

Using AI in wireless communication system for resource management and optimisation

Huda Abdalkaream Mardan¹, Suad Kakil Ahmed²

^{1,2}Computer System Department, Technical Institute of Kirkuk, Northern Technical University

ABSTRACT

The existence of Artificial Intelligence (AI) can be seen in everyday scenarios. Nowadays, the produced data by both machine and human is overwhelming in which exceeded the ability of humans to understand and digest to make decisions depending on that data. Thus, a hand of help from AI is needed to overcome such challenges. The 5G LTE communication system is a promising solution to provide a high user experience in terms of the provided speed, amount of data, and cost. However, and due to its complexity, the technology of LTE needs some improvement in terms of resource management and optimization. With the aid of AI, these two challenges can be overcome. In this paper, the AI represented by improved Q-learning algorithm with the Self-Organizing Network (SON) concept in LTE will be used to manage and optimize Handover (HO) parameters and process in the system. The ns-3 simulator result shows that AI managed to improve and optimize the LTE system performance.

Keywords: 5G,LTE, Artificial Intelligence, Self-Organizing Network, Q-learning

Corresponding Author:

Huda Abdalkaream Mardan
Computer System Department
Northern Technical University
Kirkuk, Iraq
E-mail: hudamardan@ntu.edu.iq

1. Introduction

Artificial intelligence (AI) can be used to solve many practical and everyday problems we face in our daily life. Artificial intelligence is part of computer science that includes machine learning, neural networks, and deep learning in which is designed to mimic human-like reactions[1]. Activities such as problem-solving and speech recognition and planning can be conducted with the aid of a computer equipped with AI. However, in this paper, AI represented by Reinforced Learning (RL) used to manage and optimize a communication system resources and performance with the help of extended/improved Q-learning algorithm[2]. To understand the challenges that we need to tackle in this work, we have to understand one of the problems in the 5G communication system[3, 4]. The problem is the continuous decrease in the optimal cell sizes in cellular networks where in contrast, there is an increase in cell density especially after LTE introduced[5, 6]. Consequently, the expenses for both operational and capital expenses plus network operation complexity increased as well. Yet, a solution such as Self-Organizing Networks (SON) might be a promising solution to mitigate these challenges by automating the network operations. The definition of the SON Functions (SFs) and standardization has been introduced in which every SF can be automated to work independently[7]. For instance, Coverage and Capacity Optimization (CCO), Mobility Robustness Optimization (MRO), and Cell Interference Coordination (ICIC). Functions in SFs are designed as rule-based in which required a deep understanding of the function's behavior to be designed. In this paper, we develop SFs to CCN functions to be implemented as cognitive Hyper-QL (HQL). The proposed advance is depending on agents, these agents act and behave within the network based on feedback from the network to learn how their actions affect the network[8, 9].

2-AI in Communication system network

New applications and high data rates are both next-generation's challenges for wireless networks in which required new technology for wireless radio system. The challenge we try to tackle is how to assist the radio by using decision making and intelligent adaptive learning to meet the next generation wireless networks' requirements. There are a lot of AI tools , however, Machine Learning (ML) is a promising one that can be

used for better and smarter radio terminals [10-13]. It is anticipated that 5G mobile terminals access and uses most of the available spectral band smartly throughout employing AI to learn a sophisticated method for better and efficient way to deal with spectral resources.

2.1-Reinforcement learning in wireless communications

Reinforcement Learning (RL) is a quite popular learning technique. The RL method applies the idea of system, in which this paper we call it an agent where this agent can sense its surrounding environment and current state to choose an action. However, the process that distinguishes RL from other systems is the process after the agent action. The action and its outcome will lead to whether a reward is given to the agent or not. If the agent took a good action then it receives a reward, if not, it will receive a penalty[14].

3. Self-organizing networks (SON)

In this paper, AI is represented by an improved Q-Learning algorithm[2]. This algorithm is implemented in SON environment. Network automation going to work with AI. One of the essential roles of network automation is to cope with increasing network infrastructure complexity to manage the deployed resource of the network in an optimized way[15]. The Self-Organizing concept gains popularity after the Next Generation Mobile Networks (NGMN) alliance proposed this concept to deals with the anticipated future management challenges of several radio access technologies along with the LTE network introduction. The importance of the SON concept stems from its ability to self-configuration, self-optimization, self-healing, and self-planning[16]. In this work self-planning is the used concept. This concept (self-planning) combines both configuration and optimization abilities to recomputed network parts dynamically. The main aim is to choose and improve parameters affecting service quality.

4. Hyper-Q learning

Recent Q-Learning with multi-agent extensions[2], needs to know the payoffs plus Q-functions of other agents. However, the Hyper-Q learning algorithm proposes a totally different approach. In this algorithm, instead of learning from actions, agents use mixed values strategies to learn. The algorithm estimates the strategies of other agents by using Bayesian inference to observe the agent's actions. One of the advantages of Hyper-Q is that it works effectively even if it is working dynamically with different adaptive agents. In this paper, Hyper-Q-learning is used to optimize HO in the LTE network.

4.1Formulation of hyper-Q

A state s will be observed repeatedly from an agent with ordinary Q-learning. Next, a legal action A will be taken by the agent. And then, infinite Markov Decision Process (MDP) observes immediate r and transfer to a new state. Thus, the equation for Q-learning is: $\Delta Q(s, a) = \alpha(t)[r + \gamma \max_b Q(s', b) - Q(s, a)]$ [2], where the discount parameters is γ and the $\alpha(t)$ is the learning schedule rate. Based on an appropriate method for state action, Q-learning will converge to a Q^* value function and thus an optimal π^* policy is the associated greedy policy. In MDPs every agent, in state s , selects an action a_i , while payouts r_i for agent i and state transitions are functions of the joint actions of all agents. In the inventory of a stochastic game, the agent has the job to select the best mixed strategy $\bar{x}_i = \bar{x}_i(s)$ for all other agents, instead of choosing the best action in one state. Here \bar{x}_i : The selection in this paper specifies a probability of 1 for each legal action in state s . Here (x_i) : A continuous $(N_i - 1)$ simple unit is the area of possible mixed strategies, and it is clearly more complicated to select the best mixed strategy than to select the best basis action. Now, we are considering the extensions of Q-learning for stochastic games. As a mixed strategy required dependent on the mixed strategies of other agents, the mixed strategies should be evaluated by Q function in their entirety rather than basic actions and include observation or estimation of the present mixed strategies of the other agents in the "state" designation. The formulation of the learning algorithm Hyper-Q can be considered as follows: let x indicates a mixed strategy of the Hyper-Q learner and let y indicate a combined strategy estimated by all other agents (hereinafter we call it as "opponents") for notational simplicity. The agent generates a base measured by x at the time t and the payoff r , the new states, and finally y the opponent's strategy estimated strategy

The function $Q(s, y, x)$ Hyper-Q can be set as:

$$\Delta Q(s, y, x) = \alpha(t)[r + \gamma \max_{y', x'} Q(s', y', x') - Q(s, y, x)] \quad (3)$$

For every Hyper-Q function, the greedy policy \hat{x} is determined by:

$$\hat{x}(s, y) = \arg \max (s, y, x) \text{-----} (4)$$

4.2 . Mobility and handover management HOs

In accordance with the HO conditions A3[17], a reference signal power received (RSRP) from s to t for HO trigger HOs from the serving cell to the target cell t :

$$F_t + O_t^{s,t} - H_{ys} > F_s + O_s^{s,t}$$

$F_t F_s$ is respectively the Cell Individual Offsets (CIOs) of the user in t and s cells, and no offsets; $O_t^{s,t}$, $O_s^{s,t}$, while (H_{ys}) is the Hysteresis measured in dB for the serving cell. For the case where A3 has fulfilled all the conditions for the Time To Trigger (TTT), a measurement report going to be sent by UE to initiate HO. The report contains the filtered values of both F_s and F_t in which these values have to pass through the three following filters; L1, IIR and L3[17], L3 is implemented with a filter coefficient of 6. There are 100 samples from the average window that must be updated to average quick fading every 200 ms but not so long to impact L3 results. It is obvious that the HO results mostly depend on the effects of either delaying or progressing HOs by both H_{ys} and TTT. However, Ping-Pong HO, Radio Link Failure (RLF) or HO success can lead to possible results. This modeled on timers defined in [12] by 3GPP. The (p) rate of Ping-Pong (PP) can be defined as follows: A user is registered with a PP or HO oscillation, if a successful HO takes time below the PPtime from Cell B to another Cell A Handover (HO) from cell A to cell B is successfully conducted. In other words, it is a rate at which PPs are occurring, whether in a cell or in a network per second. In this paper, we use PP-time = 5s in which it approximately the longest TTT was established.

4.3 Model of simulation

The first step in the simulation is deploying the cells, next, multiple “batches” were exacted. Each of these batches starts with the deployment of users. This redeployment enables users to get as many common radio conditions as possible to place in a different location and to track a different path. To achieve User Equipment (UE)s per cell average number, UEs distributed randomly where the per cell average number is defined as the ratio of total No. of UEs /users divided by the No. of cells. In addition, static users lactated in the center of the cell (i.e. UEs = 0 m / s). The proposed algorithm performance, i.e. HQMRO algorithm, and the performance of the Reference network algorithm (Ref. algorithm) will be compared. The Ref. algorithm is the algorithm that built-in in the network without using SON functions. The crucial simulation parameters are summarized in Table I. This work will follow the scenarios and parameters in [8]. The simulator of choice in this work is ns3[18-21], this simulator used to evaluate the HQLMRO Learning framework. The simulator provides a complete model for LTE. All the required models of SON and used algorithms were implemented in this simulator.

Table 1. Parameters of Simulation

Parameter	Value
Bandwidth of system	10 MHz
Distance between the sites	500 m
Interval of snapshot	50 ms
Users No.	240 mobile, 40 static
User speed	3, 10, 30, 60 or 120 km/h
Model for Mobility	Random walk
Pathloss equation	$A + B \cdot \log_{10}[\max(d[\text{km}], 0.035)]$; A =128.1 and B =37.6
Shadowing parameters	Standard deviation = 6 dB; Decorrelation distance = 50 m

Parameter	Value
eNB Tx power	46 dBm
eNB Tx antennas	1 per sector, gain 15 dBi, at height = 32 m
User Equipment antennas	1 Omni, gain 2 dBi, at height = 1.5m
Data rate	512 Kbps

5. Simulation results and discussion

The HO parameters are adjusted to match different mobility states by MRO. For example, the three 'normal' scenarios (10, 30, and 60 k/h) where three city areas have been represented: center, edge and an urban suburb, represent three typical urban areas: a town center, the urban edge, an urban suburb. We then take into consideration two extreme scenarios (3 and 120 km/h) that could represent an office and a road respectively. All UEs have randomly varying speeds in every mobility scenario. This done by assigning random speeds to the EU at the beginning of the simulation and adjusting the speed randomly for each load up to $\pm 40\%$ at start and throughout each load. Each 240 users have individually assigned and constantly changing speed, for example, in the town's suburb (60 km/h). The HQMRO algorithm shown below is used to optimize the HO performance in the system.

Algorithm: HQMRO - The Hybrid Q-Learning MRO Algorithm

Set $R_i = R_1$; initialize action set A_{x, R_i} for regime 1 in all states x

Repeat for each SON interval t

***if** HO action was taken at SON interval $t-1$ **do**

Identify HOAP and derive reward $r_{t-1}(x_{t-1}, a_{t-1})$

update Q-table

end if

find the state x_t (current mobility)

if state x complete learning **do**

select $a_{x,t} = a_{x,R_i}^{opt}$ the optimum value for state x

else if

regime R_i exploration is incomplete **do**

select $a_{x,t}$ (sequentially after $a_{x,t-1}$) from A_{x,R_i}

else do

select $a_{x,t} = a_{x,R_i}^{opt}$ the optimum value for state x at R_i

if all learning regimes complete for state x **do**

record all regimes complete for state x

record $a_{x,t}$ as best action in state x

end learning, indefinitely use $a_{x,t}$ in state x

else do

$R_i \leftarrow R_i + 1$

use $a_{x,t}$ to set A_{x,R_i} i.e. reconfigure A for R_i

end if

Action $a_{x,t}$ is selected by signal and allocated to all UEs in the

In the cell

$t \leftarrow t + 1$, monitor and collect statistics, continue at step *

end loop

5.1 The HO aggregate performance (HOAP)

For each velocity scenario, the HQMRO algorithm performance is compared to the Reference network algorithm (Ref). The performance is assessed according to the averages of the network metric values. Figure 1 shows the outcome of various scenarios, where the sub Fig. 1a illustrates the performance of the two algorithms HQMRO and Ref for the average HOAP values. However, two typical cases of speed (30 km/h and 120 km/h) are used to set the performance comparison. Figure 1b describes the performance differences between

HQMRO and Ref.for all the five speed cases (3,10,30, 60,and 120 km/h). The figures show that initially HQMRO is not performing well when implementing in the first learning scheme R1. Looking at R2 and R3 regions, we notice that the performance have improved as HQMRO focused on the Optimum Trigger Point (OTP) , in which it finally equivalent to the performance of Ref.In fact, HQMRO is more successful where user speeds are widespread, since it can set a correct setting for each speed range, instead of a single setting for all speeds. For the case of 120 km/h as shown in Fig. 1b and after learning, HQMRO always perform better than Ref.As the speed of users increase, each user also undertakes more HOs at a particular interval, yet, cells at higher velocity will have a shorter SON intervals as can be seen in the Fig. 1b, where the increase in speed resulted shorter convergence times.

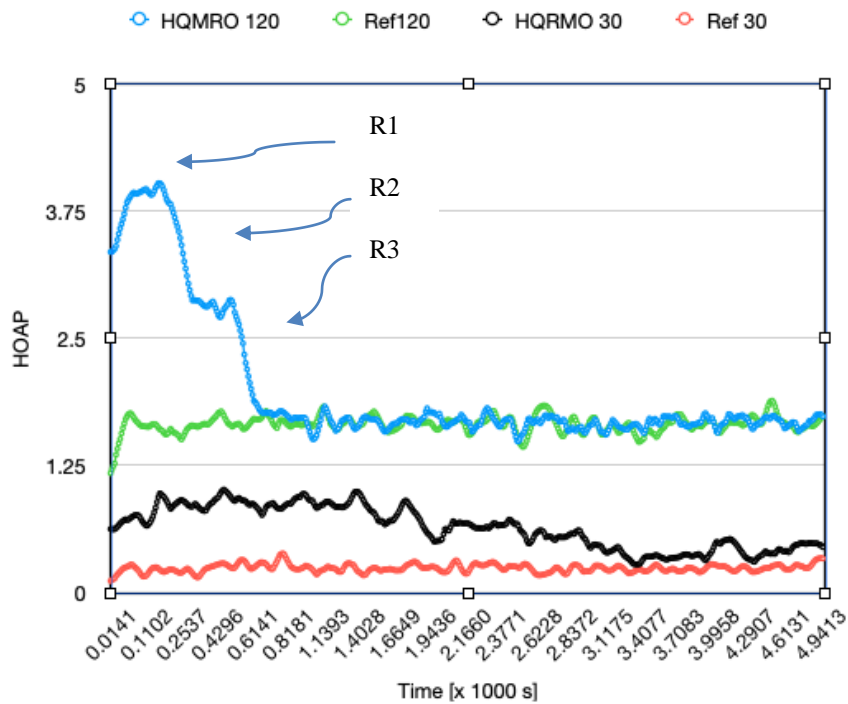


Figure 1. It comparison of two speed scenarios for HOAP

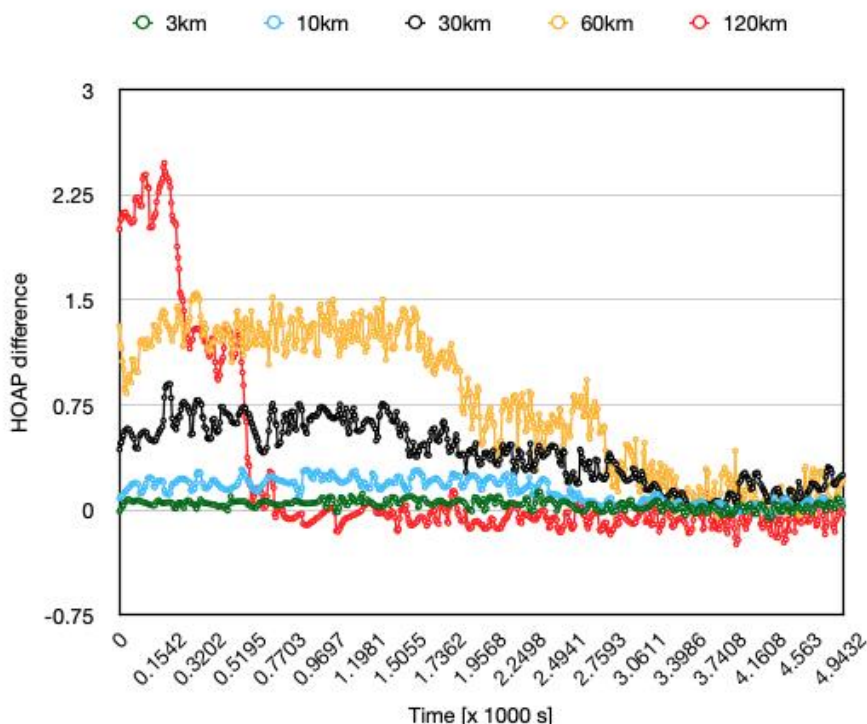


Figure 2. HQMRO performance: Average HOAP network wide compared to the Ref

6. Conclusion

Artificial Intelligence offers powerful tools to resolve very complex issues and will continue to be developed and exploited in future. Resource management and performance optimization both are a complex task to deal with. However, AI represented by RL with improved Q-learning algorithm managed to ensure better performance and resource management to the 5G LTE system. Result shows that Q-learning based MROs (HQMRO) performs well by learning most efficient setting for Hysteresis (Hys) and Trigger Time (TTT) configurations. Cells can learn a multiple policy functions by applying an approach like cooperative learning. One of the advantages of such approach is its ability to be applied in any environment as shown in the obtained results. The cells have learned a multiplicity of policy functions through cooperative learning. A good performance results with distinct and dynamically differing speeds show such an approach as robust and applicable to any environment in realistic network scenarios.

Reference

- [1] I. Suryana, A. Asrianto, D. J. J. J. o. L. Murwantono, and L. Teaching, "ARTIFICIAL INTELLIGENCE TO MASTER ENGLISH LISTENING SKILLS FOR NON-ENGLISH MAJOR STUDENTS," vol. 8, no. 1, pp. 48-59, 2020.
- [2] G. Tesauro, "Extending Q-learning to general adaptive multi-agent systems," in *Advances in neural information processing systems*, 2004, pp. 871-878.
- [3] Z. Wen et al., "Measurement and Modeling of LTE-Railway Channels in High-Speed Railway Environment," vol. 55, no. 4, p. e2019RS007050, 2020.
- [4] H. T. Alrikabi, A. H. M. Alaidi, A. S. Abdalrada, and F. T. J. I. J. o. E. T. i. L. Abed, "Analysis the Efficient Energy Prediction for 5G Wireless Communication Technologies," vol. 14, no. 08, pp. 23-37, 2019.
- [5] S. Feng and E. J. N. R. G. Seidel, Munich, Germany, "Self-organizing networks (SON) in 3GPP long term evolution," vol. 20, 2008.
- [6] H. T. S. Al-Rikabi, *Enhancement of the MIMO-OFDM Technologies*. California State University, Fullerton, 2013.
- [7] Z. Yan, P. Zhang, A. V. J. S. Vasilakos, and c. networks, "A security and trust framework for virtualized networks and software-defined networking," vol. 9, no. 16, pp. 3059-3069, 2016.
- [8] S. S. Mwanje, L. C. Schmelz, A. J. I. T. o. N. Mitschele-Thiel, and S. Management, "Cognitive cellular networks: A Q-learning framework for self-organizing networks," vol. 13, no. 1, pp. 85-98, 2016.
- [9] I. A. Aljazaery, H. T. S. Alrikabi, and M. R. J. i. Aziz, "Combination of Hiding and Encryption for Data Security," vol. 14, no. 9, p. 35, 2020.
- [10] P. Lin et al., "A west-east bridge based SDN inter-domain testbed," vol. 53, no. 2, pp. 190-197, 2015.
- [11] B. Mohammed, R. Chisab, and H. Alrikabi, "Efficient RTS and CTS Mechanism Which Save Time and System Resources," 2020.
- [12] O. H. Yahya, H. Alrikabi, I. A. J. I. J. o. O. Aljazaery, and B. Engineering, "Reducing the Data Rate in Internet of Things Applications by Using Wireless Sensor Network," vol. 16, no. 03, pp. 107-116, 2020.
- [13] F. T. Abed, H. T. S. ALRikabi, and I. A. Ibrahim, "Efficient Energy of Smart Grid Education Models for Modern Electric Power System Engineering in Iraq," in *IOP Conference Series: Materials Science and Engineering*, 2020, vol. 870, no. 1, p. 012049: IOP Publishing.
- [14] P. V. Klaine, M. A. Imran, O. Onireti, R. D. J. I. C. S. Souza, and Tutorials, "A survey of machine learning techniques applied to self-organizing cellular networks," vol. 19, no. 4, pp. 2392-2431, 2017.
- [15] S. Hämäläinen, H. Sanneck, and C. Sartori, *LTE self-organising networks (SON): network management automation for operational efficiency*. John Wiley & Sons, 2012.
- [16] C. Langton, *Artificial life: proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems*. Routledge, 2019.
- [17] E. U. T. R. J. U. h. w. g. o. Access, "Radio Resource Control (RRC) Protocol Specification," 2015.
- [18] N. T. Schmidt, "Distributed Simulation and Emulation of Mobile Nodes in Dynamic Networks," 2016.

- [19] I. Al Barazanchi, H. R. Abdulshaheed, M. Safiah, and B. Sidek, “A Survey : Issues and challenges of communication technologies in WBAN,” *Sustain. Eng. Innov.*, vol. 1, no. 2, pp. 84–97, 2020.
- [20] I. Abdulshaheed, H. R., Yaseen, Z. T., Salman, A. M., & Al_Barazanchi, “An Evaluation study of WiMAX and WiFi on Vehicular Ad-Hoc Networks (VANETs),” *IOP Conf. Ser. Mater. Sci. Eng. Pap.*, vol. 3, no. 12, pp. 1–7, 2012.
- [21] B. Durakovic and Mešetović, S., “Thermal Performances of Glazed Energy Storage Systems with Various Storage Materials: An Experimental study”, *Sustainable Cities and Society*, vol. 45, pp. 422-430, 2019.