

SUJET DE THESE	
Titre	Algorithmes d'Apprentissage pour une Stratégie de Participation à un Marché de la Connectivité Radio 5G
Contact	https://orange.jobs/jobs/offer.do?joid=71066&lang=FR
Encadrants en entreprise	Marjou Xavier, Radier Benoit Orange, Lannion, France
Encadrants académiques	Directeur : Lemlouma Tayeb Co-encadrant : Bouabdallah Ahmed Université de Rennes I (IUT de Lannion) / IMT Atlantique Laboratoire IRISA – Ecole doctorale MathSTIC tayeb.lemlouma@irisa.fr
Localisation	Orange Lannion et IUT de Lannion (Université de Rennes 1)
Résumé	
<p><i>Cette thèse permettra de découvrir les mécanismes des réseaux de télécommunications qui intégreront la 5G de demain. En outre, elle sera l'occasion de concevoir et de mettre en œuvre des algorithmes dits d'intelligence artificielle qui permettent à des agents logiciels d'apprendre par eux-mêmes, tout en utilisant les derniers outils dit d'IA existants sur le marché ainsi que de définir des approches optimales d'apprentissage, d'aide à la décision et de la prédiction.</i></p>	
Mots clés	
<p><i>Réseaux 4G ; Réseaux 5G ; Intelligence Artificielle ; Machine Learning ; Apprentissage par le renforcement; Algorithmes de prédiction ; Markov decision process ; Trading</i></p>	
Contexte	
<p>Les marketplaces (ou marchés concurrentiels) sont un endroit où se rencontrent les demandeurs et les fournisseurs pour acheter et vendre des objets ou des services [13 - 15]. Sur Internet, lieu qui abolit les distances, leur importance est encore accrue. Le monde des télécommunications et plus particulièrement de la fourniture de la connectivité réseau (IP) n'utilise généralement pas de telle marketplace mais cela pourrait voir le jour notamment dans le contexte de la 5G. Si tel était le cas, il est ainsi primordial d'étudier comment agirait un opérateur qu'il soit opérateur national ou opérateur de réseau virtuel. En effet, la participation à une marketplace devra être automatisée, ce qui nécessitera de concevoir des agents logiciels très autonomes et qui sont aptes à une prise de décision rapide en matière d'achat ou de vente de connectivité et ce en fonction des intérêts spécifiques à l'opérateur. Pour ce faire, l'agent logiciel pourra tout d'abord se baser sur des prédictions de l'évolution de l'état de son réseau (c.-à-d. des ressources disponibles, leur coût, etc.) à l'aide d'algorithmes d'apprentissage supervisés et/ou non-supervisés. Mais il lui faudra surtout apprendre à interagir avec la marketplace de manière à obtenir le plus de récompenses possibles. Idéalement, on vise que l'apprentissage de la participation soit le plus automatique possible en découvrant et en s'adaptant au fil de l'eau aux règles du jeu de la marketplace. Pour ce faire, il est primordial de concevoir un algorithme d'apprentissage par renforcement.</p>	
Sujet détaillé	
<p>L'apprentissage par renforcement (reinforcement learning) [1-2] est une branche de l'apprentissage automatique (machine learning) qui offre la particularité d'apprendre sans avoir à connaître ou implémenter explicitement les règles du jeu de l'agent, c'est-à-dire du joueur interagissant avec un</p>	

environnement. Cette forme d'apprentissage voit un engouement depuis le succès du programme *AlphaGo* de Google [3]. Parmi les clés de ce succès on peut citer l'avènement de l'apprentissage profond (deep learning) [16] qui repose sur l'utilisation conjointe de Big Data, d'algorithmes d'apprentissage et de matériel (hardware) très performant (GPU). Ces techniques sont aussi utilisées dans d'autres applications telles que pour résoudre la meilleure participation possible à d'autres jeux tels que les jeux d'échecs [4] ou le jeu vidéo Atari [5]. Des frameworks open-source sont également en train d'émerger pour permettre d'expérimenter divers algorithmes d'apprentissage par renforcement, à l'instar de Gym d'OpenAI [6]. Ce dernier intègre déjà la modélisation de nombreux environnements de problèmes à résoudre. Par exemple, on peut citer l'exemple d'un environnement où il faut chercher une ressource et la déplacer à un endroit adéquat dans un environnement géographique [7], ce qui permet d'éprouver des techniques de résolution de ces problèmes (ex : hierarchical reinforcement learning [8]).

Si des travaux existent déjà afin de résoudre certaines catégories de problème et de domaines d'applications (e.g. jeux vidéo, robotique, etc.), très peu de travaux existent encore pour le domaine des télécommunications, et encore moins dans le domaine de l'optimisation d'une participation d'opérateur à une marketplace. L'objectif de la thèse est de concevoir des algorithmes prédictifs pour une stratégie de participation d'un opérateur à une marketplace de connectivité cellulaire 5G. Ceci implique de concevoir un agent logiciel qui participera de manière quasi autonome à une marketplace en mettant en œuvre l'apprentissage par renforcement. Pour ce faire, il faudra mettre en œuvre des modèles d'investissement séquentiels à la manière de ceux utilisés dans le trading tels que les stratégies "*Buy-and-hold*" ou "*constantly rebalanced portfolios*" [12]. Il faudra également concevoir un environnement permettant de simuler les interactions des agents et les gains ou les pertes résultants ou attendus.

Le travail de thèse inclut les aspects suivants :

Première année

Etude de l'apprentissage par renforcement et conception d'un environnement marketplace, (environnement au sens Markov Decision Process [9]), environnement qui sera utilisé pour tester le potentiel d'agent logiciel représentant la participation d'un opérateur de réseau. Il faudra également étudier le dimensionnement de l'environnement de telle sorte qu'il soit réaliste entre une complexité intéressante et un temps de calcul réaliste dans le cadre de mise en œuvre de type *deep reinforcement learning* [2]; Il faudra ensuite concevoir, implémenter et expérimenter les algorithmes d'agent interagissant avec cet environnement Marketplace. Typiquement, l'environnement de prototypage sera un framework tel que le framework Gym d'OpenAI [6].

Deuxième année

En plus des récompenses externes issues de la marketplace, on vise à intégrer la notion de récompenses internes issues du réseau de l'opérateur dans les algorithmes intelligents d'un agent. Le travail de la thèse devrait également identifier et proposer des mécanismes prédictifs (e.g. le modèle Grey ou d'autres modèles candidats [10-11]) pour les ressources intra-opérateurs, voire extra-opérateurs pour enrichir les algorithmes de prise de décision à la participation à la marketplace. Selon l'avancée de la recherche, les premiers résultats doivent être valorisés sous formes de publications et/ou de brevets.

Troisième année

Rédaction de manuscrit de thèse, valorisation des résultats avec la rédaction de publications scientifiques et dépôt de brevets.

Prérequis et compétences scientifiques et techniques

Le candidat doit être titulaire d'un diplôme national de Master ou d'un autre diplôme conférant le grade de Master. Une demande de dérogation est exigée pour toutes les autres situations. Il est très

préférable que le candidat ait une expérience dans le domaine du machine Learning, et idéalement dans le sous-domaine du renforcement learning, et/ou une expérience dans le domaine des techniques/logiciels utilisés dans le domaine du trading. Les qualités/compétences suivantes sont également prises en considération :

- Autonomie
- Maîtrise de l'anglais
- Très bonnes connaissances en statistiques et en théorie des jeux
- Machine learning (supervised learning, unsupervised learning, reinforcement learning)
- Maîtrise du langage Python et de ses bibliothèques (numpy, pandas, matplotlib)
- La connaissance d'OpenAI Gym serait un plus
- La connaissance des processus de décision markovien (MDP) serait un plus

Description de l'équipe

L'équipe Architecture for Telecommunications (ARC) de Orange est une équipe en charge de concevoir les architectures des futurs réseaux d'Orange. A ce titre, elle regroupe des architectes réseaux, des délégués en standardisation (W3C, IETF, 3GPP), des chercheurs, des thésards, des apprentis et des stagiaires.

L'équipe OCIF du département D2 (Réseaux, Télécommunications et Services) de l'IRISA est une équipe qui travaille autour du protocole IP des domaines applicatifs. Ces derniers permettent de définir les besoins et les contraintes des réseaux Internet du futur et également de terrain d'expérimentation pour valider les concepts développés.

Financements / Funding bodies

CIFRE (100 %)

Bibliographie

- [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 Artificial 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, Cambridge 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.