Submodular Cost Submodular Cover with an Approximate Oracle

Victoria G. Crawford\textsuperscript{1}, Alan Kuhnle\textsuperscript{2}, My T. Thai\textsuperscript{1}

\textsuperscript{1}University of Florida

\textsuperscript{2}Florida State University
Submodular Cost Submodular Cover (SCSC)

Definition (Submodular Cost Submodular Cover (SCSC))
Let \( f, c : 2^S \rightarrow \mathbb{R}_{\geq 0} \) be monotone submodular functions defined on subsets of a ground set \( S \) of size \( n \). Given threshold \( \tau \leq f(S) \), SCSC is to find

\[
\arg\min\{c(X) | X \subseteq S, f(X) \geq \tau\}.
\]

- SCSC arises in many applications
  - Influence in a social network
  - Data summarization
- NP-hard
The Greedy Algorithm

The greedy algorithm has an approximation ratio of

\[ \rho \left( \ln \left( \frac{\alpha}{\beta} \right) + 1 \right) \]

(Soma & Yoshida 2015).

Algorithm 1: greedy\((f, c, \tau)\)

\[
\begin{align*}
&f_{\tau} = \min\{f, \tau\} \\
i = 0, \quad A_i = \emptyset \\
&\textbf{while } f(A_i) < \tau \textbf{ do} \\
&\quad u = \arg\max_{x \in S \setminus A_i} \frac{\Delta f_{\tau}(A_i, x)}{c(x)}; \\
&\quad i = i + 1, \quad A_i = A_{i-1} \cup \{u\}; \\
&\textbf{end while} \\
&\textbf{return } A_i
\end{align*}
\]
Approximate Oracle

- We analyse the greedy algorithm for SCSC given an approximate oracle to \( f \)
  - Sketch of \( f \)
  - Noisy evaluations of \( f \)

**Definition (\( \epsilon \)-Approximate Oracle)**

A function \( F : 2^S \rightarrow \mathbb{R}_{\geq 0} \) is \( \epsilon \)-approximate to \( f : 2^S \rightarrow \mathbb{R}_{\geq 0} \) if for all \( X \subseteq S \),

\[
|f(X) - F(X)| \leq \epsilon.
\]
Approximate Oracle

- \( F \) is not necessarily monotone submodular
  - Existing guarantees don’t hold

Let \( X \subseteq Y \), and \( z \notin Y \).

\[
\begin{align*}
\Delta F(X, z) & \leq \epsilon \quad \text{and} \quad \Delta F(Y, z) \leq \epsilon \\
f(X) & \leq F(X) \leq f(X \cup \{z\}) & \quad f(Y) & \leq f(Y \cup \{z\}) \leq F(Y \cup \{z\})
\end{align*}
\]
Approximation Ratios

**Theorem**

Let $A$ be the set returned by the greedy algorithm with a value oracle to $\epsilon$-approximate oracle $F$. Then $f(A) \geq \tau - \epsilon$. And if $\mu > 4\epsilon c_{\text{max}} \rho / c_{\text{min}}$, 

$$c(A) \leq \frac{\rho}{1 - \frac{4\epsilon c_{\text{max}} \rho}{c_{\text{min}} \mu}} \left( \ln \left( \frac{\alpha}{\beta} \right) + 2 \right) c(A^*).$$

- If $\epsilon = 0$, nearly reduces to existing result; $\rho \left( \ln \left( \frac{\alpha}{\beta} \right) + 1 \right)$ (Soma & Yoshida 2015)
- $\beta$ can be very small
Approximation Ratios

Theorem

Let $A$ be the set returned by the greedy algorithm with a value oracle to $\epsilon$-approximate oracle $F$. Then $f(A) \geq \tau - \epsilon$. And if $\mu > 4\epsilon c_{\max} \rho / c_{\min}$, then for any $\gamma \in (0, 1 - 4\epsilon c_{\max} \rho / c_{\min} \mu)$,

$$c(A) \leq \frac{\rho}{1 - \frac{4\epsilon c_{\max} \rho}{c_{\min} \mu} - \gamma \left( \ln \left( \frac{n\alpha \rho}{\gamma \mu} \right) + 2 \right)} c(A^*).$$

- No more $\beta$
- Incomparable
Application: Influence Threshold

- Find seed set of minimum cost such that expected propagation from seed set is at least $\tau$
- Scalable influence estimator of Cohen et al. (2014)
  - Not submodular
  - $\epsilon$-approximate
- Computed our approximation ratios
Thank you! Poster #168