Short Paper: Intelligent Dynamic Spectrum Access in Cellular Systems with Asymmetric Topologies and Non-Uniform Traffic Loads

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*—This paper assesses the robustness of the distributed reinforcement learning (RL) approach to dynamic spectrum access (DSA) in cellular systems with asymmetric topologies and non-uniform offered traffic distributions. Large scale simulations of a stadium small cell LTE network, employing a distributed Q-learning based DSA scheme, show that such asymmetries in the network environment cause no degradation of the QoS provided to any part of the network.**

***Keywords*—Distributed Reinforcement Learning; Dynamic Spectrum Access; Asymmetric Topology; Non-Uniform Traffic**

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, one of the key requirements for future 5G systems is to achieve factor-one reuse of the spectrum [1]. Therefore, there is an inherent need for DSA techniques in such systems to mitigate the effects of inter-cell interference on the system throughput and the QoS provided to the mobile subscribers.

The cellular system used for simulation experiments in this paper is designed for a stadium event scenario, where a small cell LTE network is installed in a large stadium to provide an increase in mobile data capacity to the users attending the event. The network architecture is depicted in Figure 1. There, the key feature of the problem investigated in this paper is the asymmetric topology of the network caused by a localised traffic hotspot area which requires more eNodeBs (eNBs) to serve it. Within that area the small cell eNBs are deployed three times as densely as in the rest of the network, where the offered traffic is significantly lower.

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. One of the recently investigated RL algorithms which produced excellent results in DSA scenarios similar to that depicted in Figure 1 is the distributed stateless Q-learning approach, e.g. [2]. However, its performance has only been assessed using symmetric topologies and uniform traffic loads. The purpose of this paper is to investigate the

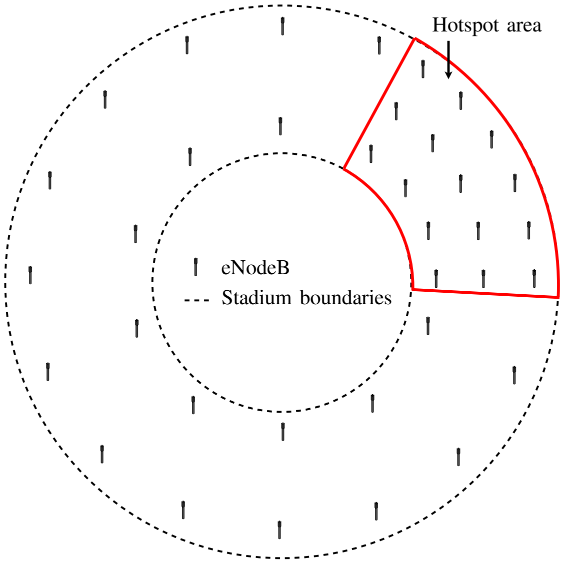


Figure 1. Stadium small cell network architecture with a hotspot area

adaptivity and robustness of this algorithm in cellular systems with asymmetric topologies and non-uniform offered traffic distributions, using the stadium small cell network from Figure 1 as an example.

The rest of the paper is organised as follows: Section II briefly introduces distributed Q-learning based DSA. The novel simulation results are discussed in Section III, and the conclusions are given in Section IV.

II. Distributed Q-Learning Based Dynamic Spectrum Access

One of the most successful and widely used RL algorithms is Q-learning. In particular, a simple stateless variant of this algorithm, as formulated in [3], has been shown to be effective for distributed DSA problems, e.g. [2]. 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 transmission. The recursive update equation for stateless Q-learning, as defined in [3], 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 discussed in [2].

III. Simulation Results and Discussion

The results presented in this section show the QoS provided to the users of the stadium small cell network from Figure 1, when it employs the distributed Q-learning DSA algorithm described in the previous section. The simulations start by having a uniform 14 Gbps/km^2 offered traffic density across the whole network with 1/3 of the eNBs switched on as depicted in the area outside of the hotspot zone in Figure 1. After 100,000 transmissions, the hotspot area shown in Figure 1 appears with the offered traffic density of 29 Gbps/km^2 followed by the activation of the additional eNBs (a detection delay of 5,000 transmissions is assumed). The overall simulation length is 1,000,000 transmissions. All other parameters and assumptions of the stadium network simulation model are described in detail in [2].

Figure 2 shows how the probability of retransmission changes throughout the duration of the simulated scenario. The plots were obtained by taking an average over 50 simulations with different random seeds, user locations and initial traffic. Firstly, it shows that there is no significant difference in the QoS provided to the whole network and the hotspot area inspected individually. A very similar learning curve is observed in both cases. Secondly, it also shows that after 100,000 transmissions, when the localised hotspot zone appears, there is no degradation of the network-wide QoS. This demonstrates how well the distributed Q-learning based DSA algorithm deals with such sudden changes in the network environment, without any centralised coordination involved.

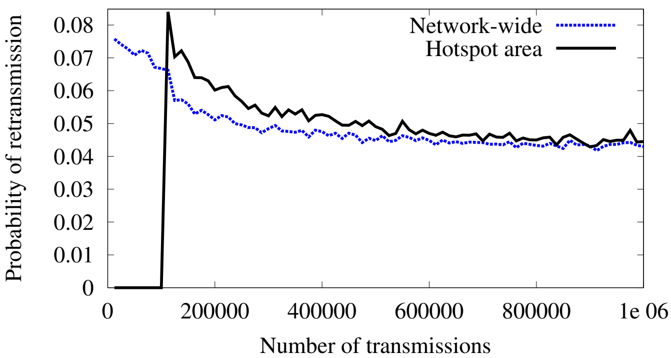
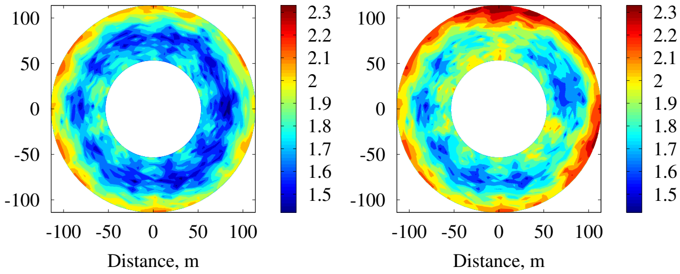


Figure 2. Probability of retransmission time response of the whole network and within the hotspot area only



(a) Without the hotspot area, Mb/s

(b) With the hotspot area, Mb/s

Figure 3. Spatial distribution of user throughput (UT) without and with the presence of a traffic hotspot area

The contour plots in Figure 3 show the spatial distribution of user throughput (UT), i.e. data rates, provided by the stadium small cell network, with and without the presence of the traffic hotspot area and the asymmetry in the network topology shown in Figure 1. Firstly, despite the introduction of the significant localised increase in the offered traffic and the change in network topology, no notable asymmetry in the UT provided to the users across the stadium is observed. Secondly, the network-wide QoS did not degrade due to the introduction of this asymmetry. On the contrary, there is a slight improvement in UT in parts of the stadium network due to a decrease in the size of the cells and the interference range reduction of the eNBs within the hotspot area. This further demonstrates that the distributed Q-learning approach to DSA is just as effective in asymmetric network scenarios, as it is in previously investigated symmetric and uniform spectrum management problems.

IV. Conclusion

The distributed stateless Q-learning approach to DSA is highly adaptable to asymmetries in the network topology and offered traffic distribution. Large scale simulations of a stadium small cell LTE network show that such an asymmetry in the radio environment causes no degradation of the QoS achieved by the intelligent DSA algorithm, with no centralised coordination required.

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] W. Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5G cellular networks," *Communications Magazine, IEEE*, vol. 52, pp. 52–60, 2014.

[2] 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.

[3] 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.