Sparsest Cut Problem

Chappidi Venkata Vamsidhar Reddy - 2021CS10557
Harsh Vora - 2021CS10548

COL754 Project Presentation

Contents

Current Section

Introduction - Sparsest Cut Problem

Given

- graph $G = (V, E)$
- costs c_e on edges, and
- demands d_i on pairs of vertices (s_i, t_i), $i = 1, \ldots, k$

Find a cut $S \subseteq V$ that minimizes the ratio

$$\rho(S) \equiv \frac{\sum_{e \in \delta(S)} c_e}{\sum_{i:|S \cap \{s_i, t_i\}|=1} d_i}.$$

Current Section

Non-Linear Program

$$\text{minimize} \quad \frac{\sum_{(i,j) \in E} c_{ij} \cdot (x_i \cdot (1 - x_j))}{\sum_{i,j \in V} d_{ij} \cdot (x_i \cdot (1 - x_j))}$$

subject to:

$$x_i \in \{0, 1\}, \quad \forall i \in V.$$

Solving the Non-Linear Program

- $x_i = 1$ if $i \in S$, and $x_i = 0$ otherwise.
- Solved using the Mixed Integer Non-Linear Programming (MINLP) solver **SCIP**.
- Solution S is the set of vertices with $x_i = 1$.

Current Section

Integer Non-Linear Program

Modify the program to use

- edge variables x_e: $x_e = 1$ if edge e is in the cut, and $x_e = 0$ otherwise.
- demand pair variables y_i: $y_i = 1$ if demand pair i is separated by the cut, and $y_i = 0$ otherwise.
- $\mathscr{P}_i$ is the set of all paths from s_i to t_i

Integer Non-Linear Program

$$\text{minimize} \quad \frac{\sum_{e \in E} c_e \cdot x_e}{\sum_{i=1}^{k} d_i \cdot y_i}$$

subject to:

$$\begin{aligned}
\sum_{e \in P} x_e \geq y_i, &\quad \forall P \in \mathscr{P}_i, 1 \leq i \leq k, \\
x_e \in \{0, 1\}, &\quad \forall e \in E, \\
y_i \in \{0, 1\}, &\quad 1 \leq i \leq k.
\end{aligned}$$

Linear Program Relaxation(Exponential Number of Constraints)

The linear program is defined as follows:

$$\text{minimize} \quad \sum_{e \in E} c_e x_e$$

$$\begin{aligned}
\text{subject to} \quad & \\
\sum_{i=1}^{k} d_i y_i &= 1, \\
\sum_{e \in P} x_e &\geq y_i, \quad \forall P \in \mathscr{P}_i,\ 1 \leq i \leq k, \\
y_i &\geq 0, \quad 1 \leq i \leq k, \\
x_e &\geq 0, \quad \forall e \in E.
\end{aligned}$$

Linear Program Relaxation(Polynomial Number of Constraints)

We introduce new variables $d_{ui} \forall u \in V, 1 \leq i \leq k$.

$$\text{minimize} \quad \sum_{e \in E} c_e x_e$$

subject to:

Linear Program Relaxation(Polynomial Number of Constraints)

$$
\begin{aligned}
\sum_{i=1}^{k} d_i \cdot y_i &= 1, \\
d_{u,i} - d_{v,i} &\leq x_e, \quad \forall e = (u,v) \in E, 1 \leq i \leq k, \\
d_{s_i,i} &= 0, \quad 1 \leq i \leq k, \\
d_{t_i,i} &\geq y_i, \quad 1 \leq i \leq k, \\
d_{u,i} &\geq 0, \quad \forall u \in V, 1 \leq i \leq k, \\
y_i &\geq 0, \quad 1 \leq i \leq k, \\
x_e &\geq 0, \quad \forall e \in E.
\end{aligned}
$$

Fretchet Embeddings

Given a metric space (V, d) and p subsets of vertices $A_1, \ldots, A_p$, a Frechet embedding $f : V \to \mathbb{R}^p$ is defined by

$$f(u) = \big(d(u, A_1), d(u, A_2), \ldots, d(u, A_p)\big) \in \mathbb{R}^p$$

for all $u \in V$.

Lemma 1: Given a metric (V, d) and the Frechet embedding $f : V \to \mathbb{R}^p$ defined above, for any $u, v \in V$,

$$\|f(u) - f(v)\|_1 \leq p \cdot d_{uv}.$$

Lemma 2: Given a metric space (V, d) with k distinguished pairs $s_i, t_i \in V$, we can pick $p = O(\log^2 k)$ sets $A_j \subseteq V$ using randomization such that a Frechet embedding $f : V \to \mathbb{R}^p$ satisfies

$$\|f(s_i) - f(t_i)\|_1 \geq \Omega(\log k) \cdot d(s_i, t_i) \quad \text{for} \quad 1 \leq i \leq k$$

with high probability.

Sparsest Cut Candidate Sets

Let (V, d) be a Frechet-embeddable metric, and let $f : V \to \mathbb{R}^m$ be the associated embedding. Then there exist $\lambda_S \geq 0$ for all $S \subseteq V$ such that for all $u, v \in V$,

$$\|f(u) - f(v)\|_1 = \sum_{S \subseteq V} \lambda_S \chi_{\nabla(S)}(u, v).$$

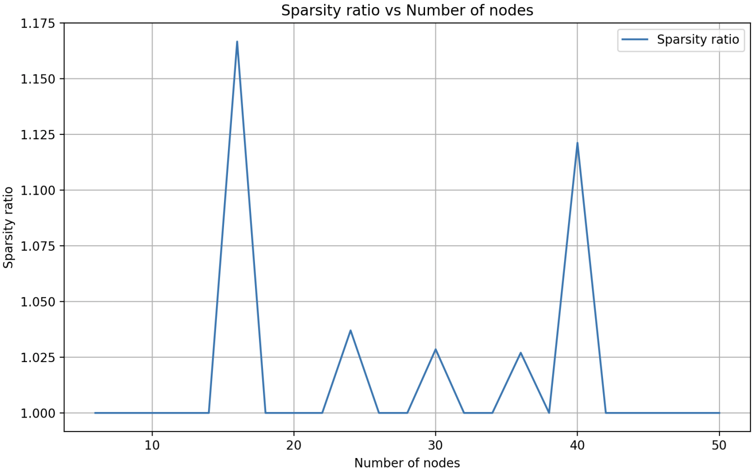
- Obtain the sets with non-zero λ_S.
- Use these sets to compute the sparsest cut.

Approximation Algorithm

- Solve the linear program relaxation.
- Generate $Qln(k)\log_2^{2k}$ $O(log^2(k))$ Frechet sets.
- Compute Frechet embeddings with obtained x_e as edge lengths.
- Compute the sparsest cut candidate sets.
- Obtain an $O(\log k)$-approximation for the sparsest cut by finding the best cut from the candidate sets.
- The algorithm runs in polynomial time.

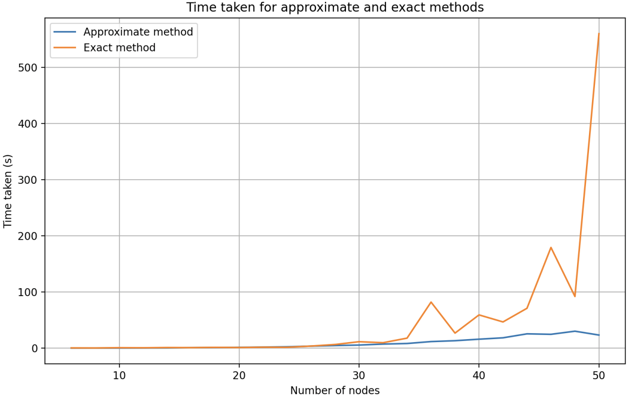
Current Section

Graphs



Sparsity Ratio vs Number of Nodes

Time vs Number of nodes



Time(s) vs Number of nodes

Current Section

Generating Candidate Sets

We can generate at most mn subsets such that λ_s is non-zero for these subsets.

- To prove(considering 1-dimensional embedding): There exist at most n non-zero λ_s.
- Construction: Sort all vertices according to their single-dimensional embedding: values $x_1 \leq x_2 \leq \cdots \leq x_n$.
- Let the order of sorted vertices be a permutation $\pi_1, \ldots \pi_n$.
- Consider the subsets $S_i = \{\pi_1, \ldots, \pi_i\}$: $1 \leq i \leq n-1$.
- $\lambda_{S_i} = x_{i+1} - x_i$.
- For any $i < j$,
  $$x_j - x_i = \sum_{k=i}^{j-1} \lambda_{S_i} = \sum_{S \subseteq V} \lambda_S \cdot \chi_{\delta(S)}(i,j)$$
- Therefore, at most n of the λ_s are non-zero.

Generating Candidate Sets

- For m dimensions, order the vertices according to the first dimension, then the second, and so on.
- $$\begin{aligned} \|f(u) - f(v)\|_1 &= \sum_{i=1}^{m} |x_u^i - x_v^i| \\ &= \sum_{i=1}^{m} \sum_{S \subseteq V} \lambda_S^i \chi_{\delta(S)}(u, v) \\ &= \sum_{S \subseteq V} \lambda_S \chi_{\delta(S)}(u, v) \end{aligned}$$
- The set of subsets can be computed in polynomial time.

Proof of approximation

There is a randomized O(log k)-approximation algorithm for the sparsest cut problem.

- $\rho(S*)$ is the minimum sparsity of computed candidate sets.
- Now the task is to prove that this value is within $O(\log k)$ of the optimal sparsity.

Proof of approximation

$$
\begin{aligned}
\rho(S^*) &= \min_{S:\lambda_S>0} \frac{\sum_{e\in\delta(S)} c_e}{\sum_{i:|S\cap\{s_i,t_i\}|=1} d_i} \\
&= \min_{S:\lambda_S>0} \frac{\sum_{e\in E} c_e \cdot \chi_{\delta(S)}(e)}{\sum_i d_i \cdot \chi_{\delta(S)}(s_i,t_i)} \\
&\leq \frac{\sum_{S\subseteq V} \lambda_S \sum_{e\in E} c_e \cdot \chi_{\delta(S)}(e)}{\sum_{S\subseteq V} \lambda_S \sum_{i=1}^{k} d_i \cdot \chi_{\delta(S)}(s_i,t_i)} \\
&= \frac{\sum_{e\in E} c_e \sum_{S\subseteq V} \lambda_S \chi_{\delta(S)}(e)}{\sum_{i=1}^{k} d_i \sum_{S\subseteq V} \lambda_S \chi_{\delta(S)}(s_i,t_i)} \\
&= \frac{\sum_{e=(u,v)\in E} c_e \|f(u)-f(v)\|_1}{\sum_{i=1}^{k} d_i \|f(s_i)-f(t_i)\|_1} \\
&\leq \frac{O(\log^2 k)\cdot \sum_{e=(u,v)\in E} c_e \cdot d_x(u,v)}{\Omega(\log k)\cdot \sum_{i=1}^{k} d_i \cdot d_x(s_i,t_i)},
\end{aligned}
$$

Proof of approximation

$$
\begin{aligned}
\rho(S^*) &\leq O(\log k)\frac{\sum_{e=(u,v)\in E} c_e \cdot d_x(u,v)}{\sum_{i=1}^{k} d_i \cdot d_x(s_i,t_i)} \\
&\leq O(\log k)\frac{\sum_{e\in E} c_e x_e}{\sum_{i=1}^{k} d_i y_i} \\
&= O(\log k)\sum_{e\in E} c_e x_e \\
&\leq O(\log k)\cdot \mathrm{OPT}.
\end{aligned}
$$

Proof of distortion

Lemma 1: Given a metric (V, d) and the Frechet embedding $f : V \to \mathbb{R}^p$ defined above, for any $u, v \in V$,

$$\|f(u) - f(v)\|_1 \leq p \cdot d_{uv}.$$

- To prove that $\sum_{e=(u,v)\in E} c_e \|f(u) - f(v)\|_1 \leq O(\log^2 k) \cdot \sum_{e=(u,v)\in E} c_e \cdot d_x(u, v)$.
- Let w be the point in the Frechet set A that is closest to v.

Proof of distortion

$$
\begin{aligned}
d(u,A) &\leq d_{uw} \leq d_{uv} + d_{vw} = d_{uv} + d(v,A),\\
d(v,A) &\leq d_{vw} + d(u,A),\\
|d(u,A) - d(v,A)| &\leq d_{uv},\\
\|f(u) - f(v)\|_1 &= \sum_{j=1}^{p} |d(u,A_j) - d(v,A_j)| \leq p \cdot d_{uv}.
\end{aligned}
$$

Proof of distortion

Lemma 2: Given a metric space (V, d) with k distinguished pairs $s_i, t_i \in V$, we can pick $p = O(\log^2 k)$ sets $A_j \subseteq V$ using randomization such that a Frechet embedding $f : V \to \mathbb{R}^p$ satisfies $\|f(s_i) - f(t_i)\|_1 \geq \Omega(\log k) \cdot d(s_i, t_i)$ for $1 \leq i \leq k$ with high probability.

Proof. To construct the embedding, we use randomization to select subsets $A_j \subseteq V$ and define the Frechet embedding $f : V \to \mathbb{R}^p$ based on distances to these subsets. Here are the steps:

- Selection of subsets:
  Let $T = \{s_1, t_1, \ldots, s_k, t_k\}$ be the set of terminals. Assume $|T|$ is a power of two (if not, pad it by duplicating elements). Let $\tau = \log_2(2k)$ so that $|T| = 2^\tau$. For each level $t = 1, \ldots, \tau$, define $2^{\tau - t}$ subsets $A_{t,\omega}$ by randomly sampling $2^{\tau - t}$ vertices from T, with replacement, for $\omega = 1, \ldots, L$, where $L = O(\log k)$.

Proof of distortion

- Frechet embedding:
  For each vertex $u \in V$, define $f(u) \in \mathbb{R}^p$, where
  $p = O(\tau L) = O(\log^2 k)$, as:

$$f(u) = \big(d(u, A_{1,1}), \ldots, d(u, A_{\tau,L})\big).$$

- Analysis of distances in the embedding:
  For a pair (s_i, t_i), we focus on ensuring that $\|f(s_i) - f(t_i)\|_1$ is large.
  Define a sequence of radii r_t for $t = 1, \ldots, \tau$ as follows:
  - $r_0 = 0$.
  - r_t is the minimum radius such that $|B(s_i, r_t)| \geq 2^t$ and $|B(t_i, r_t)| \geq 2^t$.
  - Let $\hat{t}$ be the smallest index such that $r_{\hat{t}} \geq \frac{1}{4} d(s_i, t_i)$. Then redefine $r_{\hat{t}} = \frac{1}{4} d(s_i, t_i)$.

Proof of distortion

By construction, the balls $B(s_i, r_{\hat{t}})$ and $B(t_i, r_{\hat{t}})$ do not intersect, ensuring that subsets $A_{t,\omega}$ capture meaningful separation between s_i and t_i.

- Probabilistic guarantee for subsets:
  For each subset $A_{t,\omega}$, the probability of capturing separation between s_i and t_i is analyzed using the following events:
  - $A_{t,\omega} \cap B(s_i, r_t) = \emptyset$.
  - $A_{t,\omega} \cap B(t_i, r_{t-1}) \neq \emptyset$.

  Using the properties of random sampling, we can show that the probability of these events is constant. Applying Chernoff bounds, the sum of contributions over L subsets ensures:

$$\sum_{\omega=1}^{L} |d(s_i, A_{t,\omega}) - d(t_i, A_{t,\omega})| \geq \Omega(L(r_t - r_{t-1})).$$

Proof of distortion

- Summing over all levels:
  Since the radii r_t are telescoping, we have:

$$\|f(s_i) - f(t_i)\|_1 = \sum_{t=1}^{\hat{t}} \sum_{\omega=1}^{L} |d(s_i, A_{t,\omega}) - d(t_i, A_{t,\omega})| \geq \Omega(L \cdot r_{\hat{t}}).$$

  Substituting $r_{\hat{t}} = \frac{1}{4} d(s_i, t_i)$ and $L = O(\log k)$ gives:

$$\|f(s_i) - f(t_i)\|_1 \geq \Omega(\log k) \cdot d(s_i, t_i).$$

- High probability result:
  By union bounding over all k pairs, we ensure that the desired property holds for all s_i, t_i with high probability.

Thus, we conclude that $p = O(\log^2 k)$ subsets suffice to construct the embedding f with the required property.