



統計科學研究所

INSTITUTE OF
STATISTICAL SCIENCE



S E M I N A R



S T A T I S T I C S

Seminar

Title : Exploration Through Reward Biasing in Bandits

Speaker : Prof. Ping-Chun Hsieh (Department of Computer Science, National Yang Ming Chiao Tung University)

Time : 10:30 AM~12:00 PM, Monday, June 6, 2022

Place : Presented by Webex Meeting



Abstract

Bandit learning is a classic machine learning framework that captures the explore-exploit dilemma in many sequential decision-making problems. Despite the existing variety of bandit algorithms, there is still an unsatisfactory trade-off between regret performance and computational efficiency. To tackle this, we present a new family of bandit algorithms, that are formulated in a general way based on the reward-biased maximum likelihood estimation (RBMLE) principle. This talk will cover the following three aspects:

- (i) Bandits and Regret: I will provide a brief overview of the bandit learning problems and the notion of regret.
- (ii) RBMLE: I will introduce the generic RBMLE approach in adaptive control and present how to adapt RBMLE to stochastic bandits. Through theoretical analysis and simulations, we demonstrate that the proposed RBMLE achieves regret performance comparable to the best of state-of-the-art methods while having a significant computational advantage in comparison to other best-performing methods.
- (iii) RBMLE for Contextual Bandits: Based on (ii), I will present how to extend RBMLE to contextual bandits, including the Neural Contextual Bandits that leverage the representation power of neural networks in bandit problems.

※ Lecture in English.