Pure Exploration of Multi-armed Bandit Under Matroid Constraints

Lijie Chen, Anupam Gupta, Jian Li

Tsinghua University, Carnegie Mellon University, Tsinghua University

June 23, 2016

Section

An Example

Minimum Spanning Tree (MST)

- Given an edge-weighted graph.
- Find a spanning tree with minimum weight.

MST is easy—if we know the exact weights.

What if we don't have complete information?

- Now that each edge e has an associated distribution $\mathcal{D}_e$ with mean μ_e.
- We want a spanning tree with minimum $\sum_e \mu_e$.
- Each time we can choose an edge and take a sample from that distribution.
- **Goal:** Succeed w.p. $1 - \delta$ and minimize the samples we need.

Combinatorial Multi-Armed Bandit (CMAB)

Introduced by Chen, Wang and Yuan in 2013 [CWY13].

- n stochastic arms. Each arm corresponds to an edge.
- Each one with an (unknown) distributions $\mathcal{D}_i$ with mean μ_i, and supported on $[0, 1]$.
- A family $\mathcal{S}$ of subsets of $\{1, 2, \ldots, n\}$. Each subset in $\mathcal{S}$ corresponds to a spanning tree.
- Want a subset $O \in \mathcal{S}$ maximizing $\sum_{i \in O} \mu_i$.

Combinatorial Multi-Armed Bandit (CMAB)

- CMAB is a very general framework.
- Too general: Very difficult to obtain tight lower and upper bounds for the general problem.

Matroid Constraints

We consider Matroid constraints.

- General enough to cover many interesting applications.
- Rich combinatorial structure which allows us to obtain nearly tight bounds.
- Generalizes the well studied Best-1-Arm and Best-k-Arm problems (uniform matroid).

Several Motivating Applications

The notion of matroid abstracts many combinatorial structures:

1. Uniform matroid (cardinality constraints).
2. Partition matroid.
3. Laminar matroid.
4. Transversal matroid.

Partition Matroid

- m disjoint groups of arms.
- Want the best k_i arms in the i^{th} group.
- Studied by Gabillon et al. [GGLB11] and Bubeck et al. [BWV12].

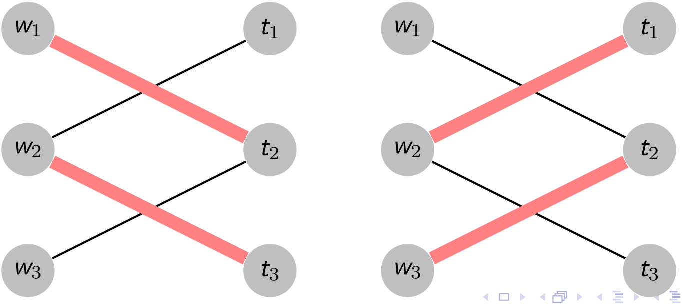
Laminar Matroid

- m disjoint groups of arms.
- Choose at most k_i arms in the i^{th} group.
- Choose at most N arms in total.

Transversal Matroid

- n workers and m tasks.
- Task i has reward distribution $\mathcal{D}_i$.
- Each worker is capable of doing a subset of tasks.
- Each worker can only do one task.
- Find the optimal subset of tasks that can be finished.
  - $\mathcal{S}$ is all the subsets of tasks admitting a matching from workers.
- Potential applications in crowdsourcing or online advertisement.

Figure: Two possible assignments for $\{t_2, t_3\}$ and $\{t_1, t_2\}$ in the family $\mathcal{S}$.



Section

Our Results

i An **optimal** PAC algorithm (for finding an ε-optimal solution).

ii An exact algorithm. Our (gap-dependent) upper bound even improves the state of art of BEST-k-ARM.

Result I: An Optimal PAC Algorithm

- A new metric: we say I is *ε-optimal*, if it becomes the optimal basis after adding ε to the mean of every arms in I.
- Stronger than all previous metrics in Best-k-Arm [KS10], [ZCL14] and [CLTL15].
- We develop an algorithm for finding an ε-optimal solution w.p. at least $1-\delta$, using at most

$$O(n\varepsilon^{-2}\cdot(\ln k+\ln\delta^{-1}))$$

  samples, where k is the rank of the matroid.
  Generalize the results by Kalyanakrishnan et al. [KTAS12] and Cao et al. [CLTL15].
- Known lower bound for Best-k-Arm:

$$\Omega(n\varepsilon^{-2}\cdot(\ln k+\ln\delta^{-1}))$$

  by Kalyanakrishnan et al. [KTAS12] (Our algorithm is indeed optimal).

Result II: An Exact Algorithm

- Output the optimal solution with probability $1 - \delta$.
- The sample complexity depends on the gaps.
  - The smaller the gaps are, the harder to identify the optimal solution.
- Assume arms have distinct means. So the optimal solution is unique.
- Δ_i (Gap) for arm i is defined as follows:
  - If $i \in \mathrm{OPT}$: the loss of the utility when you are forced not to select i.
  - If $i \notin \mathrm{OPT}$: the loss of the utility when you are forced to select i.

It is the same definition of gap in Chen et al. [CLK$^+$14], and generalizes the previous gap definition for BEST-k-ARM as in Kalyanakrishnan et al. [KTAS12].

Result II: An Exact Algorithm

- We develop an algorithm for finding the optimal solution with probability at least $1-\delta$, using at most

$$O\left(\sum_{e\in S}\Delta_e^{-2}(\ln\delta^{-1}+\ln k+\ln\ln\Delta_e^{-1})\right)$$

samples, where k is the rank of the matroid.

- For matroids, Chen et al. [CLK+14] achieves an upper bound of

$$O\left(\sum_{e\in S}\Delta_e^{-2}(\ln\delta^{-1}+\ln n+\ln\sum_{e\in S}\Delta_e^{-1})\right).$$

- Known lower bound by Chen et al. [CLK+14]:

$$O\left(\sum_{e\in S}\Delta_e^{-2}\ln\delta^{-1}\right).$$

Comparing with the state of the art of Best-k-Arm

Our result strictly improves the state-of-the-art bound for Best-k-Arm by Kalyanakrishnan et al. [KTAS12]:

$$O\left(\sum_{i=1}^{n} \Delta_i^{-2} (\ln \delta^{-1} + \ln \sum\nolimits_{i=1}^{n} \Delta_i^{-2})\right).$$

When k is much smaller than n, our algorithm achieves a significant improvement.

Section

1 Introduction

2 Results

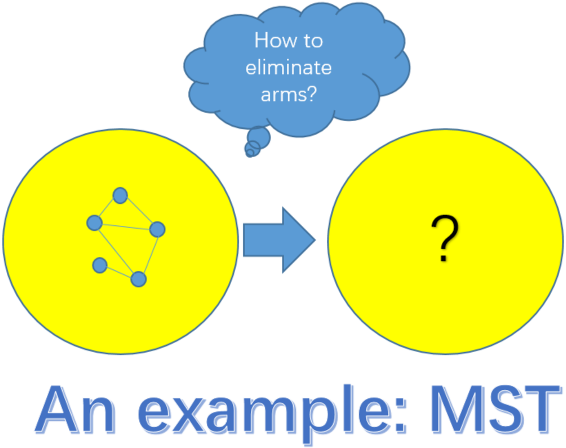
3 Techniques

High-level Idea of Our Optimal PAC Algorithms

- UCB-type algorithms usually need a $\ln n$ factor (for union bound).
- We use an elimination-based approach to get a $\ln k$ factor.
  - Recall our PAC bound: $O(n\varepsilon^{-2} \cdot (\ln k + \ln \delta^{-1}))$.

- Elimination in Best-1-Arm or Best-k-Arm is simple.
- In each round, we eliminate some arms:
  - i Find a threshold (a percentile [ZCL14], or an approximate optimal arm [KKS13]).
  - ii Eliminate arms with empirical means worse than the threshold.

Difficulty

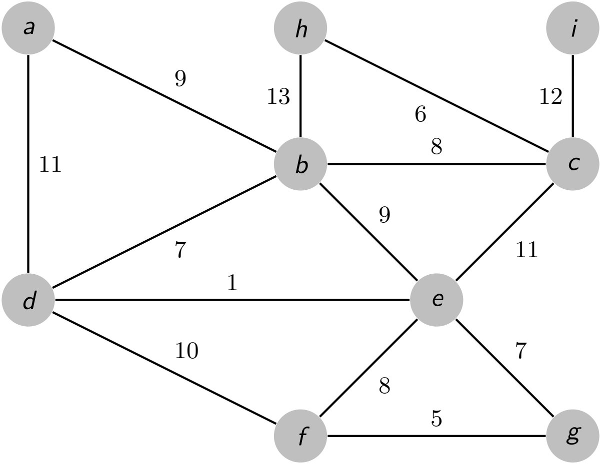


- It is unclear how to eliminate some arms in matroids: A nearly optimal solution may contain a lot of arms among the lower quarter.

Sampling and Pruning Technique

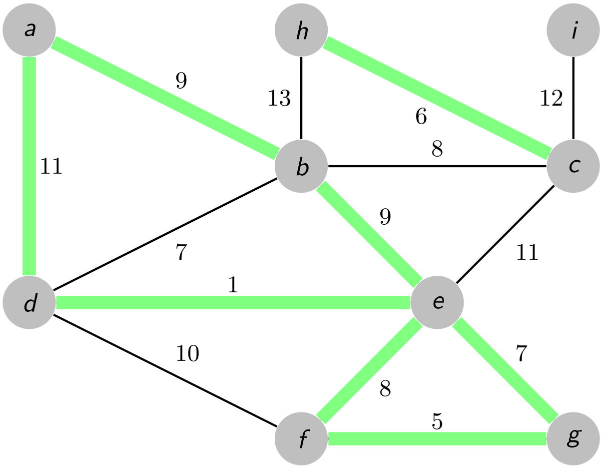
- We overcome the difficulty by using the novel sampling and pruning technique developed by Karger, Klein and Tarjan [KKT95].
  - The technique was originally used for design more efficient algorithm for MST.
- Their key idea is to use a solution for a random sampled subset to do the elimination.
- We adapt their approach to the bandit setting.

An Example: Maximum Spanning Tree



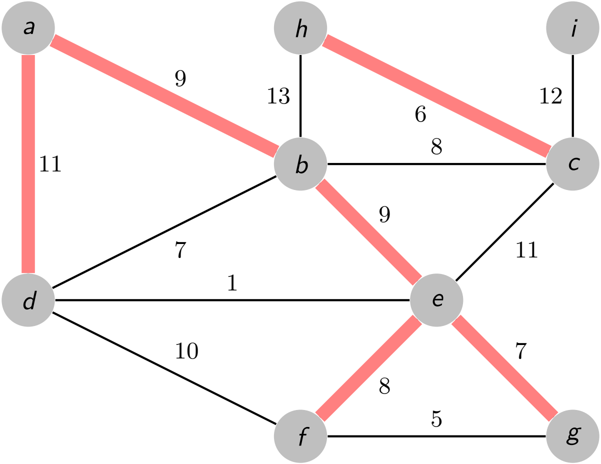
1 Want to find an ε-optimal solution.

An Example: Maximum Spanning Tree



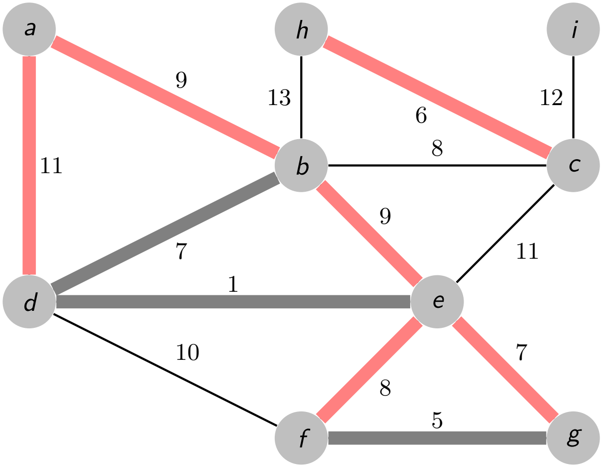
2 Select a random subset F of arms, by picking each arm with prob $p = 0.5$.

An Example: Maximum Spanning Tree



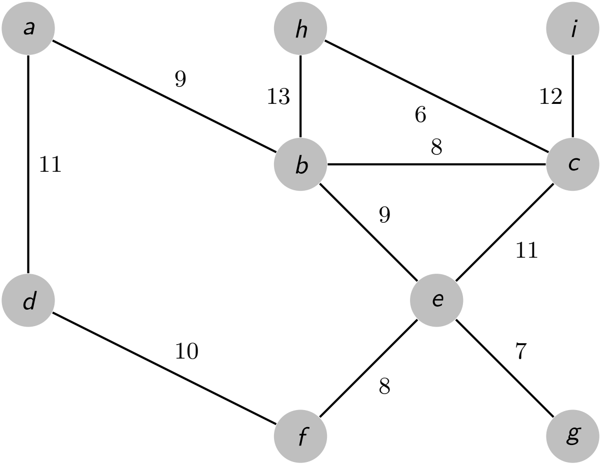
3 Recursively find an $\varepsilon/3$-optimal solution I in F.

An Example: Maximum Spanning Tree



4 Estimate the means of remaining arms, and eliminate those "bad arms"—the edges $\varepsilon/3$-approximate dominated by the corresponding path in I.

An Example: Maximum Spanning Tree



5 Recurse on the remaining arms again.

Sampling and Pruning Technique

Key point: after one elimination, w.h.p we

i Eliminate a constant fraction of arms.

ii Do not hurt the optimal solution too much.

Open Problems

Some interesting future directions:

- Apply the sampling and pruning technique in other bandit problems?
- Can we find better algorithm for other combinatorial constraints besides matroids? E.g., bipartite matching. There is an example where Chen et al.'s algorithm needs $\Omega(n^3\varepsilon^{-2})$ samples, while a simple $O(n^2\varepsilon^{-2})$ algorithms can be obtained easily for that example.
- The utility of set S is a nonlinear function of the means (rather than $\sum_{i\in S}\mu_i$), or a general function depending on the distributions (e.g., $[\max_{i\in S} x_i]$ Best-of-K Bandits by Simchowitz, Jamieson and Recht [SJR16]).

Sébastien Bubeck, Tengyao Wang, and Nitin Viswanathan.
Multiple identifications in multi-armed bandits.
arXiv preprint arXiv:1205.3181, 2012.

Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen.
Combinatorial pure exploration of multi-armed bandits.
In *Advances in Neural Information Processing Systems*, pages 379–387, 2014.

Wei Cao, Jian Li, Yufei Tao, and Zhize Li.
On top-k selection in multi-armed bandits and hidden bipartite graphs.
In *Advances in Neural Information Processing Systems*, pages 1036–1044, 2015.

Wei Chen, Yajun Wang, and Yang Yuan.
Combinatorial multi-armed bandit: General framework and applications.
In *Proceedings of the 30th International Conference on Machine Learning*, pages 151–159, 2013.

Victor Gabillon, Mohammad Ghavamzadeh, Alessandro Lazaric, and Sébastien Bubeck.
Multi-bandit best arm identification.
In *Advances in Neural Information Processing Systems*, pages 2222–2230, 2011.

Zohar Karnin, Tomer Koren, and Oren Somekh.
Almost optimal exploration in multi-armed bandits.
In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 1238–1246, 2013.

David R Karger, Philip N Klein, and Robert E Tarjan.
A randomized linear-time algorithm to find minimum spanning trees.
Journal of the ACM (JACM), 42(2):321–328, 1995.

Shivaram Kalyanakrishnan and Peter Stone.
Efficient selection of multiple bandit arms: Theory and practice.
In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 511–518, 2010.

Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone.
Pac subset selection in stochastic multi-armed bandits.
In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 655–662, 2012.

Max Simchowitz, Kevin Jamieson, and Ben Recht.
Best-of-k bandits.
arXiv preprint arXiv:1603.02752, 2016.

Yuan Zhou, Xi Chen, and Jian Li.
Optimal pac multiple arm identification with applications to crowdsourcing.
In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 217–225, 2014.