2023-06498 - Post-Doctoral Research Visit F/M
Learning Restless Bandits, with Applications to Stochastic Scheduling

Type de contrat : CDD
Niveau de diplôme exigé : Thèse ou équivalent
Fonction : Post-Doctorant

A propos du centre ou de la direction fonctionnelle

The Centre Inria de l’Université de Grenoble groups together almost 600 people in 22 research teams and 7 research support departments.

Staff is present on three campuses in Grenoble, in close collaboration with other research and higher education institutions (Université Grenoble Alpes, CNRS, CEA, INRAE, …), but also with key economic players in the area.

The Centre Inria de l’Université Grenoble Alpes is active in the fields of high-performance computing, verification and embedded systems, modeling of the environment at multiple levels, and data science and artificial intelligence. The center is a top-level scientific institute with an extensive network of international collaborations in Europe and the rest of the world.

Contexte et atouts du poste

Travel expenses (e.g. conferences) are covered within the limits of the scale in force.

Mission confiée

Multi-armed restless bandit problems (MARBP) are used to model optimal resource allocation problems. In a MARBP, there is a collection of stochastic binary-action (active/passive) projects that evolve over time, and the goal is to allocate resource to these projects in real time. A typical example of such a problem is given by stochastic scheduling, in which a machine needs to share its time between various resources. Other applications include asset management [6] or maintenance problems [3].

Until recently, MARBP's have been mostly studied from the computational point of view: for a given problem, how to compute a policy that gives a closesto-optimal solution. While these problems have been proved to be computationally hard [7], there exist computationally efficient policies, like Whittle index policies [8], which perform extremely well in practice while being computationally efficient [4].

Whittle index policy, however, assumes that the model is known. The goal of this postdoctoral project is to explore what we can do when the model information (transition probabilities and sometimes rewards) is unknown. To do that, we address this problem by deploying reinforcement learning (RL) based agents.

Although RL algorithms have been recently explored for some restless bandits models [5, 2, 1], their performance guarantee and applicability in stochastic scheduling remains limited.

We want to contribute to the following questions:

1. Starting from stochastic scheduling, propose a data-driven approach to this problem by modeling it as a restless bandit;
2. Develop reinforcement learning algorithms for the general multi-armed restless bandit problem where...
each bandit has Markovian transitions;

3. Envision application to other domains such as dynamic asset allocation, outsourcing warranty repairs, perishable inventory routing, etc.

The objective is to provide a framework for optimal control of stochastic distributed agents. This project is mostly of theoretical nature: it aims at contributing to the field of optimal control and reinforcement learning by developing new control mechanisms for distributed systems. We also plan to release prototype implementation of our optimization algorithm as open-source library.

The postdoctoral fellow will be supported by:

- A national ANR grant (JCJC REFINO) on Markovian bandits.
- A bilateral partnership between France and India via the "équipe-projet associée" AIRBEA

**Principales activités**

Main activities:

Examples of activities:

- Propose learning solutions for restless bandits
- Develop prototype implementations
- Design new models
- Write scientific papers

**Compétences**

Technical skills and level required:

Languages:

Relational skills:

Other valued appreciated:

**Avantages**

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage under conditions

**Rémunération**

2746 € gross salary / month