Adaptive submodularity of stochastic ad scheduling problem

Ajay S. Aravamudhan

October 2, 2011

We can frame the stochastic ad scheduling problem when the break size of the current break is known, as an adaptive submodular system following the description given in Golovin and Krause(2010) [2].

Our aim is not to provide a rigorous theoretical proof that the system satisfies adaptive submodularity, rather it is to describe the problem and intuitively explain how the features of the system satisfy these conditions. If we succeed, we can conclude that the intractability of the problem when scheduling ads with no prior knowledge of break lengths may be addressed by using the adaptive greedy, which is proved to perform near optimally in such systems.

Case 0: Consider the ad scheduling problem when all breaks are of equal (or known) lengths and of fixed number. Let A be the set of ads available in the inventory, and let $f : 2^A \rightarrow R_+$ be the revenue earned from scheduling ads from A. For any $A' \subseteq A$, and ad $a \notin A$,

$$f(A \cup \{a\}) - f(A) \leqslant f(A' \cup \{a\}) - f(A') \tag{1}$$

Equation (1) can be understood intuitively: set A offers a greater choice of (possibly higher value) ads to schedule, therefore the marginal gain from scheduling a (if at all) from set A is lower than the marginal gain from scheduling a from a smaller set A'. This system is thus submodular.

Case 1: Consider the stochastic ad scheduling problem. In the first case, break length of only the (current) break for which we are going to schedule ads is known, but break lengths (and number) of future breaks are unknown and are distributed according to some distribution Φ. At the start of each break of known length, we pick a set of ads to schedule, then wait for the next break, select another set of ads accoring to the next break's length, and so on. The policy tree π in this case consists of as many levels as there are breaks, each node is a set of ads that can be scheduled for that break, with as many child nodes as there are outcomes in Φ . Let O be the set of ads that have been shown. Let utility function $f : 2^A \times O^E \rightarrow R_+$ be the utlity that depends on what ads we pick and what ads have been successfully shown. We define $f_{avg}(\pi) := \mathrm{E}_\Phi[f(E(\pi, \Phi), \Phi)]$ as the expected utility of policy π. Let $\mathbb{P}(\Phi)$ be the probability distribution over realizations.

We can show adaptive monotonicity as defined in Golovin and Krause(2010) by noting that for any two policies π and π' with policy trees T^π and $T^{\pi'}$, the concatenated policy tree $T^\pi @ T^{\pi'}$ does not have a lower expected utility than T^π, because subsequent layers in a policy tree offer more scheduling opportunities that only help to increase the expected utility of π.

Also, for any policy π with policy tree T^π, and for any $0 \leqslant i < j$,

$$f_{avg}(T^\pi_{[j]}) - f_{avg}(T^\pi_{[j-1]}) \leqslant \mathbb{E}\left[f_{avg}(T^\pi_{[i]\cup\{j\}}) - f_{avg}(T^\pi_{[i]})\right] \tag{2}$$

Equation (2) can be intuitively explained. Layer $i + 1$ of $T^\pi_{[i]\cup\{j\}}$ is a distribution of items of layer j of $T^\pi_{[j]}$, and our definition of f_{avg} ensures that the expected utility of policy π does not decrease for each additional layer added to the tree. In fact, expected utility of policy π takes into account breaks of small lengths (or zero lengths) in subsequent layers into account, therefore additional layers where breaks of larger lengths occur only serve to increase the expected utility f_{avg} for a tree pruned at that layer. Thus the system can be shown to be adaptive submodular over $\mathbb{P}(\Phi)$.

While we have already found the optimal policy for this problem setting, we did find that the Greedy performed almost on par with the Perfect Information case (§5.2 in my thesis [1]). This is consistent with Golovin and Krause(2010), which proves that in adaptive submodular systems, Greedy attains a value at least $\left(1 - \frac{1}{e}\right)$ of the best policy.

Case 2: Finally, we consider stochastic ad scheduling problem when break lengths are unknown before scheduling, but where we can observe the *ex-post* realization. A key challenge here is that the sequence of ads in each item e will decide the state o for each break realization. We could redefine e to be a *permutation* of a subset of ads of O, however, it is hard to find an exact optimal policy because the number of such permutations may be intractable.

The definition of adaptive submodularity frees us from having to find the optimal policy if we could prove that such a system is adaptive submodular for *any* policy, along the lines of **Case 1**. We could then conclude that the adaptive Greedy performs near optimally, and hence is a viable policy for ad scheduling under uncertainty.

References

[1] Ajay S. Aravamudhan. Optimal ad scheduling in breaks of random lengths (graduate thesis). http://mysmu.edu/staff/ajaysa/files/thesis.pdf, July 2011.

[2] D. Golovin and A. Krause. Adaptive submodularity: A new approach to active learning and stochastic optimization. In *Proceedings of International Conference on Learning Theory (COLT)*. Citeseer, 2010.