2022 NSF CYBER-PHYSICAL SYSTEMS PRINCIPAL INVESTIGATORS' MEETING

Risk-Averse No-Regret Learning in Online Convex Games

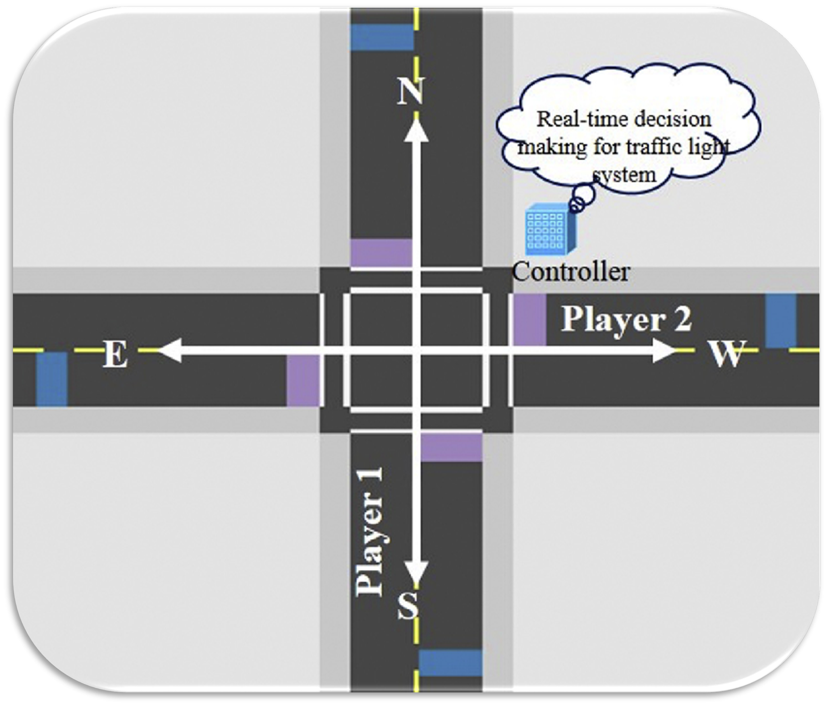
Michael M. Zavlanos

Duke University

Challenge:

- The distribution of an agent's cost function depends on other agents' actions.
- Using finite bandit feedback, it is difficult to accurately estimate the CVaR values.

Project Overview



Problem Formulation

An online convex game with N agents

- Cost function:

$$C_i(x) := \mathrm{CVaR}_{\alpha_i}[J_i(x,\xi_i)] = \mathbb{E}_F[J_i(x,\xi_i) | J_i(x,\xi_i) \geq J^{\alpha_i}],$$

where J^{α_i} is the $1-\alpha_i$ quantile of the distribution.

- CVaR-regret:

$$\mathrm{R}_{C_i}(T) = \underbrace{\sum_{t=1}^{T} C_i(\hat{x}_{i,t}, \hat{x}_{-i,t})}_{\text{actions selected by the algorithm}} - \underbrace{\min_{\tilde{x}_i \in \mathcal{X}_i} \sum_{t=1}^{T} C_i(\tilde{x}_i, \hat{x}_{-i,t})}_{\text{the best action in hindsight}}$$

Solution:

- A new risk-averse online learning algorithm that can achieve sub-linear regret in online convex games, i.e., performing as good as the best actions in hindsight.

Scientific Impact:

- Applications of the proposed algorithms to black-box CPS learning and control systems with risk-averse agents.
- The first convergence guarantees on risk-averse online convex games with bandit feedback.

Broader Impact:

- Applications to many domains, e.g., smart city, health care, etc.
- K-12, undergraduate, and graduate education
- Promoting diversity

Award ID#: CNS #1932011, Award Date: 10/01/2019