Interpretable Policies for Dynamic Product Recommendations

Marek Petrik
IBM T.J. Watson Research Center
Yorktown Heights, NY 10598
mpetrik@us.ibm.com

Ronny Luss
IBM T.J. Watson Research Center
Yorktown Heights, NY 10598
rluss@us.ibm.com

Abstract

In many applications, it may be better to compute a good interpretable policy instead of a complex optimal one. For example, a recommendation engine might perform better when accounting for user profiles, but in the absence of such loyalty data, assumptions would have to be made that increase the complexity of the recommendation policy. A simple greedy recommendation could be implemented based on aggregated user data, but another simple policy can improve on this by accounting for the fact that users come from different segments of a population. In this paper, we study the problem of computing an optimal policy that is interpretable. In particular, we consider a policy to be interpretable if the decisions (e.g., recommendations) depend only on a small number of simple state attributes (e.g., the currently viewed product). This novel model is a general Markov decision problem with action constraints over states. We show that this problem is NP hard and develop a MILP formulation that gives an exact solution when policies are restricted to being deterministic. We demonstrate the effectiveness of the approach on a real-world business case for a European tour operator's recommendation engine.