Cooperative Cognitive Network Slicing Virtualization for Smart IoT Applications

Fotis Foukalas *, Senior Member IEEE, Athanasios Tziouvaras *† and George Karetsos ‡

† Department of Digital Systems, University of Thessaly, Larissa, Greece
‡ Department of Electrical and Computer Engineering, University of Thessaly, Volos, Greece
*† Cognitive Innovations, Athens, Greece
Email: ‡ foukalas@ieee.org

Abstract—This paper proposes the cooperative cognitive network slicing virtualization solution for smart Internet of things (IoT) applications. To this end, we deploy virtualized small base stations (vSBSs) in SDR devices that offer network-slicing virtualization option. The proposed virtualized solution relies on Fed4Fire wireless experimental platform. In particular, we assume that multiple IoT devices can have access to different vSBSs, which coordinate their resources in a cooperative manner using machine learning (ML). To this end, a proactive resource management is deployed in the unlicensed band, where a cooperative solution is facilitated using the licensed band. The cooperative network slicing is managed and orchestrated using small cell virtualization offered by the Fed4Fire platform. Experimental trials are carried out for certain number of users and results are obtained that highlight the benefit of employing cooperative cognitive network slicing in future virtualized wireless networks.

Index Terms—Network slicing, cooperative machine learning, wireless network virtualization, NB-IoT.

I. INTRODUCTION

Dynamic network slicing can facilitate the efficient radio resource management among different type of devices and communication technologies [1]. A static approach on network slicing is not recommended to advanced virtualized wireless networks enabled by network function virtualization (NFV), where the dynamic radio resource management is carried out in an agile way [2]. As such a dynamic network slicing virtualization approach is encouraged, that is driven by specific quality of service (QoS) requirements per wireless communication technology [3] and is also applied on heterogeneous cloud radio access networks (C-RAN) [4]. This flexible type of radio resource allocation is provided by cognitive applications thanks to machine learning (ML) advances for wireless spectrum knowledge. Towards such a cognitive network slicing virtualization, a cooperation among the base stations is beneficial allowing the different cells to decide about the available resources collectively. This approach of Cooperative cognitive network slicing is on demand in case of unlicensed spectrum band as previous in [5] demonstrates.

Specifically, authors in [2] consider the network virtualization as enabler for network slicing, where each slice can achieve the objectives for its own virtual network performance. The dynamic network slicing has been identified in [1] and [6], where the former is enabled through fog computing collecting and analyzing data closely to the devices and the latter deals with the network slicing adaptations for different type of air interfaces such as NB-IoT and eMBB. This approach is called cross-domain network slicing, where each different domain, e.g. IoT and multimedia technologies, is managed by NFV management and orchestration entities [7]. Within the same concept, in [8], authors deal with multi-tenant network slicing for spectrum management, while in [9] deal with the statistical multiplexing at the physical layer. Spectrum management for network slicing is also important and especially in the unlicensed spectrum as highlighted recently in [10].

Moreover, in [3], authors provide a solution on network slicing to the edge. In [11], authors propose conventional radio resource management for NB-IoT devices, while in [4] authors propose the dynamic network slicing for multi-tenant heterogeneous C-RAN. In [12], authors explore network slicing for guaranteed rate services as admission control and resource allocation, through game theory. In [13] and [14], the authors deal with the network slicing in Industry 4.0 applications, while in the [15] authors propose the network multi-tenancy through a 5G network slice broker. Moreover, the authors in [16], they did not work in the network slicing virtualization applying also ML to the unlicensed band.

It is obvious that the cooperative proactive radio resource management on the unlicensed band through cooperative machine learning has not been proposed yet. In this work, we assume the virtualization of the network slicing for smart IoT applications over an experimental platform. The network slicing is considered on each virtualized eNodeB (vNB) that is deployed on an SDR infrastructure. The SDR nodes operate on the unlicensed spectrum band providing connectivity to multiple IoT devices in order to support smart agriculture and grid applications. It is already known, the 5G network slices will drive different types of vertical industries in order to isolate the communication among different applications [17]. The spectrum allocation is provided in a cooperative fashion, where the cooperation takes place over the licensed band. In order to achieve a proactive radio resource management, we employ cooperative machine learning by applying double Q learning (D-QL) over the transmit power and interference. The
cooperative network slicing is managed by network slicing instances (NSI) at the virtualization management and orchestration layer. The approach is applied to the Fed4Fire infrastructure and experimental results are available for discussion [18]. In particular, we obtain experimental results by deploying our solution on Fed4FIRE experimental large-scale platform.

The rest of this paper is organized as follows. In Sec.II, we describe the network slicing virtualization for smart IoT applications. In Sec.III, we present the cooperative network slicing virtualization for smart IoT applications. Sec.IV provides the experimental setup and results. Finally, Sec.V provides a summary of this work.

II. NETWORK SLICING VIRTUALIZATION FOR IoT APPLICATIONS

In Fig.1, we assume multiple virtualized small base stations (vSBSs) deployed using the concept of wireless network virtualization [19]. Each vSBS accommodates multiple network slices, which are deployed transparently on the physical SDR devices (i.e. SBSs). Each SBS employs one vSBS on the unlicensed band and one on the licensed band, where the former is dedicated to the ML in the unlicensed band that is giving access to the IoT devices through network slices and the latter to provide cooperation among the vSBSs. There is no network slicing on the licensed band since it is dedicated to cooperative communication only. Cooperative machine learning (CoML) is deployed over the considered network slicing virtualization dedicated to the unlicensed band that the IoT devices are connected. Machine learning (ML) is also considered a new technology for the next generation wireless access networks [20]. Thus, it provides proactive resource allocation through spectrum sensing, utility evaluation and carrier selection in order a SBS can learn the spectrum utilization history in advance [21]. The proactive resource management for 5G in unlicensed spectrum can be used for different type of IoT applications (e.g. smart agriculture and grid

1). The different SBSs can communicate with the centralized SBS via back-haul links that employ in-band wireless small cell communication [24]. Details about the modeling of the cooperative network slicing virtualization for IoT applications are given in the text below.

III. COOPERATIVE NETWORK SLICING VIRTUALIZATION FOR SMART IoT APPLICATIONS

A. Modeling and Algorithms

We consider \( i \) network slices on each vSBS, with NSI assigned \( n \) group of channels containing \( j \) component carriers each. As a result, each vSBS is allocated \( CC = i * n * j \) component carriers. In particular, we assume one vSBS transmitting over the licensed band and one over the unlicensed one as following:

\[ Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha) * Q_t(s_t, a_t) + \alpha [r(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a)] \]

with \( \alpha \) being the learning rate, \( r \) the accumulated reward received by performing action \( a_t \) when being in state \( s_t \) and \( \gamma \) the discount factor that accounts for future rewards to the current state. For the purposes of this work future states do not affect current decision making, instead we rely on the historical

- **Licensed spectrum.** The vSBSs utilize a licensed spectrum band in order to communicate and enforce cooperation among the SBSs using the licensed band.

- **Unlicensed spectrum.** The vSBSs utilize an unlicensed spectrum band to communicate with the IoT devices. To this end, listen before talk (LBT) and discontinuous transmission (DTX) mechanisms are deployed including spectrum sensing in order to obtain channel occupancy time (COT) as well according to [28].

Previous work in [25] elaborates on channel conflict among multiple IoT devices and shows that it may significantly reduce IoT throughput performance. Authors in [25] try to maximize the channel utilization rate and allocate each channel’s resources efficiently among the participating IoT devices. In this work, we also adopt such an approach and we use machine learning techniques to improve its efficiency. More specific, to account for channel resource utilization by the IoT devices we focus on minimizing IoT transmission conflicts by coordinating the channel selection process by employing the Q-Learning (QL) technique. Our solution utilizes QL so that to maximize the channel utilization according to the history of each channel’s occupancy time as in [26]. QL generally utilizes a set of actions and accumulated rewards that are updated according to the following equation:

\[ Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha) * Q_t(s_t, a_t) + \alpha [r(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a)] \]

...
behavior patterns of the channels to make our decisions. As a result the future states do not affect our decision making and thus, the future discount factor $\gamma$ is set to zero resulting in:

$$Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha) * Q_t(s_t, a_t) + \alpha (r(s_t, a_t))$$

with $s_t$ being the state under examination. An $s_t$ state contains a set of actions $a_t$ that the vSBS may perform. Such actions include the channel selection and the amount of subframes to be transmitted and are defined as follows: $a = \{a_t\}$ select channel $a$ for $i$ subframes, where $i \in \{1, 2, 3 \ldots 10\}$. Now, we search for the optimal set of actions $a_t$ for the $s_t$ set in order to achieve the maximum channel utilization while also avoiding channel signal interference. For this reason QL utilizes the accumulated knowledge on the network traffic of each channel so that it can learn its behavioral patterns. Channel utilization is a major factor that leads to throughput improvement as highlighted by previous work in [27]. In this work, authors manage to improve the throughput of an IoT network by studying and improving the channel utilization rate by the IoT devices. In our solution we adopt a similar approach of channel maximization by employing the QL algorithm. More specific, we opt to optimize the channel utilization rate for the next $Q_{t+1}$ QL state. To accomplish this, we maximize the next state $s_{t+1}$ according to the following equation:

$$s_{t+1} = \arg\max_{\forall a} Q_t(s_t, a_t)$$

Solving the maximization function 3 will lead to a $r(s_t, a_t)$ reward that will be used to update the Q-function in equation 2. We define such reward as the difference between the average and the observed channel IDLE time after the action $a$ took place:

$$r(s_t, a_t) = \text{IDLE}_n\text{avg} - \text{IDLE}_n\text{measured}$$

By measuring a small channel IDLE time, we conclude that the previous action $a$ resulted in high channel utilization rate, meaning that such decision should be rewarded well. Thus, the less the channel IDLE time, the better the channel reward. We define $n$ channel’s idle time as follows:

$$\text{IDLE}_n = (1 - \text{COT}_n)$$

with $\text{COT}_n$ being the COT of channel $n$ which is the percentage of time the channel is measured to be occupied. Such measurement is obtained by the spectrum sensing. In order also to consider transmit power control for the deployed SBSs, we extend the QL reward function to include power allocation information. To incorporate the new reward function under the ML algorithm we employ the DQL approach similarly to [28]. Thus, the reward function is formulated as follows:

$$r(s'_t, \{a_t, p_t\}) = \sigma (\text{IDLE}_n\text{avg} - \text{IDLE}_n\text{measured}) + (1 - \sigma) \beta p_{max}$$

where $p_t$ is the power selection for SBS transmission in channel $t$, the $\sigma$ is a weighting factor between the channel IDLE time and transmit power values, $\beta$ normalizes the power levels to IDLE channel values and $p_{max}$ represents the highest power measurement obtained during the previous Q iteration. The power selection $p_t$ action is defined as follows: $p = \{p_i\}$ select channel $p$ using $i$ dBm transmission power $i \in \{1, 2, 3 \ldots 10\}$, where $p_i$ is the normalized transmit power of the vSBS. We also formulate the D-QL function $Q'$ to contain power control as follows:

$$Q'_{t+1}(s'_t, p_t) \leftarrow (1 - \alpha) * Q'_t(s'_t, p_t) + \alpha (r(s'_t, \{a_t, p_t\}))$$

Our goal is to find the optimal action $p_t$ for the current set $s'_t$ in order to achieve the lowest possible transmission power, while maximizing the following formula:

$$s'_{t+1} = \arg\max_{\forall p} Q'_t(s'_t, p_t).$$

Algorithm 1 below provides non-cooperative D-QL (NC-DQL). NC-DQL does not require cooperation between SBSs and thus each SBS runs an independent instance of the algorithm. In DQL scheme, the $Q$ function is randomly updated either with power transmission values as in equation 7 focusing in maximizing the 8 or with channel IDLE time measurements as in equation 2 focusing in maximizing the function 3. In case of non-cooperative QL (NC-QL) implementation, algorithm 1 is simplified as NC-QL utilizes one $Q$ function that is related to the channel IDLE time information only. Such implementation utilizes the $Q$ function as described in 2 while focusing in 3 maximization. The final step of the algorithm is to use listen before talk (LBT) and discontinuous transmission (DTX) to the unlicensed band. LBT makes sure that the selected channel is not taken and DTX enforces the $s_t$ set of actions made by the ML algorithm.

**Algorithm 1 Non-Cooperative Double Q-Learning (NC-DQL)**

for all channels $c$ do
  Spectrum Sensing
  Calculate COT$_c$
  Randomly choose to update one of the following
  A. $Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha) * Q_t(s_t, a_t) + \alpha (r(s_t, \{a_t, p_t\}))$
     with $s_{t+1} = \arg\max_{\forall a} Q_t(s_t, a_t)$
  B. $Q'_{t+1}(s'_t, p_t) \leftarrow (1 - \alpha) * Q'_t(s'_t, p_t) + \alpha (r(s'_t, \{a_t, p_t\}))$
     with $s'_{t+1} = \arg\max_{\forall p} Q'_t(s'_t, p_t)$
end for
  Perform LBT
  Perform DTX

Algorithm 2 provides the proposed cooperative Double Q-Learning (C-DQL). In contrast with individual learning, cooperative learning utilizes channel and power information by every available SBS while also incorporating such information in one global learning function as in [29]. In this work authors utilize a global QL function which aggregates the information collected by the various agents. We employ a similar approach but we also focus on the distributed QL similar to [30] as each
SBS updates its local Q-states individually and thus, integrating all previously selected actions and their corresponding rewards. Under this premise each SBS-i transmits its local Q state \( Q_i(s_t, a_t) \) (or \( Q'_i(s'_t, p_t) \) in case of transmission power control) information to the centralized SBS where the maximization of the global Q function takes places as follows:

\[
s_{t+1} = \sum_{n=0}^{n=i} \arg \max_{s_t} Q_n(s_t, a_t) \tag{9}
\]

or when considering transmit power control:

\[
s'_{t+1} = \sum_{n=0}^{n=i} \arg \max_{s'_t} Q'_n(s'_t, p_t) \tag{10}
\]

Finally, the centralized SBS transmits the final decision to all remaining SBSs in order for them to allocate the resources accordingly. In cooperative QL scheme (C-QL) the same procedure with C-DQL follows except that the function 2 is used as global Q function and the following equation is maximized:

\[
Q_{t+1}(s_t, a_t) = \sum_{n=0}^{n=i} \arg \max_{s_t} Q_n(s_t, a_t).
\]

**Algorithm 2 Cooperative Double Q-Learning (C-DQL)**

```plaintext
for all channels c do
  Spectrum Sensing
  Calculate COT
  Receive \( Q_t \) information from all i-SBSs
  randomly choose to update one of the following
  A. \( Q_{t+1}(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q_t(s_t, a_t) + \alpha \cdot r(s_t, \{a_t, p_t\}) \)
  with \( s_{t+1} = \sum_{n=0}^{n=i} \arg \max_{s_t} Q_n(s_t, a_t) \)
  B. \( Q'_{t+1}(s'_t, p_t) \leftarrow (1 - \alpha) \cdot Q'_t(s'_t, p_t) + \alpha \cdot r(s'_t, \{a_t, p_t\}) \)
  with \( s'_{t+1} = \sum_{n=0}^{n=i} \arg \max_{s'_t} Q'_n(s'_t, p_t) \)
end for

Broadcast decisions made to other SBSs
Perform LBT
Perform DTX
```

**B. Cooperative NSI Implementation**

We explain below the implementation of the cooperative cognitive network slicing using the proposed virtualization framework to deploy the C-QL and C-DQL algorithms. Fig.2 depicts the cooperation scheme among the available SBSs. Each SBS-i performs spectrum sensing on its allocated i group of channels and then calculates the corresponding COT and IDLE values as described in equation 5. In the sequel, it transmits the \( Q_i(s_t, a_t) \) and \( Q'_i(s'_t, p_t) \) sets to the centralized SBS (cSBS) that acts as a coordinator as described in the previous section. The coordinator cSBS updates the global Q-sets according to the sets received by the rest of the SBSs and utilizes QL to make decisions for each SBSi transmission. The QL outputs contain information about channel selection, transmit power control and the amount of subframes to be transmitted on each channel. Next, that information is transmitted back to the rest of the SBSs. In order to ensure the autonomous operation of each SBS, we deploy LBT and DTX, which enable the transmission of a specific amount of subframes thus, enforcing the QL decisions made by the coordinator.

**IV. Experimental Setup and Results**

**A. Experimental Setup**

Fig.3 depicts the detailed set-up of the experiment. We deploy three B210 USRPs considered as SBSs with each device composed of two virtual instances veNB1 and veNB2 as in [19].

The veNB1 is capable of transmitting on the unlicensed spectrum, specifically in band 20 (791-821 MHz) on the downlink (DL) channel. The veNB1 is composed of three network slices each of them is assigned to a group of transmission channels.
For baseband processing, we used the srsLTE [31] software in order to create an NB-IoT implementation. Our choice of utilizing the LTE baseband is not arbitrary as previous work in [32] shows that NB-IoT devices use LTE design either as standalone operation mode or within the LTE spectrum. Our choice of LTE baseband for NB-IoT is also compatible with the 3GPP standardization as discussed in [33]. Finally, the veNB2 operates on the LTE licensed band (e.g. Band 38 of 2570-2620 MHz) to cooperate with other SBSs using the in-band communications capabilities [24].

Table I provides the srsLTE parameters and the configuration options selected for the experimental set-up. The baseband utilizes the sub-1GHz frequency band-20 with channel bandwidth of 180 KHz and carrier spacing of 15KHz. Sub-1GHz bands are used for low cost IoT devices due to ideal propagation conditions [5], while the 180KHz bandwidth is selected when considering IoT as highlighted in [34]. The aforementioned configuration results in 10 channels with 1 resource block (RB) each and 1 component carrier (CC). We also employ the standalone operation mode as described in [34] and [32] with single tone transmission mode and QPSK modulation type. QPSK modulation is utilized by the IoT downlink channel as the 3GPP standard overview of the release 14 shows in [33]. The maximum theoretical bandwidth is 2.5 Mbp/s per SBS, and thus 7.5 Mbp/s for the whole network. In order to satisfy the requirements for the cooperative D-QL algorithm, we specify a centralized SBS (cSBS) playing the role of controller, where the learning decision is made as described in Sec.II.

The coverage level within the IoT devices depends on the channel conditions. The extreme coverage level corresponds to a low power received value and a normal coverage level corresponds to a high power received value Each selected coverage class determines the transmission parameters including the number of repetitions. Such a deployment allows each IoT device in different coverage conditions characterized by different ranges of path loss. Depending on the coverage level, the serving cell indicates to the UE to repeat the transmission 1, 2, 4, 8, 16, 32, 64, 128 times, using the same transmission power on each retransmission. Combining the different retransmissions allows a coverage extension. For the purposes of this work, we make sure that the coverage level 0 and 1 criteria are met. As such we select the appropriate sub-carrier spacing configuration, i.e. 15 KHz and we guarantee that the maximum coupling loss (MCL) levels are between 144dB and 154dB. Thus, we assume three coverage classes for the NB-IoT devices linked to three SDR devices:

- CE level 0: normal coverage with MCL 144 dB and 15 kHz sub-carrier spacing.
- CE level 1: robust coverage with MCL 154 dB and 15 kHz sub-carrier spacing.
- CE level 2: extreme coverage with MCL 164 dB and 3.75 kHz sub-carrier spacing.

### Table I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>veNB1 Band</td>
<td>Band 38</td>
</tr>
<tr>
<td>veNB1 Frequency</td>
<td>2570 - 2620 MHz</td>
</tr>
<tr>
<td>veNB2 Band</td>
<td>Band 20</td>
</tr>
<tr>
<td>veNB2 Frequency (DL)</td>
<td>791 - 821 MHz</td>
</tr>
<tr>
<td>Channel Bandwidth</td>
<td>180 KHz</td>
</tr>
<tr>
<td>Carrier spacing</td>
<td>15 KHz</td>
</tr>
<tr>
<td># Channels</td>
<td>10</td>
</tr>
<tr>
<td># RBs</td>
<td>1</td>
</tr>
<tr>
<td># Component carriers</td>
<td>1</td>
</tr>
<tr>
<td>IoT Operation mode</td>
<td>Standalone</td>
</tr>
<tr>
<td>Transmission mode</td>
<td>Single tone</td>
</tr>
<tr>
<td>Modulation</td>
<td>QPSK</td>
</tr>
<tr>
<td>Transmit power</td>
<td>14 dBm</td>
</tr>
</tbody>
</table>

B. Experimental Results

We implement all QL alternatives in the experimental setup described above under various traffic conditions. We provide below the network traffic conditions and the different QL implementations considered in the course of the experimentation:

- Uniform traffic: the IoT traffic is equally distributed across the available channels.
- Non-uniform traffic: the IoT traffic is randomly distributed across the available channels.
- Low traffic: the IoT traffic occupies 25% of the channels.
- Medium traffic: the IoT traffic occupies 50% of the channels.
- High traffic: the IoT traffic occupies 75% of the channels.
- Non cooperative Q-Learning (NC-QL): the QL algorithm runs individually on each vSBS. In this setup the SBSs compete for the available network resources.
- Non cooperative Double Q-Learning (NC-DQL): the QL runs individually on each vSBS, also taking decisions for transmit power control under the D-QL paradigm.
- Cooperative Q-Learning (C-QL): the QL runs in a distributed way on each vSBS with the cSBS acting as the centralized controller which is responsible for cooperation between SBSs. Thus SBSs do not compete for network resources, instead they cooperate in order to maximize the resource utilization of the whole network.
- Cooperative Double Q-Learning (C-DQL): the QL runs on each vSBS in a distributed way as in C-QL, while also taking decisions for transmit power control by employing the D-QL algorithm.

To evaluate the performance of our setup, we measure the achievable throughput considering the different ML implementations. To this end, we assess the performance of each QL implementation by measuring the throughput achieved by the three SBSs collectively namely network throughput on the unlicensed band. Fig.4 depicts such results, where the following observations are made:

- C-DQL algorithm achieves the best throughput. This is expected as distributed learning takes into account inputs from each SBS and thus it fosters cooperation instead of SBSs competition. The result is an efficient coexistence scheme that maximizes the network throughput.

---

TABLE I

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>veNB1 Band</td>
<td>Band 38</td>
</tr>
<tr>
<td>veNB1 Frequency</td>
<td>2570 - 2620 MHz</td>
</tr>
<tr>
<td>veNB2 Band</td>
<td>Band 20</td>
</tr>
<tr>
<td>veNB2 Frequency (DL)</td>
<td>791 - 821 MHz</td>
</tr>
<tr>
<td>Channel Bandwidth</td>
<td>180 KHz</td>
</tr>
<tr>
<td>Carrier spacing</td>
<td>15 KHz</td>
</tr>
<tr>
<td># Channels</td>
<td>10</td>
</tr>
<tr>
<td># RBs</td>
<td>1</td>
</tr>
<tr>
<td># Component carriers</td>
<td>1</td>
</tr>
<tr>
<td>IoT Operation mode</td>
<td>Standalone</td>
</tr>
<tr>
<td>Transmission mode</td>
<td>Single tone</td>
</tr>
<tr>
<td>Modulation</td>
<td>QPSK</td>
</tr>
<tr>
<td>Transmit power</td>
<td>14 dBm</td>
</tr>
</tbody>
</table>
NC-QL is outperformed by other QL schemes, especially the C-DQL. NC-QL does not utilize channel information by other SBSs, instead it relies on the information provided by the host SBS only. As a result the three SBSs compete for the available channels, thus preventing the measured throughput.

To study the impact of the QL training to the SBS throughput, we measure the achievable throughput during subsequent QL iterations. Fig.5 depicts the results over 1000 QL training iterations. It is clear that C-DQL achieves the best performance per iteration when compared to other QL schemes. Also QL performance is analogous to the amount of training steps undertaken, and thus it requires some time to maximize its performance. Another observation is related to QL throughput improvement rate which is lowered when the algorithm closes to its convergence point. As a result, QL iterations from 900 to 1000 do not offer significant throughput increase compared to previous training steps.

In order to verify the efficiency of DQL, we measure the power consumption on the transmitter side on each vSBS under three different QL configurations (Fig.6). We observe that the C-DQL performs better in terms of power saving when compared to NC-DQL. C-DQL, due to its cooperative nature exploits spectrum power information obtained by every vSBS and then properly adjusts the transmission power, resulting in lowering the power consumption of each vSBS. Further, comparison between the NC-QL and the C-QL demonstrates the significant power reduction achieved by the latter. Fig.6 depicts the transmit power of the NC-QL algorithm as comparison baseline as it remains static during run time. Our choice for 14 dBm baseline power transmission is compliant with 3GPP release 14 as described in [33]. The 14dBm power class is introduced in release 14 and is ideal for low power NB-IoT devices.
In this work, we present the cooperative cognitive network slicing virtualization for smart IoT applications such as smart grid and agriculture. We deploy a cooperative machine learning over virtualized base stations that operate on the unlicensed band and cooperate on the licensed band. The radio resources are monitored and observed through spectrum sensing and the results are collectively made known to the different base stations. A centralized type of networking management is taken place using a cooperative double Q learning algorithm. Specifying the requirements of the different IoT devices and the wireless networking setup, experiments are carried out using the Fed4Fire wireless experimental platform. Experimental results are presented and discussed highlighting the impact of deploying a cooperative scheme among the virtualized network slicing.

Acknowledgement
This work has been funded by the Fed4Fire+ H2020 project with Grant agreement ID: 732638 under the Open Call 6.

References