



<https://latribunelibre.com/emploi/post-doc-algorithmes-de-bandits-pour-la-recherche-de-plus-court-chemin-f-h-3>

Post doc-Algorithmes de bandits pour la recherche de plus court chemin F/H

Description

Votre rôle est d'effectuer un travail de Post doc sur le sujet » Algorithmes de bandits pour la recherche de plus court chemin dans un environnement multi-agents, stochastique ou adverse » dans le cadre du projet ANR CANCUN (Time-Sensitive Vehicular Control and Collaborative Perception via AI-augmented Sustainable Networks) démarré en 2025.

Dans un environnement industriel, comme un port, une chaîne de montage..., des véhicules autonomes et d'autres pilotés par des humains doivent parcourir des trajets correspondant à leurs différentes tâches. Les véhicules partagent le même réseau de circulation et peuvent se gêner mutuellement, modifiant les temps de parcours et donc la consommation énergétique. L'objectif est de minimiser l'énergie nécessaire à la réalisation des parcours des véhicules autonomes. Ce problème peut se modéliser comme une recherche de plus court chemin stochastique dans un Processus de Décision de Markov [1], où les transitions entre les noeuds (croisements) connaissent les actions (choix des routes) sont déterministes et où l'énergie nécessaire au parcours d'une route serait une variable aléatoire. Des premiers travaux considérant un seul véhicule autonome dans un trafic mixant véhicules autonomes et véhicules pilotés par des humains ont montré que dans le cadre stochastique, un algorithme de bandits [2] permettait d'explorer quasi-optimalement l'espace des chemins [3]. L'objet de ces travaux de recherche est d'étendre cette approche dans plusieurs directions :

- Multi-agent : contrôle de plusieurs véhicules autonomes partageant le même réseau de circulation en utilisant par exemple une approche bandits multi-joueurs [4].
- Communication : utilisation de messages entre les agents pour accélérer la convergence [5].
- Adversité : les véhicules pilotés par des humains ont des buts déterministes, et ont donc un effet non stochastique (en pire cas adverse) sur les temps de parcours des véhicules autonomes [6].

Les algorithmes seront analysés, puis testés sur un simulateur (SUMO), puis sur une plateforme de véhicules commandés à distance en milieu industriel.

Qualifications

- Doctorat informatique / mathématiques / réseaux de télécommunication.
- Compétences en apprentissage et notamment apprentissage par renforcement, processus de décision markovien, bandits.
- Compétences en mathématiques, statistiques, optimisation.
- Compétence en réseau de télécommunication mobile.
- Compétences en informatique (Python, C++), la connaissance du simulateur

Organisme employeur
Orange SA

Type de poste
Temps plein

Secteur
TÉLÉCOMMUNICATIONS
FILAIRE

Lieu du poste
22113, LANNION, LANNION,
France

Salaire de base
40000 € - **Salaire de base**
45000 €

Date de publication
10 janvier 2026 à 10:03

Valide jusqu'au
09.02.2026

de réseaux routiers SUMO serait un plus.

- Maitrise de l'anglais écrit et oral.
- Bonnes qualités relationnelles.

Les références

- [1] Dynamic Programming and Optimal Control: Volume I, D. Bertsekas, Belmont, MA, USA: Athena Scientific, 4th ed., 2012.
- [2] Bandit Algorithms, T. Lattimore and C. Szepesvari, Cambridge University Press, 2020.
- [3] Multi-armed Bandit for Stochastic Shortest Path in Mixed Autonomy, Yu Bai, Yiming Li, and Xi Xiong, arXiv, 2025.
- [4] Multi-Player Bandits – a Musical Chairs Approach, Jonathan Rosenski, Ohad Shamir, Liran Szlak, NeurIPS, 2016.
- [5] Decentralized Exploration in Multi-Armed Bandits, Raphaël Féraud, Réda Alami, Romain Laroche, ICML, 2019.
- [6] Stochastic Shortest Path with Adversarially Changing Costs, Aviv Rosenberg and Yishay Mansour, IJCAI, 2021.