Les mathématiques de l'apprentissage machine
« Atelier sur les défis modernes en théorie d'apprentissage »
23 au 26 avril 2018

Mathematics of Machine Learning
"Workshop on Modern Challenges of Learning Theory"
April 23-26, 2018

Nonstochastic bandits with anonymous feedback

Nicolò Cesa-Bianchi *

cesa.bianchi@gmail.com

A common pattern in content recommendation is that the response elicited in a user by the system is typically not instantaneous, and might occur well after the recommendation was issued. This delay, which depends on several unknown factors, implies that the reward obtained by the recommender can actually be seen as the combined effect of many previous recommendations. We investigate this phenomenon in a nonstochastic bandit setting where the loss of an action is not immediately charged to the player, but rather spread over at most d consecutive steps in an adversarial way. Unlike the standard bandit setting with delayed feedback, here the player cannot observe the individual delayed losses, but only their sum. We show a general reduction transforming a standard bandit algorithm into one that can operate in this harder setting. We also show how the regret of the transformed algorithm can be bounded in terms of the regret of the original algorithm. Our reduction cannot be improved in general: we prove a lower bound on the regret of any bandit algorithm in this setting that matches (up to log factors) the upper bound obtained via our reduction. Finally, we show how our reduction can be extended to more complex bandit settings, such as combinatorial linear bandits and online bandit convex optimization.

Joint work with Yishay Mansour (Tel Aviv and Google) and Claudio Gentile (INRIA).

*Department of Computer Science, Università degli Studi di Milano, via Comelico 39, 20135 Milano, Lombardia, ITALY