The Traveling Salesman: Classical Tools and Recent Advances

Lecture 5: Dynamic Programming for s-t Path TSP

Anke van Zuylen
XV Summer School in Discrete Mathematics
Valparaiso, January 6-10, 2020

WILLIAM & MARY
CHARTERED 1693

What We Know So Far

Recall: s-t Path TSP

Input:

- A complete undirected graph $G = (V, E)$;
- Start vertex $s \in V$, end vertex $t \in V$;
- Edge costs $c(e) \equiv c(i,j) \geq 0$ for all $e = (i,j) \in E$;
- Edge costs satisfy the **triangle inequality**: $c(i,j) \leq c(i,k) + c(k,j)$ for all i, j, k.

Goal: Find a min-cost path from s to t that visits all other vertices in between.

Recall: LP relaxation

Let $\delta(S)$ be the set of edges with exactly one endpoint in S, and $x(E') \equiv \sum_{e \in E'} x(e)$.

Call $\delta(S)$ an <u>s-t cut</u> if $s \in S, t \notin S$ (or $t \in S, s \notin S$). Call $\delta(S)$ a <u>non s-t cut</u> if $s, t \notin S$ (or $s, t \in S$).

$$
\begin{aligned}
\text{Min} \quad & \sum_{e \in E} c(e)x(e) \\
\text{subject to:} \quad & x(\delta(i)) = \begin{cases} 1, & \forall i = s, t, \\ 2, & \forall i \neq s, t, \end{cases} \\
& x(\delta(S)) \geq \begin{cases} 1, & \forall s\text{-}t \text{ cuts } \delta(S), \\ 2, & \forall \text{non } s\text{-}t \text{ cuts } \delta(S), \end{cases} \\
& 0 \leq x(e) \leq 1, \qquad \forall e \in E.
\end{aligned}
$$

Recall: Minimum Spanning Tree LP

The minimum spanning tree can be found as the solution to the following LP (Edmonds 1971):

$$\text{Minimize} \quad \sum_{e \in E} c(e) z(e)$$

subject to:

$$\begin{aligned} & z(E) = |V| - 1, & \\ & z(E(S)) \leq |S| - 1, & \forall |S| \subseteq V, |S| \geq 2, \\ & z(e) \geq 0, & \forall e \in E, \end{aligned}$$

where $E(S)$ is the set of all edges with both endpoints in S.

If x^* is an optimal solution for the s-t path TSP LP, it is also feasible for the spanning tree LP.

W_T- Matching LP

Let T be a spanning tree. If edge costs are non-negative and satisfy the triangle inequality, a minimum-cost matching of W_T can be found as a solution to the following LP:

$$\text{Minimize} \quad \sum_{e \in E} c(e) z(e)$$

subject to:

$$\begin{aligned} z(\delta(S)) &\geq 1 \quad \forall s\text{-}t \text{ cuts } \delta(S) : |\delta(S) \cap T| \text{ even}, \\ z(\delta(S)) &\geq 1 \quad \forall \text{non } s\text{-}t \text{ cuts } \delta(S) : |\delta(S) \cap T| \text{ odd}, \\ z(e) &\geq 0 \quad \forall e \in E. \end{aligned}$$

Recall: No Even Narrow Cuts – No Problem

Let x^* be an optimal solution for the s-t path TSP LP.

Definition

A cut $\delta(S)$ is narrow if $x^*(\delta(S)) < 2$.

Observation

Let T be a spanning tree, and $W_T = Odd_T \triangle \{s, t\}$.

If $c(T) \leq OPT_{LP}$, and there is no narrow cut $\delta(S)$ for which $|\delta(S) \cap T|$ is even, then adding a minimum-cost W_T-matching to T gives an s-t traveling salesman path of cost at most $\frac{3}{2}OPT_{LP}$.

Reason: Narrow cuts are s-t cuts. The W_T-matching LP has no constraint for such a cut if $|\delta(S) \cap T|$ is odd.

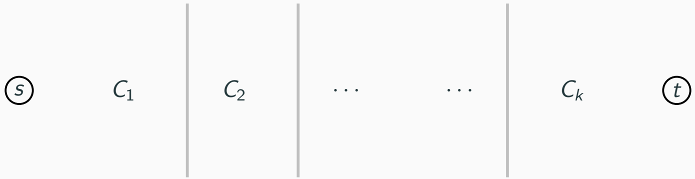
Since $\frac{1}{2}x^*(\delta(S)) \geq 1$ for all other cuts, $\frac{1}{2}x^*$ is feasible for the W_T-matching LP if there are no narrow cuts in which the tree is even.

Recall: Narrow Cuts Are Nested

Theorem (An, Kleinberg, Shmoys (2012))

If $\delta(S_1), \delta(S_2)$ are narrow cuts, $S_1 \neq S_2$, then either $S_1 \subset S_2$ or $S_2 \subset S_1$.

So the narrow cuts look like $s \in S_1 \subset S_2 \subset \cdots \subset S_k \subset V$.



Each narrow cut $\delta(S_i)$ is indicated by a gray line; $S_i = \bigcup_{j=1}^{i} C_j$ is all nodes to the left of the line.

What We Tried So Far

In Lecture 3, we saw that in Graph-TSP instances, there exists a tree T with $c(T) \leq OPT_{LP}$ that has no even narrow cuts.

In Lecture 4, we saw that x^* is a convex combination of spanning trees, which on average have cost OPT_{LP}, and which have "few" even narrow cuts on average.

Today, we will do something different. Instead of improving the algorithm, we improve the lower bound. We will find an LP solution that is only narrow on cuts $\delta(S)$ in which we know the (unique) edge T selects!

Dynamic Programming to "Guess" Lonely Cuts and Lonely Edges

Some Wishful Thinking

Given an optimal LP solution x^*, let $\mathcal{N} = \{\delta(S) : 1 \leq x^*(\delta(S)) < 2\}$.

Suppose we knew for each cut in $\mathcal{N}$, whether optimal s-t traveling salesman path P has

$|P \cap \delta(S)|$ $\begin{cases} = 1 \text{ and the unique edge } e \text{ in } P \cap \delta(S) \quad \text{OR} \\ \end{cases}$ ~~≥ 2~~ ≥ 3

for every $\delta(S) \in \mathcal{N}$.

Call the cuts in $\mathcal{N}$ with $|P \cap \delta(S)| = 1$ the lonely narrow cuts, and let $\mathcal{L}^P$ be the set of lonely narrow cuts.

Call the unique edge e in $P \cap \delta(S)$ for $\delta(S) \in \mathcal{L}^P$ a lonely edge, and let $E_{\mathcal{L}^P}$ be the set of lonely edges.

An Improved Bound on OPT Given $\mathcal{L}^P$ and $E_{\mathcal{L}^P}$

$$
\begin{aligned}
\text{Min} \quad & \sum_{e \in E} c(e) y(e) \\
\text{subject to:} \quad & y(\delta(i)) = \begin{cases} 1, & i = s, t, \\ 2, & i \neq s, t, \end{cases} \\
& y(\delta(S)) \geq \begin{cases} 1, & \forall s\text{-}t \text{ cuts } \delta(S), \\ 2, & \forall \text{non } s\text{-}t \text{ cuts } \delta(S), \end{cases} \\
& y(\delta(S)) = 1, \qquad \forall \delta(S) \in \mathcal{L}^P, \\
& y(\delta(S)) \geq 3, \qquad \forall \delta(S) \in \mathcal{N} \setminus \mathcal{L}^P, \\
& y(e) = 1, \qquad \forall e \in E_{\mathcal{L}^P}, \\
& 0 \leq y(e) \leq 1, \qquad \forall e \in E.
\end{aligned}
$$

Let OPT_{LP*} be the optimal value of the strengthened LP. Then

$$OPT_{LP} \leq OPT_{LP*} \leq OPT.$$

Finding the "Cheapest Choice" for $\mathcal{L}^P$ and $E_{\mathcal{L}^P}$

Consider the following problem *StrengthenLP*($\mathcal{N}$).

$$
\begin{aligned}
\text{Min}_{\substack{\mathcal{L}\subseteq\mathcal{N},E_{\mathcal{L}}\subseteq E:\\ |E_{\mathcal{L}}\cap\delta(S)|=1 \text{ for all } \delta(S)\in\mathcal{L}}} \quad \text{Min}_y \quad & \sum_{e\in E} c(e)y(e) \\
\text{subject to:} \quad & y(\delta(i)) = \begin{cases} 1, & i = s,t, \\ 2, & i \neq s,t, \end{cases} \\
& y(\delta(S)) \geq \begin{cases} 1, & \forall s\text{-}t \text{ cuts } \delta(S), \\ 2, & \forall \text{non } s\text{-}t \text{ cuts } \delta(S), \end{cases} \\
& y(\delta(S)) = 1, \qquad \forall \delta(S) \in \mathcal{L}, \\
& y(\delta(S)) \geq 3, \qquad \forall \delta(S) \in \mathcal{N}\setminus\mathcal{L}, \\
& y(e) = 1, \qquad \forall e \in E_{\mathcal{L}}, \\
& 0 \leq y(e) \leq 1, \qquad \forall e \in E.
\end{aligned}
$$

Let OPT_{LP**} be the optimal value of this problem. Then

$$OPT_{LP} \leq OPT_{LP**} \leq OPT_{LP*} \leq OPT.$$

Solving *StrengthenLP*$(\mathcal{N})$

If we have decided $\delta(S') \in \mathcal{L}$ and (u, v) is the lonely edge in $\delta(S')$ then the LP cannot have any other edges in $\delta(S') \rightarrow$ problem decomposes into two subproblems (subproblem 1 is on S' from $s_1 = s$ to $t_1 = u$ and subproblem 2 is on $V \setminus S'$ from $s_2 = v$ to $t_2 = t$).

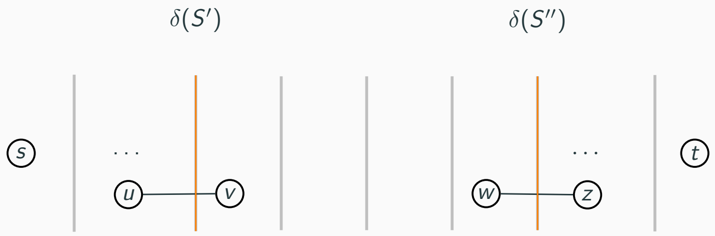
Solving *StrengthenLP*$(\mathcal{N})$

Suppose we have already decided $\delta(S') \in \mathcal{L}$ and (u, v) is the lonely edge in $\delta(S')$. If we now decide $\delta(S'')$ is the next cut in $\mathcal{L}$ after $\delta(S'')$, and that (w, z) is the lonely edge in $\delta(S'')$ then the cost of this decision is:

- $c(w, z)$ plus the cost of a subproblem on $S'' \setminus S'$ from v to w (see next slide),

plus the cost of the subproblem on $V \setminus S''$ from z to t.

Subproblem



$$c(w,z) + \text{Min} \quad \sum_{e \in E} c(e)y(e)$$

subject to:

$$y(\delta(i)) = \begin{cases} 1, & i = v, w \\ 2, & i \in S'' \setminus S', i \neq v, w, \end{cases}$$

$$y(\delta(S)) \geq \begin{cases} 1, & \forall \text{ } v\text{-}w\text{-cuts } \delta(S), S \subset S'' \setminus S' \\ 2, & \forall \text{ non } v\text{-}w\text{-cuts } \delta(S), S \subset S'' \setminus S' \end{cases}$$

$$y(\delta(S)) \geq 3, \qquad \text{for } \delta(S) \in \mathcal{N}, S' \subset S \subset S'',$$

$$0 \leq y(e) \leq 1, \qquad \forall e \in E.$$

Solving $StrengthenLP(\mathcal{N})$ as a Shortest Path Problem

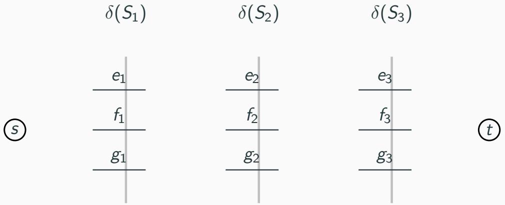
The previous observations can give a dynamic programming formulation to solve $StrengthenLP(\mathcal{N})$. We can also view the dynamic program as a shortest path instance:

- a vertex (S', u, v) for every $\delta(S') \in \mathcal{N}, (u, v) \in \delta(S')$;
- a start vertex $(\emptyset, \emptyset, s)$ and target vertex $(V, t, \emptyset)$;
- arc from (S', u, v) to (S'', w, z) if $S' \subset S''$;
  cost of the arc is the solution to the subproblem from preceding slide.

Interpretation: If the shortest path from the start vertex to the target vertex visits vertex (S', u, v), then $\delta(S') \in \mathcal{L}$ and (u, v) is the lonely edge in $\delta(S')$.

Example: Solving *StrengthenLP*($\mathcal{N}$) as a Shortest Path Problem

Narrow cuts, and their edges in G:



Shortest path instance (all arcs from left to right exist, cost is given by subproblem as in previous slide):

Solving *StrengthenLP*$(\mathcal{N})$ as a Shortest Path Problem

Shortest path instance:

- a vertex (S', u, v) for every $\delta(S') \in \mathcal{N}, (u, v) \in \delta(S')$;
- a start vertex $(\emptyset, \emptyset, s)$ and target vertex $(V, (t, \emptyset))$;
- arc from (S', u, v) to (S'', w, z) if $S' \subset S''$;
  cost of the arc is $c(w, z)$ plus the cost of an optimal v-w path TSP LP solution from defined on $S'' \setminus S'$ with added $y(\delta(S)) \geq 3$ constraints for "intermediate" narrow cuts.

Theorem

Any path from $(\emptyset, \emptyset, s)$ to $(V, t, \emptyset)$ corresponds to a solution to StrengthenLP$(\mathcal{N})$ of cost equal to the cost of the path, and vice versa.

$\mathit{StrengthenLP}(\mathcal{N})$ **Can Be Solved in Polynomial Time**

Let $|V| = n, |E| = m$. Because the narrow cuts are nested, there are at most $n - 1$ of them. So the shortest path instance has $O(nm)$ nodes and $O(n^2 m^2)$ edges.

Thus, we can construct the shortest path instance in polynomial time, and we can find the shortest path in polynomial time as well.

Approximating s-t Path TSP Using $\mathit{StrengthenLP}(\mathcal{N})$

Bad News

For the new LP solution y, there may be $\delta(S) \notin \mathcal{N}$ with $y(\delta(S)) < 2$.

How to deal with new narrow cuts??

Traug and Vygen: Recurse!

Traub and Vygen propose a recursive method, that repeats the process we described for the narrow cuts for y, getting a new LP solution y', repeating the process for the narrow cuts for y', etc. They show how to combine these LP solutions to prove the following result.

Theorem (Traub, Vygen (2018))

For any fixed ϵ, there exists a polynomial-time algorithm that returns a solution of cost at most $(\frac{3}{2} + \epsilon)OPT$.

Zenklusen: Extend the Definition of Narrow Cuts!

Given an LP solution x^*, let

$$\tilde{\mathcal{N}} = \{\delta(S) : s \in S, t \notin S, 1 \leq x^*(\delta(S)) < 3\}.$$

Determine[1] $\mathcal{L}^* \subseteq \tilde{\mathcal{N}}$, $E^*_{\mathcal{L}^*} \subseteq E$, and an LP solution y^* minimizing $StrengthenLP(\tilde{\mathcal{N}})$.

[1] The narrow cuts in $\tilde{\mathcal{N}}$ are not nested, but the shortest path reduction still works "as is". It follows from a result of Karger (1993) that there are at most $|V|^4$ cuts in $\tilde{\mathcal{N}}$, and that they can be found efficiently, so the shortest path instance can be constructed and solved in polynomial-time.

Easy Narrow Cuts

Let $z = \frac{1}{2}x^* + \frac{1}{2}y^*$. Then z is a feasible solution to the s-t path TSP LP, and

$$\sum_{e \in E} c(e)z(e) \leq OPT.$$

Claim

The only cuts $\delta(S)$ for which $z(\delta(S)) < 2$ are the cuts in $\mathcal{L}^*$.

Proof:

Zenklusen's Algorithm

- Compute x^*. Let

$$\tilde{\mathcal{N}} = \{\delta(S) : s \in S, t \notin S, 1 \leq x^*(\delta(S)) < 3\}.$$

- Determine $\mathcal{L}^* \subseteq \tilde{\mathcal{N}}$, $E^*_{\mathcal{L}^*} \subseteq E$, and an LP solution y^* minimizing $\mathit{StrengthenLP}(\tilde{\mathcal{N}})$.
- Let T be a minimum-cost tree such that for the cuts in $\mathcal{L}^*$, the tree contains the edge in $E^*_{\mathcal{L}^*}$ (and no other edges in the cuts in $\mathcal{L}^*$).
- Let $\mathrm{W}_T = \mathrm{Odd}_T \triangle \{s, t\}$. Compute a minimum-cost W_T-matching M.
- Shortcut an Eulerian path in $(V, T \sqcup M)$.

Theorem (Zenklusen (2018))

This algorithm returns an s-t traveling salesman path of cost at most $\frac{3}{2}OPT$.

Analysis

The theorem follows from:

Lemma

$$c(T) \leq \sum_{e \in E} c(e) y^*(e) \leq OPT.$$

Lemma

$$c(M) \leq \frac{1}{2} \sum_{e \in E} c(e) z(e) \leq \frac{1}{2} OPT,$$

where $z = \frac{1}{2}(x^* + y^*)$.

Summary and Open Questions

The Bigger Picture

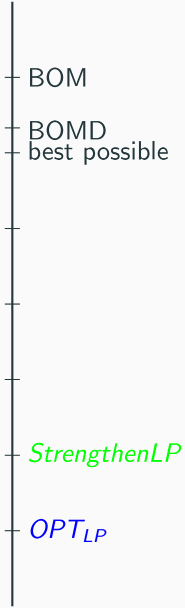
Goals:

- improving the constant in the approximation guarantee?
- improving our understanding of fundamental concepts (e.g. the integrality gap of the LP relaxation)?
- develop powerful new techniques?

Summary: Approximating the s-t Path TSP

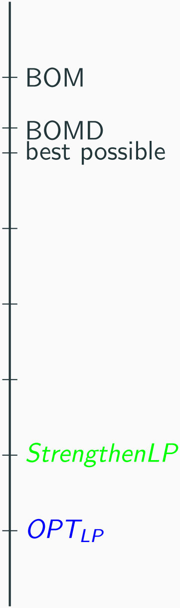
In lectures 3-5, we saw that

- For Graph-TSP instances, there is an algorithm that finds an s-t traveling salesman path of cost at most $1.5 OPT_{LP}$. This settles the integrality gap for these instances.
- For arbitrary metrics, the Best-of-Many-with-Deletion algorithm finds an s-t traveling salesman path of cost $< 1.53 OPT_{LP}$.
- For arbitrary metrics, Zenklusen's analysis implies the Best-of-Many algorithm finds an s-t traveling salesman path of cost at most $1.5 OPT$.



Open Questions: Approximating the s-t Path TSP

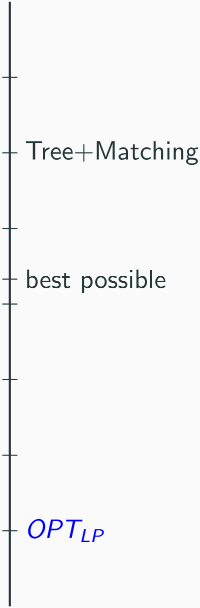
- Is the integrality gap for the LP equal to 1.5? there an algorithm that finds an s-t traveling salesman path of cost at most $1.5OPT_{LP}$?
- Does the Best-of-Many algorithm find an s-t traveling salesman path of cost at most $1.5OPT_{LP}$??? Or is there an example where it does not?
- Is there an algorithm that finds an s-t traveling salesman path of cost strictly less than $1.5OPT$?
  - Traub and Vygen (2018) and Traub, Vygen and Zenklusen (2019) show the answer is yes for Graph-TSP instances.
  - Traub, Vygen and Zenklusen (2019) also show that if the answer is yes for arbitrary metrics, then there also is an algorithm that finds a traveling salesman tour of cost strictly less than $1.5OPT$.



Summary: Approximating the TSP

In lectures 1-2, we saw that

- For Graph-TSP instances, there is an algorithm that finds a traveling salesman tour of cost at most $\frac{4}{3}|V|$ if the graph is 2-edge-connected and subcubic. This settles the integrality gap for these instances.
  If the graph is arbitrary, there is an algorithm that finds a tour of cost at most $1.4 OPT_{LP}$.
- For arbitrary metrics, the best known is the classical Christofides-Serdyukov Tree+Matching algorithm that finds traveling salesman tour of cost at most $1.5 OPT_{LP}$.
- Main open question (that has defied progress many years now!): devise an algorithm that finds a tour of cost less than $1.5 OPT_{LP}$.



One Possible Way Forward

Conjecture (Schalekamp, Williamson, vZ (2014))

The worst case integrality gap instances for the Subtour LP have optimal LP solutions in which $x^*(e) \in \{0, \frac{1}{2}, 1\}$ *for all* $e \in E$*, and the edges with* $x^*(e) = \frac{1}{2}$ *form vertex-disjoint cycles of odd length.*

Given such an LP solution, can you find a tour of cost less than $1.5 \sum_{e \in E} c(e)x^*(e)$?

Very recent result:

Theorem (Karlin, Klein, Oveis Gharan (2019))

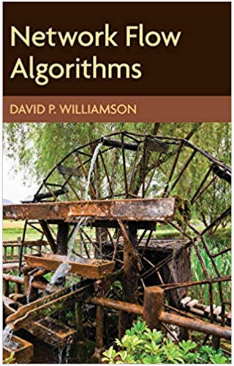
There exists a 1.49993-approximation algorithm for TSP on instances for which the optimal Subtour LP solution is half-integral.

That's All Folks!

Acknowledgements

Slides for the first three lectures were adapted from slides by David Williamson.

David also authored the book that will be given to the best student in this course:



Many thanks to the organizers + students for making this summer school a great experience!