Applications to Other Proximity Problems

Euclidean min spanning tree (EMST) Exact alg's: # edges

- Kruskal $O(n \log n)$
- Prim $O(n \log n + mn)$
- Karger-Klein-Tarjan $O(n)$ ran
- Chazelle '97 $O(n \alpha(n))$ det.

$d = 2, O(n \log n)$
$d = 3, \tilde{O}(n^{4/3})$
$d \geq 3, \tilde{O}(n^{2 - 2/(5d/2 + 1)})$

Idea: find sparse subgraph containing EMST

Def: Given n pts $P$ in $\mathbb{R}^d$
a c-spanner is a subgraph $H$ of complete graph
st: $\forall p, q \in P, d(p, q) \leq d_H(p, q) \leq c \cdot d(p, q)$

EMST alg'n:
- Compute c-spanner $H$ with $O(n)$ edges
- Compute MST $T$ of $H$ by Prim
  $\Rightarrow O(n \log n)$ time

Analysis:
- Let $T^*$ be the EMST.
- Replace each edge $pq \in T^*$ w. shortest path from $p \to q$
  in $H$
  $\Rightarrow w(T) \leq w(T^*) \leq c \cdot w(T^*) \Rightarrow$ factor $\leq c$
Spanner Method 1: Yao graph (’82)

Idea: rounding dir

for each $p \in P$

form $O\left(\frac{1}{\sin \theta}\right)$ disks/cones at $p$ of angle $\theta$

for each cone

add edge from $p$ to nearest neighbor in cone.

$\Rightarrow O\left(\frac{1}{\theta \sin \theta}\right)$ edges

Analysis:

Fix $p, q \in P$.

Let $s$ be nearest neighbor of $p$ in cone containing $q$.

Know $d(p, s) \leq d(p, q)$

$d(s, q) < d(p, q)$ assuming $\theta < \frac{\pi}{3}$

$$d_H(p,q) \leq d(p,s) + d_H(s,q)$$

by induction on edge length

$$\leq d(p,s) + c d(s,q)$$

$$\leq (1 + 2c \sin \frac{\theta}{2}) d(p,s') + c d(s,q)$$

$$\leq c d(p,s') + c d(s,q)$$

$= cd(p,q)$ by setting $c = \frac{1}{1 - 2c \sin \frac{\theta}{2}}$

Runtime?

$O\left(\frac{1}{\theta \sin \theta} \log n\right)$ with data structures

for range searching (and some modification...).
Def (Callahan-Kosaraju '92)

Given n pts P in Rd,
an ε-well-separated pair decomposition (WSPD)
is a collection of pairs (S₁, T₁), ..., (Sᵱ, Tᵱ) s.t:

a) ∀ p, q ∈ P (p≠q),
   ∃ unique i s.t. p ∈ Sᵱ, q ∈ Tᵱ or vice versa

b) ∀ i, Sᵱ and Tᵱ are well-separated
   i.e. ∃ 2 balls enclosing Sᵱ and Tᵱ of radius r
      that are of dist ≥ \frac{r}{ε}

Thm: \exists O(ε²) WSPD with m = O(\frac{1}{ε^d} n) pairs.

Pf: idea - quadtree (compressed, but unbalanced)
WSPD(A, B): // given quadtree cells A, B of same side length
if |PN A| ≤ ε & |PN B| ≤ ε, done
shrink A, B into quadtree cells Â, ̂B of same side length r
(with PN Â = PN A, PN ̂B = PN B)
if Â, ̂B have dist ≥ r/ε,
output pair (PN Â, PN ̂B) & return
divide Â into 2^d subcells \{A_i\}
\hat{B}\vdots\vdots\vdots\{B_j\}
for i = 1, ..., 2^d
for j = 1, ..., 2^d
WSPD(A_i, B_j)

Analysis:
- note that Â or ̂B corresponds to a node in compressed quadtree
- fix a node Â of compressed quadtree, say Â has side length r
- # quadtree cells ̂B of side length r at dist ≤ r/ε = O\left(\frac{1}{\epsilon^d}\right)

⇒ # recursive calls = O\left(\frac{1}{\epsilon^d} n\right)

runtime O( n log n + \frac{1}{\epsilon^d} n)

time to build compressed quadtree
\(T(n) = n \left(\log n + \frac{1}{\epsilon^d}\right)\)

Rmk: other trade-offs (max deg, hop diameter, weight, ...)
Rmk: for EMST, Arya, Mount '16 \(O(n \log n + \frac{1}{\epsilon^d})\)
Obs. Let \((S_i, T_i), \ldots, (S_m, T_m)\) be an \(\varepsilon\)-WSPD.

Pick an arbitrary \(s_i \in S_i, t_i \in T_i\).

Then \(\{s_i, t_i, \ldots, s_m, t_m\}\) is a \((1 + O(\varepsilon))\)-spanner.

**Proof.** Fix \(p, q \in P\).

Say \(p \in S_i, q \in T_i\).

\[
d_H(p, q) \leq d_H(p, s_i) + d(s_i, t_i) + d_H(t_i, q)
\]

\[
\leq c \cdot d(p, s_i) + d(s_i, t_i) + c \cdot d(t_i, q)
\]

by induction

\[
\leq (c+1) d(p, s_i) + d(p, q) + (c+1) d(t_i, q)
\]

\[
\leq (1 + 2(c+1) \varepsilon) d(p, q)
\]

\[
\leq c \cdot d(p, q) \text{ by setting } c = \frac{1 + 2\varepsilon}{1 - 2\varepsilon}
\]

\[
= 1 + O(\varepsilon)
\]

\(\square\)