Beyond Adaptive Submodularity: Approximation Guarantees of Greedy Policy with Adaptive Submodularity Ratio

Kaito Fujii (UTokyo) & Shinsaku Sakaue (NTT)

The 36th International Conference on Machine Learning
Jun. 12, 2019
Application: Influence maximization

Select a subset of ads to influence as many people as possible
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Adaptive setting
Select ads one by one
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Adaptive setting
Select ads one by one
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Adaptive setting
Select ads one by one
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Adaptive setting
Select ads one by one
Application: Influence maximization

Select a subset of ads to influence as many people as possible.

Adaptive setting
Select ads one by one.
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance

Adaptive setting
Select ads one by one

Q1 When does the greedy policy work well?
Application: Influence maximization

Select a subset of ads to influence as many people as possible

Non-adaptive setting
Select a subset in advance

Adaptive setting
Select ads one by one

Q1 When does the greedy policy work well?

Q2 How different are the non-adaptive and adaptive policies?
We propose a new concept called adaptive submodularity ratio
We propose a new concept called adaptive submodularity ratio.
We propose a new concept called **adaptive submodularity ratio**

\[ \gamma_{\ell,k} = 1 \quad \text{submodular functions} \]

\[ \gamma_{\ell,k} = 0 \quad \text{arbitrary monotone functions} \]

**Submodularity ratio**

[Das–Kempe’11]
We propose a new concept called **adaptive submodularity ratio**

$$\gamma_{\ell,k} = 1$$

- **Submodularity ratio**
  - [Das–Kempe’11]
  - $$\gamma_{\ell,k} = 0$$

- **Submodular functions**
- **Adaptive submodular functions**
  - [Golovin–Krause’11]

- **Arbitrary monotone functions**
We propose a new concept called **adaptive submodularity ratio**

\[ \gamma_{\ell,k} = 1 \quad \text{submodular functions} \]

\[ \gamma_{\ell,k} = 0 \quad \text{arbitrary adaptive monotone functions} \]

**Submodularity ratio**

[Das–Kempe’11]

**Adaptive submodularity ratio**

[Golovin–Krause’11]

[this study]
Adaptive submodularity ratio $\gamma_{\ell, k} \in [0, 1]$ is a parameter that measures the distance to adaptive submodular functions.

$$\gamma_{\ell, k} \triangleq \min_{|\psi| \leq \ell, \pi \in \Pi_k} \frac{\sum_{v \in V} \Pr(v \in E(\pi, \Phi) | \Phi \sim \psi) \Delta(v|\psi)}{\Delta(\pi|\psi)}$$
Adaptive submodularity ratio \( \gamma_{\ell,k} \in [0,1] \) is a parameter that measures the distance to adaptive submodular functions.

\[
\gamma_{\ell,k} \overset{\Delta}{=} \min_{|\psi| \leq \ell, \pi \in \Pi_k} \sum_{v \in V} \Pr(v \in E(\pi, \Phi) | \Phi \sim \psi) \Delta(v|\psi) / \Delta(\pi|\psi)
\]

the expected marginal gain of policy \( \pi \)
Adaptive submodularity ratio $\gamma_{\ell,k} \in [0, 1]$ is a parameter that measures the distance to adaptive submodular functions:

$$\gamma_{\ell,k} \triangleq \min_{|\psi| \leq \ell, \pi \in \Pi_k} \frac{\sum_{v \in V} \Pr(v \in E(\pi, \Phi)|\Phi \sim \psi) \Delta(v|\psi)}{\Delta(\pi|\psi)}$$

the expected marginal gain of single element $v$

the probability that element $v$ is selected by policy $\pi$
Adaptive submodularity ratio $\gamma_{l,k} \in [0, 1]$ is a parameter that measures the distance to adaptive submodular functions.

$$\gamma_{l,k} \triangleq \min_{|\psi| \leq l, \pi \in \Pi_k} \sum_{v \in V} \Pr(v \in E(\pi, \Phi)|\Phi \sim \psi) \Delta(v|\psi) / \Delta(\pi|\psi)$$

Q1: When does the greedy policy work well?
Adaptive submodularity ratio $\gamma_{\ell,k} \in [0, 1]$ is a parameter that measures the distance to adaptive submodular functions.

$$\gamma_{\ell,k} \triangleq \min_{|\psi| \leq \ell, \pi \in \Pi_k} \sum_{v \in V} \Pr(v \in E(\pi, \Phi)|\Phi \sim \psi) \Delta(v|\psi)$$

Q1 When does the greedy policy work well?

**Theorem** Adaptive Greedy is $(1 - \exp(-\gamma_{k,k}))$-approximation.
A non-adaptive policy approximates an optimal adaptive policy

\[ \text{GAP}_k(f, p) \triangleq \Delta \]

An optimal non-adaptive policy

An optimal adaptive policy
Bounds on adaptivity gaps

A non-adaptive policy approximates an optimal adaptive policy

\[ \text{GAP}_k(f, p) \triangleq \min_{S \subseteq V : |S| \leq k} \frac{\mathbb{E}[f(S, \phi)]}{\sum_{v \in S} \mathbb{E}[f(\{v\}, \phi)]} \]

Q2 How different are the non-adaptive and adaptive policies?

Theorem \[ \text{GAP}_k(f, p) \geq \beta_{0,k} \gamma_{0,k} \]
Application: Influence maximization

Non-adaptive setting
Select a subset in advance

Adaptive setting
Select ads one by one
Application: Influence maximization

Theorem: $\gamma_{l,k} \geq \frac{k + 1}{2k}$ on bipartite graphs with the triggering model.
Select a subset of features to be observed precisely

\[ y \approx A(\phi) \]

Non-adaptive setting
Select a subset in advance

Adaptive setting
Observe features one by one

\[ \min_{\phi} \min_S \phi_j \ell + \frac{7}{8} \lambda \min (A(\phi)^\top S A(\phi)) S \]
Select a subset of features to be observed precisely

Non-adaptive setting
Select a subset in advance

Adaptive setting
Observe features one by one

Theorem
\[ \gamma_{l,k} \geq \min_{\phi} \min_{S \subseteq V : |S| \leq l + k} \lambda_{\min}(A(\phi)^T A(\phi)_S) \]
Adaptive Submodularity Ratio is applied to

Theorem 1  Bounds on approximation ratio of Adaptive Greedy

Theorem 2  Bounds on adaptivity gaps

Application 1  Influence maximization on bipartite graphs

Application 2  Adaptive feature selection

Poster #163 at Pacific Ballroom, Wen 6:30–9:00 PM