Well-Separated Pair Decomposition for the Unit-disk Graph Metric and its Applications

Jie Gao
Department of Computer Science
Stanford University

Joint work with
Li Zhang
Systems Research Center
Hewlett-Packard Labs
Geometric well-separated pair decomposition

Point sets $A, B$ are $c$-well-separated if $d(A, B) \geq c \cdot \max(D(A), D(B))$.

- Diameter $D(A) = \max_{a, b \in A} d(a, b)$.
- Distance $d(A, B) = \min_{a \in A, b \in B} d(a, b)$.

For a set $S$ of points, a set of pairs $\mathcal{P} = \{(A_i, B_i)\}$, is a $c$-well-separated pair decomposition (WSPD) of $S$ if

- $(A_i, B_i)$ is $c$-well-separated.
- For any two points $a \neq b \in S$, there exists $i$ such that $a \in A_i$, $b \in B_i$. 

Geometric well-separated pair decomposition

Callahan and Kosaraju [1992] showed that a $c$-WSPD of $O(n)$ pairs for $n$ points in $\mathbb{R}^d$ can be computed in $O(n \log n)$ time.

Observation: if $(A, B)$ is $c$-well-separated, then for any $a_1, a_2 \in A, b_1, b_2 \in B$, $d(a_1, b_1) \leq (1 + 2/c)d(a_2, b_2)$.

WSPD approximates the $\Theta(n^2)$ pair-wise distances by $O(n)$ pairs.

$\Rightarrow$ Many problems involving all-pairs distances can be approximated by WSPD in almost linear time.

Apps: N-body problem, $k$-nearest neighbor, geometric spanners, etc.
Extend WSPD to other metrics?

Any metric admits WSPD of size $O(n^2)$: just take all the pairs of points.

Fact: general graph metric may not admit subquadratic WSPD. :-(

Example: $c$-WSPD of a star graph has size $\Omega(n^2)$, for $c > 1$.

In this talk, we show that for $n$-points unit-disk graph metric, a WSPD of size $O(n \log n)$ can be found in $O(n \log n)$ time.
Overview

• Unit-disk graph (UDG) metric.
  – UDG is used to model wireless communication networks.
• Well-separated pair decomposition for unit-disk graph metric:
  – Construction.
  – Bound of the size of the WSPD.
• Applications.
• Extensions to higher dimensions, unweighted unit-disk graph.
• Conclusion and future work.
Unit-disk graph metric

$I(S) = (S, E)$ is a weighted graph where an edge $e = (p, q) \in E$ if $d(p, q) \leq 1$. The weight of $e$ is the Euclidean distance $d(p, q)$.

Unit-disk graph metric $(S, \pi)$:

- $\pi(s_1, s_2)$ is the length of the shortest path in $I(S)$.
- Diameter $D(S')$ for $S' \subseteq S$ is $\max_{s_1, s_2 \in S'} \pi(s_1, s_2)$.
- Distance $\pi(S_1, S_2)$ for $S_1 \subseteq S$, $S_2 \subseteq S$ is $\min_{s_1 \in S_1, s_2 \in S_2} \pi(s_1, s_2)$. 
WSPD for unit-disk graph

Theorem 1 For unit-disk graph metric, a \( c \)-WSPD of size \( O(c^4 n \log n) \) can be computed in time \( O(c^4 n \log n) \).

Proof. Define the density of the points to be the maximum number of points covered by any unit disk.

1. Nodes with constant bounded density.
   - Apply packing argument, which is similar with the collision detection algorithm of a necklace[Guibas02].

2. Nodes with arbitrary density.
   - Find a minimal cover (or clustering) with radius \( O(1/c) \), cluster-heads have bounded density.
WSPD construction

1. Construct a bounded degree spanning tree $T$.
2. Recursively decompose $T$ balancedly by removing an edge once at a time.
3. Queue $Q \leftarrow \{(S, S')\}$
   
   While $\{Q \neq \emptyset\}$ do
   
   Take $(A, B)$ from $Q$, $a, b$ are arbitrary points from $A, B$;
   
   if $d(a, b) \geq (c + 2) \cdot \max(|A| - 1, |B| - 1)$
   
   then output $(A, B)$ to $P$;
   
   else
   
   if $|A| = |B| = 1$, discard it, since they contain the same point.
   
   take the children of the larger one, say $A_1, A_2$ of $A$,
   
   put $(A_1, B), (A_2, B)$ into $Q$. 
Example

T

decomposition tree T’

({123}, {45678})

Recurse

({123}, {456})

({123}, {78})

OK! output

9
Bound the size of WSPD

Lemma 2  \( P \) is a \( c \)-WSPD of \( S \). Furthermore, each ordered pair of distinct points \((p, q)\) is covered by exactly one pair in \( P \).

Proof.  \( d(a, b) \leq \pi(a, b); D(A) \leq |A| - 1. \)

Lemma 3  Each pair \((A, B)\) that ever appears in the queue satisfies \( 1/\beta \leq |A|/|B| \leq \beta, \beta \) is a constant.

Proof.  By induction.

Lemma 4  If \((A, B_i) \in P, i = 1, \cdots m(A), \) then \( B_i \cap B_j = \emptyset, \) and \( m(A) = O(c^2|A|). \)

Proof.  Constant bounded density and packing argument.
Bound the size of WSPD, cont.

Lemma 5 \(|P| = O(c^2n \log n)|.

Proof.

- The number of subsets with size in \([2^i, 2^{i+1}]\) is \(O(n/2^i)\).
- They generate at most \(O(c^2 2^i) \cdot O(n/2^i) = O(c^2 n)\) pairs.
- We sum up over \(1 \leq i \leq \log n\).

Also, running time: \(O(c^2 n \log n)\).

We are all set with constant density case!
**Arbitrary density**

1. Find minimal cover, the clusterheads have bounded density $O(1/c^2)$.
2. Construct the WSPD on clusterheads.
3. Include for each clusterhead the other points it covers.

- For a far-away pair, including the other points still makes a well-separated pair.
- For a nearby pair, the distance between any points is smaller than 1, we apply the geometric WSPD.

(i) Far-away pairs  
(ii) Nearby pairs: geometric WSPD.
Estimate the distances

To use the WSPD, we need to compute $O(n \log n)$ pairs of distances.

- Preprocessing and distance queries.

1. Approximate distance query $\Rightarrow 2.42$-approx. distance for any 2 points.

- Find a subgraph $G$, $G$ is a planar $\frac{4\sqrt{3}}{9}\pi \approx 2.42$-spanner.
- Apply the distance oracle of [Thorup01] on the planar graph $G$.
  - Preprocessing time $O(n \log^3 n)$.
  - Each distance query takes $O(1)$ time.
- Running time: $O(n \log^3 n)$. 
Estimate the distances, cont.

2. Exact distance query $\Rightarrow (1 + \varepsilon)$-approximate distance for any 2 points.

- Apply [Eppstein et.al.93] algorithm to find balanced separators.
- Extend the shortest distance algorithm by [Arikati et.al.96].
  - Preprocessing time $O(n^2/(\varepsilon r))$.
  - Each distance query takes $O(r/\varepsilon)$ time, $1 \leq r \leq \sqrt{n}$.
- Running time: $O(n\sqrt{n \log n}/\varepsilon^3)$. 
Applications

- (1 + \varepsilon)-distance oracle with size $O(n \log n/\varepsilon^4)$ and $O(1)$ query time.
- For any $A, B \in S$, we get algorithms for:
  - Diameter of $A$.
  - Furthest neighbor of $s$ with respect to $A$: $\max_{a \in A} \pi(s, a)$.
  - (Bichromatic) closest pair of $A, B$: $\min_{a \in A, b \in B} \pi(a, b)$.
  - Center with respect to $A$: $\min_{s \in S} \max_{a \in A} \pi(s, a)$.
  - Median with respect to $A$: $\min_{s \in S} \sum_{a \in A} \pi(s, a)$.
  - Stretch factor of a subgraph $G_1$: $\max_{s, t \in S} \frac{\pi_1(p, q)}{\pi(p, q)}$.

with

- (1 + \varepsilon)-approximation in time $O(n \sqrt{n \log n/\varepsilon^3})$.
- 2.42-approximation in time $O(n \log^3 n)$.
Approximate distance oracle

**Corollary 6**  For a unit-disk graph on $n$ points and for any $\varepsilon > 0$, we can preprocess it into a data structure with $O(n \log n / \varepsilon^4)$ size so that for any query pair, a $(1 + \varepsilon)$-approximate distance can be answered in $O(1)$ time.

- Build $c$-WSPD. $c = 4/\varepsilon$.
- For each pair $(A, B)$, compute the distances between $(p, q)$, $p \in A$, $q \in B$.
- Query the 2.42-approximate distance for $u, v$.
- There is only $O(1)$ candidate pairs.

Note: Thorup’s $(1 + \varepsilon)$-distance oracle for planar graph has size $O(n \log n)$ and query time $O(1/\varepsilon)$. 
Example: bichromatic closest pair

Define the closest pair between \( A, B \in S \) as \( \min_{a \in A, b \in B} \pi(a, b) \).

- Mark the node in \( T' \) that contains a point of \( A (B) \) accordingly.
- Mark a pair \( (A_i, B_i) \) of WSPD iff \( A_i \) and \( B_i \) are both marked.
- Find the smallest distance between all the marked pairs.

The others are similar.
Higher dimensions

For a unit-ball graph metric of $n$ points in $\mathbb{R}^k$, $k \geq 3$, with constant bounded density and any constant $c \geq 1$, we can compute a $c$-WSPD with $O(n^{2-\frac{2}{k}})$ pairs in time $O(n^{2-\frac{2}{k}})$.

Note: This bound is tight in the worst case.
Every edge in the unit-disk graph has weight 1.

- For points with constant-bounded density,
  - the bound on WSPD still works.
- For points with arbitrary density,
  - there may not be subquadratic WSPD, e.g., a complete graph.
  - for applications, we find the minimal cover first.
Conclusion and future work

- We extend WSPD to unit-disk graph metric and apply to many proximity problems.
- Open problems:
  - Size of the WSPD in the plane: $O(n \log n)$ and $\Omega(n)$. Close the gap?
  - Is there an almost linear-time $(1 + \varepsilon)$-approximation algorithm for $O(n \log n)$ shortest path queries?