A Multi-Arm Bandit Approach To Subset Selection Under Constraints
Ayush Deva, Kumar Abhishek and Sujit Gujar

But What If Qualities Are Unknown?
We consider a setting where the qualities of the agents are unknown to the planner beforehand and needs to be estimated through sequential selection. We model this as a Multi Arm Bandit problem and leverage the popular UCB algorithm to design an abstract algorithm SS-UCB.

The algorithm takes in the available agents, their costs, quality threshold ($\alpha$), tolerance parameter ($\epsilon_2$) and a suitable offline subset selection algorithm, SSA.

Algorithm 2 SS-UCB
1. Inputs: $N$, $\alpha$, $c_k$, $R$, costs $c = \{c_k\}_{k \in N}$
2. For each agent $i$, maintain: $w_i$, $q_i^*$, $(q_i)^*$
3. $r = \frac{3n R}{2 \epsilon_2}$; $t = 0$
4. while $r \leq t$ (Explore Phase) do
5. Play a super-arm $S^* = N$
6. Observe qualities $Q^*_i$, $V_i \in S^*$ and update $w_i$, $q_i$
7. $t \rightarrow t + 1$
8. while $r \leq T$ (Explore-Exploit Phase) do
9. For each agent $i$, set $(q_i)^* = q_i^* + \frac{3n R}{2 \epsilon_2}$
10. $S^* = SSA(\{(q_i)^*\}_{i \in N}, c_k, \alpha + \epsilon_2 R)$
11. Observe qualities $Q^*_i$, $V_i \in S^*$ and update $w_i$, $q_i$
12. $t \rightarrow t + 1$

Key Results
Using DPSS as our SSA in Algorithm 2 (DPSS-UCB), we show that:
1. DPSS-UCB returns a subset which approximately satisfies the quality constraint with a high probability after $t$ rounds, where $t \sim O(ln T)$

Approximate but Faster Solution
The time complexity of DPSS is of $O(2^n)$, which makes it difficult to scale when $n$ is large. We propose an approximate, greedy-based, polynomial time, $O(n \log n)$, algorithm, GSS, to our ILP. Further, we empirically show that by using GSS as the SSA in Algorithm 2 (GSS-UCB), we achieve similar results to DPSS-UCB.

Applications

References