Better than 1.5-apx for TSP (Metric)

Note Title 11/9/2020

History: Christofides / Serdyukov 76? 1.5

(Improvements for case when metric = unweighted graph distances) best 1.4 Sebo-Vygen

Karlin Klein OveisGharan '19 $(1.5-\varepsilon)$ for half integral LP soln case

'20 $(1.5-\varepsilon)$ for general case!

Setting: $G = (V, E)$ $\binom{V}{2}$ edge costs c_e from metric $c_{ij} \le c_{ik} + c_{kj}$ $\forall i,j,k$

- Find tour visits each vtx exactly once

- Find an Eulerian (multigraph) that spans each vertex at least once (sufficient).

(Chr-Serd)

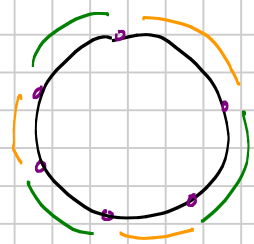
① Find MST T^*. $c(T^*) \le c(OPT)$

② $O \leftarrow$ odd degree nodes in T^*.

$M^* \leftarrow$ min cost perfect matching on O $c(M^*) \le c(OPT)/2$

$(T^* \cup M^*)$ Eulerian submultigraph.

$c(Alg) \le 1.5\ c(OPT)$



Claim: $c(Alg) \le 1.5\ c(\text{LP relaxation}) \le 1.5\ c(OPT)$

What relaxation? "Subtour Elimination LP"

Sometimes called Held-Karp LP but maybe knew before H-K?

$$\boxed{\begin{aligned} \min\ & \textstyle\sum c_e x_e \\ \text{st}\ & x(\partial v) = 2 \quad \forall v \in V \\ & x(\partial S) \ge 2 \quad \forall S \subseteq V, (S \ne \emptyset, V) \\ & x \ge 0 \end{aligned}}$$

guess edge e_0 in OPT

set $x^*(e_0) = 1$.

Easy: LP value $\le$ OPT.

Just use the OPT solution as a 0-1 solution.

Thm [Wolsey]: $c(A_G) \le 1.5$ LP value. (x^* = OPT LP sol^n $\Rightarrow$ LP value = $c^T x^*$)

$$\boxed{\begin{aligned} &\min \; c^T x \\ &\text{st} \;\; x(\text{Edges within } S) \le |S|-1 \quad \forall S \subseteq V \\ &x(E) = |V|-1 \\ &x \ge 0 \end{aligned}}$$

Pf: ① x^* (after dropping edge e_0) is feasible to spanning tree polytope.

- $x^*(\text{Edges within } S) = \dfrac{\sum_{v\in S} x^*(\partial v) - \sum x^*(\partial S)}{2} \le \dfrac{2|S|-2}{2} = |S|-1$

- $x^*(\text{non } e_0 \text{ edges}) = \dfrac{2|V|-2}{2}$

every edge counted twice, so divide by 2.

$\Rightarrow$ x^* soln to Sp. tree polytope. $\Rightarrow$ $c(MST) \le c^T x^* = $ LP value

② $x^*/2$ feasible to the "O-join" problem — min cost set of edges with odd degree on O, even degree elsewhere

$\Rightarrow$ convert to perfect matching of no extra cost using metric costs

O-join LP

$$\boxed{\begin{aligned} &\min \; \sum c_e x_e \\ &\text{st} \;\; x(\partial S) \ge 1 \quad \forall S \text{ st } |S \cap O| \text{ odd.} \\ &x \ge 0 \end{aligned}}$$

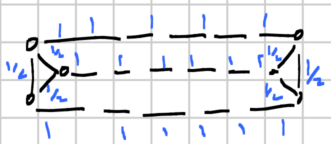
$\frac{x^*}{2}$ is feasible for this LP. (regardless of O)

$\Rightarrow$ cost of $M^* \le c^T x^*/2 = \frac{1}{2}$ LP value

$\Rightarrow$ $c(A_G) \le \frac{3}{2}$ LP value.

Q: do better? Ideally: $c(A_G) \le 1.1$ LP value. (or even better?)

Not possible with this LP.



LP value $= 3\ell + 2$

Any integer soln $\ge 4\ell$

$c(A_G) \ge (\frac{4}{3} - \varepsilon)$ LP value as $\ell \to \infty$

Integrality gap.

Conjecture: is worst gap.

i.e. $\forall$ metric instance, $\exists$ integer sol^n with cost $\le \frac{4}{3} \cdot$ LP sol^n.

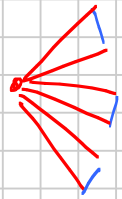
Big open problem

OK: how to get even ε better than $3/2$? (Will try for $4/3$ another day ☺)

Algorithmic Gap: Say complete graph, all edge lengths $= 1$.

$OPT = n$.

LP solution: set $x_e = \frac{2}{n-1}$ $\forall$ edges. $\Rightarrow$ $c^T x = \binom{n}{2} \frac{2}{n-1} = n$. ☺

But algo: Sps. $T^* =$  star.

$\Rightarrow$ all leaves have odd degree $\Rightarrow$ Matching has cost $\frac{n-1}{2}$

$\Rightarrow$ overall cost $(n-1) + \frac{n-1}{2}$ $= \frac{3}{2}(n-1)$ ☹

Got to be smarter on even trivial instances!!!

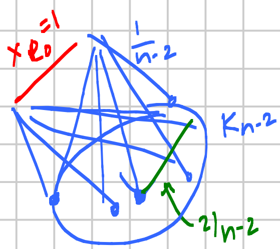
> **Idea:** [Oveis Gharan Saberi Singh 11]
>
> Pick a tree at random.
>
> Hope it has a cheap matching!

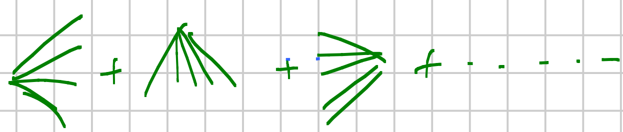
What kind of Randomness?

Solution x^* (minus edge e_0) is in the spanning tree polytope. ← its vertices are spanning trees.

$\Rightarrow$ can write $x^* = \sum_T p_T \chi_T$ ← indicator vectors of spanning trees

↑ convex combo (probability distribution)



essentially breaks into 

So maybe same problem again!!

Need trees to stop having a lot of odd degree vertices

Need them to be more disordered $\Rightarrow$ have a lot of Entropy!!

Note: $E[\text{degree of most nodes}] = 2$ **Problem** is: these trees give them degree either 1 or $(n-1)$.

One Solution: Max Entropy distribution

Given x^* (without e_0) $\in$ Spanning tree polytope, can write $x^* = \sum p_T \chi_T$

vertices of polytope

probability distrib

How to choose $p_1, p_2, \dots, p_N$?

$N = $ # of spanning trees?

Find distribution that maximizes entropy !!

$$\max H(p) = \sum_T p_T \log 1/p_T$$

$$\text{st} \quad \sum p_T \chi_T = x^*$$

p is probability distribution over spanning trees.

← concave function, so maximizing is tractable upto additive eps.

← exponential # of vars, but poly constraints ⇒ go to dual and solve that

Fact: can get "weights" λ_e $\forall e \in E(G)$ s.t. $p_T = \dfrac{\prod_{e \in T} \lambda_e}{Z}$ $\forall$ spanning tree T of G

= "partition function"

= normalizer $= \sum_{T' \text{ sp tree}} \left(\prod_{e \in T'} \lambda_e \right)$

and sampling from this distribution (almost) can be done in poly time

Feel the picking unif ⇒ random tree of G

For now: assume can sample from MaxEnt distribution above.

Great. But is it any good for TSP?

What about our starting example (the complete graph)?

lets forget about the minor effects of e_0, detail

There, essentially picking a uniformly random spanning tree of K_n.

(Recall: $\exists\ n^{n-2}$ spanning trees of K_n, by Cayley's formula, pick one of them).

(labeled)

Then cost $= (n-1)$.

And matching cost $= \frac{1}{2}$ (# odd degree vertices of T)

[Thm: # of odd degree vertices of random spanning tree of K_n $= 0.57n$

$\Rightarrow \mathbb{E}[\text{cost of Christofides}] = (1 + \frac{0.57}{2})\, n = 1.285n$]

for this instance!!

How do you prove this? And how does it extend to more general graphs?

Magic of Random Spanning Trees

Theorem: For a graph $G = (V, E)$, let $F \subseteq E$ be a subset of edges.

Let T be a random spanning tree of G. (also holds true for trees picked from the MaxEnt distribution)

then #(edges in $T \cap F$) behaves like a sum of independent Bernoulli's.

Corollary: Fix any vertex v, degree of v in T is some sum of Bernoullis. $Z_1 + Z_2 + \cdots + Z_{n-1}$

So cannot be 1 with high prob, and $(n-1)$ with rest.

this distribution cannot be achieved by a sum of Bernoullis!

What do we know about the biases of these variables? Say $p_1 \geq p_2 \geq \cdots \geq p_{n-1}$

(a) $E[Z_1 + \cdots + Z_{n-1}] = 2$.

(b) $\Pr[Z_1 + Z_2 \cdots + Z_{n-1} \geq 1] = 1$ since we pick a (connected) spanning tree T.

$\Rightarrow$ must have $p_1 = 1$ $\Rightarrow$ $p_2 + \cdots + p_{n-1} = 2 - p_1 = 1$.

What about the rest? No idea.

But sps we want to upper bound the probability that degree is odd.

$$\Pr[\text{degree of } v \text{ is odd}] = \Pr[Z_1 + \cdots + Z_{n-1} = \text{odd}]$$

$$= \Pr[Z_2 + \cdots + Z_{n-1} = \text{even}] \quad \text{where } E[Z_2 + \cdots + Z_{n-1}] = 1.$$

Fact [Hoeffding 1956] for any indep Bernoullis $Z_1\, Z_2 \ldots Z_m$ and any function $g: \mathbb{R} \to \mathbb{R}$, the value of $E[g(Z_1 + Z_2 + \cdots + Z_m)]$ is maximized when the biases $p_1, p_2 \ldots p_m$ all take on values in $\{0, p, 1\}$ for some p.

!

So $\Pr(\sum Z_i = \text{even})$ is max when $p_2 = 1$ $p_3 = p_4 \cdots = 0$. ← $\Pr(\,) = 0$ since always odd

or $p_2 = p_3 = 1/2$, $p_4 = p_5 = \cdots = 0$ etc. $\Pr(\,) = 1/2$

$P_2 = P_3 = P_4 = 1/3$ . .

Turns out it is when we set $P_2 = P_3 = P_4 \cdots = P_{n-1} = 1/n-2$

$\Pr(\ ) = \frac{8}{27} + 3 \cdot \frac{2}{3} = \frac{14}{27}$

⋮

Now: $\Pr\left(\sum Z_i = \text{even}\right) \leq \Pr\left(\text{Poi}(\lambda = 1) = \text{even}\right) = \sum_{k \text{ even}} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-\lambda}\left(\frac{e^{\lambda} + e^{-\lambda}}{2}\right)$

$= \frac{1 + e^{-2\lambda}}{2}$

As $n \to \infty$, Behaves like a Poisson variable with mean 1.

but $\lambda = 1 \Rightarrow \frac{1 + e^{-2}}{2} = 0.57$.

So: $\mathbb{E}[\text{odd degree vertices}] \cong 0.57n$

$\Rightarrow$ for complete graph, min cost matching cost $= \frac{0.57}{2} n = 0.285n$.

Much better than Christofides-Serdyukov!

General Graphs? Analysis complicated by many factors. Here is main idea of the proof

1. Show a fractional solution to the matching polytope for the odd degree vertices.

$z_e = \frac{x_e^*}{2}$ is a solution for <u>any</u> spanning tree

but gives cost $\frac{\text{tree cost} + \text{LP cost}}{2} \leq 1.5 \text{ OPT}$

not good.

So get a better solution.

2. Suppose all non-trivial cuts (i.e. cuts that are not "degree" cuts) all are $\geq (2+\varepsilon)$ in value.

then define $z_e = \frac{x_e^*}{(2+\varepsilon)}$ as a first candidate for the matching solution

Note: $z(\delta S) = \sum_{e \in \partial S} x_e^* / (2+\varepsilon) \geq 1$ if S is a nontrivial cut.

So need to "fix" this for the degree cuts.

Attempt #1: if $\underset{T}{\text{degree}}(v)$ is odd (happens w.p. at most 57%, remember)

then charge $z_e = \frac{x_e^*}{2}$ $\forall e \in \partial v$.

Recall — the only tricky sets are singleton sets, so this suffices.

Problem: edge $\overset{(u,v)}{e}$ may be raised by either u or v

So if u and v are odd in opposite parts of sample space (if they are anti-correlated)

then e may be raised w.p. 1.

☹

Attempt #2: show that $\forall$ edges e, it is "good" (i.e. both endpoints are even)

w.p. some constant >0.

① [Oveis Gharan Saberi Singh 2011] Showed for all edges that have $x_e \notin [0.49, 0.51]$

② [Karlin Klein Oveis Gharan 2021] Showed for most edges even if all $x_e \in \{0, \frac{1}{2}, 1\}$

Main idea: the endpoints u,v cannot be perfectly anticorrelated

Use more "negative correlation" properties of random spanning trees

[See Shayan's notes]

Other pieces:

- Remove requirement that all non-trivial cuts are $\geq 2+\varepsilon$.
  - use "cactus" hierarchy of near-min cuts

- not all edges are "good", so charging argument.

  Lower some edges because they are good. Now other cuts are not happy, so raise them.

Many more ideas — see talks by Nathan, Shayan (links on webpage).