Non-monotone Adaptive Submodular Maximization

IJCAI-15
24TH INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE
BUENOS AIRES
JULY 25TH–31TH, 2015

Alkis Gotovos
ETH Zurich

Amin Karbasi
Yale University

Andreas Krause
ETH Zurich

Motivation

- Many AI problems boil down to selecting a number of elements from a large set of options
- Sequentially make "smart" choices based on past observations
- Fundamental goal: Find classes of objective functions that are amenable to efficient sequential optimization with theoretical approximation guarantees

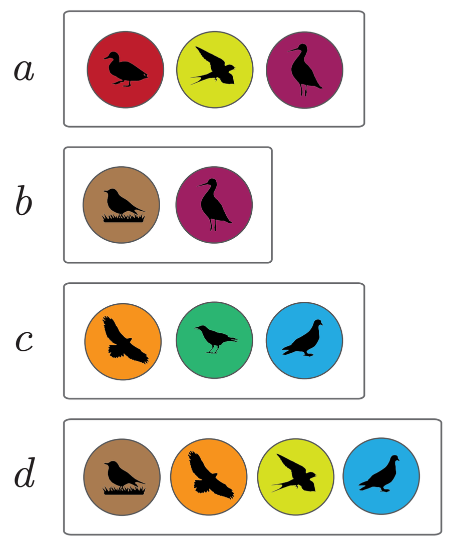
Example applications

- Active learning for medical diagnosis
- Viral marketing in social networks

Running Example: Birdwatching

Visit locations and observe different bird species (max cover problem)

- Ground set: $V = \{a, b, c, d\}$
- Objective: $f : 2^V \to \mathbb{R}_{\geq 0}$
- Example: $f(\{d\}) = 4$
  $f(\{c, d\}) = 5$



Monotonicity and Submodularity

- f is *monotone*
  Visiting a location provides non-negative benefit
- f is *submodular*
  Locations have "diminishing returns"; the more of them we have already visited, the less benefit we get from visiting a new one
- Example: $f(\{c\}) = 3$
  $f(c \mid \{d\}) = f(\{c, d\}) - f(\{d\}) = 5 - 4 = 1$

Monotone Submodular Maximization

Want to maximize f—observe as many bird species as possible

- Unconstrained problem → Trivial OPT $= f(V)$
- Cardinality-constrained problem (visit up to k locations) → NP-hard

Greedy algorithm

- Start with empty set of locations
- Keep adding the location that provides the largest benefit—the most new bird species)
- Stop as soon as we have visited k locations

Theorem [Nemhauser et al., 1978]

If f is monotone submodular, then greedy gives a $(1 - 1/e)$-approximation.

Non-monotone Objectives

- Assume each set A of locations has an associated cost $c(A)$
- New objective: $g(A) = f(A) - c(A)$
- For example, uniform cost term: $c(A) = \lambda|A|$
- Visiting a location may cost more than the benefit it provides
  ↓
  g is *non-monotone*
- Greedy has no guarantees for non-monotone functions

Random greedy algorithm

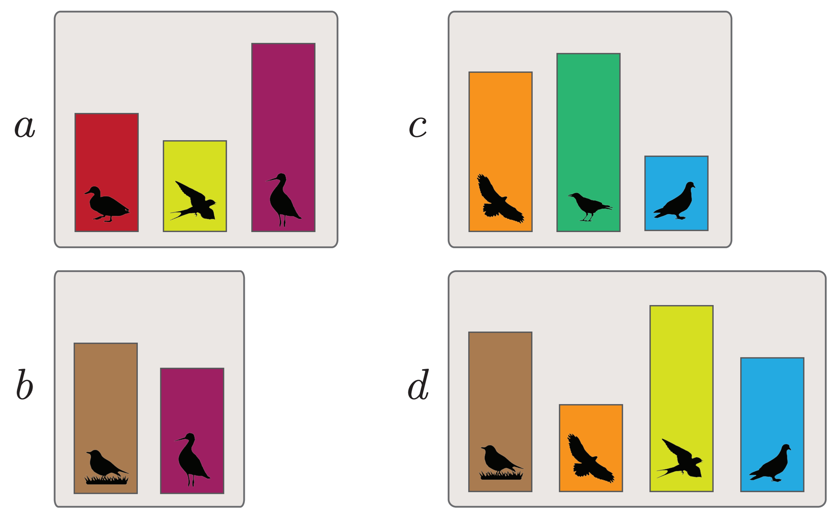
Idea: At each step, uniformly at random add one of the k most beneficial locations

Theorem [Buchbinder et al., 2014]

If f is submodular, then random greedy gives a $(1/e)$-approximation (in expectation).

If f is also monotone, then random greedy gives a $(1 - 1/e)$-approximation (in expectation).

The Adaptive Setting



- In practice, the number of observed bird species will vary according to some distribution per location
- Two-argument objective: $f(A, \phi)$
  - A: Set of visited locations
  - ϕ: Random realization of the environment
- Non-adaptive: Commit to set A before observing any outcomes (e.g., take expectation over ϕ)
- *Adaptive*: Take past outcomes into account to make better decisions at each step

Monotonicity → Adaptive monotonicity
Submodularity → Adaptive submodularity
Greedy → Adaptive greedy (policy)

Theorem [Golovin and Krause, 2011]

If f is adaptive monotone submodular, then adaptive greedy gives a $(1 - 1/e)$-approximation (in expectation).

Adaptive Random Greedy

	Non-adaptive	Adaptive
Monotone	Greedy	Adaptive greedy
Non-monotone	Random greedy	?

- No known algorithm with theoretical guarantees for non-monotone adaptive submodular objectives
- We propose the *adaptive random greedy* policy to fill this gap

Input: $V,\ f,\ p(\phi),\ k$

$A \leftarrow \varnothing$
$\psi \leftarrow \varnothing$
for $i = 1$ **to** k
  Compute marginal gains $\Delta(v \mid \psi)$, for all $v \in V \setminus A$
  $\mathcal{M}_k \leftarrow$ set of k elements with the largest marginal gains
  Sample element m from $\mathcal{M}_k$ uniformly at random
  $A \leftarrow A \cup \{m\}$
  Observe outcome $\phi(m)$
  Update history ψ
return A

Theoretical Guarantees

- We require a slightly stronger condition than adaptive submodularity, which holds for the majority of practical objectives
- The expectation here is taken over both the randomization of the algorithm, as well as the randomness of the environment

Theorem [Our contribution]

If f is adaptive submodular, and, additionally, $f(\cdot, \phi)$ is submodular for any realization ϕ, then adaptive random greedy gives a $(1/e)$-approximation (in expectation).

If f is also adaptive monotone, then adaptive random greedy gives a $(1 - 1/e)$-approximation (in expectation).

Non-monotone Objectives

We present two classes of objective functions that naturally arise in practice, and are adaptive submodular but *not* monotone.

1. Objectives with a modular cost term

$g(A, \phi) = f(A, \phi) - c(A)$

- $f(A, \phi)$: Monotone adaptive submodular function
- $c(A)$: Modular cost term, $c(A) = \sum_{a \in A} c_a$

Example: Network influence maximization

- Select a subset of nodes to maximize spread of influence
- Ground set: Nodes of the graph
- $f(A, \phi)$: classic network influence objective (ϕ captures the random outcomes of the independent cascade model)
- c_a: cost of choosing node a (e.g., proportional to its degree)

2. Objectives with factorial realizations

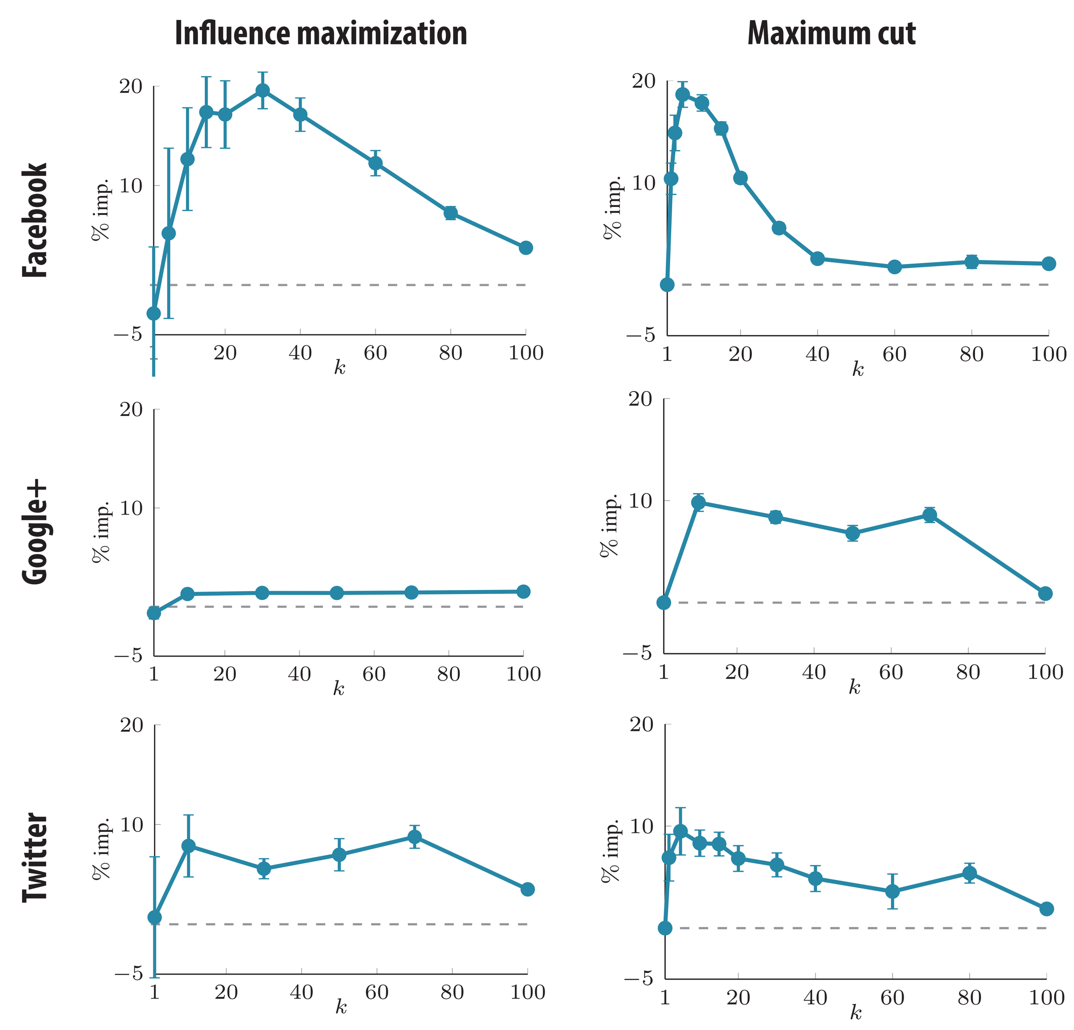
- The dependence of $f(A, \phi)$ on ϕ is constrained to the outcomes of the selected elements
- $f(\cdot, \phi)$ is submodular, for any realization ϕ
- The distribution of realizations ϕ factorizes over V

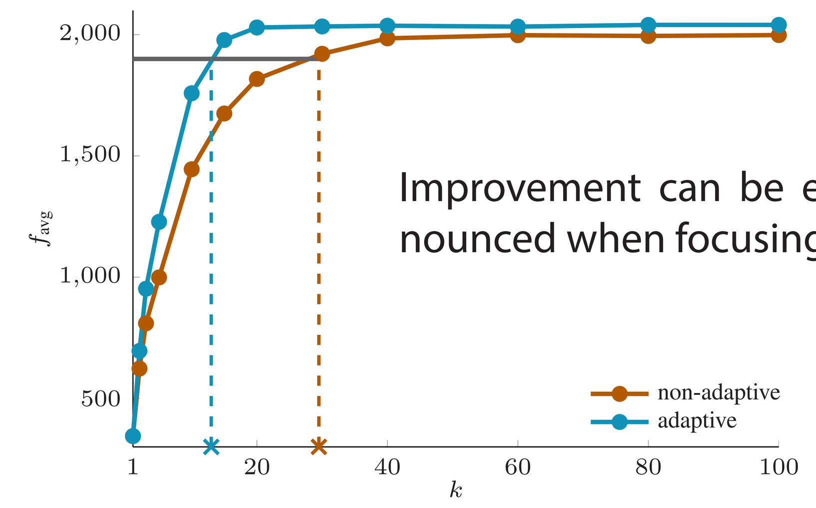
Example: Maximum graph cut

- Select a subset of nodes to maximize the weight of the edges cut
- When picking a node, either that node or a random neighbor theoreof is added to the cut
- Ground set: Nodes of the graph
- Easy to check that the above properties hold

Experiments

- Three network data sets from the KONECT database, representing ego networks of Facebook, Google+, and Twitter
- Subsample each of them down to 2000 nodes
- Ground set: 100 randomly sampled nodes
- Repeat experiments over random ground sets and realizations
- Compare adaptive random greedy to non-adaptive version





Improvement can be even more pronounced when focusing on "coverage".

References

Niv Buchbinder, Moran Feldman, Joseph Naor, and Roy Schwartz. *Submodular maximization with cardinality constraints.* SODA, 2014.

Daniel Golovin and Andreas Krause. *Adaptive submodularity: theory and applications in active learning and stochastic optimization.* JAIR, 2011.

George L. Nemhauser, Laurence A. Wolsey, and Marshall L. Fisher. *An analysis of approximations for maximizing submodular set functions.* Mathematical Programming, 1978.