Title: Phase Transitions and Cyclic Phenomena in Online and Offline Learning

Speaker: David Simchi-Levi, Professor of Engineering Systems, MIT

Abstract:

We consider the classical stochastic multi-armed bandit problem with constraints on the total switching cost and total inventory consumption. We prove matching upper and lower bounds on regret and provide near-optimal algorithms for this problem. Surprisingly, we discover phase transitions and cyclic phenomena of the optimal regret. That is, we show that associated with the inventory constrained multi-armed bandit problem, there are phases defined by the number of arms and switching costs, where the regret upper and lower bounds in each phase remains the same and drop significantly between phases. The results enable us to fully characterize the trade-off between regret and incurred switching cost in the stochastic inventory constrained multi-armed bandit problem, contributing new insights to this fundamental problem. Finally, we discover similar phenomena in a classical revenue management problem, the dynamic pricing model, where pre-existing offline data is applied to improve the online learning process.