



**HAL**  
open science

## Adaptative Artificial Intelligence for Efficiency Schedulers Provider in Wireless Networks

Guillaume Terrier, Cédric Gueguen, Yassine Hadjadj-Aoul

► **To cite this version:**

Guillaume Terrier, Cédric Gueguen, Yassine Hadjadj-Aoul. Adaptative Artificial Intelligence for Efficiency Schedulers Provider in Wireless Networks. ISNCC 2024 - 11th International Symposium on Networks, Computers and Communications, Oct 2024, Washington, United States. pp.1-6. hal-04749287

**HAL Id: hal-04749287**

<https://inria.hal.science/hal-04749287v1>

Submitted on 22 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Adaptative Artificial Intelligence for Efficiency Schedulers Provider in Wireless Networks

Guillaume Terrier, Cédric Gueguen and Yassine Hadjadj-Aoul

University Rennes 1/IRISA, Rennes, France

guillaume.terrier@irisa.fr, cedric.gueguen@irisa.fr, yassine.hadjadj-aoul@irisa.fr

**Abstract**—With the increased demands for 5G networks and the limited radio resources, providing high spectral efficiency, low delay, low energy consumption, and other Key Performance Indicators (KPIs) is a challenging task. Extensive research has been conducted to propose efficient solutions for specific objectives and contexts. Although these solutions (often heuristic-based) are highly effective in specific contexts, their performance diminishes when applied in different conditions. This implies difficulties in adapting to environmental variations and/or changes in objectives. In order to overcome this problem, we propose an approach employing reinforcement learning to dynamically derive the formula for a scheduler that can be adapted to any context and objective. The proposed solution is validated with a Proof of Concept (PoC), which highlights the Artificial Intelligence (AI) ability to identify the adequate scheduler to optimize spectral efficiency in different traffic loads contexts.

**Index Terms**—Wireless Network, Artificial Intelligence, Scheduling, Machine Learning, Reinforcement Learning

## I. INTRODUCTION

In wireless networks, radio resources are scarce and subject to diverse channel attenuation due to signal phenomena such as shadowing, path loss, and multi-path fading. Additionally, the number of mobiles is expanding. To optimize spectral efficiency remains a priority, all the more so as it is the main tool for reducing delay, ensuring fairness, and minimizing energy consumption. Guaranteeing this objective is the crucial issue in radio resource allocation algorithms (schedulers).

Traditional schedulers like Round Robin (RR) or Random Access (RA) originate from wired networks [1], [2]. They operate equitably regarding the Resource Unit (RU) allocated to each mobile but lack opportunistic capabilities. Indeed, while simple to implement, they do not consider channel conditions. Consequently, they are not well-suited to wireless networks and offer suboptimal spectral efficiency, failing to correctly optimize other KPIs. Despite its status as an outdated scheduling technique, RR retains its relevance as a valuable reference for performance evaluation purposes.

Numerous research efforts have been undertaken to enhance spectral efficiency, leading to the development of several opportunistic scheduler approaches. In contrast to RR, opportunistic algorithms take into consideration the channel state when allocating RUs. One of the most renowned opportunistic schedulers is MaxSNR [3], widely acknowledged and referenced in numerous papers. Its fundamental principle is to leverage multi-user diversity. The greater the number of mobile devices in the network, the more combinations arise, resulting

in higher spectral efficiency gains. Indeed, with the MaxSNR the allocation of RUs goes to the mobile with the highest Signal-to-Noise Ratio (SNR). It allocates the RU  $n$  to the mobile  $k \in \{1, 2, 3, \dots, k_{max}\}$ , such as:

$$mobile = \underset{k}{\operatorname{argmax}}(SNR_{k,n}). \quad (1)$$

When each mobile device consistently has packets to transmit, such as during congestion or when dealing with full buffer sources, effectively managing multi-user allocation is not problematic. So, MaxSNR offers optimal spectral efficiency. However, under typical traffic loads within the system's capacity, where all packets from each mobile device are expected to be transmitted sooner or later, a limitation of the MaxSNR scheduling algorithm becomes evident. This limitation arises from its inherent unfairness characteristic. In the presence of varying signal qualities among mobiles, MaxSNR prioritizes those located close to the base station with low path loss, effectively clearing their buffers rapidly. Consequently, this diminishes the system's inherent multi-user diversity and reduces the number of possible combinations available for further allocation, particularly where distant mobile devices are yet to be scheduled. As a result, this leads to suboptimal overall system throughput performance.

The Proportional Fair (PF) scheduler and PF-based schedulers serve two primary objectives: spectral efficiency and fairness [1], [4], [5]. In contrast to MaxSNR's approach that assigns RUs based solely on the highest SNR at any given moment (Instantaneous Signal-to-Noise Ratio (ISNR)), PF relies on the better utilization of multi-user diversity. Unlike MaxSNR, PF specifically allocates RU  $n$  to a mobile  $k$  that possesses the highest ratio between instantaneous Signal-to-Noise Ratio and Mean Signal-to-Noise Ratio (MSNR) such as:

$$mobile = \underset{k}{\operatorname{argmax}} \left( \frac{ISNR_{k,n}}{MSNR_k} \right). \quad (2)$$

Through this mechanism, it delivers better throughput performance than MaxSNR when the system's load remains within sustainable limits, primarily due to its superior fairness. Indeed, more fairness permits a better conservation of multi-user diversity and allows for further combinations, which is not the case after congestion when buffers are full and multi-user diversity conservation is not a concern.

These schedulers rely on heuristics. They have proven to be more or less efficient in specific scenarios with focused objectives and lack of adaptability for various network conditions.

With the advent of Artificial Intelligence (AI), new solutions have emerged, often based on Reinforcement Learning (RL) technologies such as Q-Learning and Deep Q-Networks. In

[6], researchers address the Internet of Things (IoT) context and attempt to predict optimal channel assignment. [7] proposes an AI spectrum allocation for an integrated Access and Backhaul network (IAB). Other research focuses on radio resource allocation in the 5G cell context, as seen in [8] and [9]. In [9], the authors propose an AI scheduler based on a Q-learning approach that results in network rate optimization. The solution increases throughput by 11% compared to Round Robin (RR). While this direction is promising, it requires further investigation because each proposal targets a fixed objective and cell context. Furthermore, it is widely acknowledged that MaxSNR and PF outperform RR by more than 11% in several contexts, leaving room for further improvement.

In this paper, our objective is to develop an AI capable of determining the appropriate resource allocation strategy to achieve any predefined objective in any context. In contrast to the majority of literature proposals, our contribution is not to propose a scheduler that aims for a fixed goal in a specific context, leading to a particular resource allocation method. The idea is to determine, for any given objective, regardless of the context, which policy to employ for scheduling from a set of possibilities. This approach allows us to find the appropriate scheduler for the desired KPI and changing contexts. Called the Artificial Intelligence Scheduler Provider (AISP), this solution demonstrates its capacity to identify the scheduler that optimizes specified objectives in diverse contexts. We will showcase this in a Proof of Concept (PoC): for a well-known objective such as maximizing Spectral Efficiency, our AI adeptly discerns the algorithm to employ in accordance with the literature's expectations. As we have chosen to prioritize spectral efficiency as the central objective for this PoC, we will employ the well-known MaxSNR and PF as benchmark algorithms.

The paper is organized as follows: Section II proposes the system description. Section III describes the AISP solution. Section IV presents the PoC results, and the final section concludes the paper.

## II. SYSTEM DESCRIPTION

This study focuses on the radio resource allocation problem for the set of mobiles located in the coverage zone of a access point. A centralized approach is assumed since this allows efficient opportunistic scheduling approaches. The physical layer is considered to operate using a Time Division Duplex mode (TDD) which allows a good compatibility with the Orthogonal Frequency Division Multiplexing (OFDM) based transmission mode. The bandwidth is divided by frequency range in subcarriers. Time is also divided in frame, each one being composed of Time Slots (TS) of constant length (which is an integer multiple of the OFDM symbol duration). A Resource Unit (RU) is defined as a (subcarrier, time slot) pair that can be allocated to any mobiles with a specific modulation order.

Each frame duration is assumed equal to a value inferior to the coherence time of the channel, allowing the transmission on each sub-frequency to experience flat fading during each

frame. Therefore, channel's state is assumed static for the frame's duration for a mobile on a subcarrier. However channel state differs for a same mobile on different subcarrier and also for on a same RU between mobiles. Thanks to SNR measurement of the signal sent by each mobile (for instance during the uplink contention sub frame), the access point is thus, able to estimate their channel state attenuation at a given TS on each subcarrier. According to [10], channel state can be assumed stable on a scale of 50 ms allowing to consider a full knowledge of radio conditions available at the receiver [11]. This requires that mobiles have to transmit their control information alternatively on each subcarrier once for several frames (in order for the access point to successfully refresh the channel state information). On each RU, a modulation order adapted to the channel state (between the access point and the selected mobile) is assumed. This provides the flexible resource allocation allowing opportunistic scheduling.

The channel gain between the access point and the mobile  $k$  on the subcarrier  $n$  is given by:

$$G_{k,n} = h 10^{\frac{\chi\sigma}{10}} \left( \frac{d_{ref}}{d_k} \right)^\tau, \quad (3)$$

where  $h$  represents the Rayleigh multi-path fading, which is modeled by an exponential distribution,  $X$  is a standard Gaussian random variable,  $\sigma$  is the standard derivation of shadowing in dB,  $d_{ref}$  is the reference distance,  $d_k$  is the distance between mobile  $k$  and access point while  $\tau$  is path loss exponent. The SNR computation of mobile  $k$  on subcarrier  $n$  associated to access point is given by:

$$\gamma_{k,n} = \frac{P_n G_{k,n}}{B_{sub} N_0}, \quad (4)$$

where parameter  $P_n$  is the the transmitted power on subcarrier  $n$  of access point. Parameter  $N_0$  is the thermal noise power density and parameter  $B_{sub}$  is the subcarrier spacing. To compute the spectral efficiency  $\eta_{k,n}$  of mobile  $k$  on subcarrier  $n$  associated to access point, the Shannon's formula is used such as:

$$\eta_{k,n} = \log_2 \left( 1 + \frac{\gamma_{k,n}^i}{\Gamma} \right), \quad (5)$$

where parameter  $\Gamma$  is a SNR correction factor that takes into account the difference between the information-theoretic performances and the practical implementation of the MCS [12] defined as follows:

$$\Gamma = - \frac{\ln(5E)}{1.5}, \quad (6)$$

where  $E$  is a BER Target.

## III. AISP

The AI described in this paper aims to find the adequate scheduler among a set of large possibilities able to reach focused objectives in a specific context. For one defined objective and different contexts, the AI could select different resource allocation solution (likewise for different focused objectives in a same context). The AI does not directly perform resource allocation but try to find the efficient scheduler formula that should be used.

The basic principle is that the AI (describe in Fig. 1) dynamically test and evaluate different scheduling formulas

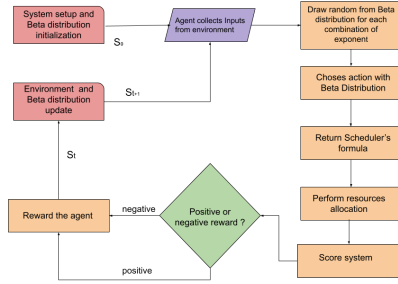


Fig. 1: AI life cycle.

in order to maximize the targeted score. If the current selected formula is efficient, consequently the score is good, the number of positive rewards grows. This will increase the probability to be selected afterwards (for this scheduler formula). On the contrary, if the selected scheduler formula is inefficient, the score will be considered as bad and the number of negative rewards grows that will decrease the probability to continue to select this solution in the future. Thereby, during a training phase inefficient solution will be quickly discarded and good solution conserve. Then, only the best solutions stay and AI finish its training by elected the best remaining one for the targeted objective and experimented context. If the aimed objective and/or experimented context change, the past selected formula could provide inefficient performances resulting in score dropping and the AI may evolved towards another scheduler.

The schedulers are expressed as formulas based on inputs (depending on  $k$  and/or  $n$ ) that are collected by AI. The generic scheduler formula consist in allocating the RU  $n$  to the mobile  $k$  that has the greatest value such as :

$$mobile = \underset{k}{\operatorname{argmax}} (Input_1^\gamma * Input_2^\delta * \dots * Input_i^\omega). \quad (7)$$

The exponent's values  $\{\gamma, \delta, \dots, \omega\}$  are defined among a set of possibilities  $\in \mathbb{R}$  and several  $i$  input can be considered (e.g.  $ISNR_{k,n}$ ,  $MSNR_k$ , delay of mobile  $k$ , buffer occupancy of mobile  $k$ , energy consumption of mobile  $k$ , residual energy of the battery of mobile  $k$ , Jain's fairness index of the system, etc.). Modifying the value of these exponents widely changes how the scheduling works and the resulting performance. A scheduler defined by two inputs with the following pair of exponent values  $\{\gamma, \delta\}: \{1, -1\}$ , will perform the scheduling as such:

$$mobile = \underset{k}{\operatorname{argmax}} \left( \frac{Input_1}{Input_2} \right). \quad (8)$$

To fallback on classical and opportunistic well-known scheduler (MaxSNR and PF),  $Input_1$  could be  $ISNR_{k,n}$  and  $Input_2$  the  $MSNR_k$  value.

Any desired targeted score function can then be used. For example, if spectral efficiency is targeted:

$$\frac{\sum_{n=1}^{n_{max}} \sum_{k=1}^{k_{max}} ISNR_{k,n} * c_{k,n}}{\sum_{n=1}^{n_{max}} \sum_{k=1}^{k_{max}} c_{k,n}}, \quad (9)$$

where  $c_{k,n} = 1$  if mobile  $k$  receive the RU  $n$  and equal 0 otherwise,  $n_{max}$  the total number of RU and  $k_{max}$  the total number of mobile in the system.

During the training phase, in order to progress in the differentiation between inefficient/efficient scheduler's formulas we use Reinforcement Learning with Thompson Sampling and Beta distribution. Reinforcement Learning is a Machine Learning's model, with essential features such as agent, environment, actions, reward and policies. The agent or decision-maker retrieves information from the environment. Then it will select an action (scheduling formula) based on the Thompson Sampling policy. The number of available actions is defined by the number of input retrieved by the agent and the number of value that each exponent can take and can be calculated as follow:  $Nb_e^{Nb_i}$  (with  $Nb_e$  the number of values available for each exponent and  $Nb_i$  the number of inputs used for scheduling).

Thompson sampling is employed to address the trade-off between exploration and exploitation. [13] exhibit that it is a probabilistic method that is efficient to solve the multi-armed bandit problem. Thompson sampling operates based on Bayesian distribution. Each time the AI needs to decide which type of scheduling to execute (take an action), it selects a random number  $x$  for each action using the associated Beta distribution and chooses for the one with the highest random number. The Beta Distribution's shape and probability density function (PDF) are defined by two parameters commonly referred to as alpha and beta:

$$f(x, \alpha, \beta) = x^{\alpha-1} \cdot (1-x)^{\beta-1} \overline{B(\alpha, \beta)}, \quad (10)$$

where  $x$  is a random variable, between 0 and 1.  $\alpha$  and  $\beta$  are the shape parameter greater than 0.  $B(\alpha, \beta)$  is the Beta function, ensuring the PDF integrates to 1 in the interval [0,1]. Beta distribution is commonly employed in statistics and probability. When  $\alpha = \beta$ , the distribution is centered at 0.5, and its shape resembles a symmetric bell curve. When the number of negative rewards ( $\beta$ ) exceeds the positive ones ( $\alpha$ ), it becomes right-skewed, with a longer tail on the right-hand side and concentration on the left side (towards 0). Consequently the associated solution to this distribution will be less likely chosen. In opposition, if  $\alpha > \beta$ , it exhibits a left-skewed shape and associated solution will be more likely chosen. If  $\alpha = \beta = 1$ , the distribution becomes uniform.

Each time a scheduler formula is fixed thanks Thompson Sampling (a combination of exponent values for each input), the scheduler performs an amount of resource allocations<sup>1</sup>. The AI computes the score obtained. After calculating the score, it is checked whether the score has increased or decreased and the Thompson Sampling associated distribution updated. If it increases, the agent receives a positive reward of 1 and  $\alpha$  increase. Conversely, if it does not increase, the agent receives a negative reward of 1 and  $\beta$  increase. In fine, all possible scheduler formula are tested and the more accurate to reach the defined objective emerged whatever the context is.

<sup>1</sup>This value can be set as a number of frame duration. If it is set too low, a problem could appear: some elected formula could benefit of previous concurrent scheduling behavior and AI convergence would be delayed. If too high, simulation duration is also increased to reach AI convergence. Consequently a trade-off as to be done (12 frames offers good compromise).

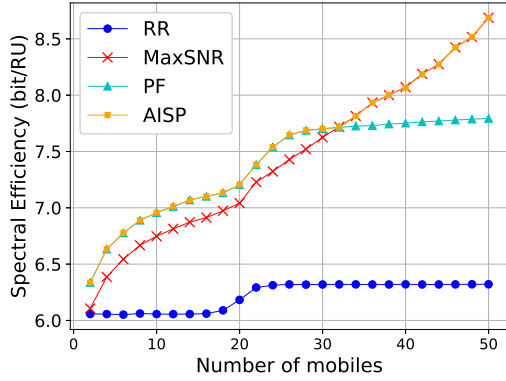


Fig. 2: Mean value of bit/RU for a number of mobile given.

#### IV. PoC'S RESULTS

The purpose of this paper is to propose a new approach that employs reinforcement learning to dynamically find the formula for an adapted scheduler to the experimented context in order to reach a defined objective. A whole performance evaluation on multiple contexts and for multiple objectives will be conducted in future works. This paper proposes a Proof of Concept (PoC) for a specific targeted objective that is spectral efficiency optimization in different context of traffic load. Focusing on spectral efficiency optimization for different traffic loads is interesting for a PoC since it is known that PF provides excellent results before congestion due to its fairness and multi-user diversity preservation. On the contrary, after congestion, when buffers are always full, MaxSNR is more effective since close mobiles, with the best ISNR, stay always with data to transmit. Consequently, these scheduling strategies, extensively studied in the literature, represent a very good benchmark for the PoC.

In this PoC we use only two inputs in the scheduler formula: ISNR and MSNR. Although this avoids discovering new potential scheduler formulas that could outperform existing solutions (the object of the future works), restricting the breathing space in the scheduler formula to these two inputs allows us to verify if our AI comes to the same conclusion as literature: PF must be used before congestion and MaxSNR when buffers of close mobiles are full. The set of values that can be assigned to the exponents is  $\{-1, 0, 1\}$  that combine to the number of inputs to offer nine distinct schedulers (actions) possible for the election. If the action chosen by our AI returns the best spectral efficiency among the other actions, or if it succeeds in improving its own spectral efficiency, the agent receives a positive reward. Conversely, if the action fails to enhance spectral efficiency or leads to a decrease, the agent receives a negative reward.

Performance simulations are conducted thanks to discrete time event simulator using parameters done in table I. We study the impact of different traffic loads and solutions performances. Mobiles use Variable Bit Rate (VBR) [14]. The mobiles are constituted of 2 groups that are located at 2 different distances from the access point, half close and the other half 2 times

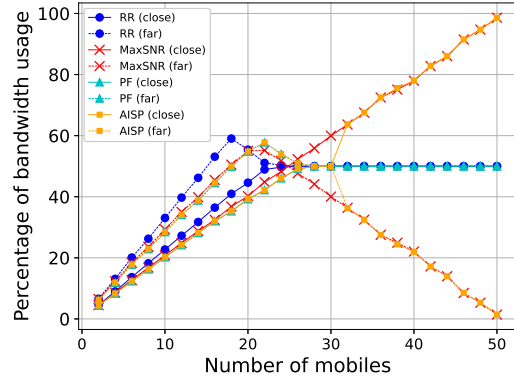


Fig. 3: Bandwidth usage.

Parameters	Value
Cell Radius	500 m
Number of subcarriers	128
Number of Time Slots	10
RRH transmit power	20 W (43 dBm)
Standard deviation of shadowing	$\sigma = 8$ dB
Path-loss exponent ( $\tau$ )	3.5 (urban context)
Target BER	$5 \times 10^{-5}$
Subcarrier spacing	15 kHz
Thermal noise power density ( $N_0$ )	-174 dBm/Hz
Simulation duration	300 000 TS
Possible value for exponents	$\{-1, 0, 1\}$
Number of allocation performed by AI before drawing new formula	12 frames

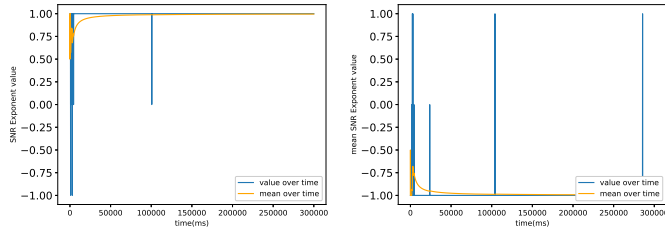
TABLE I: Simulations parameters.

further. Mobiles are added 2 by 2 respecting an equal proportion in terms of distance. Different distances allow us to underline possible fairness issues and this exhibits PF's abilities to distinguish itself from MaxSNR thanks to its superior multi-user preservation capacity.

#### A. General results

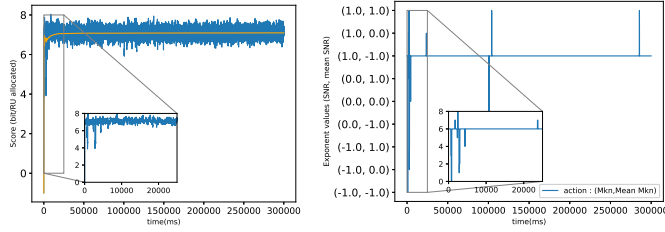
Figure 2 displays the spectral efficiency obtained at the end of the simulation with different loads. Figure 3 underlines how RUs are distributed among mobiles based on their distance from the antenna and the system's load. These two figures are related, and when linked, they explain what is happening in our system in terms of spectral efficiency.

First of all, it is important to identify (on Fig. 2) the traffic load key points where RR, MaxSNR, and PF encounter congestion: 18 for RR, 22 for the 2 opportunistic schedulers. RR is the first solution that experiments congestion since it has the lowest spectral efficiency (slightly higher than 6 bit/RU) due to its non-opportunistic behavior. This result (18) can be determined theoretically with the calculation: number of RU in a frame (1280) multiplied by the mean spectral efficiency (6.05) then divided by wanted throughput for a mobile during a frame duration (420 bits). Same calculus can be done in order to theoretically retrieve MaxSNR and PF congestion using their respective spectral efficiency at this moment (7.2 bit/RU for MaxSNR and 7.38 bit/RU for PF).



(a)  $\gamma$  value over time. (b)  $\delta$  value over time.

Fig. 4:  $\gamma$  and  $\delta$  values with 16 mobiles over time.



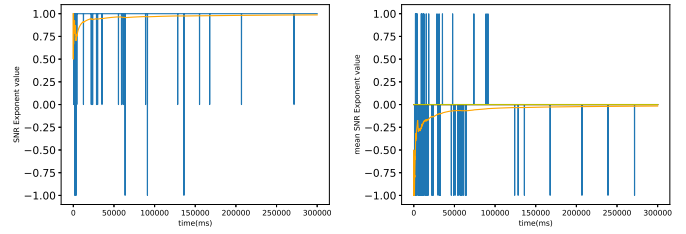
(a) System's score over time. (b) Actions over time.

Fig. 5: Score's values and actions with 16 mobiles taken over time.

From 2 mobiles to 16, RR does not experience congestion, and its spectral efficiency remains flat (Fig. 2) since it does not take advantage of the growing multi-user diversity. Close mobile that have better spectral efficiency than far mobiles due to lower path loss use less RU (Fig. 3). These vacant RU are used by far mobile. When the system reaches a load of 18 mobiles, RR struggles to empty all buffers and RR that allocates the RU turn by turn begins to reallocate RUs from far-away mobiles to close mobiles (Fig. 3). By allocating more RUs to close mobiles with better signal to noise ratio, global system spectral efficiency increases (Fig. 2) until it stabilizes due to the fairness in the allocation of RUs, with fifty-fifty distribution between close and far-away mobiles (Fig. 3).

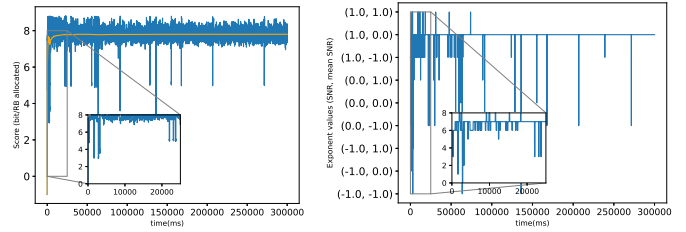
In contrast to RR, MaxSNR takes benefit to multi-user diversity to improve spectral efficiency (Fig. 2). By making more efficient allocations than RR, it uses fewer RUs for the same number of mobiles (Fig. 3). Before congestion RU not used by close mobile are left to far mobile that are statistically serve after. We can see on figure 3 the start of congestion that is underlined when the mean number of allocated RUs for far-away mobiles starts to decrease (since MaxSNR formula prioritized close mobile). When congestion is overpass, close mobile tends to receive all the RU (Fig. 3) at the expense of far mobile (with boost spectral efficiency (Fig. 2)).

As previously explain, PF, with its better fairness and multi-user management, obtains better spectral efficiency than MaxSNR when the system is not severely congested (Fig. 2). Indeed, PF is less often in the situation to manage just a few amount of mobiles (low multi-user diversity and low combination) than MaxSNR due to hasty emptying of close mobile buffer. When the system's load reaches 22 mobiles, the mean number of allocated RUs for mobiles far away from the antenna starts to decrease, as seen in figure 3. RUs lend to far



(a)  $\gamma$  value over time. (b)  $\delta$  value over time.

Fig. 6:  $\gamma$  and  $\delta$  values with 34 mobiles over time.



(a) System's score over time. (b) Actions over time.

Fig. 7: Score's values and actions with 34 mobiles taken over time.

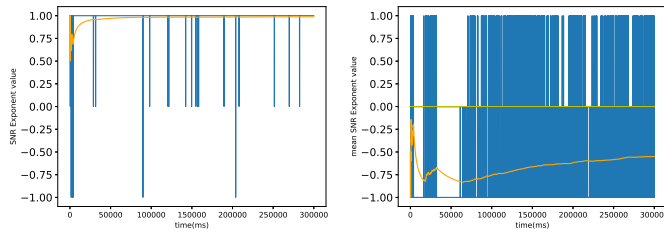
by close mobile begin to be retrieved. However when traffic load highly overpass congestion (32 mobiles), PF continue to maintain a certain amount of fairness by serving far mobile and its spectral efficiency ends up outperformed by MaxSNR (Fig. 2).

To conclude these results analysis. PF appears to be the best solution of the literature from 2 to 30 mobiles while MaxSNR outshine it on higher system load. As shown in figure figure 2 and 3, the proposed solution (AISP) selects a scheduling formula corresponding to PF when it is the best (from 2 to 30 mobiles) and MaxSNR when it appears better.

### B. Working without congestion

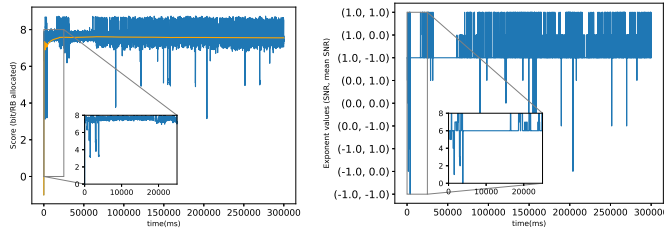
Figure 4a shows (in blue) the ISNR values choose across time by the proposal AI when traffic load is low. The yellow curve represent the mean value chose by the solution. Figure 4b is the same results for MSNR input. These figures show that when the the system is not experiencing congestion, the exponent associated with the ISNR values (Fig. 4a) and the MSNR (Fig. 4b) selected by the AI proposal quickly trends towards a value of 1/-1. This correspond to the best solution as depicted in figure 2 and 3: PF. This signifies that the AI has discerned the most effective approach for increasing throughput in this context.

Figure 5a exhibits the score obtained other the time for an acceptable traffic load of 16 mobiles and figure 5b underlines the actions tested by our proposal over the time for each exponent parameters couple  $\{\gamma, \delta\}$ . As shown in figure 5b, the AI proposal swiftly converges to PF heuristic in few tries that quickly provides high and stable global score on spectral efficiency (Fig. 5a). In the zoomed part of the figure we can distinguish that the abrupt deterioration in the score are relied to the test of inefficient solutions.



(a)  $\gamma$  value over time. (b)  $\delta$  value over time.

Fig. 8:  $\gamma$  and  $\delta$  values for 26 mobiles over time.



(a) System's score over time. (b) Actions over time.

Fig. 9: Score's values and actions with 26 mobiles taken over time.

### C. Working in during congestion

Figure 6 and 7 are identical to figure 4 and 5, except that they represent 34 mobiles instead of 16. During high congestion (after 30 mobiles), the best algorithm available for maximizing spectral efficiency is MaxSNR. Results show that the values chosen by AI for the ISNR and MSNR exponents converge to 1 and 0, respectively. This means that the AI has decided to use a MaxSNR allocation (Fig. 7b), allowing it to increase the spectral efficiency score (Fig. 7a). However, we can observe in figures 6a, 6b, and 7b that the AI tests more solutions than in a context with low traffic load. This is due to more hesitation between two schedulers that exhibit close, good performance on targeted spectral efficiency, particularly between MaxSNR (1,0) and an unnamed scheduler (1,1), which could be defined as highly unfair, similar to MaxSNR.

### D. In between

Between 22 and 32 mobiles the system is in transition and PF stays the best solution while the difference with MaxSNR reduce more and more. Figure 8 and 9 are identical with figures 4 and 5 except that is is for 26 mobiles instead of 16. As expected, the AI select PF scheduler in accordance with figure 2 that underlines this stay the best solution in this context. However, figures 8a, 8b and 9b underline the choice is more complicated since scheduler performances for the targeted objective are very close and difficult to detect. This requires more tests.

### E. Discussion

The PoC results seems to reveal the good potential of an AI approach that try to discover the adequate scheduling strategy to apply for reach a targeted objective in multiple context. This work succeed to find the adequate scheduler among a panel

of available strategy for the spectral efficiency optimization objective. Enlarge this panel by increasing the number and type of inputs and/or the set of considered exponents values will increase the AI training time but may provide efficient solution currently undiscovered.

## V. CONCLUSION

In conclusion, this paper introduce an Artificial Intelligence solution that uses Reinforcement Learning (RL). It possesses the capability to identify and employ the adequate scheduler in order to reach a focused objective within various wireless networks cell contexts. Our innovative solution has demonstrated its ability in a Proof of Concept and has succeed to adaptability find the adequate scheduler to used, in diverse traffic load contexts, while targeting objective such as spectral efficiency optimization. Looking ahead, our future work will integrate and consider more environment inputs in order to enlarge the schedulers panel that could possibly lead to new innovative solution or confirms/corroborates the existing one. In addition other objective and system context will be studied.

## REFERENCES

- [1] M. Minelli, M. Ma, M. Coupechoux, and P. Godlewski, "Scheduling impact on the performance of relay-enhanced LTE-a networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, 2016.
- [2] A. Kuurne and A. Miettinen, "Weighted round robin scheduling strategies in (e)gprs radio interface," in *IEEE 60th Vehicular Technology Conference. VTC2004-Fall.*, vol. 5, pp. 3155–3159 Vol. 5, 2004.
- [3] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser ofdm with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, 1999.
- [4] M. I. Elhadad, E.-S. M. El-Rabaie, and M. Abd-Elnaby, "Capacity enhanced scheduler for LTE downlink system based on PF algorithm," in *2016 Fourth International Japan-Egypt Conference on Electronics, Communications and Computers (JEC-ECC)*, pp. 5–8, 2016.
- [5] X. Ge, H. Jin, and V. C. M. Leung, "Cdf-based scheduling algorithm for proportional throughput fairness," *IEEE Communications Letters*, vol. 20, pp. 1034–1037, May 2016.
- [6] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah, and N. Nasser, "A deep learning method for predictive channel assignment in beyond 5g networks," *IEEE Network*, vol. 35, no. 1, pp. 266–272, 2021.
- [7] J. Park, H. Jin, J. Joo, G. Choi, and S. C. Kim, "Double deep q-learning based backhaul spectrum allocation in integrated access and backhaul network," in *2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pp. 706–708, 2023.
- [8] A. Giannopoulos, S. Spantideas, N. Capsalis, P. Gkonis, P. Karkazis, L. Sarakis, P. Trakadas, and C. Capsalis, "Wip: Demand-driven power allocation in wireless networks with deep q-learning," in *2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 248–251, 2021.
- [9] S. Suknum, C. Thoasiri, and N. Jinaporn, "Q-learning-based resource allocation in heterogeneous cellular networks," in *2022 International Electrical Engineering Congress (iEECON)*, pp. 1–3, 2022.
- [10] T. E. Truman and R. W. Brodersen, "A measurement-based characterization of the time variation of an indoor wireless channel," in *Proc. IEEE Int. Universal Personal Communications Record (ICUPC)*, 1997.
- [11] Y. G. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461 – 471, 1999.
- [12] H. Seo and Byeong Gi Lee, "A proportional-fair power allocation scheme for fair and efficient multiuser OFDM systems," in *IEEE Global Telecommunications Conference, 2004*, vol. 6, Nov 2004.
- [13] J. Zhu, X. Huang, X. Gao, Z. Shao, and Y. Yang, "Multi-interface channel allocation in fog computing systems using thompson sampling," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13542–13554, 2021.
- [14] G. Horvath and P. Fazekas, "Modelling of youtube traffic in high speed mobile networks," in *Proceedings of European Wireless 2015; 21th European Wireless Conference*, pp. 1–6, 2015.