Concept and Existence of Good Vertex Sparsifiers

Radu Vintan

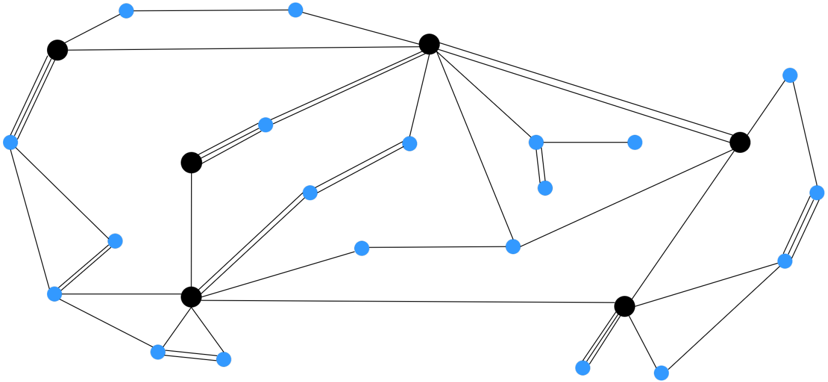
Technische Universität München

July 12, 2021

What is given

Internet: satisfy demands between different pairs of black nodes.

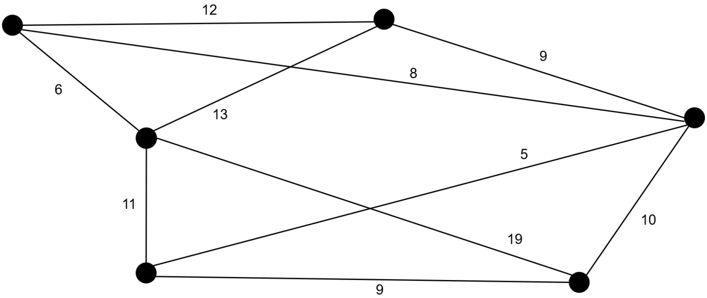
Routing commodities without creating bottlenecks is hard.



What is given

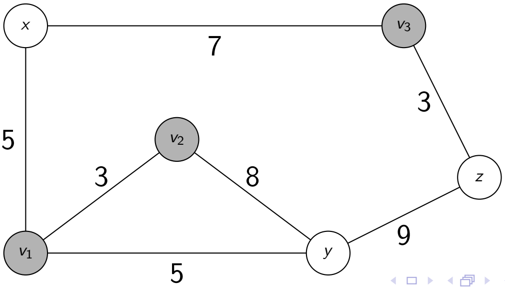
Sparsifying the graph

Routing here should be similar to routing in the big graph: same bottlenecks.



What is given

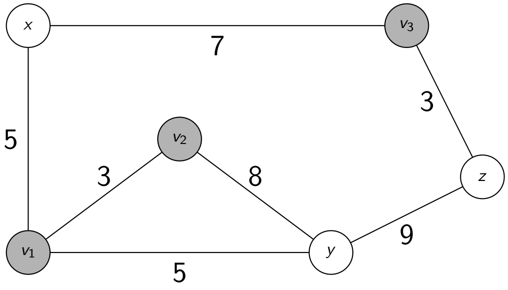
- $G := (V, E)$ undirected graph with capacities $c : E \to \mathbb{R}_{\geq 0}$
- $n := |V|$ and $m := |E|$
- $K \subset V$ set of terminals, $k := |K|$
- Running example:



Basic tools

Define *cut function* $h : 2^V \to \mathbb{R}_{\geq 0}$ by $h(A) := \sum_{u \in A,\ v \notin A} c(u, v)$.

E.g. $h(\{v_1, v_2, y\}) = 9 + 5 = 14$.



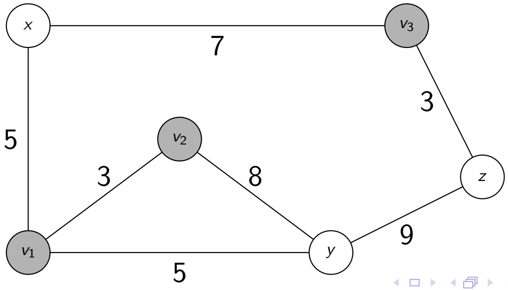
Basic tools

Define *terminal cut function* $h_K : 2^K \to \mathbb{R}_{\geq 0}$ by

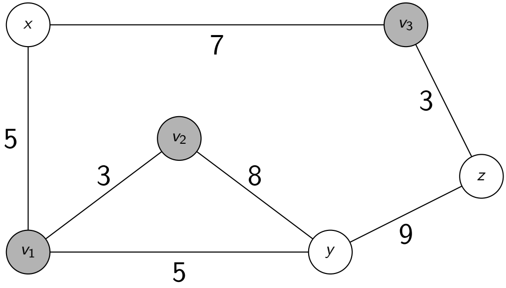
$$h_K(U) := \min_{A \subset V,\ A \cap K = U} h(A)$$

i.e. minimum cut that isolates U from $K \setminus U$.

How to compute e.g. $h_K(\{v_1, v_2\})$?

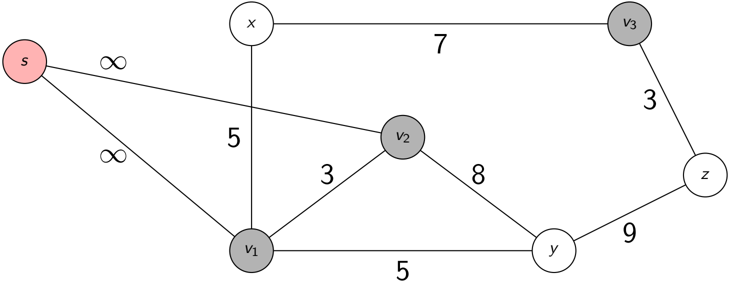


Computing $h_K(\{v_1, v_2\})$



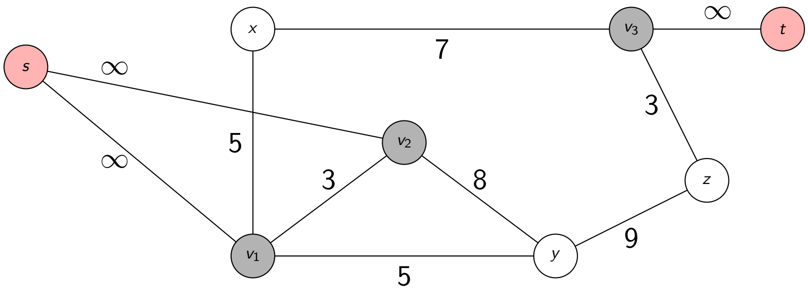
Computing $h_K(\{v_1, v_2\})$

Add source node s and connect it to v_1 and v_2, i.e. the set U.



Computing $h_K(\{v_1, v_2\})$

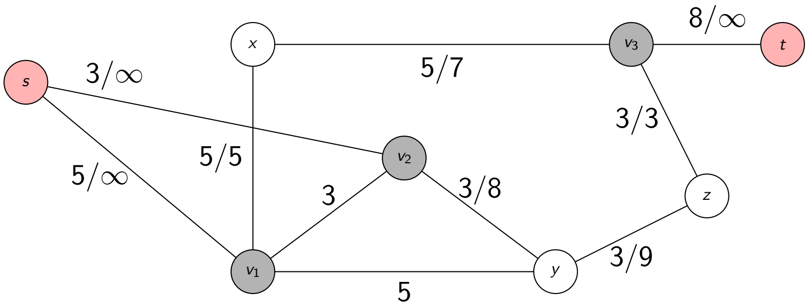
Add target node t and connect it to v_3, i.e. the set $K \setminus U$.



Computing $h_K(\{v_1, v_2\})$

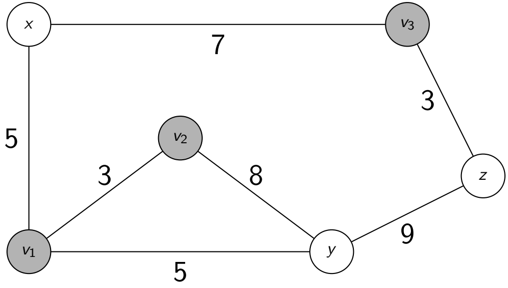
Compute *Max Flow*!

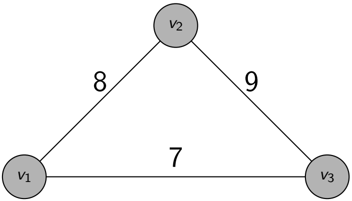
Obtain $h_K(\{v_1, v_2\}) = 8$ for the cut $A := \{x, v_1, v_2, y, z\}$.



Vertex sparsifier

Define a graph G' only on K and use its cut function h' as an approximation of h_K.





Vertex sparsifier

$G' := (K, E')$ is a *(cut-)sparsifier* of G if

$$h_K(U) \leq h'(U) \text{ for all } U \subset K$$

So it is in our interest to find a sparsifier that *minimizes* the *quality*

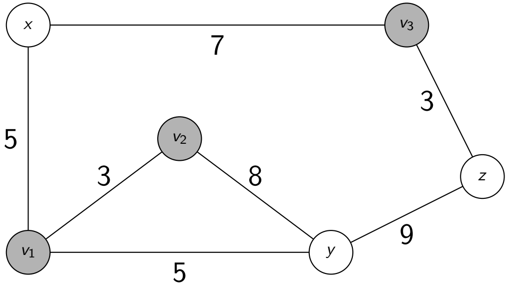
$$\max_{U \subset K} \frac{h'(U)}{h_K(U)}$$

Basically find a sparsifier with h' as small as possible.

Finding a sparsifier

First ansatz: partition nodes around terminals.

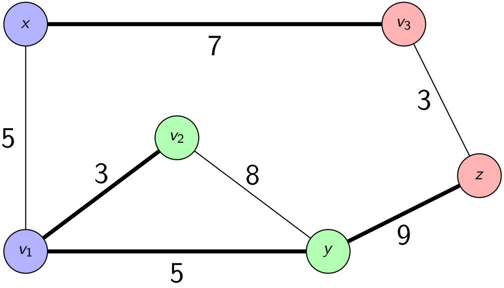
$f : V \to K$, s.t. $f(t) = t$, for all $t \in K$



Contracting

$f(x) = v_1, f(y) = v_2, f(z) = v_3$

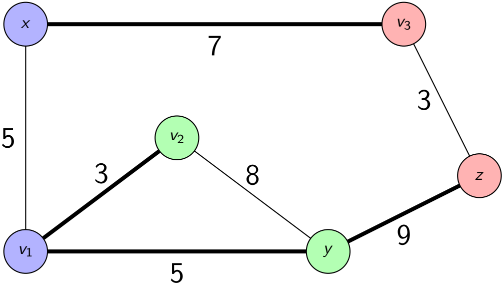
Contract non-bold edges.



Contracting

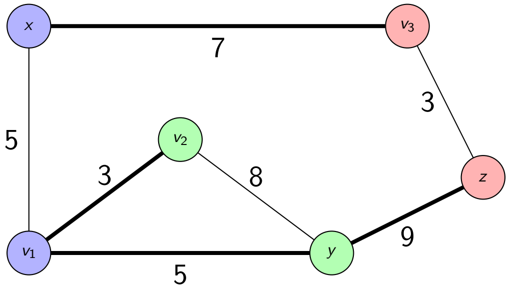
For the "supernodes" v_1, v_2, v_3, new costs c_f:

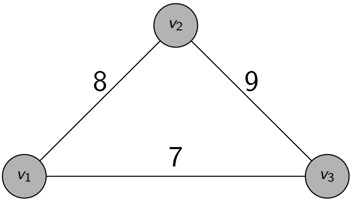
$c_f(v_1, v_2) = 3 + 5 = 8$, $c_f(v_2, v_3) = 9$ and $c_f(v_3, v_1) = 7$.



Partitioning gives sparsifiers!

Obtain G_f defined on K by contracting:





$h_f(\{v_1, v_2\}) = 16 \geq 8 = h_K(\{v_1, v_2\})$. In fact, $h_f \geq h_K$.
G_f is a sparsifier

Partitioning gives sparsifiers!

Can we always find partitioning f, s.t. G_f has good quality?

Unfortunately no. Idea for fix: combine different partitionings

i.e. good sparsifier $G' := \sum_{\text{partitioning } f} \text{weight}(f) G_f$

Main Idea

Want to prove there exists a sparsifier of good quality.

Meat of the proof consists of a zero-sum game:

- We are P_1. We provide a partitioning/sparsifier f.
- P_2 is our adversary and provides a subset $U \subset K$. His objective: $\frac{h_f(U)}{h_K(U)}$ should be as big as possible.
- $N(f, U) := \frac{h_f(U)}{h_K(U)}$ is P_2-s win if P_1 plays f and P_2 plays U

Randomized strategies for game

Strategies are randomized:

- P_1 allowed to choose distribution γ on partitionings f
- P_2 allowed to choose distribution λ on subsets $U \subset K$

This will allow us to combine different partitionings.

Neumann's Minimax Theorem

Assume randomized strategies, i.e. each player allowed to choose a distribution. Then:

- $\exists$ optimal distribution γ on partitionings for P_1
- $\exists$ optimal distribution λ on subsets $U \subset K$ for P_2
- No player can improve by changing distribution.
- $v := E_{(f \leftarrow \gamma, U \leftarrow \lambda)} N(f, U)$ is the *game value*

Lower bound for v

Assume P_1 plays the optimal distribution γ on partitionings. Then P_2 cannot win more than v points. So, for any subset $U \subset K$:

$$E_{(f \leftarrow \gamma)}\ N(f, U) \leq v$$

Lower bound for v

Take $G' := \sum_f \gamma(f) G_f$, i.e. a convex combination of sparsifiers given by partitionings (like we planned!)

$$E_{(f \leftarrow \gamma)} N(f, U) = \sum_f \frac{\gamma(f) h_f(U)}{h_K(U)} = \frac{h'(U)}{h_K(U)} \leq v \text{ for all } U \subset K$$

$\implies$ quality of G' better than v

Upper bound for v

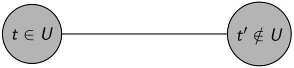
Assume P_2 plays the optimal distribution λ on subsets $U \subset K$. Then P_2 guarantees a win of at least v points. So, for any partitioning f:

$$v \leq E_{(U \leftarrow \lambda)} \; N(f, U)$$

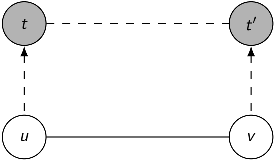
What is $N(f, U)$?

$$N(f, U) := \frac{h_f(U)}{h_K(U)}$$

G_f:



G:



$$h_f(U) = \sum_{\substack{\{u,v\} \in E \\ f(u), f(v) \text{ separated by } U}} c(u, v)$$

Upper bound for v

$$v \le E_{(U \leftarrow \lambda)}\, N(f,U) = \sum_{U \subset K} \lambda(U) \frac{h_f(U)}{h_K(U)}$$

Substitute formula for $h_f(U)$ and get:

$$E_{(U \leftarrow \lambda)}\, N(f,U) = \sum_{\{u,v\} \in E} c(u,v) \underbrace{\sum_{\substack{U \subset K \\ f(u), f(v) \text{ separated by } U}} \frac{\lambda(U)}{h_K(U)}}_{D_\lambda(f(u), f(v))}$$

What is D_λ ?

$$D_\lambda : K^2 \to \mathbb{R}_{\geq 0}, \;\; D_\lambda(t, t') = \sum_{\substack{U \subset K \\ t, t' \text{ separated by } U}} \lambda(U) \cdot \frac{1}{h_K(U)}$$

Fix U: metric! (*non-negative, symmetric, triangle inequality*)

D_λ convex combination of metrics, so metric

What is D_λ ?

We want to find f that makes the following small:

$$v \leq E_{(U \leftarrow \lambda)} \, N(f, U) = \sum_{\{u,v\} \in E} c(u, v) D_\lambda(f(u), f(v))$$

D_λ metric on K^2

0-extension problem

For metric $\delta : K^2 \to \mathbb{R}_{\geq 0}$

Find partitioning $f : V \to K$, s.t.

- $f(t) = t$ for all $t \in K$
- $\displaystyle\sum_{\{u,v\}\in E} c(u,v)\delta(f(u), f(v))$ is minimized

Geometrically: *glue* non-terminal v to $f(v)$.
How to solve this?

0-extension problem

For metric $\delta : K^2 \to \mathbb{R}_{\geq 0}$

Find partitioning $f : V \to K$, s.t.

- $f(t) = t$ for all $t \in K$
- $\displaystyle\sum_{\{u,v\} \in E} c(u,v)\delta(f(u), f(v))$ is minimized

Geometrically: *glue* non-terminal v to $f(v)$.
How to solve this? No idea...

Relaxed 0-extension problem

Generalize:
For metric $\delta : K^2 \to \mathbb{R}_{\geq 0}$

Extend δ to a metric on V^2, s.t.

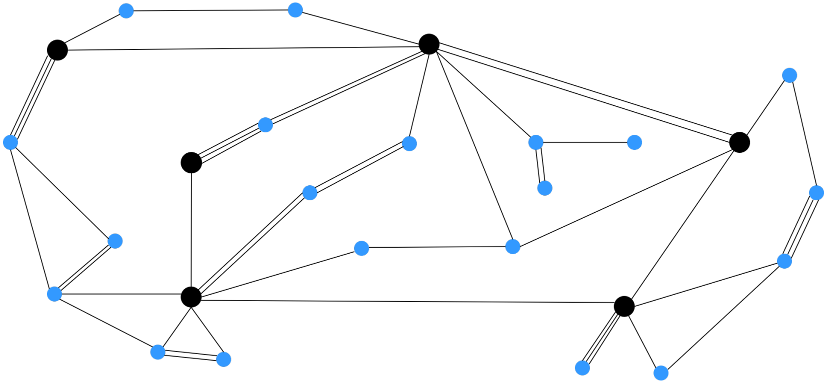
- $\displaystyle\sum_{\{u,v\} \in E} c(u,v)\delta(u,v)$ is minimized

Allowed but not *forced* to glue non-terminals to terminals
Can be solved because LP

Visualizing the relaxed 0-extension problem

Sum of lengths of all edges $= \sum_{\{u,v\} \in E} c(u,v)\delta(u,v)$

Choose positions of blue nodes to minimize this



Use of relaxed problem

Using results of [Fakcharoenphol et al., 2003] relating to the 0-extension problem, we can show:

If there exists a metric β on V, s.t.

$$\sum_{\{u,v\}\in E} c(u,v)\beta(u,v) = \mathcal{O}(1)$$

then there is a partitioning f, s.t.

$$\sum_{\{u,v\}\in E} c(u,v)\beta(f(u),f(v)) \leq \mathcal{O}\left(\frac{\log k}{\log\log k}\right)$$

Our plan

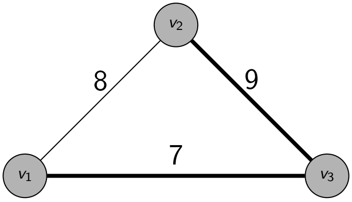
$$D_\lambda : K^2 \to \mathbb{R}_{\geq 0},\ \ D_\lambda(t,t') = \sum_{\substack{U \subset K \\ t,t' \text{ separated by } U}} \frac{\lambda(U)}{h_K(U)}$$

Extend D_λ to some metric β on V^2, s.t. $\sum_{\{u,v\} \in E} c(u,v)\beta(u,v)$ is small.

Then use previous result to find good f.

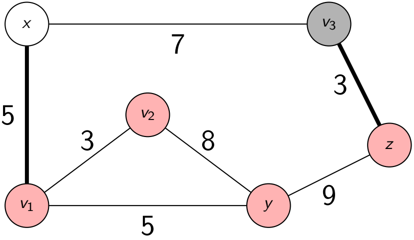
Extending D_λ

D_λ for some $U \subset K$
e.g. $U := \{v_1, v_2\}$



increase distance by $\frac{\lambda(U)}{h_K(U)}$

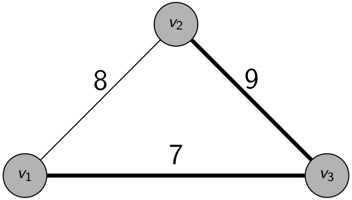
$d : E \to \mathbb{R}_{\geq 0}$ for corresponding min-terminal-cut A of size $h_K(U)$



increase distance by $\frac{\lambda(U)}{h_K(U)}$

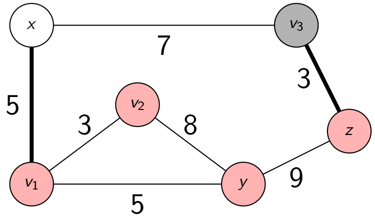
Extending D_λ

D_λ for some $U \subset K$
e.g. $U := \{v_1, v_2\}$



increase distance by $\frac{\lambda(U)}{h_K(U)}$

$d : E \to \mathbb{R}_{\geq 0}$ for corresponding min-terminal-cut A of size $h_K(U)$



increase distance by $\frac{\lambda(U)}{h_K(U)}$

$\sum_{\{u,v\}\in E} c(u,v)d(u,v)$ increases by

$$\sum_{\{u,v\} \text{ crosses } A} c(u,v)\frac{\lambda(U)}{h_K(U)} = h_K(U)\frac{\lambda(U)}{h_K(U)} = \lambda(U)$$

Properties of β

λ distribution $\implies$ at end we have $\sum_{\{u,v\}\in E} c(u,v)d(u,v) = 1$

Take $\beta : V^2 \to \mathbb{R}_{\geq 0}$ the distance function induced by d, i.e. shortest-path-metric.

$$\implies \sum_{\{u,v\}\in E} c(u,v)\beta(u,v) = 1$$

Also not hard to see that $\beta(t,t') \geq D_\lambda(t,t')$ for all $t, t' \in K$.

Remember!

Using results of [Fakcharoenphol et al., 2003] relating to the 0-extension problem, we can show:

If there exists a metric β on V, s.t.

$$\sum_{\{u,v\}\in E} c(u,v)\beta(u,v) = \mathcal{O}(1)$$

then there is a partitioning f, s.t.

$$\sum_{\{u,v\}\in E} c(u,v)\beta(f(u),f(v)) \leq \mathcal{O}\left(\frac{\log k}{\log\log k}\right)$$

Idea: we have a nice response for P_2!

Upper bound for v

Using the partitioning f we found before, we get:

$$v \leq E_{(U \leftarrow \lambda)}\, N(f, U) = \sum_{\{u,v\} \in E} c(u, v) D_\lambda(f(u), f(v))$$

$$\leq \sum_{\{u,v\} \in E} c(u, v) \beta(f(u), f(v))$$

$$\leq \mathcal{O}\left(\frac{\log k}{\log \log k}\right)$$

Main result

Theorem ([Moitra, 2013])

For any $G := (V, E)$ with terminals $K \subset V$, there exists a sparsifier $G' := (K, E')$, s.t.

$$1 \leq \max_{U \subset K} \frac{h'(U)}{h_K(U)} \leq \mathcal{O}\left(\frac{\log k}{\log \log k}\right)$$

where $k := |K|$.

Observations

- Proved existence; construction can be done using LP of game, but extremely expensive
- Factor $\mathcal{O}\left(\frac{\log k}{\log \log k}\right)$ can be improved for some particular cases

Conclusion

- There exist sparsifiers of quality $\mathcal{O}\left(\frac{\log k}{\log \log k}\right)$ for any graph G with terminals K
- Proof uses 0-extension problem and game theory.
- Obtained sparsifier combines different partitionings.
- Application: routing, but... not so clear, because we don't know how to construct a sparsifier efficiently.

References I

Fakcharoenphol, J., Harrelson, C., Rao, S., and Talwar, K. (2003).
An improved approximation algorithm for the 0-extension problem.
Symposium on Discrete Algorithms, pages 257–265.

Karzanov, A. V. (1998).
Minimum 0-extensions of graph metrics.
European Journal of Combinatorics, pages 71–101.

Moitra, A. (2013).
Vertex sparsification and oblivious reductions.
SIAM Journal on Computing 42, no. 6.

Räcke, H. (2002).
Minimizing congestion in general networks.
In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 43–52.