jointly to satisfy diverse requirements. 5G networks will be with multiple layers and RATs, and operate on multiple bands scattering at diverse frequencies with different spectrum access modes. Exclusive use of dedicated spectrum will continue to be the preferred way of spectrum usage by 5G operators and other spectrum access modes are possibly complementary to it.

So, b<>com gambles on multi-RAT as a potential technologies for 5G to address the market drivers, the use cases and the targeted requirements.

4.1.Multi-RAT integration and management

The ever-increasing number of RATs to be supported in a given deployment makes it crucial to consider multi-RAT integration and management issues. The objective is to facilitate uniform multi-RAT management and convergence among disparate technologies, both 3GPP and non-3GPP, such as Wi-Fi. New radio access technologies should also be considered (Mm Waves, broadcast, etc.) as integrated RATs. Operation efficiency and user experience would be dramatically improved by automatically steering devices to the most suitable RAT in a seamless way. While multi-RAT management has been an important aspect in previous mobile generations, it will be more critical in 5G, particularly for services such as ultra-high-definition video or tactile Internet.

Given the likelihood of having multiple, heterogeneous wireless access points available in ultra-dense scenarios (e.g., 5G, LTE, 3G and Wi-Fi), some kind of decoupling between the user and control planes should be provided in order to separate the user payload from the necessary signaling. Multi-RAT integration may also consider simultaneous connections to multiple RATs (simultaneous transmit and receive signals) in an opportunistic manner as proposed by the method introduced in §3. The system should be "intelligent' enough for detecting data issued from a RAT and to retransmit (in a free spectrum band for instance) via another RAT for network coverage increase and/or interference mitigation purpose.

The expected impact on the network of such schemes is the introduction of a logical entity that coordinates resources among multiple RATs. So, evolutions at network but also at the terminal levels are essential. That's why, b<>com is working on multi-RAT by using agile spectrum management for optimizing network capacity, reducing energy power consumption while increasing network coverage. The main objective is to work on the mutualization of common processing (RF, filtering, demodulation, channel coding, etc.) to generically address several multi-RAT architectures (common digital RF front-end, common LDPC (low Density parity Check) and Turbo decoding architectures). The detection and selection of the best RAT is depending on specific criteria via metrics that have to de defined: SINR[v]/RAT, Energy Efficiency/RAT, Spectral Efficiency/RAT, latency/RAT, link budget/RAT, etc. minimizing the latency in the decision process, minimizing the power consumption and/or optimizing the spectrum sharing, etc.

The generic and transparent approach to implement in hardware a reconfigurable digital modem is carried out for selecting the best RAT regarding the above metrics (SINR, coverage, energy efficiency …) impacting directly on the choice of the "free/unused" spectrum band (and "best" RAT). This reconfigurable digital modem design will be fast-reconfigurable (<100µs), low power consumer with high sensibility for multipurpose applications (connectivity, cellular and broadcast) including geo-localization feature with arbitrary waveform (mainly for M2M applications).

4.2. ETSI Reconfigurable Radio Systems

Since July 2014, b<>com is active in ETSI RRS and want to introduce multi-RAT for a better usage in spectrum management and green challenge by using the Software Defined Radio (SDR) approach. Among the key topics, b<>com wants to:

- anticipate 5G and agile spectrum management focused on Multi-RATs architectures;
- propose green criteria integrated in 5G /agile spectrum management combined with multi-RAT Management in multi-technology-HetNets (Heterogeneous networks);
- propose a new work item dealing 5G green oriented spectrum management;
- contribute to the Working Group 2 (WG2) on Multi-RAT Mobile devices;
- and participate to WG1 activities and establish links between mid-term and long-term evolutions.

5. <Conclusion>

In this paper, after introducing the main requirements to which future 5G system should answer, we introduced an opportunistic spectrum access method that allows a better spectrum sharing that could be applied in Multi-RAT environment. Indeed, b<>com believes in multi-RAT as a key technology that should (at least) play a role in 5G definition. However, Multi-RAT needs evolutions of network and User Equipment sides to mutualize almost as possible the processing functions to select dynamically the most appropriate RAT with the lowest latency. We saw, otherwise said, that

Multi-RAT re-introduces the SDR approach and allows a better usage in spectrum management and resource sharing. Also, we introduced the idea that b<>com would like to address at ETSI RRS at middle and long terms.

<Bibliography>

[1] Q. Zhao, A. Swami, "A Survey of Dynamic Spectrum Access: Signal Processing and Networking Perspectives", in IEEE ICASSP: special session on Signal Processing and Networking for Dynamic Spectrum Access, April, 2007

[2] J. Mitola, "Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio", Licentiate proposal, KTH, Stockholm, Sweden, 8 May 2000.

[3] J. Palicot, "Radio Engineering: From Software radio to Cognitive Radio", Wiley 2011; ISBN: 978-1-84821-296-1

[4] W. Jouini, C. Moy, J. Palicot, "Decision making for cognitive radio equipment: analysis of the first 10 years of exploration", *EURASIP Journal on Wireless Communications and Networking 2012*, 2012:26.

[5] R. Agrawal, "Sample mean based index policies with o(log(n)) regret for the multi-armed bandit problem," Advances in Applied Probability, 27:1054–1078, 1995

[6] W. Jouini, D. Ernst, C. Moy, J. Palicot, "Upper confidence bound based decision making strategies and dynamic spectrum access", ICC, Cape Town, South Africa, May 2010

[7] Ettus Research "Products" - accessed 02/04/2012 http://www.ettus.com/products

[8] W. Jouini, C. Moy, J. Palicot, "Upper Confidence Bound Algorithm for Opportunistic Spectrum Access with Sensing Errors", CrownCom'11, 1-3 June 2011, Osaka, Japan

[9] C. Robert, C. Moy, and C.-X. Wang, "Reinforcement learning approaches and evaluation criteria for opportunistic spectrum access," in IEEE International Conference on Communications (ICC), June 2014.

[10] C. Moy, "Reinforcement Learning Real Experiments for Opportunistic Spectrum Access" Karlsruhe Workshop on Software Radios, WSR'14, Karlsruhe, Germany, 12-13 March 2014

[i] E2E : End to End
[ii] V2V : Vehicular to Vehicular
[iii] D2D : Device to Device
[iv] M2M : Machine to Machine
[v] SINR : Signal to Noise plus Interference Ratio