

## **Titre : Algorithmes de Bandits Multi-Joueurs pour les Réseaux de l'Internet des Objets**

**Mots clés :** Internet des Objets (IdO), Radio Intelligente, Théorie de l'apprentissage, Apprentissage séquentiel, Apprentissage par renforcement, Bandits multi-bras (BMB), Apprentissage décentralisé, Bandits multi-bras multi-joueurs, Détection des points de changement, Bandits multi-bras non stationnaires

**Résumé :** Dans cette thèse de doctorat, nous étudions les réseaux sans fil et les appareils reconfigurables qui peuvent accéder à des réseaux de radio intelligente, dans des bandes non licenciées et sans supervision centrale. Nous considérons des réseaux actuels ou futurs de l'Internet des Objets (IdO), avec l'objectif d'augmenter la durée de vie de la batterie des appareils, en les équipant d'algorithmes d'apprentissage machine peu coûteux mais efficaces, qui leur permettent d'améliorer automatiquement l'efficacité de leurs communications sans fil. Nous proposons différents modèles de réseaux de l'IdO, et nous montrons empiriquement, par des simulations numériques et une validation expérimentale réaliste, le gain que peuvent apporter nos méthodes, qui se reposent sur l'apprentissage par renforcement.

Les différents problèmes d'accès au réseau sont modélisés avec des Bandits Multi-Bras (BMB), mais l'analyse de la convergence d'un grand nombre d'appareils jouant à un jeu collaboratif sans communication ni aucune coordination reste délicate, lorsque les appareils suivent tous un modèle d'activation aléatoire. Le reste de ce manuscrit étudie donc deux modèles restreints, d'abord des bandits multi-joueurs dans des problèmes stationnaires, puis des bandits mono-joueur non stationnaires. Nous détaillons également une autre contribution, la bibliothèque Python open-source SMPyBandits pour des simulations numériques de problèmes MAB, qui couvre les modèles étudiés et d'autres.

## **Title : Multi-Players Bandit Algorithms for Internet of Things Networks**

**Keywords:** Internet of Things (IoT), Cognitive Radio, Learning Theory, Sequential Learning, Reinforcement Learning, Multi-Armed Bandits (MAB), Decentralized Learning, Multi-Player Multi-Armed Bandits, Change Point Detection, Non-Stationary Multi-Armed Bandits

**Abstract:** In this PhD thesis, we study wireless networks and reconfigurable end-devices that can access Cognitive Radio networks, in unlicensed bands and without central control. We focus on Internet of Things networks (IoT), with the objective of extending the devices' battery life, by equipping them with low-cost but efficient machine learning algorithms, in order to let them automatically improve the efficiency of their wireless communications. We propose different models of IoT networks, and we show empirically on both numerical simulations and real-world validation the possible gain of our methods, that use Reinforcement Learning.

The different network access problems are modeled as Multi-Armed Bandits (MAB), but we found that analyzing the realistic models was intractable, because proving the convergence of many end-devices playing a collaborative game without communication nor coordination is hard, when end-devices all follow random activation patterns. The rest of this manuscript thus studies two restricted models, first multi-players bandits in stationary problems, then non-stationary single-player bandits. We also detail another contribution, SMPyBandits, an open-source Python library for numerical MAB simulations, covering all the studied models and more.