# Case-Based Cognitive Cellular Systems for Temporary Events

Nils Morozs, Tim Clarke and David Grace
Department of Electronics, University of York
Heslington, York YO10 5DD, United Kingdom
E-mail: {nm553, tim.clarke, david.grace}@york.ac.uk

***Abstract*—In this study we investigate the use of case-based reinforcement learning (RL) for dynamic secondary spectrum sharing in cognitive cellular systems for temporary events. The performance of the proposed case-based RL scheme is evaluated using system level simulations that involve a stadium small cell network, an eNB on an aerial platform and a local primary LTE network. Compared to classical RL, the case-based RL approach results in increased adaptivity of the cognitive cellular system to sudden changes in its environment caused by the aerial eNB being dynamically switched on and off. We also show that a cognitive cellular system that employs the proposed dynamic spectrum access scheme is able to accommodate a 51-fold increase in offered traffic with no need for additional spectrum and with no degradation in the quality of service of the primary users.**

***Keywords*—*Reinforcement Learning; Case-Based Reasoning; Dynamic Spectrum Sharing; Temporary Event Networks***

## I. Introduction

One of the fundamental tasks of a cellular system is spectrum management, concerned with dividing the available spectrum into a set of resource blocks or subchannels and assigning them to voice calls and data transmissions in a way which provides a good quality of service (QoS) to the users. Flexible dynamic spectrum access (DSA) techniques play a key role in utilising the given spectrum efficiently. For example, cognitive cellular systems employ intelligent opportunistic DSA techniques that allow them to access licensed spectrum underutilized by the incumbent users [1].

An emerging state-of-the-art technique for intelligent DSA is reinforcement learning (RL); a machine learning technique aimed at building up solutions to decision problems only through trial-and-error [2]. The most widely used RL algorithm in both artificial intelligence and wireless communications domains is Q-learning [3]. Therefore, the algorithm developed in this study employs distributed Q-learning based DSA.

The purpose of this paper is to propose a way of improving the stability of RL based DSA algorithms for temporary event networks with dynamic topologies that use secondary LTE spectrum sharing. The technique investigated for solving this problem is case-based RL, a combination of RL and case-based reasoning (CBR). CBR is broadly defined as the process of solving new problems by using the solutions to similar problems solved in the past [4]. In case-based RL these solutions are learned by using an RL algorithm. The only example of applying this methodology in the wireless communications domain is [5], where a DSA scheme is designed for a small generic cellular network with its own dedicated spectrum, i.e. without the presence of the primary users.

The rest of the paper is organised as follows: Section II introduces the temporary event scenario investigated in this study. Section III explains distributed Q-learning based DSA. In Section IV we introduce the concept of case-based RL and propose a case-based Q-learning scheme for dynamic secondary spectrum sharing. Simulation results are discussed in Section V, and the conclusions are given in Section VI.

## II. Temporary Event Scenario

The spectrum sharing problem investigated in this paper and currently considered in the EU FP7 ABSOLUTE project is depicted in Figure 1. It is designed for a stadium event scenario and involves a temporary cognitive cellular infrastructure that is deployed in and around a stadium to provide extra capacity and coverage to the users and event organizers involved in a temporary event, e.g. a football match or a concert.

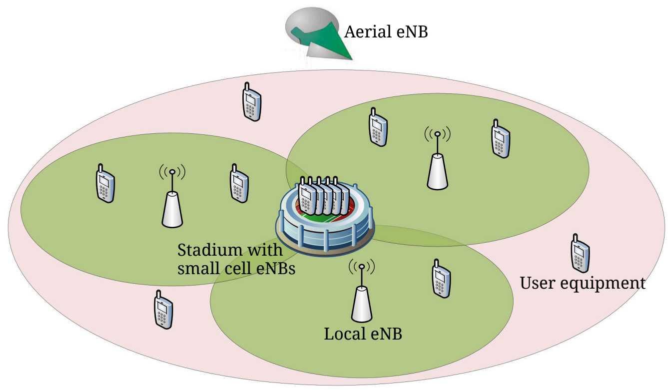



Figure 1. Stadium temporary event scenario

The cognitive small cells and the AeNB have secondary access to a 20 MHz LTE channel, also used by a network of 3 local primary eNBs (PeNBs). Furthermore, in this paper we consider a dynamic topology case, where the AeNB can be switched on and off several times throughout the duration of the event. For example, it can be switched on for providing the event organizers with a dedicated access network when required, and switched off to have its batteries recharged or to minimise the energy consumption in general.

## III. Distributed Q-Learning Based Dynamic Spectrum Access

One of the most successful and widely used RL algorithms is Q-learning [3]. In particular, a simple stateless variant of this algorithm, as formulated in [6], has been shown to be effective for several distributed DSA problems, e.g. [5][7].

Each eNB maintains a Q-table $Q(a)$ such that every subchannel $a$ has an expected reward or Q-value associated with it. Upon each file arrival, the eNB either assigns a subchannel to its transmission or blocks it if all subchannels are occupied. It decides which subchannel to assign based on the current Q-table and the greedy action selection strategy described by the following equation:

$$\hat{a} = \underset{a}{\operatorname{argmax}}(Q(a)),\ a \in A', A' \subset A \tag{1}$$

where $\hat{a}$ is the subchannel chosen for assignment out of the set of currently unoccupied subchannels $A'$, $Q(a)$ is the Q-value of subchannel $a$, and $A$ is the full set of subchannels.

The values in the Q-tables are initialised to zero, so all eNBs start learning with equal choice among all available subchannels. A Q-table is updated by an eNB each time it attempts to assign a subchannel to a file transmission. The recursive update equation for stateless Q-learning, as defined in [6], is given below:

$$Q(a) \leftarrow (1 - \alpha)Q(a) + \alpha r \tag{2}$$

where $r$ is the reward associated with the most recent trial and is determined by a reward function, and $\alpha \in [0, 1]$ is the learning rate parameter which weights recent experience with respect to previous estimates of the Q-values. The choice of these parameters is described in [7].

## IV. Case-Based Reinforcement Learning

Case-based RL is a combination of RL and case-based reasoning (CBR), where the solutions to previously known problems are used for helping to learn solutions to new problems [4]. For example, in [5] we apply this technique to make the base stations of a small cellular network learn appropriate spectrum assignment policies for three distinct network topology phases. Algorithm 1 shows our proposed adaptation of this case-based Q-learning scheme to the dynamic secondary spectrum sharing scenario described in Section II. The functionality afforded by CBR, as an extension to classical RL, is described by steps 4, 5, and 9 of Algorithm 1.

**Algorithm 1** Subchannel assignment using case-based Q-learning for dynamic secondary spectrum sharing

1: **if** all subchannels are occupied **then**
2:     Block transmission
3: **else**
4:     *Identify current case (AeNB is on/off)*
5:     *Choose Q-table most suitable for the identified case*
6:     Assign the best subchannel using Equation (1)
7:     Observe the outcome, calculate the reward $r = \pm 1$
8:     Update $Q(a)$ using Equation (2)
9:     *Store Q-table and associate it with current case*
10: **end if**

## V. Simulation Results and Discussion

The spectrum sharing problem described in Section II involves an AeNB and a network of small cell eNBs that have to share spectrum among themselves and with a primary system of local eNBs operating in the area.

The primary system is assumed to employ a dynamic ICIC scheme, where all three eNBs exchange their current spectrum usage as RNTP messages every 20 ms, and exclude the subchannels currently used by the other two eNBs from their available subchannel list [8].

### A. Simulation Setup

The stadium small cell network architecture [7] is such that the users are located in a circular spectator area 53.7 - 113.7 m from the centre of the stadium. The spectator area is covered by 78 eNBs arranged in three rings at 1 m height, e.g. with antennas attached to the backs of the seats or to the railings between different row levels. Seat width is assumed to be 0.5 m, and the space between rows is 1.5 m, which yields the total capacity of 43,103 seats. 25% of the stadium capacity is filled with randomly distributed wireless subscribers, i.e. $\approx$ 10,776 user equipments (UEs). 500 UEs are randomly distributed outside the stadium in a circular area from the stadium boundary out to 1.5 km from the stadium centre point. The offered traffic is 20 Mb/s outside of the stadium and 1 Gb/s inside. The other key parameters and assumptions of the simulation model are listed in Table I.

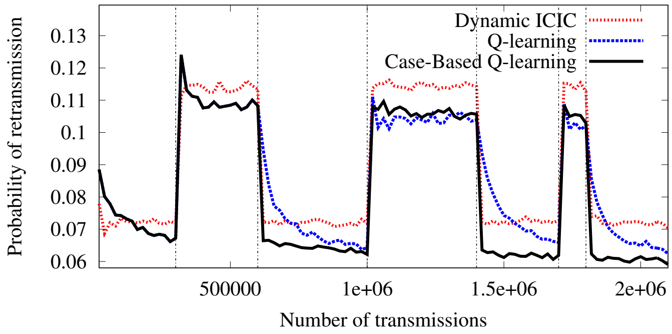
TABLE I. Network model parameters and assumptions

| Parameter | Value |
|---|---|
| Channel bandwidth | 20 MHz: 100 LTE virtual resource blocks (VRBs) |
| Subchannel bandwidth | 4 VRBs: 4 $\times$ 180 kHz |
| Frequency band | 2.6 GHz |
| UE receiver noise floor | 94 dBm |
| Stadium propagation model | WINNER II B3 [7] |
| Outdoor propagation model | WINNER II C1 [7] |
| Propagation model between stadium and outdoors | Combined WINNER II C4 with C1 term [7] |
| Propagation model between AeNB and the ground | Free space + 8dB log-normal shadowing |
| Traffic model | 3GPP FTP Traffic Model 1, file size - 4.2 Mb ($\approx$0.5 MB) [7] |
| Retransmission scheduling | Uniform random back-off between 0 and 960 ms |
| Link model | 3GPP Truncated Shannon Bound model [7] |
| Primary eNB Tx power | 10 dBW |
| **Assumptions** | |
| Cognitive small cell and aerial eNBs employ open loop power control, using a constant Rx power of -74 dBm (20 dB Signal-to-Noise Ratio) | |
| The minimum Signal-to-Interference-plus-Noise Ratio (SINR) allowed to support data transmission is 1.8 dB | |
| One subchannel (4 VRBs) is allocated to every data transmission | |

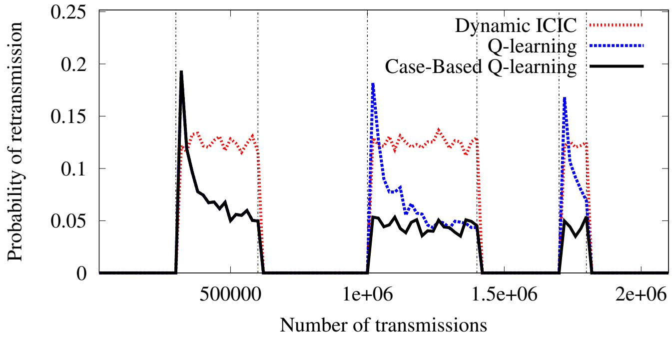
The cognitive small cell network and the AeNB which is located above the stadium centre point at 300 m altitude have secondary access to a 20 MHz LTE channel also used by the primary network. The latter consists of 3 primary eNBs (PeNBs) whose coordinates, with respect to the stadium centre point, are $(-600, -750)$, $(100, 750)$ and $(750, -800)$ m.

### B. Temporal Performance

Figure 2 shows the average temporal performance of the secondary network in terms of its probability of retransmission ($P(retransmission)$). The plots are obtained by averaging every data point using the results from 50 simulations with




(a) Stadium small cell network




(b) Aerial eNB

Figure 2. Probability of retransmission in the secondary cognitive network

different randomly generated UE locations and initial traffic. All simulations start with the AeNB being switched off. The vertical dash-dot lines in Figure 2 represent the times at which the AeNB is switched on and back off again.

Figure 2a shows how well the stadium small cell network adapts to the sudden irregular changes in its environment caused by the AeNB being switched on/off. It demonstrates that the "case-based Q-learning" scheme proposed in this study can seamlessly switch between the two different cases of the environment, compared with the classical Q-learning approach which has to adapt its policies anew every time.

The difference in performance between the scheme proposed in this paper and the two baseline schemes is even more substantial in Figure 2b, which shows the average $P(retransmission)$ temporal response of the AeNB. Firstly, both learning schemes significantly outperform the purely heuristic "dynamic ICIC" approach, same as that used in the primary system. Secondly, compared with classical Q-learning, the novel CBR functionality implemented in all cognitive eNBs results in a 70% reduction in $P(retransmission)$ experienced by the AeNB users shortly after the AeNB is switched on for the second time and all subsequent times.

### C. Primary User Quality of Service

An essential requirement for cognitive cellular systems is to ensure that they do not have a harmful effect on the QoS in the primary system. Table II compares the QoS provided to the users outside of the stadium with and without the presence of the stadium users and the secondary network.

The results in Table II show that it is possible to develop a temporary heterogeneous cognitive network that is capable of servicing a dramatic increase in the offered traffic (1 Gb/s in addition to the original 20 Mb/s, i.e. by a factor of 51), but with no need for additional spectrum and with no degradation in the primary user QoS.

TABLE II. PRIMARY USER QUALITY OF SERVICE (QOS) WITH AND WITHOUT THE PRESENCE OF THE SECONDARY NETWORK (SN)

| **QoS metric** | **Without SN** | **With SN** |
|---|---|---|
| Probability of retransmission | $3.0 \times 10^{-3}$ | $3.4 \times 10^{-3}$ |
| Mean user throughput (UT), Mb/s | 3.04 | 3.07 |
| 95th percentile UT, Mb/s | 3.16 | 3.16 |
| 5th percentile UT, Mb/s | 2.70 | 2.89 |
| Mean UT 0-100 m from the stadium, Mb/s | 2.96 | 2.89 |

## VI. CONCLUSION

The case-based RL method proposed in this study is an effective and feasible approach to dynamic secondary spectrum sharing in temporary cognitive cellular systems with dynamic topologies. System level simulations that involve a stadium small cell network, an eNB on an aerial platform and a local primary LTE network show that augmenting RL with the CBR functionality results in increased adaptivity of the cognitive cellular system to sudden changes in its radio environment, caused by the aerial eNB being dynamically switched on and off. For example, it is capable of achieving a 70% reduction in the number of retransmissions of the aerial eNB shortly after being switched on, compared to a classical RL approach. Furthermore, the cognitive cellular system, that employs the proposed DSA scheme with only secondary access to an LTE channel, is shown to accommodate a 51-fold increase in the offered traffic with no need for additional spectrum and with no degradation in the QoS of the primary users.


## ACKNOWLEDGMENT

This work has been funded by the ABSOLUTE Project (FP7-ICT-2011-8-318632), which receives funding from the 7th Framework Programme of the European Commission.


## REFERENCES


[1] J. Sachs, I. Maric, and A. Goldsmith, "Cognitive cellular systems within the TV spectrum," in *IEEE Symposium on New Frontiers in Dynamic Spectrum*, 2010.

[2] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA, 1998.

[3] C. Watkins, "Learning from Delayed Rewards," Ph.D. dissertation, University of Cambridge, England, 1989.

[4] I. Watson, "Case-based reasoning is a methodology not a technology," *Knowledge-Based Systems*, vol. 12, no. 56, pp. 303 – 308, 1999.

[5] N. Morozs, D. Grace, and T. Clarke, "Case-based reinforcement learning for cognitive spectrum assignment in cellular networks with dynamic topologies," in *Military Communications and Information Systems Conference (MCC)*, 2013.

[6] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," in *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*, 1998.

[7] N. Morozs, D. Grace, and T. Clarke, "Distributed Q-learning based dynamic spectrum access in high capacity density cognitive cellular systems using secondary LTE spectrum sharing," in *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2014.

[8] S. Sesia, M. Baker, and I. Toufik, *LTE-The UMTS Long Term Evolution: From Theory to Practice.* John Wiley & Sons, 2011.