On the Detection of Spectrum Irregularities through Deep Learning in Dense IoT architectures

Katerina Kyriakou, Kostas Chounos, Thanasis Korakis
Dept. of Electrical and Computer Engineering
University of Thessaly, Greece
{aikyriakou, hounos, korakis}e-ce.uth.gr

2023 IEEE Conference on Standards for Communications and Networking (CSCN) | 979-8-3503-9538-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/CSCN60443.2023.10453180

***Abstract*—In this work, a novel Deep Learning model that combines Q Learning and Neural Networks is proposed and experimentally evaluated. The proposed scheme is developed in order to detect and tackle the effects of spectrum anomalies, which are unexpectedly appeared in ultra-dense Internet of Things (IoT) architectures. The protocol which is taken into consideration in this work, is the IEEE 802.11ah (Wi-Fi HaLow) and it is tested under strong interference, caused between wireless links which transmit in non-overlapping frequencies. The proposed approach trains a model that constantly observes the wireless environment and obtains an optimal policy for the transmission time parameter re-configurations of the participating devices. The experimental evaluation showcases that the proper training of the proposed Deep Q Learning model, leads to remarkable increased Packet Delivery Ratio (PDR) and throughput in the examined scenarios. Apart from the improvements (+35%) observed on a PDR and throughput basis, the proposed algorithm also achieves overall higher channel utilization, increased transmission opportunities, and fairness in terms of channel access.**

***Index Terms*—IEEE 802.11ah, Iot Ultra-dense Networks, Interference Mitigation, Deep Q Networks**

I. Introduction

In recent years, the Internet of Things (IoT) has appeared, well-established and revolutionized the way that modern devices interact with each other. More specifically, the major goal for these architectures is to harmonically orchestrate a wide array of heterogeneous wireless and wired devices, by facilitating seamless connections and data exchange among them. These network topologies envision that all everyday objects, devices, and machines are interconnected and can communicate with each other over the Internet. In such a way, huge enhancements and improved efficiency may be achieved. From a networking aspect, some of the significant benefits which can be retrieved through the use of IoT architectures, are enhanced connectivity, reduced latency and energy efficiency for the involved devices. However, based on the large number of wireless devices which are involved in the IoT ultra-dense architectures nowadays, several major challenges arise as well. Thus, strategic planning is required in order to deal with the increased network management complexity.

As the number of devices which are accommodated within smaller geographical areas keeps increasing, the available frequency spectrum becomes more congested. This congestion can lead to interference, reduced data rates, and compromised performance, necessitating advanced spectrum management techniques. It is worth to be noted that both licensed (NB-IoT) and unlicensed spectrum bands (Wi-Fi HaLow, LoRaWAN, Sigfox, Zigbee) are utilized as a home for several IoT wireless protocols. Unlike the former case in which there are distinct ways of managing the spectrum, in the latter there is a high probability of reduced network performance, due to the parallel access of the same spectrum from many devices at the exact same time. In some protocols operating at the sub-1 GHz unlicensed bands like IEEE 802.11ah (Wi-Fi HaLow), mechanisms for the parallel channel's access have been already developed. Specifically, a MAC layer access control mechanism called Restricted Access Window (RAW) [1] pursues contention elimination since it divides the IEEE 802.11ah stations (STAs) into groups and performs medium access only for STAs belonging to the same RAW group over a specific time period. Further channel access control is applied for each individual RAW group by leveraging the enhanced distributed channel access (EDCA) mechanism, which supports differentiated and distributed access by using four QoS categories with varying priorities. As for the detection of interference, IEEE 802.11ah also performs the conventional Clear Channel Assessment (CCA), in order to determine whether there is an ongoing transmission or the channel is idle. However, in several cases, it is proved that the aforementioned mechanisms are not working as expected, and crucial performance degradations are caused in the involved wireless devices.

II. Related

A significant amount of research for interference mitigation techniques has been investigated in the literature for the IoT ultra-dense architectures. Regarding collisions elimination, most protocols utilize random access approaches such as Pure Aloha for channel access control and Slotted Aloha. LoRaWAN and Sigfox use Pure Aloha for channel access control, while NB-IoT uses Slotted Aloha. Initially, Bankov *et al.* [2] points out the performance weaknesses of the specific approach in LoRaWAN in highly dense networks, caused by the absence of a channel access control scheme. The author's purpose is to evaluate the performance through simulations and to propose potential solutions for the case of dense networks where performance is decreased. The results of concurrent transmissions without channel access mechanism denote negligible packet loss in small device load but impose significant success performance degradation as the amount of nodes increases. On the general comparison and evaluation of the ALOHA access scheme in different IoT protocols (LoRaWAN, Sigfox), the authors of [3] aim to obtain valuable information regarding the performance and its dependence on the total number of nodes through modeling and simulation. Through these simulations, the authors observe that both schemes suffer significantly in terms of collisions and packet error rate as the amount of IoT devices increases, similar to the previously mentioned research. However, the specific evaluation denotes that Sigfox is slightly improved compared to LoRaWAN simulation results. Despite that the channel access in the licensed IoT frequencies and protocols is a much more straightforward procedure, there are also works in the literature that are dealing with this problem. Indicatively, Harwahyu *et al.* [4] modeled random access in NB-IoT protocol as multi-channel slotted ALOHA, in order to analyze its behavior in terms of channel access success probability and average access delay. For results verification, the authors performed simulations that show significantly similar behavior between model analysis and simulation results. As for the evaluation of the channel access scheme, the simulations lead to the conclusion that it suffers both from high access delay and low access success probability as the number of devices increases, due to the complexity of interaction between the different coverage enhancement (CE) levels in NB-IoT.

Research community has recently also focused attention on the emerging IEEE 802.11ah protocol, in order to investigate interference mitigation schemes. Initially, Pandya *et al.* [5] propose a dynamic frequency allocation algorithm based on neighbor contention detection, that aims at interference elimination observed in dense environments. The specific model is evaluated in a developed Matlab simulator and the results of the paper denote higher throughput, lower energy consumption and, overall, improved channel utilization in dense topologies. Additionally, in [6] the authors suggest an efficient hierarchical MAC layer approach for multiple channel allocation in multiple relay nodes in IEEE 802.11ah. The results from the ns-3 simulator appear to maximize throughput and decrease interference, thus, allowing a large amount of devices to transmit concurrently without collisions and with larger transmission opportunities. Other interference mitigation approaches include grouping approaches. Naghzali *et al.* [7] proposes a non-orthogonal multiple access based grouping technique that utilizes successive interference cancellation, in order to improve scalability of IEEE 802.11ah networks by reducing potential collisions and improving throughput. The simulation results evaluate that the proposed grouping scheme benefits dense network cases significantly.

Finally, investigations have started focusing their attention on utilization of Machine Learning algorithms to optimize performance in IEEE 802.11ah. Mondal *et al* . [8] provides Machine Learning approaches using K-Means algorithm both to predict optimal RAW configuration parameters for current channel state and to acquire optimal policy selection for RAW configurations in variable traffic scenarios and heterogeneous stations. The proposed ML scheme is evaluated in simulator ns-3 and the results denote significant increase in throughput and PDR compared to traditional RAW application.

The comparison of existing approaches for interference mitigation in ultra-dense IEEE 802.11ah environments with this paper's proposed technique leads to the observation that all the research works mentioned above are solely evaluated through simulations, which include the use of either ns-3 or Matlab applications, while this paper's algorithm is evaluated in real hardware. Moreover, none of the existing research solutions examine the configuration of transmission time parameters for throughput optimization, but rely on optimal RAW parameters prediction. As for the ML approach requirements in data, the above mentioned paper uses K-Means where a large data set is required for training, while this paper benefits from Reinforcement Learning where data is collected through environment learning. To the best of our knowledge, this is the first work to do so, in which additionally the experiments are performed by utilizing real devices. In such a way, we are able to discover and examine potential spectrum irregularities which do not occur in simulation environments. Our main challenge is the optimal policy selection of transmission time parameters for back-off and carrier sense using ML techniques. This paper contributes in the following aspects:

- Detecting spectrum irregularities caused by densely deployed STAs, even they are configured to transmit in non-overlapping channels.
- Discovering the optimal TX time parameters for Carrier Sense and Backoff time determination to improve PDR and throughput in cases of high interference.
- Implementing and evaluating of a Reinforcement Learning approach that combines Q Learning and Neural Networks for real-time dynamic adjustment of TX time parameters in order to improve PDR performance.
- Observing significant improvements in PDR and throughput when the proposed model is applied for optimal policy discovery.

III. Framework

In the context of this work, we focus on maximizing the achieved performance of the IEEE 802.11ah (Wi-Fi HaLow) protocol in ultra-dense IoT architectures. The proposed framework consists of a centralized model where a Deep Reinforcement Learning unit endeavors to discover optimal transmission time configuration parameters for the involved IEEE 802.11ah nodes. The proposed architecture aims at the overall Packet Delivery Ratio (PDR) performance acceleration and fair spectrum sharing in dense topologies suffering from inexplicable collisions. Following at this section, a description of the proposed framework is given and thoroughly analyzed. More specifically, this includes the problem statement, the developed mechanism, as well as some of the key parameters for the examined hardware.

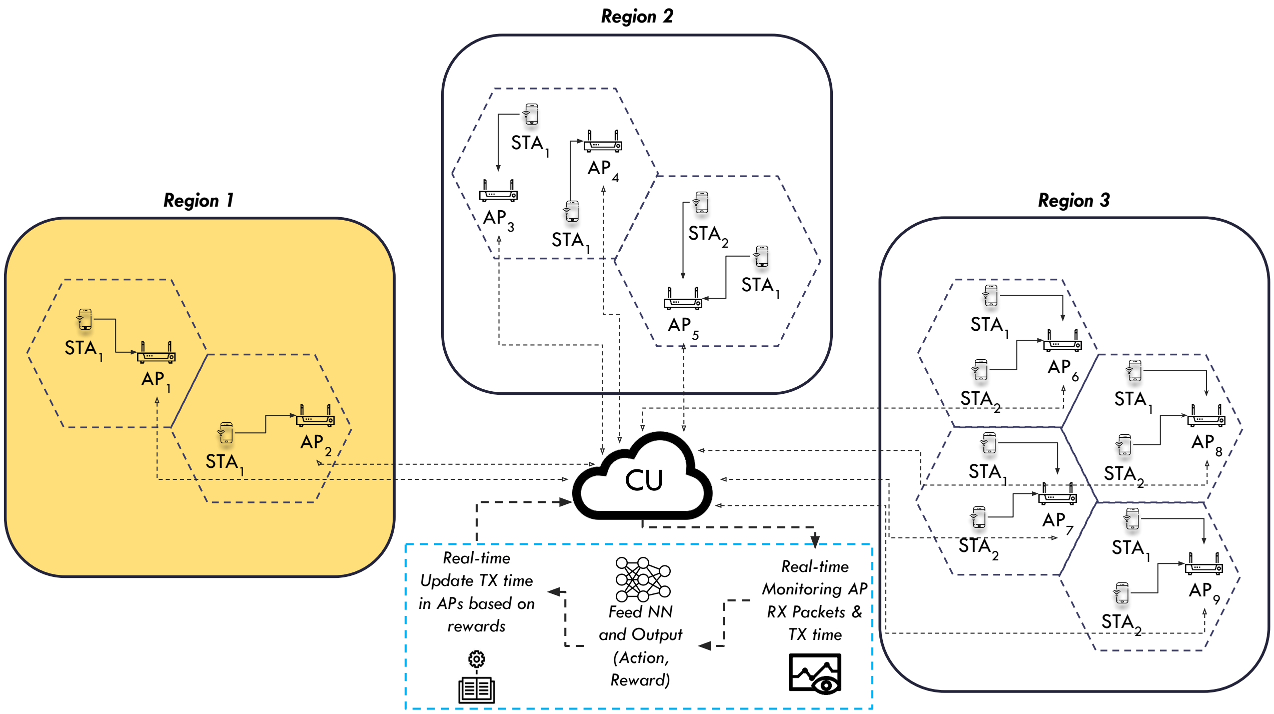


Fig. 1: Representation of the proposed model consisting of a central Reinforcement Learning unit and multiple stations and access points. The combination of stations and access points in the orange part is the architecture we focus on in this paper.

A. Problem Statement

Although the IEEE 802.11ah standard is designed to support a large number of devices operating at the same geographical regions and frequencies, severe interference scenarios may be observed, even in cases where only two pairs of nodes are transmitting simultaneously and close to each other. It is worth to be noted, that the impact of the interference which is examined in this work, is catastrophic even though the Access Points (APs) are configured to operate in non-overlapping wireless channels of the sub-1GHz unlicensed band as shown in Figure 2. There, it is also obvious that the total throughput is significantly decreased for both wireless links.

In this paper, a centralized Deep Q-Learning algorithm that addresses the problem described above is developed and experimentally evaluated. The proposed framework maximizes the transmission opportunities for both links and increases the social welfare of the network. Figure 1 shows the representation of the examined architecture, where a central remote unit, gathers the information about all stations and access points participating in the network. All the functionalities of the central unit may be easily executed from an everyday conventional PC (laptop, desktop etc). The considered system model includes N Station nodes, M Access points (More than one stations can transmit to the same access point) and 1 Deep Q Network Agent.

B. Deep Q Networks

Given the system model, our algorithm proposes a Reinforcement Learning approach [9], and specifically Deep Q Networks, in a centralized processing unit, which acts as a coordinator for the optimal configuration of transmission time parameters in IEEE 802.11ah nodes. Deep Q Networks rely on Q Learning, which is a Reinforcement Learning algorithm and the Neural Networks area of Machine Learning. Reinforcement learning is one of the main Machine Learning types and its objective focuses on developing self-trained agents targeting towards an optimal policy selection by interacting with the environment [10]. The learning procedure in this Machine Learning area consists of a series of actions for given environment states that aim for the maximization of a calculated reward that determines whether the action applied had a positive or a negative influence.

Q-Learning is an iterative Reinforcement Learning method for the discovery of an optimal or near-optimal policy, given states and corresponding possible actions, at discrete time. In Q-Learning, a sequential policy selection problem can be represented as $S_0, A_0, R_1, S_1, A_1, R_2, ..., S_t, A_t, R_{t+1}, t \in T$, where $S_t \in S$ at time step t is a state of the finite state set of observations of the environment situation, $A_t \in A$ is an action of the finite action set of updates to be performed for given state, and R_{t+1} is the reward calculated for the state S_{t+1} which is a result of the previously selected action A_t. At each iteration, the rewards are calculated and the Q-values stored in a data entity named Q-Table are updated until convergence. In cases of enormous amount of states and actions, Q-Table could lead to increased capacity demands, that would make Q Learning not applicable to real data scenarios.

Since the potential massive storage requirement of Q-Table is considered to be inefficient for many use cases, we adopt the Deep Q Network approach which is proposed at [11]. Deep Q Network maintains Q-Learning's features and propose the usage of Neural Networks instead of Q-Table for the approximation of Q-Values. The neural network accepts an environment state as an input and provides an output of every reward for each possible action that can be applied. The number of the proposed DQN model outputs is equal to the amount of actions. The weights are updated until convergence by selecting one of the output actions based on ϵ Greedy Exploration Strategy, measuring the corresponding reward, and forwarding it to the Neural Network.

C. IEEE 802.11ah Chipset Parameters

A key feature that we exploit during the formation of our algorithm is the transmission time parameters available in IEEE 802.11ah. Specifically, the NRC7292 devices [12] which are utilized in the context of this work, offer a command line application which allows users to monitor statistics and quality, and configure parameters for the transmissions in real time. Since our goal is the interference mitigation in dense networks, we focus on a specific configuration command which is related to Carrier Sense (CS) definition in terms of time. This parameter is the transmission (TX) time, configured only in station nodes, and consists of two sub-parameters:

- Carrier sense time $\in [0, 12480]$ in $\mu seconds$, the amount of time a station senses the environment for other signals detection before actually transmitting
- Blank time $\in [0, 4294967295]$ in $\mu seconds$, the amount of time a station waits before transmitting again, after a successful transmission

D. Proposed Model

A deep Q Network training algorithm is developed for our case study since the amount of possible actions is large due to the various combinations of Carrier Sense time and Blank time, and it could scale up even more for future scenarios with more complex deployments. The purpose of our algorithm is the formation of a Deep Q Network unit that, given finite sets of states and actions, is trained iteratively until convergence to the selection of TX time parameters for IEEE 802.11ah, that maximize PDR and impose fairness in channel utilization for dense networks.

Below, the definition of the Q-Learning's state, action and reward in our case study is given, where the number of stations $N = 2$:

- State set S, the combinations of TX time parameters for all stations participating in the network.
- Action set A, the combinations of increasing or decreasing TX time for all stations participating in the network. The TX time parameters are increased or decreased in time steps of T $\mu seconds$. A size is $|A|^N$.
- Reward R, the average of successfully received packets $\overline{RX}$ in all access points divided by their percentage difference P:

$$R_{t+1} = \frac{\overline{RX}}{P} \tag{1}$$

where rx_i is the number of received packets at access point i, $\overline{RX} = \frac{1}{N}\sum_{i=1}^{N} rx_i$ and $P = \dfrac{|rx_1 - rx_2|}{(\dfrac{rx_1 + rx_2}{2})} \times 100$

In the actions definition, the TX time step is introduced since small TX time variations do not lead to significant performance difference for the specific network topology. The reward is a value that determines whether total throughput has been improved by the previously selected action and, moreover, if the current selection imposes fairness in the participating stations. Throughput improvement is controlled by the average amount of successfully received packets $\overline{RX}$. Fairness is determined by the percentage difference of the received packets P. Therefore, the reward R_{t+1} should be improved for higher average of received packets $\overline{RX}$ and for lower percentage difference P.

During each iteration of training, the DQN agent collects information about the current state of all stations, selects an action to perform and the next state is determined based on the selected action. The algorithm is summarized in the following steps:

1) Agent collects the state of TX time values of each station and feeds this information as input to the Neural network, which generates all possible pairs of action - Q-values as output
2) The agent selects the action to perform either randomly with probability ϵ or by the maximum value with probability $1 - \epsilon$
3) The selected action is performed and the reward of the environment is calculated by observing the amount of successfully received packets for T seconds
4) The reward is forwarded backwards to the Neural Network for further weights update, until convergence

In Table I, the configuration parameters used for Q-Learning and Neural network in our implementation are summarized.

Deep Q Network Learning Parameters	
Activation Function	ReLu
Optimizer	SGD
Epochs	100
Epoch Length	60 iterations
Learning Rate	0.01
Number of Layers	4 (1 Input, 2 Hidden, 1 Output)
Number of Neurons per Hidden Layer	(10, 5)
Error Function	RMSE
Discount factor γ	0.99
Exploration Probability Start Value	1.0
Exploration Probability Decay	0.005
Carrier Sense Time Step	$300 \in [0, 1500]$
Blank Time Step	$3000 \in [0, 15000]$

TABLE I: Parameters of Neural Network and Q Learning for Deep Q Network in our implementation

IV. Evaluation

This section, contains an analytical description for both the experimental setup and the results which are obtained. More specifically, the considered experimental topology of this work is depicted at Fig. 1 (yellow squared), and it is conducted in an environment of 4 IEEE 802.11ah nodes. The IEEE 802.11ah

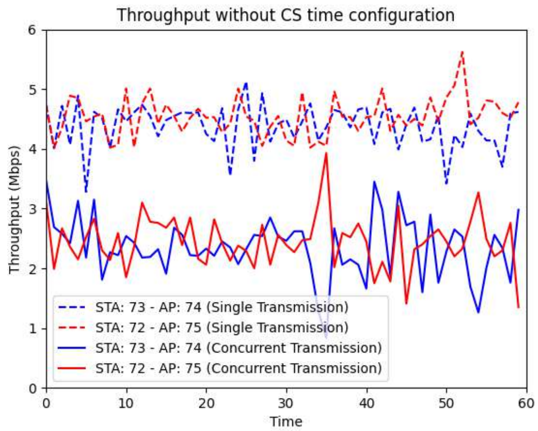


Fig. 2: Comparison of throughput in single transmission and concurrent transmissions for two pairs of nodes configured in non-overlapping channels. Single and concurrent transmissions execution for 60 seconds in 2MHz bandwidth at MCS 7. Station 72 transmits in frequency 864MHz and station 73 in frequency 866 MHz.

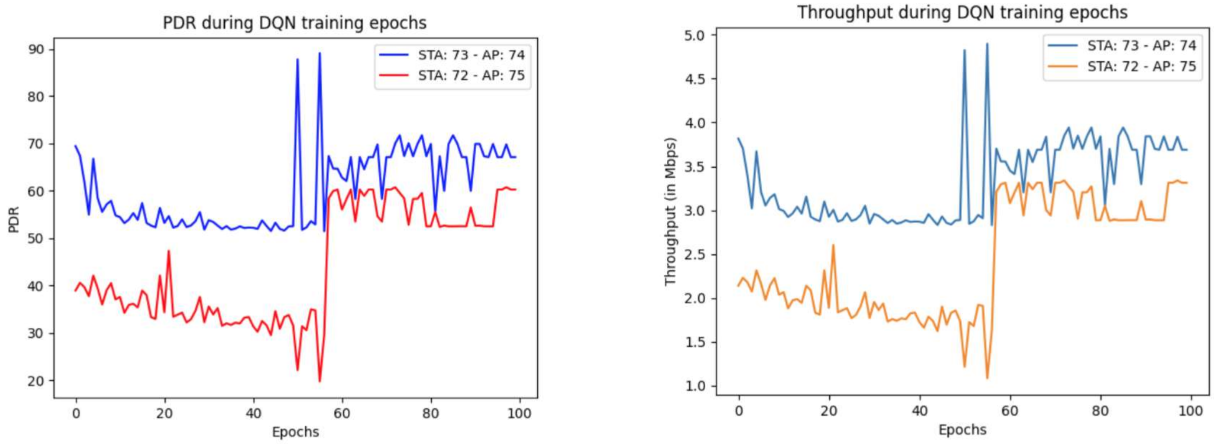
devices which are used, consisting of Newracom NRC7292 chipsets and are running on Raspberry Pi 3 Model B+. Two devices are operating in Station mode (STA) while the other two in Access Point (AP) mode, and a single remote machine (laptop) is responsible for the training of DQN agent. The participating devices are located in the same floor and divided in three separate spaces with walls and doors between. More specifically, AP 74 is located in the first space, STA 73 and AP 75 are in the second space and STA 72 is in the third space. STA 73 transmits to AP 74 in a distance of 4m and STA 72 transmits to AP 75 in a distance of 6m, while the two devices (STA 73 and AP 75) both located in the second space are divided by 2m distance. The proposed framework developed by using the Python programming language for the agent and Keras [13] for the Neural Network infrastructure. Finally, the Neural Network is constructed from Sequential Keras Model and Dense layers.

In the evaluation scenarios examined, each station is configured to communicate with one access point and both pairs are set in non-overlapping wireless channels in distinctly separated frequencies. The experiments are evaluated in bandwidth of 2MHz at MCS 7. The first pair of nodes is configured to transmit in the central frequency 864MHz (863 - 865MHz) and the second one in frequency 866MHz (865 - 867MHz). All the data traffic demands created in the context of this work were generated through the iperf [14] software.

The initial set of experiments aims to determine the data rate behavior for the previously described specifications both in case of single transmission and concurrent transmissions of IEEE 802.11ah in dense topologies. As observed in Fig. 2, the single transmission of each station leads to an average throughput of 4.5 Mbps, which is close to the theoretically expected data rate of 6.5 Mbps [15]. Moreover, when it comes to concurrent transmission at distant frequencies, it is expected for the throughput behavior to be similar with the one of single transmissions, since the wireless links are non-overlapping. However, this is not the case based on what it is experimentally observed, as the data rates in this case decrease significantly by almost 45%. This denotes, that in this environment there are strong effects of collisions between the two wireless links. As shown in Fig. 2, in an experiment of 60 seconds there is a significant throughput degradation for both wireless AP-STA pairs.

The next set of experiments that has been conducted, primarily focuses on the observation of PDR and, also the throughput differentiation over the training epochs of the Deep Q Network agent model. Figure 3a, depicts the PDR during the training epochs. At the beginning of the training, when the exploration probability is high, the agent mostly explores the environment to collect as many information as possible, and therefore, the PDR behaviour is low. However, as the epochs increase and the exploration probability becomes smaller, the agent mostly exploits its knowledge. Therefore, after almost 55 epochs, the training shows convergence at an overall increased PDR and fair spectrum sharing, by increasing first node's PDR from 35% to 55% and second node's PDR from 55% to 65%. The initial model training starts to converge at almost 3 hours and completes at 4 hours. The model is designed to dynamically adapt to demand changes by observing the successfully received packets in APs in real-time and updating the TX time parameters accordingly until model's convergence to an optimal policy. The experimental demonstration on this paper focuses on the observation of the worst scenario where STAs transmit concurrently on full demand of resources. After initial training, the model and its trained weights are saved and can be re-trained in different transmission scenarios and channel conditions. In the case of detecting the interference scenario that has been mitigated by the already trained DQN agent, convergence can occur only after a few training epochs. Each training epoch lasts on average 2 minutes and at most 3.5 minutes, providing us with the optimal transmission time parameters to eliminate collisions. The low training and retraining times required, make the proposed system fully applicable in modern architectures, in which there are constant changes in the spectral conditions. Furthermore, a similar behaviour occurs in terms of throughput evaluation. When observing data rate during DQN training epochs in Figure 3b, at first the values of both nodes are respectively at 2 Mbps and 3 Mbps because of high exploration probability. As the learning proceeds, the throughput starts to converge after a specific amount of epochs to improved and fair values of 3 Mbps and 3.5 Mbps respectively.

In this work the experiments conducted, proved that significant improvements may be obtained through the utilization of the proposed framework. The key considered metrics of this work were PDR and throughput. The results show that all involved wireless links simultaneously profit from the proposed algorithm as they both crucially improve their key examined metrics, after an amount of training epochs. It is worth mentioning that by improving both throughput and



(a) PDR observation

(b) Throughput observation

Fig. 3: PDR and Throughput differentiation over training epochs of DQN agent

PDR in such dense topology scenarios, we achieve increased transmission opportunities for all participating station nodes, and thus the social welfare is also remarkably improved.

V. Conclusions

In this paper we investigate problematic scenarios in terms of interference for dense network topologies of IEEE 802.11ah devices. A specific topology of two stations close to each other transmitting concurrently in non-overlapping frequencies leads to significant decrease of PDR, almost 45%. To mitigate the observed performance degradation caused by transmission collisions, we propose a Deep Learning approach that combines Q Learning and Neural Networks. The proposed approach trains a model that observes the environment and obtains an optimal policy of transmission time parameter configurations for participating STAs. The evaluation experiments show that training such a DQN model, leads to remarkable increased PDR and throughput (+35%) in the examined scenarios while converging to the optimal policy after 55 training epochs. Apart from improving PDR and throughput, the proposed algorithm achieves the overall higher channel utilization, increased transmission opportunities, and fairness in terms of channel access. As a future work, we intend on the detection and resolve of additional spectrum irregularities, the evaluation of the proposed model in various topologies with larger amount of devices and, also, the examination of more configuration parameters such as duty cycle or CCA threshold for performance optimization.

VI. Acknowledgments

The research leading to these results has received funding from the European Horizon 2020 Programme for research, technological development and demonstration under Grant Agreement Number No 101008468 (H2020 SLICES-SC). Additionally, this work has received funding from the European Horizon Europe Research and innovation funding programme under Grant Agreement Number No 101079774 (Horizon Europe SLICES-PP).The European Union and its agencies are not liable or otherwise responsible for the contents of this document; its content reflects the view of its authors only.

References

[1] Sri Pavan Badarla and V. P. Harigovindan. Restricted access window-based resource allocation scheme for performance enhancement of ieee 802.11ah multi-rate iot networks. *IEEE Access*, 9:136507–136519, 2021.

[2] Dmitry Bankov, Evgeny Khorov, and Andrey Lyakhov. On the limits of lorawan channel access. In *2016 International Conference on Engineering and Telecommunication (EnT)*, pages 10–14, 2016.

[3] Niemah I. Osman and Esra B. Abbas. Simulation and modelling of lora and sigfox low power wide area network technologies. In *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)*, pages 1–5, 2018.

[4] Ruki Harwahyu, Ray-Guang Cheng, and Chia-Hung Wei. Investigating the performance of the random access channel in nb-iot. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–5, 2017.

[5] Bhoomek pa and Tzi-Dar Chiueh. Interference aware coordinated multiuser access in multi-band wlan for next generation low power applications. *Wireless Networks*, 25, 05 2019.

[6] Saurabh Kumar, HoChul Lim, and HyungWon Kim. Hierarchical mac protocol with multi-channel allocation for enhancing ieee 802.11ah relay networks. In *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1458–1463, 2015.

[7] Motahareh Naghzali, Mahdi Kazeminia, and Mehri Mehrjoo. Scalability improvement of ieee 802.11ah iot networks. *Wireless Personal Communications*, 129(1):729–746, Mar 2023.

[8] Md. Arifuzzaman Mondal and Md. Iftekhar Hussain. Station grouping mechanism using machine learning approach for ieee 802.11ah. *Ad Hoc Networks*, 149:103238, 2023.

[9] Andrew G. Barto. Chapter 2 - reinforcement learning. In Omid Omidvar and David L. Elliott, editors, *Neural Systems for Control*, pages 7–30. Academic Press, San Diego, 1997.

[10] Seyed Sajad Mousavi, Michael Schukat, and Enda Howley. Deep reinforcement learning: An overview. In Yaxin Bi, Supriya Kapoor, and Rahul Bhatia, editors, *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016*, pages 426–440, Cham, 2018. Springer International Publishing.

[11] Long-Ji Lin. Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine learning*, 8:293–321, 1992.

[12] *"Nrc7292 product specifications"*, https://newracom.com/products/nrc7292, [Accessed: 7 Aug 2023].

[13] *"Keras"*, https://github.com/fchollet/keras, [Accessed: 22 Aug 2023].

[14] *"Iperf"*, https://iperf.fr/, [Accessed: 29 Aug 2023].

[15] Weiping Sun, Munhwan Choi, and Sunghyun Choi. Ieee 802.11ah a long range 802.11 wlan at sub 1 ghz. *Journal of ICT Standardization*, 1:83–107, 01 2013.