Mathematics > Optimization and Control

Adaptive Learning of Uncontrolled Restless Bandits with Logarithmic Regret

Cem Tekin, Mingyan Liu

(Submitted on 20 Jul 2011)

In this paper we consider the problem of learning the optimal policy for the uncontrolled restless bandit problem. In this problem only the state of the selected arm can be observed, the state transitions are independent of control and the transition law is unknown. We propose a learning algorithm which gives logarithmic regret uniformly over time with respect to the optimal finite horizon policy with known transition law under some assumptions on the transition probabilities of the arms and the structure of the optimal stationary policy for the infinite horizon average reward problem.

Subjects: **Optimization and Control (math.OC)**; Learning (cs.LG)
Cite as: **arXiv:1107.4042v1 [math.OC]**

Submission history

From: Cem Tekin [view email]
[v1] Wed, 20 Jul 2011 17:33:43 GMT (19kb)

Which authors of this paper are endorsers?