Multi-Armed bandit Learning in IoT Networks (MALIN)

Rémi Bonnefoi¹, Lilian Besson¹ and Christophe Moy²
¹ CentraleSupélec/IETR, F-35576, Cesson-Sévigné Cedex, France. Remi.Bonnefoi@CentraleSupelec.fr, Lilian.Besson@CentraleSupelec.fr
² Univ Rennes, CNRS, IETR – UMR 6164, F-35000, Rennes, France. Christophe.Moy@Univ-Rennes1.fr

Goal

- With the advent of the Internet of Things (IoT), unlicensed band are going to be shared by a large number of devices with dissimilar characteristics. In such context, solutions are required to allow the coexistence of devices and to avoid performance drop due to interference.
- In this demonstration, we show that reinforcement learning algorithms and in particular Multi-Armed Bandit algorithms can be used as a means of improving the performance of IoT communications.

MAB learning

- A user faces N choices (e.g. N channels)
- The channels provide him a reward with a given probability

How to identify the best channel?

- MAB learning algorithms are proved to be optimal to solve this problem.
- Any MAB algorithm can be used for channel selection (UCB, Thompson Sampling or others)
- With the UCB algorithm, the channel with the highest index is chosen for each transmission

\[
B_j(t) = \sum_{i=0}^{t-1} r_i(t) \mathbb{I}(a_i = j) + \sqrt{\frac{\alpha \ln(t)}{T_j(t)}}
\]

Acknowledgement

This work is supported by the European Union through the European Regional Development Fund (ERDF), and by Ministry of Higher Education and Research, Brittany and Rennes Métropole, through the CPER Project SOPHIE / STIC & Ondes. Part of this work is also funded by the ANR projects SOGREEN and BADASS, by Région Bretagne, France, by the French Ministry of Higher Education and ENS Paris Saclay.

References