Mach Learn (2013) 92:377–401
DOI 10.1007/s10994-013-5359-2

Exploration and exploitation of scratch games

Raphaël Féraud · Tanguy Urvoy

Received: 10 January 2013 / Accepted: 12 April 2013 / Published online: 1 May 2013

Abstract We consider a variant of the multi-armed bandit model, which we call *scratch games*, where the sequences of rewards are finite and drawn in advance with unknown starting dates. This new problem is motivated by online advertising applications where the number of ad displays is fixed according to a contract between the advertiser and the publisher, and where a new ad may appear at any time. The *drawn-in-advance* assumption is natural for the adversarial approach where an oblivious adversary is supposed to choose the reward sequences in advance. For the stochastic setting, it is functionally equivalent to an urn where draws are performed without replacement. The *non-replacement assumption* is suited to the sequential design of non-reproducible experiments, which is often the case in real world. By adapting the standard multi-armed bandit algorithms to take advantage of this setting, we propose three new algorithms: the first one is designed for adversarial rewards; the second one assumes a stochastic urn model; and the last one is based on a Bayesian approach. For the adversarial and stochastic approaches, we provide upper bounds of the regret which compare favorably with the ones of EXP3 and UCB1. We also confirm experimentally that these algorithms compare favorably with EXP3, UCB1 and Thompson Sampling by simulation with synthetic models and ad-serving data.

Keywords Adversarial multi-armed bandits · Stochastic multi-armed bandits · Finite sequences · Scratch games

1 Introduction

In its most basic formulation, the multi-armed bandit problem can be stated as follows: there are K arms, each having an unknown, and infinite sequence of bounded rewards. At each step, a player chooses an arm and receives a reward issued from the corresponding sequence

Editors: Zhi-Hua Zhou, Wee Sun Lee, Steven Hoi, Wray Buntine, and Hiroshi Motoda.

R. Féraud (✉) · T. Urvoy
Orange Labs, 2, avenue Pierre Marzin, 22307, Lannion, France
e-mail: raphael.feraud@orange.com