Analysis of Greedy Approximation with Non-submodular Potential Function

Ding-Zhu Du, Ron Graham, P.M. Pardalos
Pengjun Wan, Weili Wu, Wenbo Zhao

The Aim

of this talk is to present a little technique, like the secret of optical glass, a simple and valuable technique.

Organization

- Background (why valuable?)

  1) submodularity

  2) relationship with greedy approximation
- The technique (how to deal with nonsubmodular potential function)

Background

- There exist many greedy algorithms in the literature.
- Some have theoretical analysis. But, most of them do not.
- A greedy algorithm with theoretical analysis usually has a submodular potential function.

What is a submodular function?

Consider a function f on all subsets of a set E.
f is submodular if

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

Set-Cover

Given a collection *C* of subsets of a set *E*, find a minimum subcollection *C′* of *C* such that every element of *E* appears in a subset in *C′*.

Example of Submodular Function

For a subcollection A of C, define

$f(A) = | \cup_{S \in A} S |.$

Then

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

Greedy Algorithm

$C' \leftarrow \emptyset;$

while $|E| > f(C')$ do

choose $S \in C$ to maximize $f(C' \cup \{S\})$ and

$C' \leftarrow C' \cup \{S\};$

Analysis

Suppose $S_1, S_2, \ldots, S_k$ are selected by Greedy Algorithm. Denote $C_i = \{S_1, \ldots, S_i\}$. Then

$$f(C_{i+1}) \geq f(C_i) + (|E| - f(C_i)) / opt$$

$$(|E|-f(C_i))(1-1/opt) \geq |E|-f(C_{i+1})$$

$$|E|-f(C_{i+1}) \leq (|E|-f(C_i))(1-1/opt)$$

$$\leq (|E|-f(C_{i-1}))(1-1/opt)^2$$

$$\leq \cdots$$

$$\leq |E|(1-1/opt)^{i+1}$$

Choose i to be the largest one satisfying

$$opt \leq |E|-f(C_i).$$

Then $\quad k-i \leq opt$

$$opt \leq |E|(1-1/opt)^i$$

$$
\begin{aligned}
opt &\le |E|(1 - 1/opt)^i \\
&\le |E| e^{-i/opt} \\
i &\le opt \ln(|E|/opt)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
k &\le opt + i \\
&\le opt\,(1 + \ln(|E|/opt))
\end{aligned}
$$

Analysis

Suppose $S_1, S_2, \ldots, S_k$ are selected by Greedy Algorithm. Denote $C_i = \{S_1, \ldots, S_i\}$. Then

$$f(C_{i+1}) \geq f(C_i) + (|E| - f(C_i))/opt$$

Denote $\Delta_X f(A) = f(A \cup \{X\}) - f(A)$.

Consider an optimal solution $C^* = \{X_1, \ldots, X_{opt}\}$

Denote $C_j^* = \{X_1, \ldots, X_j\}$.

By greedy rule, $\Delta_{S_{i+1}} f(C_i) \geq \Delta_{X_{j+1}} f(C_i)$

for all $0 \leq j \leq opt - 1$

$$\text{Thus, } \Delta_{S_{i+1}} f(C_i) \geq \left(\sum_{0 \leq j \leq opt-1} \Delta_{X_{j+1}} f(C_i)\right) / opt$$

$$\geq \left(\sum_{0 \leq j \leq opt-1} \Delta_{X_{j+1}} f(C_i \cup C_j^*)\right) / opt$$

$$= (f(C_i \cup C^*) - f(C_i)) / opt$$

$$= (|E| - f(C_i)) / opt$$

Where we need submodularity ?

$$\Delta_{X_{j+1}} f(C_i) \geq \Delta_{X_{j+1}} f(C_i \cup C_j^*)$$

$$A \subset B \Rightarrow \Delta_X f(A) \geq \Delta_X f(B)$$

Actually, this inequality holds if and only if *f* is submodular and

A is a subset of *B* => $f(A) \leq f(B)$

(monotone increasing)

$(f$ is submodular$)$ implies

$$A \subset B \Rightarrow \Delta_x f(A) \geq \Delta_x f(B) \text{ for } x \notin B$$

$(f$ is monotone increasing$)$ implies

$$A \subset B \Rightarrow \Delta_x f(A) \geq \Delta_x f(B) \text{ for } x \in B$$

Meaning of Submodular

- The earlier, the better!
- Monotone decreasing gain!

Theorem

Greedy Algorithm produces an approximation within ln n +1 from optimal.

The same result holds for weighted set-cover.

Weighted Set Cover

Given a collection *C* of subsets of a set *E* and a weight function *w* on *C*, find a minimum total-weight subcollection *C′* of *C* such that every element of *E* appears in a subset in *C′*.

Greedy

while $|E| > f(C')$ do

choose $S \in C$ to maximize $f(C' \cup \{S\}) / w(S)$

and $C' \leftarrow C' \cup \{S\}$;

A General Problem

Consider a set E, a monotone increasing, submodular function f on all subsets of E and a weight function w on E. Define

$T = \{A \mid \forall x \in E, f(A \cup \{x\}) = f(A)$.

Find minimum total-weight A in T.

A General Theorem

If $f(\emptyset) = 0$, f is an integer function

and $w(x) \geq 0, \forall x \in E$,

then Greedy gives a $(\ln \gamma + 1)$-approximation

where $\gamma = \max_{x \in E} f(\{x\})$.

Is it true?

- Every previously known one-stage greedy approximation with theoretical analysis has a submodular (or supermodular) potential function.
- Almost, only one exception which is about Steiner tree.

Steiner Tree

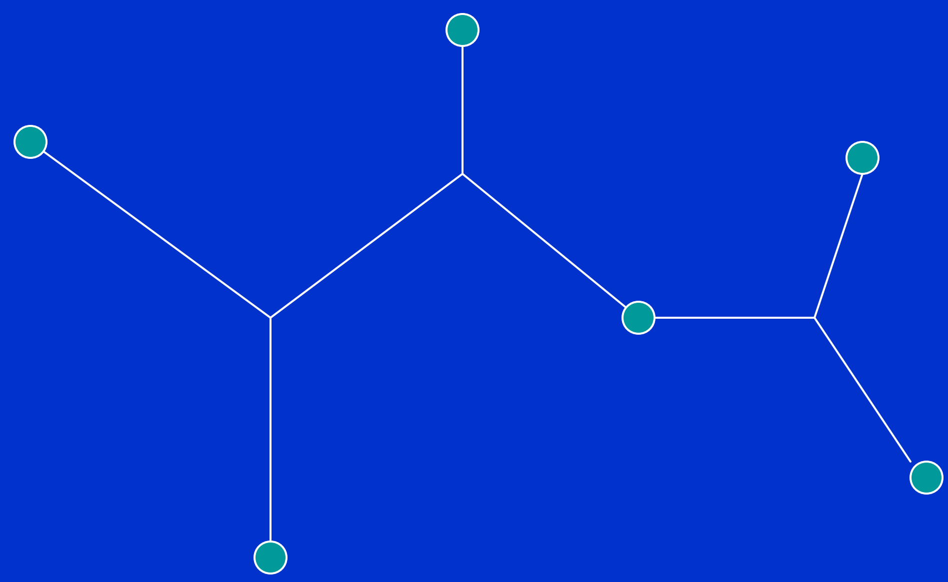
Given a finite set of points, call terminals, in a metric space, find a minimum length tree interconnecting them.

- Euclidean plan
- Rectilinear plan
- Network

Full Components

- A Steiner tree is full if every terminal is a leaf.
- Every Steiner tree can be decomposed into small full Steiner subtrees, call full component.
- A full component with k terminals is called a k-component.

Full Components

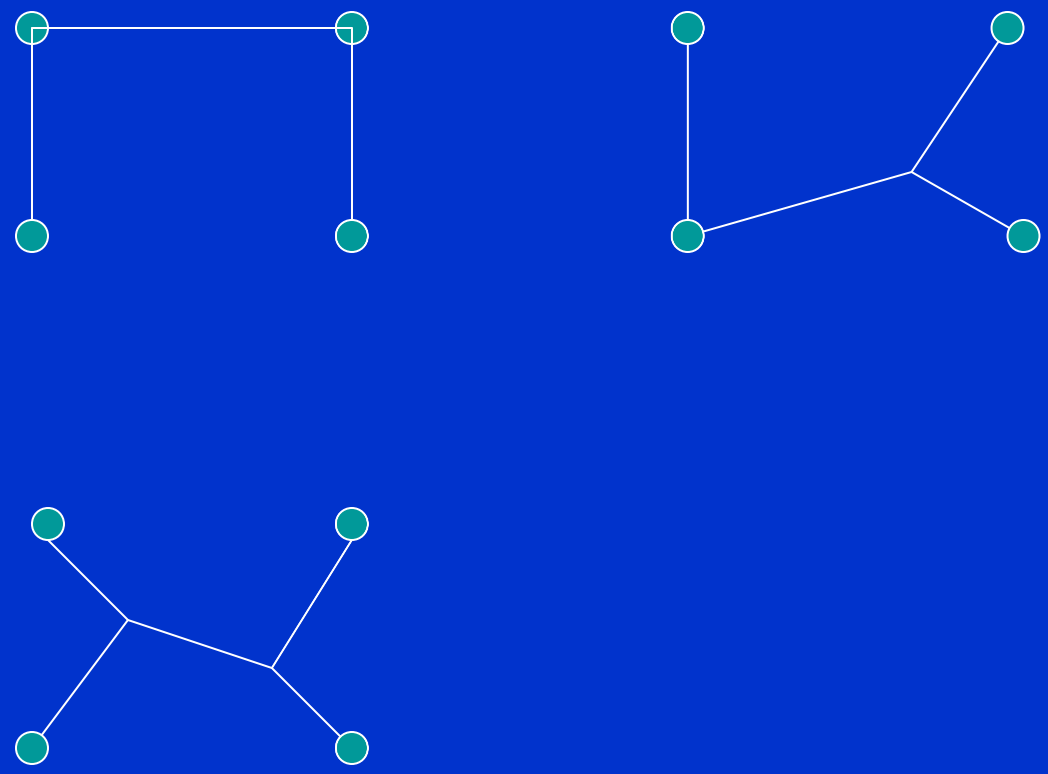


Approximation for Network ST

- Minimum spanning tree (submodular)
- iterated 1-Steiner tree (non-submodular)
- 3-restricted Steiner tree (submodular)
- K-restricted Steiner tree (submodular)

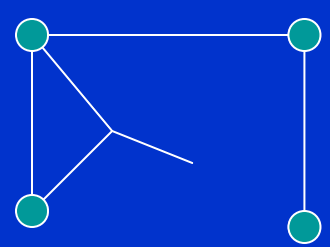
Iterated 1-Steiner Tree

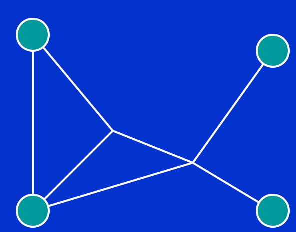
- At each iteration, add a Steiner node to maximize the reduction of the total length.
- A new Steiner node can connect to an old Steiner node. (This is the main difference from 3-restricted Steiner tree.)



Why non-submodular?

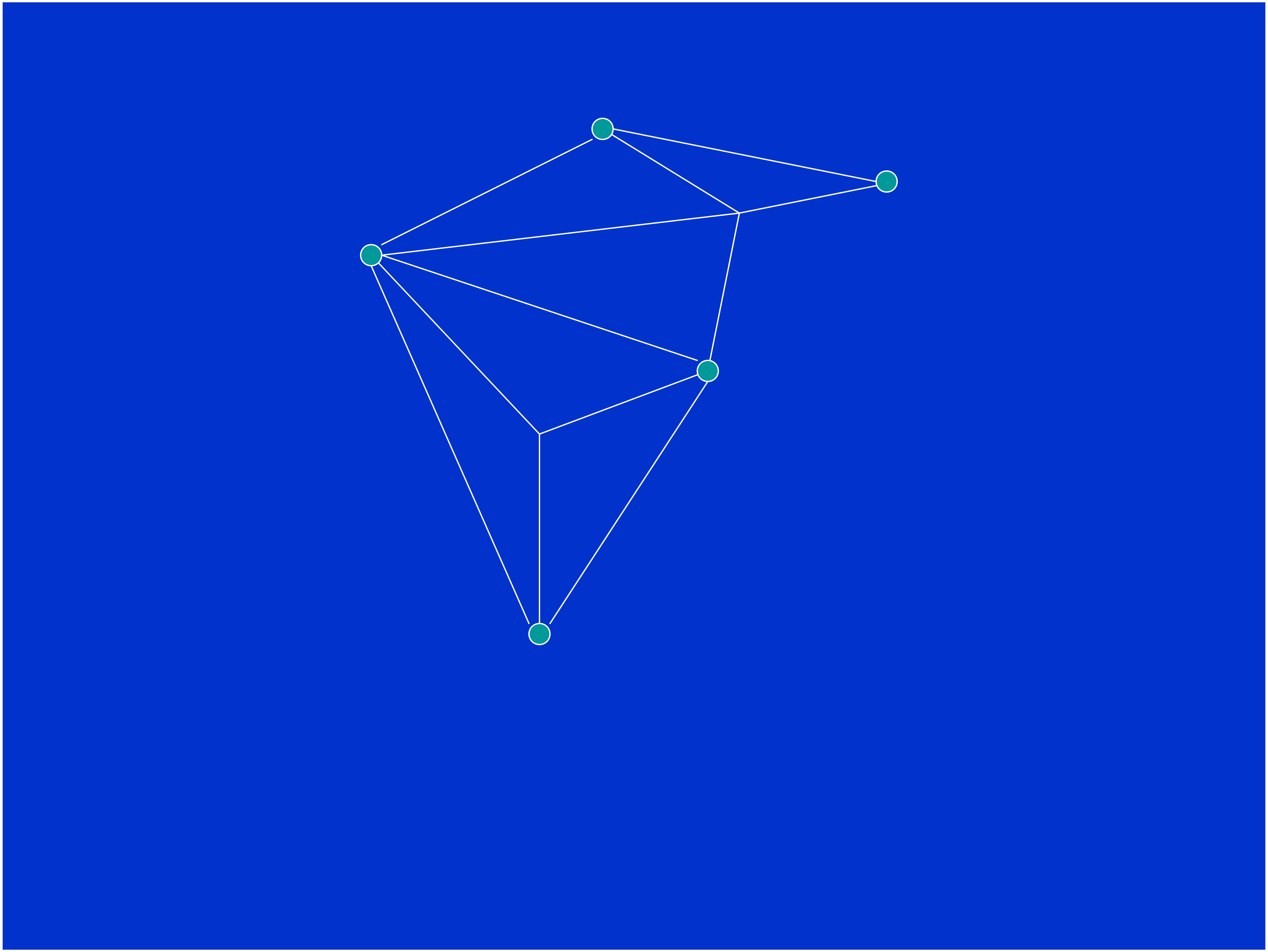
- After the 1st one is added, the gain of the 2nd one is increasing.





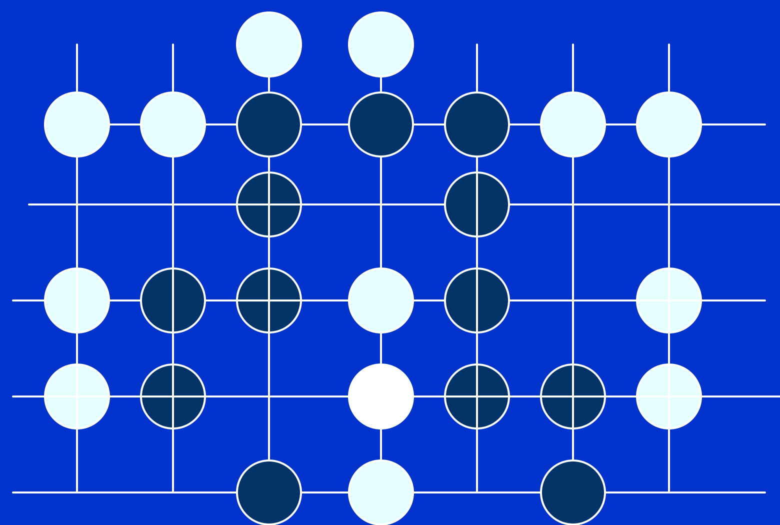
History

- S.-K. Chang (1972)
- J.M. Smith, D.T. Lee and J.S. Liebman (1981)
- A. B. Kahng and G. Robin (1992)
- G. Robin and A. Zelikovsky (2000) gave an theoretical analysis to iterated 1-Steiner tree for pseudo-bipartite graphs.



How should we do with nonsubmodular functions?

Find a space to play your trick



Where is the space?

Suppose $S_1, S_2, ..., S_k$ are selected by Greedy Algorithm. Denote $C_i = \{S_1, ..., S_i\}$. Then

$$f(C_{i+1}) \geq f(C_i) + (|E| - f(C_i)) / opt$$

Why the inequality true?

Because

$\Delta s_{i+1} f(C_i) \geq \Delta s f(C_i)$ for every S,

including those S in the optimal solution C^*.

Note that $|E| = f(C^*)$ and $opt = |C^*|$.

Suppose $C^* = \{X_{1}, ..., X_{opt}\}$ and $C_i^* = \{X_{1}, ..., X_i\}$.

Then $f(C^*) = \sum \Delta x_{j+1} f(C_i \cup C_j^*)$.

Moreover, by submodularity of f,

$\Delta x_{j+1} f(C_i) \geq \Delta x_{j+1} f(C_i \cup C_j^*)$.

Observations

- The submodularity has nothing to do with sequence chosen by the greedy algorithm.

  It is only about *X1, …, Xopt*
- The ordering of *X1, …, Xopt* is free to choose.

When f is nonsubmodular

(1) We may choose $X_1, \ldots, X_{opt}$ such that f is submodular on $\{ X_1, \ldots, X_{op} \}$.
(1-iterated Steiner tree)

(2) We may choose certain ordering of $X_1, \ldots, X_{op}$ to make $\Delta_{x_{j+1}} f(C_i \cup C_j^*) - \Delta_{x_{j+1}} f(C_i)$ smaller.
(Connected dominating set)

Iterated k-Steiner Tree

Theorem. Iterated k-Steiner tree has the approximation performance same as that of k-restricted Steiner tree.

Connected Dominating Set

Theorem. Connected dominating set in graph has polynomial-time $a(1+\ln \Delta)$-approximation for any $a > 1$, where Δ is the maximum node degree.

Applications

- Iterated k-Steiner trees
- ($\ln n + 1$)-approximation for minimum connected dominating set
- Minimum energy topological control in wireless networks, etc.

Connected Dominating Set

Given a graph, find a minimum node-subset such that

- each node is either in the subset or adjacent to a node in the subset and
- subgraph induced by the subset is connected.

Thank you!