Title:
Optimal Learning and Experimentation in Bandit Problems

Author(s):
Monica Brezzi and Tze Leung Lai

Technical Report number (Dept. of Statistics, Stanford Univ.):
bf 2000-31

Date:
bf September 2000

Abstract:

This paper studies how and how much active experimentation is used in discounted or finite-horizon optimization problems of an agent who chooses actions sequentially from a finite set of actions, with rewards depending on unknown parameters associated with the actions. Closed-form approximations are developed for the optimal rules in these "multi-armed bandit" problems. Some refinements and modifications of the basic structure of these approximations also provide a nearly optimal solution to the long-standing problem of incorporating switching costs into multi-armed bandits.