



Deterministic Sequencing of Exploration and Exploitation for Multi-Armed Bandit Problems

Keqin Liu, Qing Zhao

(Submitted on 30 Jun 2011)

In the Multi-Armed Bandit (MAB) problem, there are a given set of arms with unknown reward distributions. At each time, a player selects one arm to play, aiming to maximize the total expected reward over a horizon of length T . The essence of the problem is the tradeoff between exploration and exploitation: playing a less explored arm to learn its reward statistics for future benefit or playing the arm with the largest sample mean at the current time. An approach based on a Deterministic Sequencing of Exploration and Exploitation (DSEE) is developed for constructing sequential arm selection policies. Under this approach, the key issue is the cardinality of the exploration sequence, which balances the accuracy and the cost of learning and defines a lower bound on the regret. It is show that when the moment-generating functions of the arm reward distributions are properly bounded, the optimal logarithmic order of the regret can be achieved by DSEE. The condition on the reward distributions can be gradually relaxed at a cost of a higher (nevertheless, still sublinear) regret order: for any positive integer p , $O(T^{1/p})$ regret can be achieved by DSEE when the moments of the reward distributions only exist up to the p th order. The proposed DSEE approach complements existing work on MAB by providing corresponding results under a set of relaxed conditions on the reward distributions. Furthermore, with a clearly defined tunable parameter-the cardinality of the exploration sequence, the DSEE approach is more easily extendable to variations of MAB, as demonstrated by its generalization to a decentralized MAB with multiple players under corrupted reward observations. Potential applications include dynamic spectrum access, multi-agent systems, Internet advertising and Web search.

Comments: 19 pages, 2 figures

Subjects: **Optimization and Control (math.OC)**; Learning (cs.LG); Systems and Control (cs.SY); Probability (math.PR); Statistics Theory (math.ST)

Cite as: [arXiv:1106.6104v1](https://arxiv.org/abs/1106.6104v1) [math.OC]

Download:

- [PDF](#)
- [PostScript](#)
- [Other formats](#)

Current browse context:

math.OC

[< prev](#) | [next >](#)

[new](#) | [recent](#) | [1106](#)

Change to browse by:

cs

[cs.LG](#)

[cs.SY](#)

math

[math.PR](#)

[math.ST](#)

stat

References & Citations

- [NASA ADS](#)

Bookmark(what is this?)



Submission history

From: Keqin Liu [[view email](#)]

[v1] Thu, 30 Jun 2011 02:12:32 GMT (24kb)

[Which authors of this paper are endorsers?](#)

Link back to: [arXiv](#), [form interface](#), [contact](#).