

**Off-Policy Evaluation in Reinforcement Learning via a Factor-Embedding MDP**

**Degui Li**

**Faculty of Business Administration at the University of Macau**

**Email: [deguili@um.edu.mo](mailto:deguili@um.edu.mo)**

In this talk, the speaker will introduce a factor-embedding Markov decision process (MDP), a general framework for dynamic systems that captures complex heterogeneity. Unlike conventional factor analysis, a factorized structure is incorporated into the transition dynamics of an MDP, allowing evolving and context-specific heterogeneity across both individuals and time. Within the developed framework, an iterative algorithm is proposed to estimate the expected reward at a given time for a specific subject. The theoretical statistical guarantees of the proposed estimators are established under some regularity conditions, particularly with regard to the maximum error rate of a low-rank matrix. Both simulation and empirical studies are provided to examine the numerical performance of the developed model and methodology.