THESIS	
Title	**Participation Strategies using Learning Algorithms in a Marketplace of 5G Radio Connectivity**
Contact	https://orange.jobs/jobs/offer.do?joid=71066&lang=FR
Supervisors (Orange)	**Marjou Xavier, Radier Benoit** Orange, Lannion, France
Supervisors (University)	**Director: Lemlouma Tayeb** **Co-Supervisor: Bouabdallah Ahmed** University of Rennes I (IUT of Lannion) / IMT Atlantique IRISA Labs –MathSTIC Doctoral School tayeb.lemlouma@irisa.fr
Localization	Orange Lannion and IUT of Lannion (University of Rennes 1)

Abstract

The topic of this thesis focuses on the study of telecommunication mechanisms related to future 5G networks. The main goal is to design and implement AI algorithms that allow autonomous software agents to acquire the required knowledge, for instance using available AI tools, and define the optimal approaches for learning, decision support, and prediction in a marketplace environment.

Keywords

4G Networks; 5G Networks; Artificial Intelligence; Machine Learning; Reinforcement Learning; Predictive Algorithms; Markov Decision Process; Marketplace; Trading

Context

Marketplaces (or competitive markets) are a meeting point where consumers and providers can buy and sell items or services. An open marketplace encourages competition between the service providers by allowing the services to be standardized and compared. The marketplace also allows the resellers of a service to thrive by making no distinction between a seller and a reseller and focusing only on the service [13]. Although the concept of "marketplace" and techniques for advertising and purchasing services are not new and first standardization can be traced back to 2000 with UDDI, UPnP, OSGi, MINT (a market for INternet transit), virtualized network functions, etc. [13-15], the fields of telecommunication and IP connectivity provision were not concerned by such marketplace to date. However, this situation could significantly change particularly with the context of future 5G networks. In such cases, it is therefore essential to study how national and virtual network operators should act, interact, and take profit from such environments. Such targeted participation in a marketplace have to be automated, which requires the design of highly autonomous software agents that should be able of rapid decision-making in terms of connectivity purchasing and selling based on the specific objectives of the operator. Hence, the software agent could first rely on predictive mechanisms and the evolution of the state of the network (i.e. available resources, costs, etc.) using (un)supervised learning algorithms. Moreover, agents have to learn how to optimally interact with the marketplace in order to increase the possible benefits or scores. Ideally, we aim that the learned participation behavior becomes as automatic as possible by discovering the rules of the "game" and to be adapted to these marketplace rules. Reinforcement and learning algorithms are identified as the best candidates to reach this objective.

Subject

Reinforcement learning is a branch of Machine Learning that offers the ability to learn without a prior

knowledge about the game's rules and how to implement these rules in order to enable the software agent (i.e. the player) to interact with the environment [1-2]. This form of learning is becoming increasingly popular in particular since the great success of the Google's *AlphaGo* program [3]. Among the keys to this success, we can cite the emergence and advent of Deep Learning [16], which allows computational models that are composed of multiple processing layers to learn Big Data with various levels of abstraction and the use of learning algorithm with high-performance hardware. These techniques are also used in other applications such as the determination of the best possible participation in other games such as Chess [4] or Atari video game [5]. Open-source frameworks, such as OpenAI Gym [6], are also emerging and allow the experimentation of various reinforcement learning algorithms. The OpenAI Gym toolkit, for instance, already integrates the modeling of several environmental problems to be solved. For instance, an environment where the objective is to find a resource and move it to a suitable location in a geographic environment [7] which allows experimenting solving techniques such as the hierarchical reinforcement learning [8].

As previously highlighted, a number of works have been proposed and applied in order to solve some categories of problems and application domains (e.g. video games, robotics, etc.). However, very few works have been conducted in the telecommunications field and specifically in the field of optimizing the operator's participation in a marketplace. The objective of this thesis is to investigate and design predictive algorithms that lead to the determination of the best operator's participation strategies in a 5G cellular connectivity marketplace. This involves the design of a software agent that will participate in a near-autonomous way in the marketplace by implementing a Reinforcement Learning based approach. To do this, it will be necessary to implement sequential investment models in the manner of those used in Trading such as "*Buy-and-hold*" or "*Constantly rebalanced portfolios*" strategies [12]. The design of an environment is also required to simulate agent interactions and resulting or expected gains or losses.

The main stages of this Ph.D. thesis are:

First year
The study of the Reinforcement Learning and design of a marketplace environment based on the Markov Decision Process [9]. This environment will be used to experiment the abilities of a software agent representing the participation of a network operator in a marketplace. It will also be necessary to study the scalability and dimensioning of the environment in order to find a tradeoff between the complexity and the required processing time when using an implementation framework of deep reinforcement learning [2]. It will be then necessary to design, implement and evaluate the agent algorithms interacting with this Marketplace environment. Typically, the prototyping environment will be a framework similar or based on the OpenAI Gym framework [6].

Second year
In addition to the external gains (benefits/rewards) obtained in the marketplace, we aim to integrate, into the proposed smart algorithms of the software agent, the concept of internal rewards from the operator's network. The work should also identify and propose predictive mechanisms (e.g. using the Grey model or other candidate models [10-11]) regarding the intra-operator or even extra-operator resources in order to enrich the decision-making of the participation in the marketplace. Depending on the progress of the research, the first results must be valorized through studies, publications and/or patents.

Third year
Writing of the thesis manuscript and the valorization of the results with well-ranked publications and/or patents.

Candidate profile

The candidate must hold an MSc diploma (national Master's degree) or another degree conferring the French Master's degree (Master 2). An exemption request is required for all other situations. It is very preferable that the candidate has a good background in the field of machine learning, and ideally in the sub-fields of reinforcement learning, and/or techniques/software used in the field of trading.

The following qualities/skills are also considered:

- Autonomy and motivation for research
- Good communication skills in English (oral and written)
- Very good knowledge of statistics and game theory
- Machine learning (supervised learning, unsupervised learning, reinforcement learning)
- Good programming skills in Python language and related libraries (e.g. numpy, pandas, matplotlib)
- OpenAI Gym and/or Markov decision process (MDP) skills would be a plus

Team Description

The Orange team of Architecture for Telecommunications (ARC) is in charge of designing the architectures of Orange's future networks. In this respect, it brings together network architects, standardization members (W3C, IETF, 3GPP), researchers, Ph.D. students, apprentices, and trainees.

The OCIF team of the D2 (Networks, Telecommunications, and Services) department of the IRISA works on the IP protocol and various application domains. The goal is to define and manage the needs and constraints related to future Internet networks and experimental environments required to validate the proposed concepts.

Funding bodies

CIFRE (100 %)

Bibliography

[1] R. S. Sutton and A. G. Barto, Reinforcement Learning, An introduction, MIT Press Cambridge, MA, UK, 2017
[2] K. Arulkumaran, M. P. Deisenroth, M. Brundage and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," in IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 26-38, Nov. 2017.
[3] D. Silver et al. Mastering the Game of Go without Human Knowledge, International Journal of Science Nature, vol. 550, pp 354–359, 2017. https://www.nature.com/articles/nature24270
[4] D. Silver et al. Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, arXiv preprint, arXiv:1712.01815, 2017a. https://arxiv.org/abs/1712.01815
[5] V. Mnih et al. Playing Atari with deep reinforcement learning. Deepmind Technologies, 2013, arXiv:1312.5602, https://arxiv.org/abs/1312.5602
[6] Gym Project: toolkit for developing and comparing reinforcement learning algorithms, 2018, https://gym.openai.com/
[7] Gym Project: Modélisation d'environnement sous Gym, 2018, https://github.com/openai/gym/blob/master/gym/envs/toy_text/taxi.py
[8] G. D. Thomas, Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition, Journal of Arti cial Intelligence Research, Vol. 13, pp 227-303, 2000, http://jmvidal.cse.sc.edu/library/dietterich00a.pdf
[9] W. T. Scherer, S. Adams and P. A. Beling, "On the Practical Art of State Definitions for Markov Decision Process Construction," in IEEE Access, vol. 6, pp. 21115-21128, 2018.
[10] N. E. Frigui, T. Lemlouma, S. Gosselin, B. Radier, R. Le Meur and J-M. Bonnin Optimization of the Upstream Bandwidth Allocation in Passive Optical Networks Using Internet Users' Behavior Forecast, 22nd ONDM 2018, Dublin, Ireland, 14-17 May, 2018.

[11] Z. Liouane, T. Lemlouma, P. Roose, F. Weis and H. Messaoud, An Improved Extreme Learning Machine Model for the Prediction of Human Scenarios in Smart Homes, Springer Journal of Applied Intelligence, September, 2017.
[12] N. Cesa-Bianchi and G. Lugosi, Prediction, Learning, and Games, Cambrige University Press, 2006
[13] S. Bhat, R. Udechukwu, R. Dutta and G. N. Rouskas, Network Service Orchestration in Heterogeneous 5G Networks Using an Open Marketplace, IET Networks Journal, vol. 6, no. 6, pp. 149-156, 11 2017.
[14] D. Yu, L. Mai, S. Arianfar et al., Towards a Network Marketplace in a Cloud. Proc. of the 8th USENIX Conf. on Hot Topics in Cloud Computing, ser. HotCloud'16, Berkeley, CA, USA, USENIX Association, 2016, pp. 84– 89.
[15] G. Xilouris, E. Trouva, F. Lobillo, F. et al., T-nova: A Marketplace for Virtualized Network Functions. 2014 European Conf. on Networks and Communications (EuCNC), June 2014, pp. 1–5
[16] Y. LeCun, Y. Bengio and G. Hinton, Deep learning. International Journal of Science Nature, vol. 521, pp 436-444, 2015.