32nd URSI GASS, Montreal, 19–26 August 2017

Proof-of-Concept: Spectrum and Energy Efficient Multi-User CR Network via Vacancy and Quality based Channel Selection

Navikkumar Modi*[(1)], Philippe Mary[(2)], Christophe Moy[(1)] and Sumit Darak[(3)]
(1) CentraleSupélec, IETR, UMR CNRS 6164 France
(2) INSA de Rennes, IETR, UMR CNRS 6164, France
(3) Department of Electronics and Communications Engineering, IIIT Delhi, India

Abstract

This demonstration presents a proof-of-concept for opportunistic spectrum access in the decentralized network consisting of multiple secondary users (SUs). SUs need to independently characterize these frequency bands based on estimated primary user (PU) occupancy statistics and channel quality and avoid collision with other SUs. The proposed, quality of service- upper confidence bound (QoS-UCB) policy, as well as existing state-of-the-art multi-armed bandit (MAB) learning approaches have been implemented on the proposed testbed consisting of USRPs from Ettus Research, GNU radio companion for PUs and Matlab/Simulink for SUs. The experimental results show that the proposed QoS-UCB policy outperforms existing learning approaches in terms of SU throughput, number of SU collisions and the number of frequency band switchings making SU terminals energy efficient.

1 Introduction

One of the most promising dynamic spectrum access (DSA) approach to mitigate spectrum scarcity is opportunistic spectrum access (OSA) [9]. OSA paradigm allows unlicensed users, i.e. secondary users (SUs), to access the vacant licensed frequency bands given that SUs do not interfere with transmission of primary users (PUs). The core of OSA problem is to learn which channels are the best in terms of chosen criterions, e.g. availability and quality [7, 8]. In [5], multi-armed bandit (MAB) framework has been proven to cope with OSA analytically with only availability criterion, then in [7], it has been extended to be optimal for both availability and channel quality criterions. The energy efficiency is just as much dependent on channel propagation conditions, e.g. channel quality, as the spectral efficiency and hence it is necessary to adapt the system to the changing environment. Furthermore, it is necessary to minimize the number of SU collisions as well as number of channel switching to improve SU throughput and energy efficiency of battery operated SU terminals. Such decision making procedures are usually designed using online learning algorithms and reader may refer to [7, 3, 4] for details.

In this paper, we present a demonstrator which focuses on learning for MAB framework that models OSA in multi-player settings in order to enable SUs to identify and transmit over vacant channels, and keep the channel switching cost (CSC) as minimum as possible. Here, CSC stands for the total penalty incurred in terms of delay, power, hardware reconfiguration and protocol overhead when SU switches from one frequency channel to another. Design of such learning policies for cognitive radio (CR) networks is a challenging task and one of the objective of the work presented in this demonstration on real radio signals. We propose a universal software radio peripheral (USRP) based testbed for analyzing the performance of learning policies for OSA in CR networks. To the best of our knowledge, the proposed testbed is the first worldwide proof-of-concept which compares the performance of various learning policies using real radio signals.

The objective of the proposed demonstration is to experimentally compare the performance of existing state-of-the-art MAB learning approaches [2, 6, 1] and our proposed QoS-UCB approach [7] in real radio environment. The design details of the proposed testbed is discussed in next section followed by experimental results and conclusions in Section III and IV, respectively.

2 USRP Testbed

The USRP testbed is shown in Fig. 1 and is a significant extension of the testbed in [8, 3, 4]. It consists of two units: 1) Left hand side unit is the primary user traffic generator, and 2) Right hand side unit acts as SUs. Both the units are discussed in detail next.

2.1 Primary User Traffic Generator

The chosen design environment for the PU traffic generator is GNU Radio Companion (GRC) and the hardware platform is made of a laptop and a USRP from Ettus Research. The main reason for choosing GRC is the precise control on each parameter of the transmission chain compared to other environments. In the beginning, the number of frequency bands, rewards of MAB framework which models PU traffic and corresponding channel statistics, i.e. independent

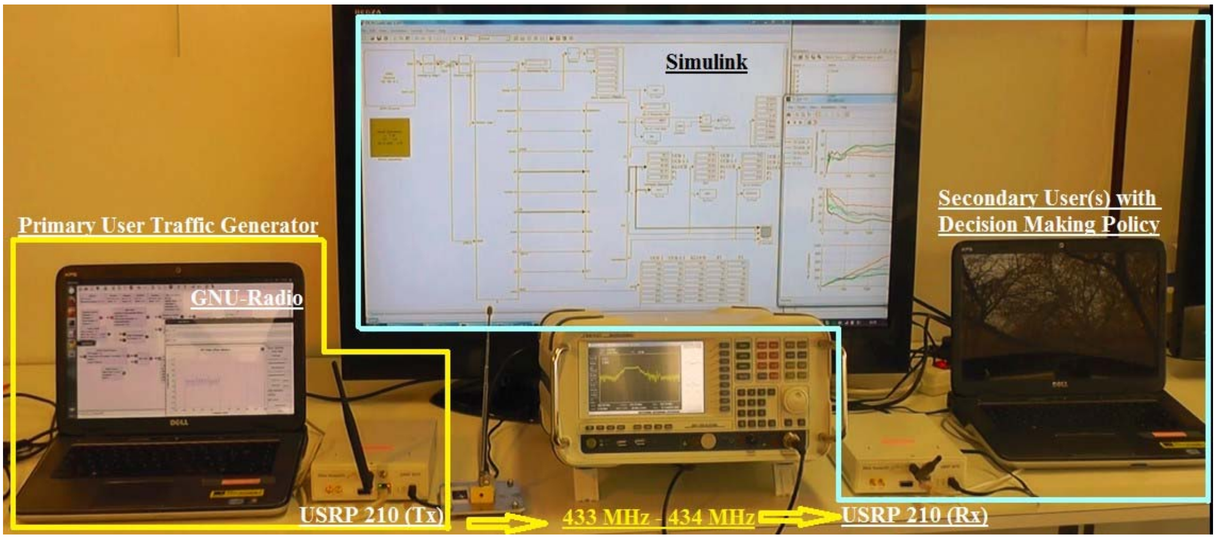

Figure 1. Left hand side (laptop with GNU Radio + USRP) is generating primary user traffic on 8 channels (Tx). Right hand side (laptop with Simulink + USRP) is a secondary user employing energy detector for channel sensing and online learning algorithm based learning policies (Rx). A spectrum analyzer monitors the RF signals.

identically distributed (i.i.d.) traffic or Markovian traffic, are fixed. The transmission bandwidth, which is restricted by bandwidth of analog front-end of USRP, is divided into $K = 8$ uniform bandwidth frequency bands. Multi-channels occupation pattern is digitally managed thanks to an orthogonal frequency division multiplexing (OFDM) approach, using inverse fast Fourier transform (IFFT). In each time slot, masking vector of size K is generated by Traffic Model block based on given frequency band statistics. This masking vector can have 1 or 0 values where 1 and 0 indicate that corresponding band is occupied and vacant, respectively. Next step is mapping data to be transmitted on sub-carriers of occupied bands. The data modulation used is a quadrature phase shift keying (QPSK) modulation. This is followed by sub-carrier mapping using IFFT and transmission via USRP. In the proposed testbed, number of sub-carriers, center frequency and transmission bandwidth are 256, 433.5 MHz and 1 MHz, respectively. The 256 carriers are grouped by 32 in order to make finally 8 channels. For demonstration purpose, each time slot duration is one second so that it can be followed by human eye. However, it can be reduced to the order of milliseconds and will have no direct effect on the performance of learning policies.

2.2 Secondary User with Decision Making

The chosen design environment for the SU terminal is MATLAB/Simulink and USRP from Ettus Research. USRP is tuned to receive signal of bandwidth 1 MHz centered at 433.5 MHz. The received signal is then down-sampled, digitized and passed to the learning policy implemented using Simulink. Then, an on-line learning policy selects one frequency band at each time slot. The chosen channel is sensed using a spectrum sensing detector, i.e. energy detector. Note that energy detector is not ideal and sensing errors may occur [5], however this is supported by the learning policy with no risk of divergence from optimal channels but with slower convergence behavior. If the band is sensed as vacant, it is assumed that SU transmits over the chosen band. If multiple SUs choose the same frequency band, then all users suffer from collision, and transmission fails. SU transmission is not implemented in this demonstrator, as this demonstration largely focuses and analyzes the learning and decision making process, both in single-player and multi-player cases. In case of multi-player, each SU is independently implemented in Simulink with their respective learning policy. In existing work, sensing is assumed as imperfect which is the case in real radio conditions. Thus, the proposed testbed with non-ideal energy detectors will enable to study the performance of learning policies in presence of sensing errors. However, performance comparisons of various detectors and their effect on MAB learning policies is not discussed here, and is the future steps of this work.

In the proposed testbed, the receiver, which acts as an SU, consists of reinforcement learning algorithm for decision making process. We analyze and compare the performance of various MAB learning approaches, such as QoS-UCB with random rank selection ([7, 1]), UCB1 with random rank selection ([5, 1]), Kullback-Leibler UCB with random rank selection ([2, 3]), Bayesian UCB (BUCB) with random rank selection ([6, 3]) and KLUCB-BUCB ([3]) algorithms. Note that our proposed QoS-UCB is the only policy which takes into an account availability and quality, i.e. measured energy level with energy detector, for decision making while others only considered availability.

3 Experimental Results and Analysis

In this section, the performance of various MAB learning approaches in terms of number of transmission opportunities, number of collisions and channel switching cost (CSC) is compared on the proposed testbed. We consider $K = 8$ channels in the proposed testbed. Scenario 1 and 2 correspond to the case where the channel occupancy is i.i.d. and Bernoulli distributed with vacancy probability (P_{vac}) and scenario 3 and 4 model the channel occupancy as a Markovian process with the transition probabilities (P), as given in Table 1.

As discussed before and in [7], the measured spectrum level is recorded and used as a quality information metric for QoS-UCB policy. Indeed, instead of having only a free or occupied state information at the output of energy detector, we get a soft-metric representing the measured power level which will be used later to rate the band quality. Empirical average of quality information for 8 channels after 1000 time slots and 10 experiment is also presented in Table 1.

Each numerical result presented hereafter in this section is the average value achieved by iterating over 10 independent experiments on USRP testbed. Each experiment on USRP testbed consists of a time horizon of 1000 time slots for each SU and one time slot corresponds to one second, however time slot could be considered as 1ms for result analysis in terms of convergence speed. In multi-player case, we assume that all SUs employ the same MAB learning approach but do not exchange any information with other SUs.

For Scenario 1 to 4, Fig. 2 shows the percentage of transmission opportunities exploited by various MAB learning approaches in multi-player with 4 SUs. It can be observed from Fig. 2 that the proposed QoS-UCB policy exploits similar or higher percentage of transmission opportunities compared to state-of-the-art decision making approaches for Scenario 1, 2 and 4. This is due to the fact that the proposed QoS-UCB approach learns on dual optimization criteria and hence is able to quickly find an optimal channel. In Scenario 3, we consider slowly varying PUs activity (channel state changes less frequently), and thus it is beneficial for algorithms like KL-UCB which is more explorative compared to our proposed QoS-UCB and others.

As discussed above, the number of CSC should be as minimum as possible for making SU terminals energy efficient. In Fig. 3, number of CSCs of different MAB learning approaches are compared in Scenarios 1, 2, 3 and 4 for multi-player case. It can be observed that the proposed QoS-UCB policy offers the lowest number of CSC. Numerically, average number of CSC of the proposed QoS-UCB is significantly lower than that of UCB1, KL-UCB and BUCB policies, due to learning on two criterions, i.e. availability and quality, compared to only availability in UCB1, KL-UCB and BUCB, which helps QoS-UCB to quickly find an optimal channel compared to UCB1, KL-UCB and BUCB.

In additions to CSC, number of collisions should also be as minimum as possible. This is because, collision leads to waste of the energy required for data preprocessing and

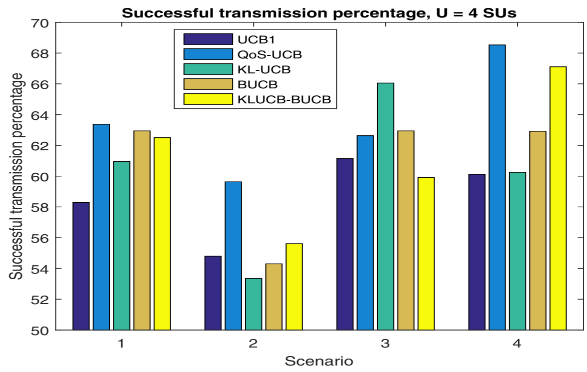

Figure 2. Comparisons of successful transmission percentage for various P_{vac} distributions in Scenario 1 to 4 of different learning approaches for multi-player case with 4 SUs.

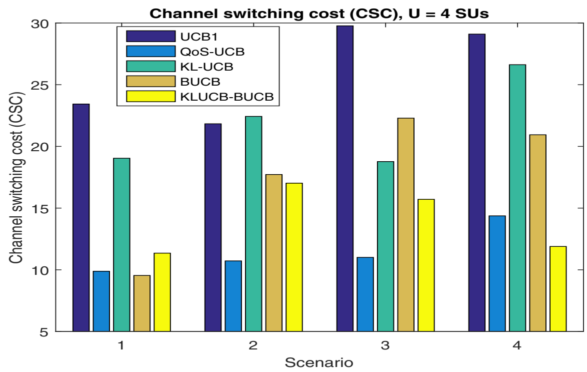

Figure 3. Comparisons of channel switching cost of different learning approaches in Scenario 1 to 4 for multi-player case with 4 SUs.

transmission and it may be higher than the energy required for CSC. In Fig. 4, the number of collisions suffered by all SUs is compared in Scenarios 1, 2, 3 and 4 for multi-player case. Numericaly, SUs employing the QoS-UCB policy suffers from fewer number of collisions than SUs employing other MAB learning approaches. Thus, lower number of CSC as well as collisions make the QoS-UCB policy comparatively more energy efficient and suitable for battery operated SU terminals. From the presented experimental analysis, we argue that the proposed QoS-UCB policy using dual criterion (e.g. availability and quality) optimization for OSA in multi-player decentralized network is not only superior in terms of spectrum utilization but also energy efficient.

4 Conclusion

A testbed has been developed to evaluate various MAB learning policies for opportunistic spectrum access in the decentralized networks in real radio environment. It consists of USRPs from Ettus Research, GNU radio companion for PUs and Matlab/Simulink for SUs. Experimental results show that the proposed QoS-UCB policy offers superior performance over existing state-of-the-art learn-

Scenario	Channels	1	2	3	4	5	6	7	8
Scenario 1	P_{vac}	0.50	0.30	0.40	0.50	0.60	0.70	0.80	0.90
Scenario 2	P_{vac}	0.50	0.90	0.80	0.70	0.60	0.50	0.40	0.30
Scenario 3	P_{01}	0.05	0.03	0.04	0.05	0.06	0.07	0.08	0.09
	P_{10}	0.05	0.07	0.06	0.05	0.04	0.03	0.02	0.01
Scenario 4	P_{01}	0.50	0.74	0.78	0.83	0.85	0.86	0.88	0.90
	P_{10}	0.50	0.26	0.22	0.17	0.15	0.14	0.12	0.10
Empirical quality information	R_1	1.33	1.53	1.41	0.92	1.17	0.70	0.89	1.36

Table 1. Several scenarios to verify the proposed approach on testbed.

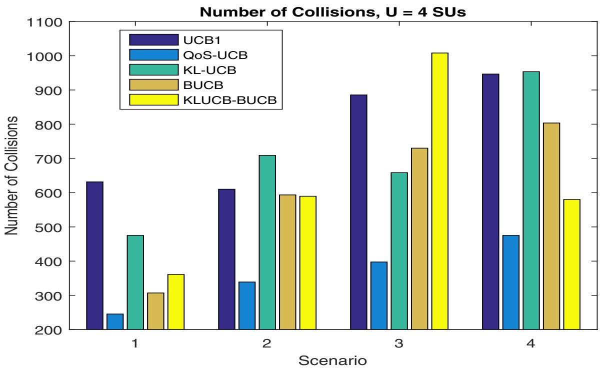

Figure 4. Comparisons of number of collisions of different learning approaches in Scenario 1 to 4 for multi-player case with 4 SUs.

ing approaches for decentralized network with known as well as unknown number of active secondary users. In additions, the proposed QoS-UCB policy achieves significantly lower number of collisions and channel switching cost while maintaining similar transmission percentage, and hence it wastes significantly less energy required for hardware reconfiguration, data preprocessing and retransmission.

5 Acknowledgements

This work has received a French government support granted to the CominLabs excellence laboratory and managed by the National Research Agency in the "Investing for the Future" program under reference No. ANR-10-LABX-07-01. The authors would also like to thank the Region Bretagne, France, for its support of this work.

References

[1] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami. Distributed algorithms for learning and cognitive medium access with logarithmic regret. *IEEE Journal on Selected Areas in Communications*, 29(4):731–745, April 2011.

[2] Olivier Cappé, Aurélien Garivier, Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz. Kullback-Leibler Upper Confidence Bounds for Optimal Sequential Allocation. *Annals of Statistics*, 41(3):1516–1541, 2013.

[3] Sumit Darak, Navikkumar Modi, Amor Nafkha, and Christophe Moy. Spectrum utilization and reconfiguration cost comparison of various decision making policies for opportunistic spectrum access using real radio signals. In *11th EAI International Conference on Cognitive Radio Oriented Wireless Networks, Grenoble, France, May 2016*, June 2016.

[4] Sumit J. Darak, Christophe Moy, and Jacques Palicot. Proof-of-concept system for opportunistic spectrum access in multi-user decentralized networks. *EAI Endorsed Transactions on Cognitive Communications*, 16(7), 9 2016.

[5] Wassim Jouini, Christophe Moy, Jacques Palicot, et al. Upper confidence bound algorithm for opportunistic spectrum access with sensing errors. *6th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications, Osaka, Japan*, page 17, 2011.

[6] Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On bayesian upper confidence bounds for bandit problems. In Neil D. Lawrence and Mark A. Girolami, editors, *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS-12)*, volume 22, pages 592–600, 2012.

[7] Navikkumar Modi, Philippe Mary, and Christophe Moy. QoS driven channel selection algorithm for opportunistic spectrum access. In *IEEE Globecom 2015 Workshop on Advances in Software Defined Radio Access Networks and Context-aware Cognitive Networks*, San Diego, USA, December 2015.

[8] Clément Robert, Christophe Moy, and Honggang Zhang. Opportunistic Spectrum Access Learning Proof of Concept. In *SDR-WinnComm'14*, page 8, Schaumburg, United States, March 2014.

[9] Q. Zhao and A. Swami. A survey of dynamic spectrum access: Signal processing and networking perspectives. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 4, April 2007.