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

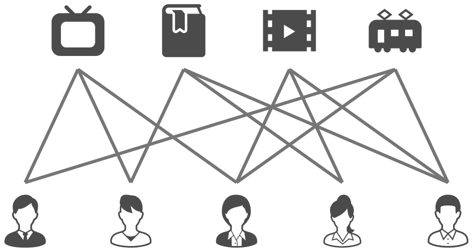
Kaito Fujii (UTokyo) & **Shinsaku Sakaue** (NTT)

The 36th International Conference on Machine Learning

Jun. 12, 2019

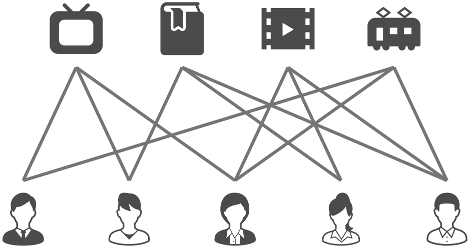
Application: Influence maximization

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

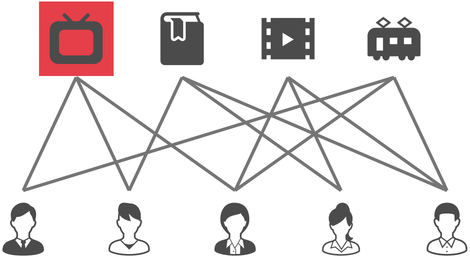


Non-adaptive setting

Select a subset in advance

Application: Influence maximization

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

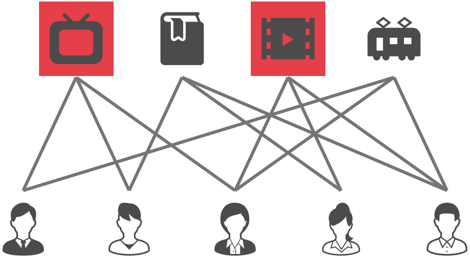


Non-adaptive setting

Select a subset in advance

Application: Influence maximization

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

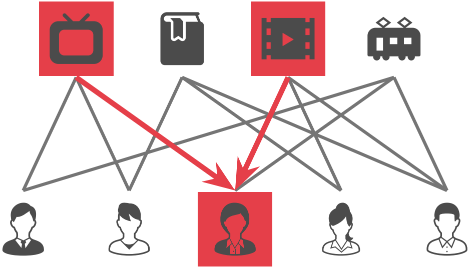


Non-adaptive setting

Select a subset in advance

Application: Influence maximization

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

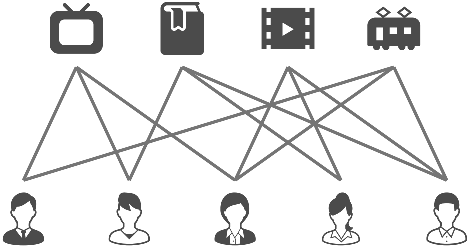


Non-adaptive setting

Select a subset in advance

Application: Influence maximization

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

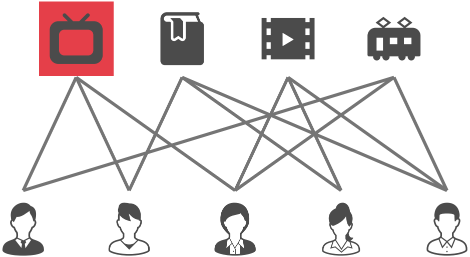


Adaptive setting

Select ads one by one

Application: Influence maximization

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

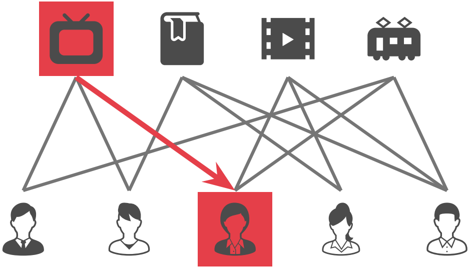


Adaptive setting

Select ads one by one

Application: Influence maximization

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

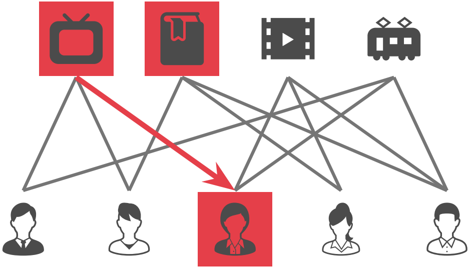


Adaptive setting

Select ads one by one

Application: Influence maximization

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

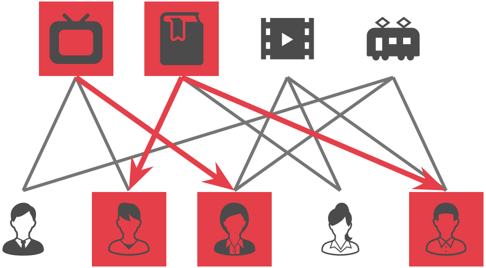


Adaptive setting

Select ads one by one

Application: Influence maximization

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

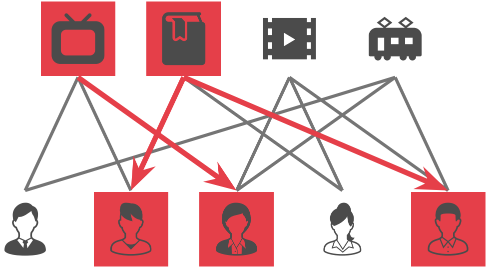


Adaptive setting

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

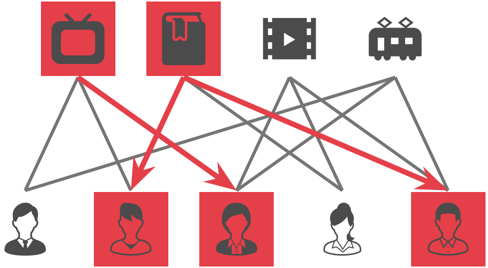
Adaptive setting

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?**

Adaptive submodularity ratio

We propose a new concept called adaptive submodularity ratio

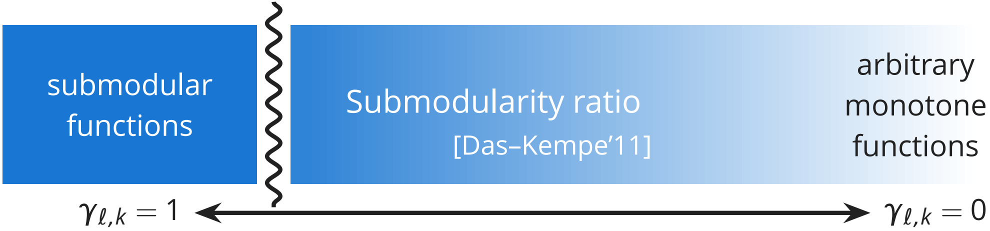
Adaptive submodularity ratio

We propose a new concept called adaptive submodularity ratio

submodular functions

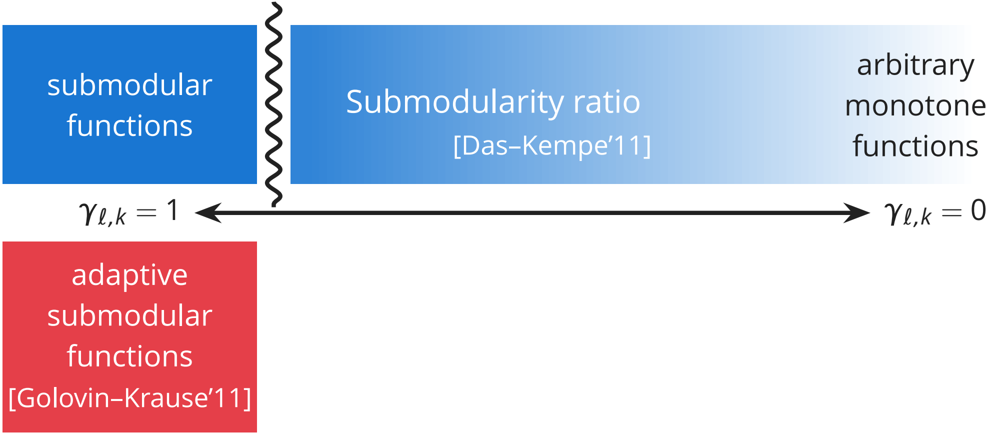
Adaptive submodularity ratio

We propose a new concept called adaptive submodularity ratio



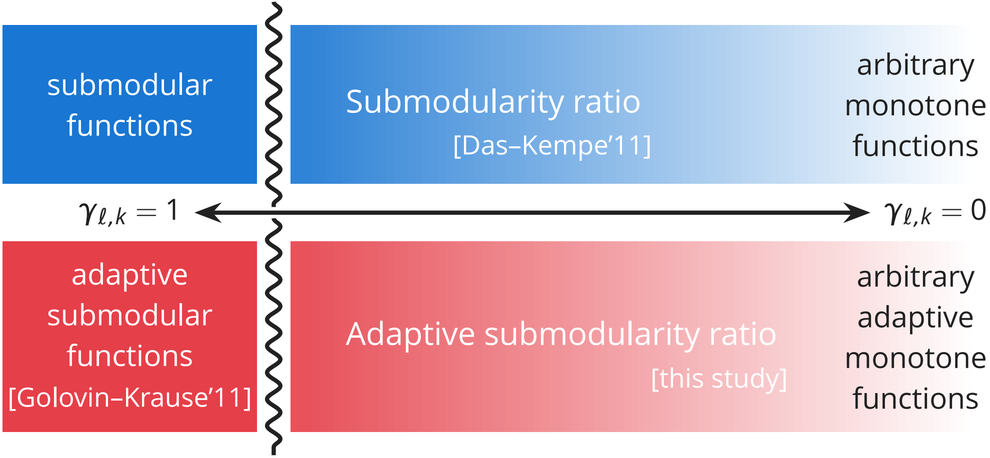
Adaptive submodularity ratio

We propose a new concept called adaptive submodularity ratio



Adaptive submodularity ratio

We propose a new concept called adaptive submodularity ratio



Adaptive submodularity ratio

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

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

$$\gamma_{\ell,k} \triangleq \min_{|\psi| \leq \ell,\ \pi \in \Pi_k} \frac{\sum_{v \in V} \Pr(v \in E(\pi, \Phi) | \Phi \sim \psi) \Delta(v|\psi)}{\underbrace{\Delta(\pi|\psi)}_{\text{the expected marginal gain of policy } \pi}}$$

Adaptive submodularity ratio

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

the expected marginal gain of single element v

the probability that element v is selected by policy π

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

Adaptive submodularity ratio

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

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

Q1 When does the greedy policy work well?

Adaptive submodularity ratio

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

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

Q1 When does the greedy policy work well?

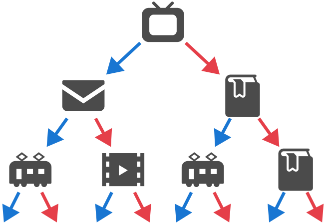
Theorem **Adaptive Greedy** is $(1 - \exp(-\gamma_{k,k}))$-approximation

Bounds on adaptivity gaps

A non-adaptive policy approximates an optimal adaptive policy

$\mathrm{GAP}_k(f, p) \triangleq$

An optimal
non-adaptive policy

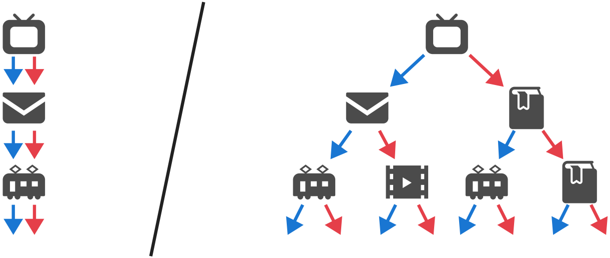


An optimal
adaptive policy

Bounds on adaptivity gaps

A non-adaptive policy approximates an optimal adaptive policy

$\mathrm{GAP}_k(f, p) \triangleq$

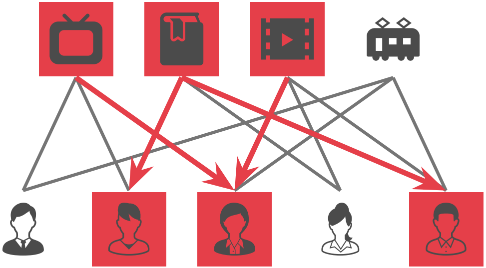


Q2 How different are the non-adaptive and adaptive policies?

Theorem $\mathrm{GAP}_k(f, p) \geq \beta_{0,k}\gamma_{0,k}$

$$\beta_{0,k} \triangleq \min_{S \subseteq V:\, |S| \leq k} \frac{\mathbb{E}[f(S, \Phi)]}{\sum_{v \in S} \mathbb{E}[f(\{v\}, \Phi)]}$$

Application: Influence maximization



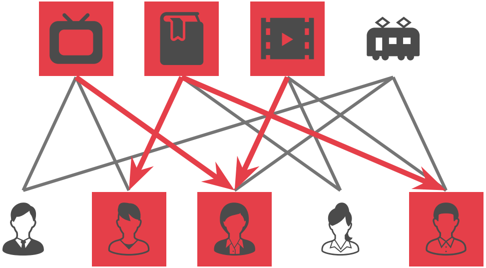
Non-adaptive setting

Select a subset in advance

Adaptive setting

Select ads one by one

Application: Influence maximization



Non-adaptive setting

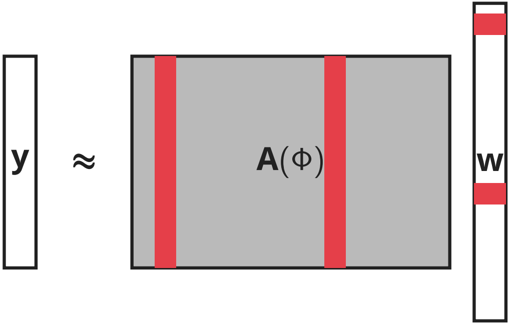
Select a subset in advance

Adaptive setting

Select ads one by one

Theorem $\gamma_{\ell,k} \geq \dfrac{k+1}{2k}$ on bipartite graphs with the triggering model

Select a subset of features to be observed precisely



Non-adaptive setting

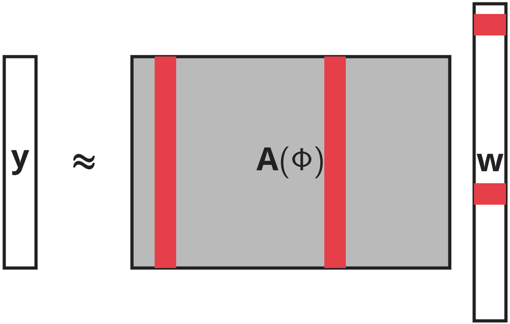
Select a subset in advance

Adaptive setting

Observe features one by one

Application: Adaptive Feature Selection

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 $\gamma_{\ell,k} \geq \min_{\phi} \min_{S \subseteq V:\, |S| \leq \ell+k} \lambda_{\min}(\mathbf{A}(\phi)_S^\top \mathbf{A}(\phi)_S)$

Summary

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