In Favour of Posting

THE CHINESE UNIVERSITY OF HONG KONG

Department of Information Engineering

Seminar

Online Learning in Dynamic Spectrum Access: Restless Bandits, Equilibrium and Social Optimality

by

Professor Mingyan Liu

Electrical Engineering and Computer Science

University of Michigan

Date : 2 March, 2011 (Wed.)

Time : 3:00p.m. - 4:00p.m.

Venue : Room 1106, William M.W. Mong Engineering Building
The Chinese University of Hong Kong

Abstract

We consider a dynamic spectrum access problem where the time-varying condition of a channel (e.g., as a result of random fading or certain primary users' activities) is modeled as an arbitrary finite-state Markov chain. At each instance of time, a (secondary) user selects and uses a channel and receives a certain reward as a function of the state of the channel (e.g., good channel condition results in higher data rate for the user). Each channel has potentially different state space and statistics, both unknown to the user, who tries to learn which one is the best so it can maximize its usage of the best channel. The objective is to construct good online learning algorithms so as to minimize the difference between the user's performance in total reward and that of using the best channel (on average) had it known which one is the best from a priori knowledge of the channel statistics (also known as the regret). This is an instance of the multiarmed bandit problem, and is well studied when each reward process is iid over time. In our case the reward processes are Markovian, and furthermore, restless, in that the channel conditions will continue to evolve independent of the user's actions. This leads to a restless bandit problem, for which there exists relatively few results on either algorithms or performance bounds in this learning context.

We introduce an algorithm that utilizes regenerative cycles of a Markov chain to compute a sample-mean based index policy, and show that under mild conditions on the state transition probabilities of the Markov chains this algorithm achieves logarithmic regret uniformly over time, and that this regret bound is also optimal. We also show that this result can be easily extended to the case when the user is allowed to use multiple channels at a time. We numerically examine the performance of this algorithm along with a few other algorithms with Gilbert-Elliot channel models, and discuss how this algorithm may be further improved (in terms of its constant) and how this result may lead to similar bounds for other algorithms. We then consider this type of online learning in a multiuser setting where simultaneous access to the same channel by multiple users may lead to collision and reduced reward. We show how such multiuser learning converges to a Nash equilibrium of an equivalent game, and how appropriate modifications to the learning algorithms can induce socially optimal channel allocation.

Biography

Mingyan Liu received her Ph.D in electrical engineering from the University of Maryland, College Park, in 2000 and joined the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, in September 2000, where she is currently an Associate Professor. Her research interests are in optimal resource allocation, performance modeling and analysis, and energy efficient design of wireless, mobile ad hoc, and sensor networks. She is the recipient of the 2002 NSF CAREER Award, and the University of Michigan Elizabeth C. Crosby Research Award in 2003. She serves on the editorial board of IEEE/ACM Trans. Networking, IEEE Trans. Mobile Computing, and ACM Trans. Sensor Networks.

**** ALL ARE WELCOME ****

Host: Professor Jianwei Huang (Tel: 2609-8353, Email: jwhuang@ie.cuhk.edu.hk)
Enquiries: Information Engineering Dept., CUHK (Tel.: 2609-8385)