國立陽明交通大學電信工程研究所
系統組論文研討

DISTRIBUTED NO-REGRET LEARNING FOR MULTI-STAGE SYSTEMS WITH END-TO-END BANDIT FEEDBACK

主講人：美國德州農工大學 電機與電腦工程學系 侯弈宏 教授

日期：114 年 10 月 08 日（星期三）時間：13:20~15:20 地點：工程四館219室

Abstract

This talk focuses on multi-stage systems with end-to-end bandit feedback. In such systems, each job needs to go through multiple stages, each managed by a different agent, before generating an outcome. Each agent can only control its own action and learn the outcome of the job. It has neither knowledge nor control on actions taken by agents in the next stage. The goal of this talk is to develop distributed online learning algorithms that achieve sublinear regret in adversarial environments.

This multi-stage setting significantly expands the traditional multi-armed bandit problem, which considers only one agent and one stage. In addition to the exploration-exploitation dilemma in the traditional multi-armed bandit problem, we show that the consideration of multiple stages introduces a third component, education, where an agent needs to choose its actions to facilitate the learning of agents in the next stage. To solve this newly introduced exploration-exploitation-education trilemma, we propose a simple distributed online learning algorithm, ϵ-EXP3. We theoretically prove that the ϵ-EXP3 algorithm is a no-regret policy that achieves sublinear regret. Simulation results show that the ϵ-EXP3 algorithm significantly outperforms existing no-regret online learning algorithms for the traditional multi-armed bandit problem.

Biography

Dr. Hou is a Professor in the ECE Department of Texas A&M University. His research interests include cloud/edge computing, wireless networks, and machine learning.

Dr. Hou received the B.S. in Electrical Engineering from National Taiwan University in 2004 and his M.S. and Ph.D. in Computer Science from the University of Illinois, Urbana-Champaign in 2008 and 2011, respectively. He received the 2025 IEEE Communications Society William R. Bennett Prize, the Best Paper Awards in ACM MobiHoc 2020 and ACM MobiHoc 2017, the Best Student Paper Award in WiOpt 2017, and the C.W. Gear Outstanding Graduate Student Award from the University of Illinois at Urbana-Champaign.