ICCG 2019

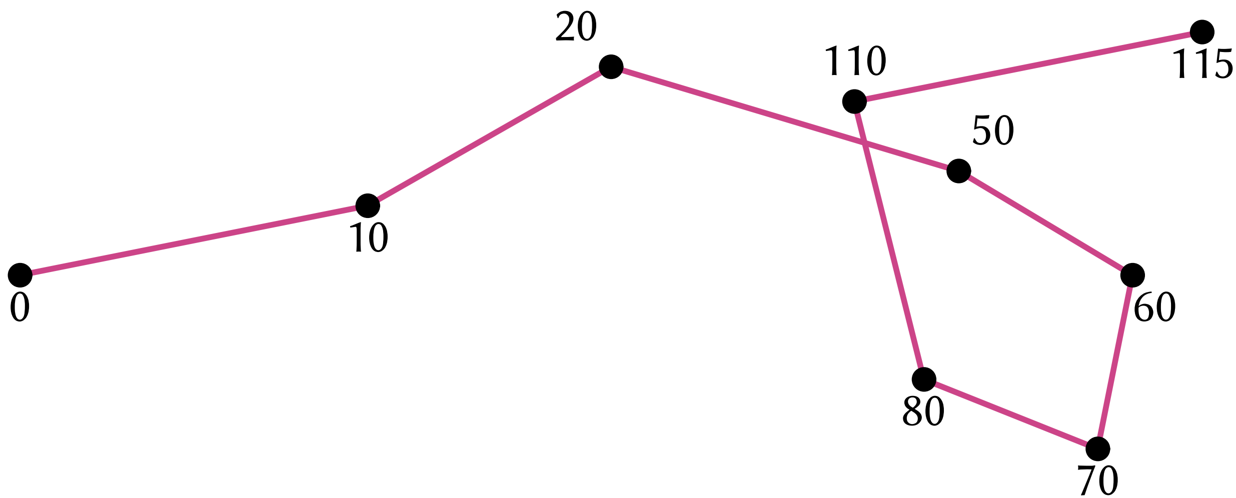
Approximate Discontinuous Trajectory Hotspots

Ali Gholami Rudi

Babol Noshirvani University of Technology

Trajectory:

- Polygonal curve + timestamps for the vertices



Analysis: finding regions in which the entity remains long

Trajectory Hotspots

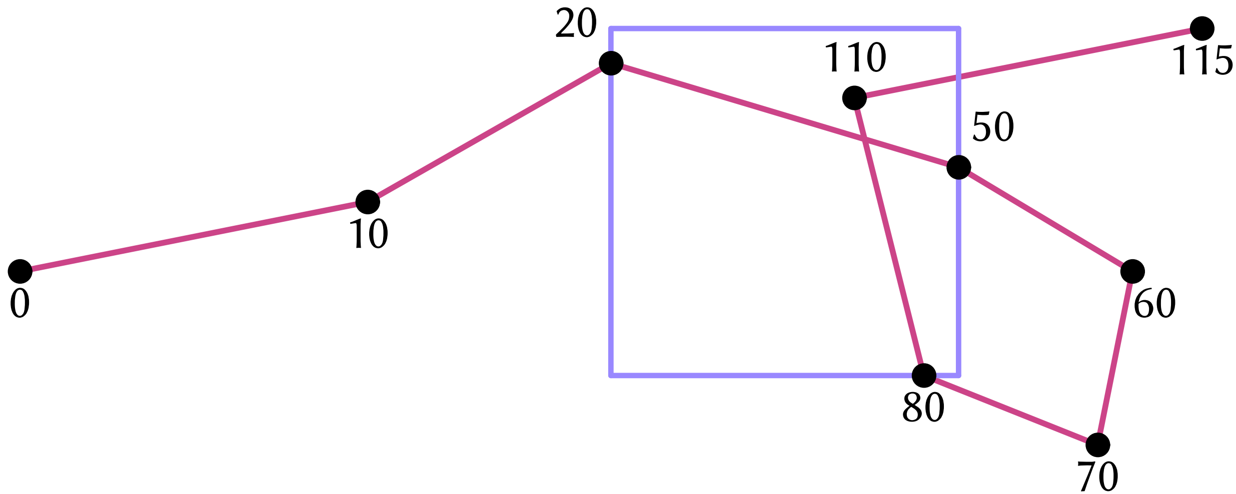
Different shapes and definitions:

- The number of visits: Benkert et al. (2010)
  - Visits from different entities
  - An $O(n^2)$ time algorithm
- The duration of the visits: Gudmundsson et al. (2013)
  - Continuous: the duration of a continuous visit
  - Discontinuous: entities may leave and return later
  - An $O(n^2)$ time algorithm

Square Weights

Weight of a square *r*:

- $w(r)$: the duration of the presence of entity inside it



Exact Hotspots

An exact hotspot:

- An axis-aligned square of fixed side length s
- With the maximum weight
  - s is an input and fixed through the algorithms

Approximate Hotspots

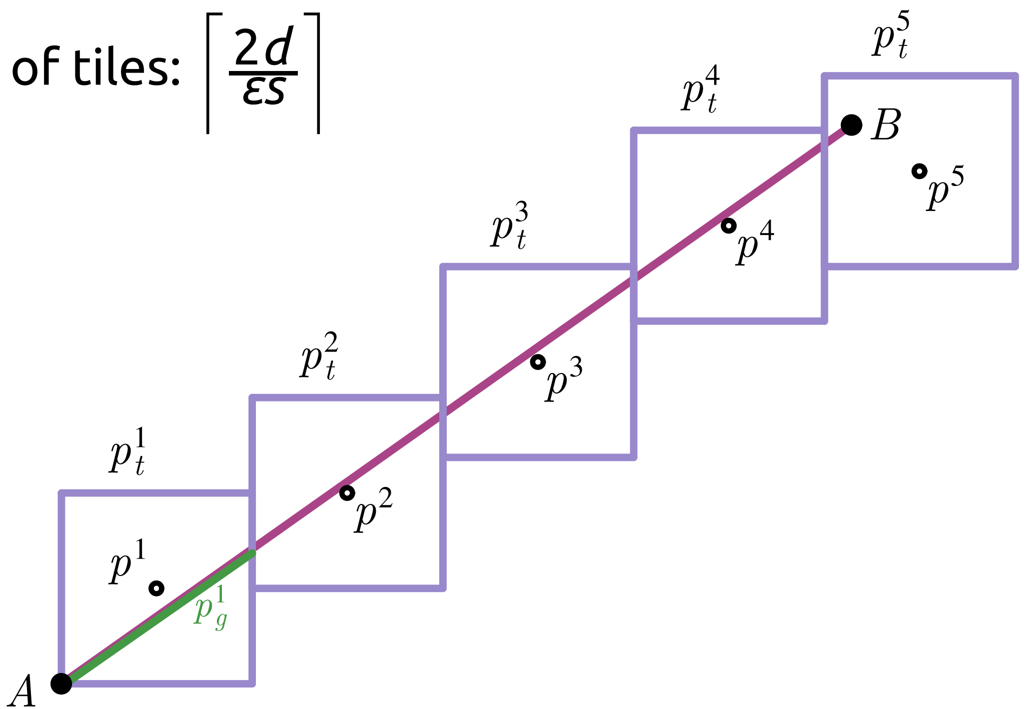
- Size-approximate (or simply approximate) hotspot:
  - the hotspot can be slightly larger
  - its weight ≥ the weight of an exact hotspot
  - c-size-approximate = side length is cs ($c > 1$)
- Duration-approximate
  - its weight can be slightly smaller
  - c-duration-approximate = the weight is at least ch
  - where h is the weight of an exact hotspot ($c < 1$)

Goal

- Present a $(1 + \varepsilon)$-size-approximate hotspot
  - ε is an input ($\varepsilon > 0$)
- Use it to obtain a 1/4-duration approximate hotspot

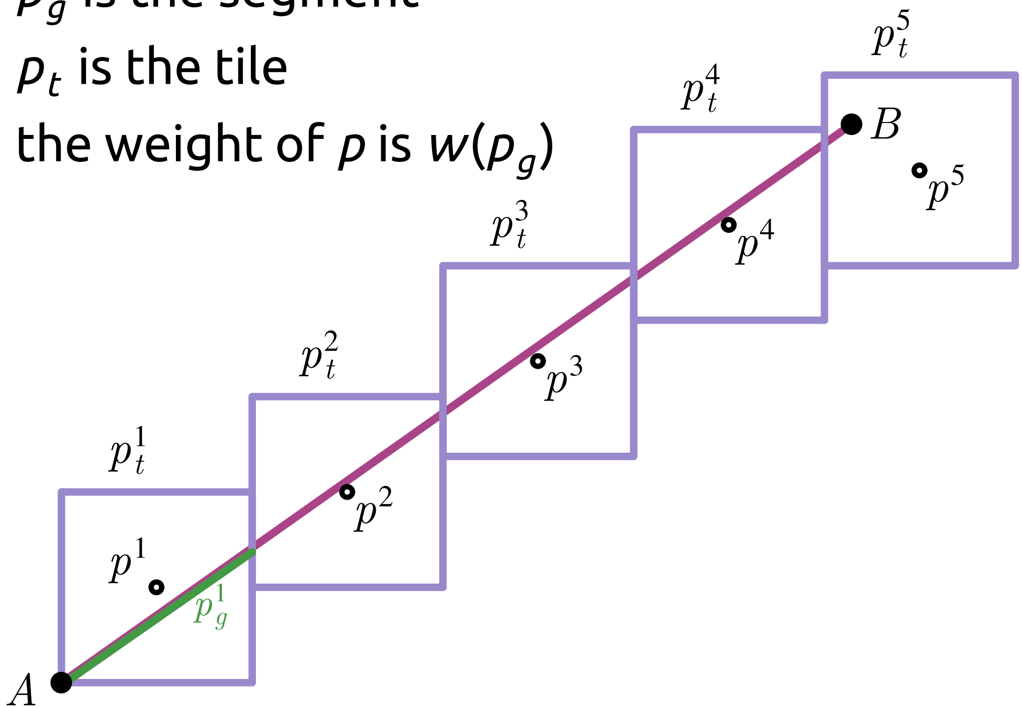
Tiling

- Cover each edge with $O(n)$ non-overlapping tiles
  - tile = a square of side length $\varepsilon s/2$
- # of tiles: $\left\lceil \frac{2d}{\varepsilon s} \right\rceil$



Tiling

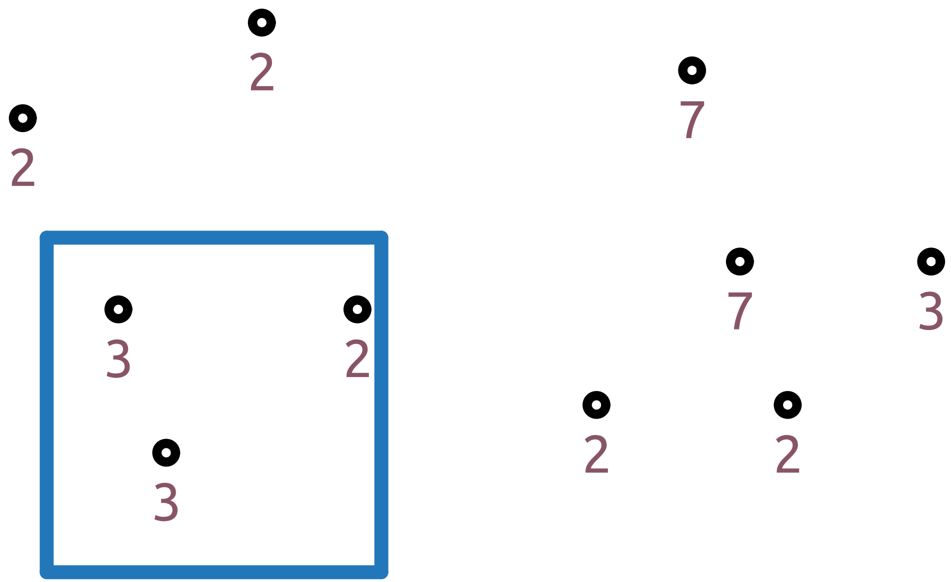
- A weighted point p at the center of each tile
  - p_g is the segment
  - p_t is the tile
  - the weight of p is $w(p_g)$



Point-Weight

The point-weight of a square r

- Denoted as $w'(r)$
- The sum of the weights of the points inside it

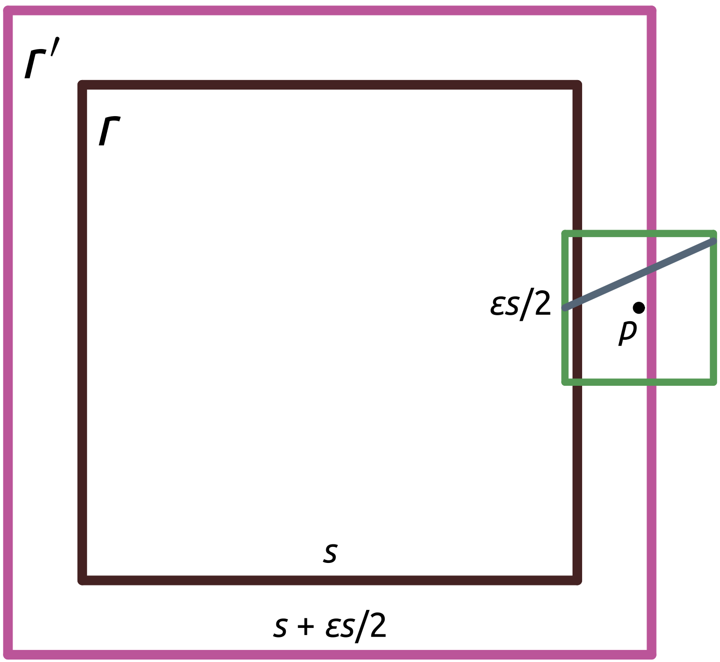


$w'(r) = 8$

Lemma 1: Nested Squares

Lemma: $w(r) \leq w'(r')$

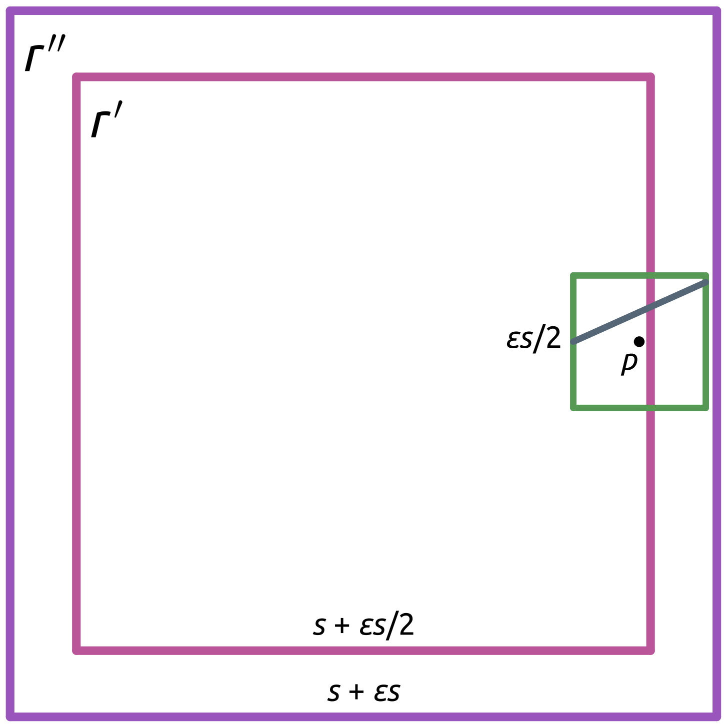
- segment p_g (partially) inside r implies p in r'



Lemma 2: Nested Squares

Lemma: $w'(r') \leq w(r'')$

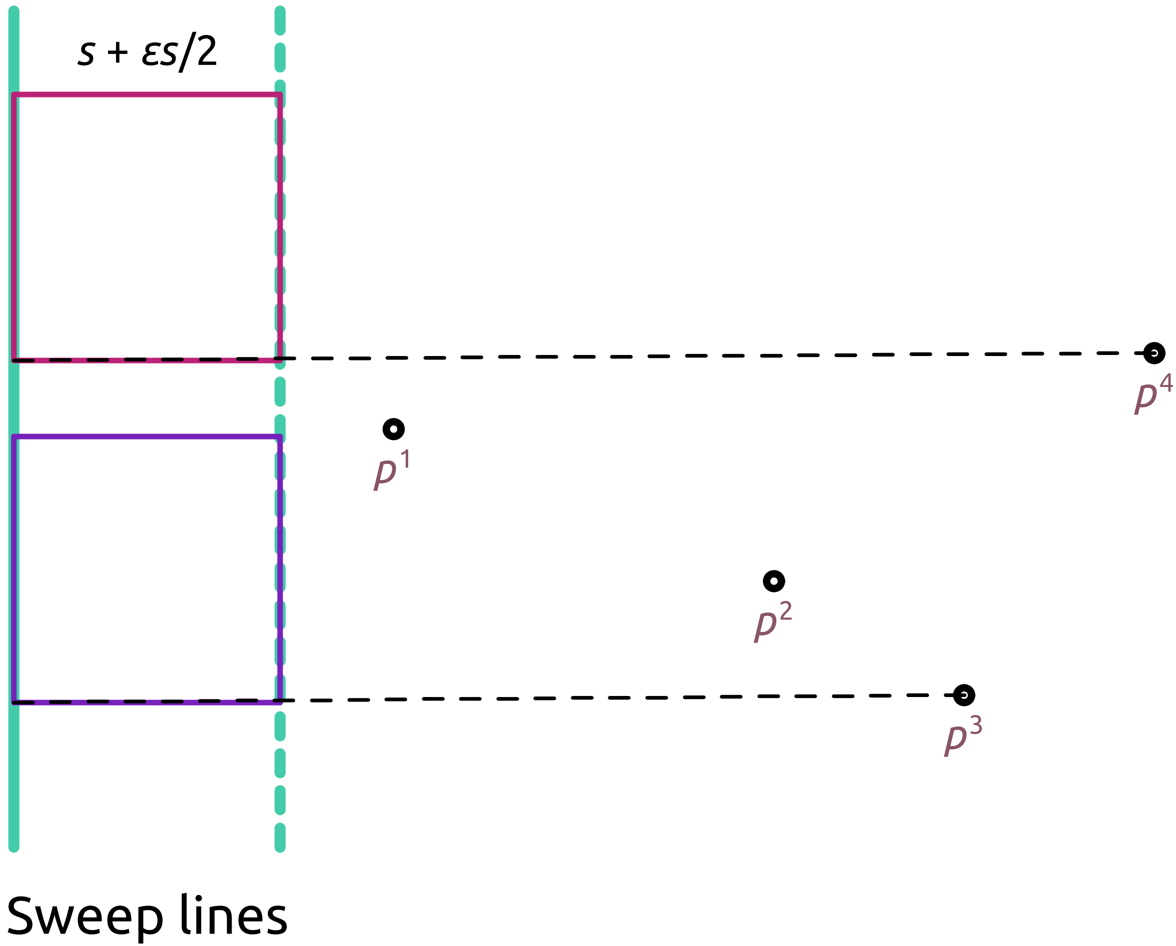
- p inside r' implies p_g totally contained in r''



Finding a square of side length $s + \varepsilon s/2$ with the maximum point-weight in $O(m \log m)$ for m points

- A plane sweep algorithm
  - Squares whose lower side has the height of a point
  - Keep track of m squares
  - The i-th square has the same height as the i-th point
- A data structure W:
  - The i-th entry is the weight of i-th square

Max Point-Weight Square



Sweep Line Performance

The time complexity of the sweep line algorithm

- $O(m \log m)$ with suitable data structure for W
- m is the number of weighted points

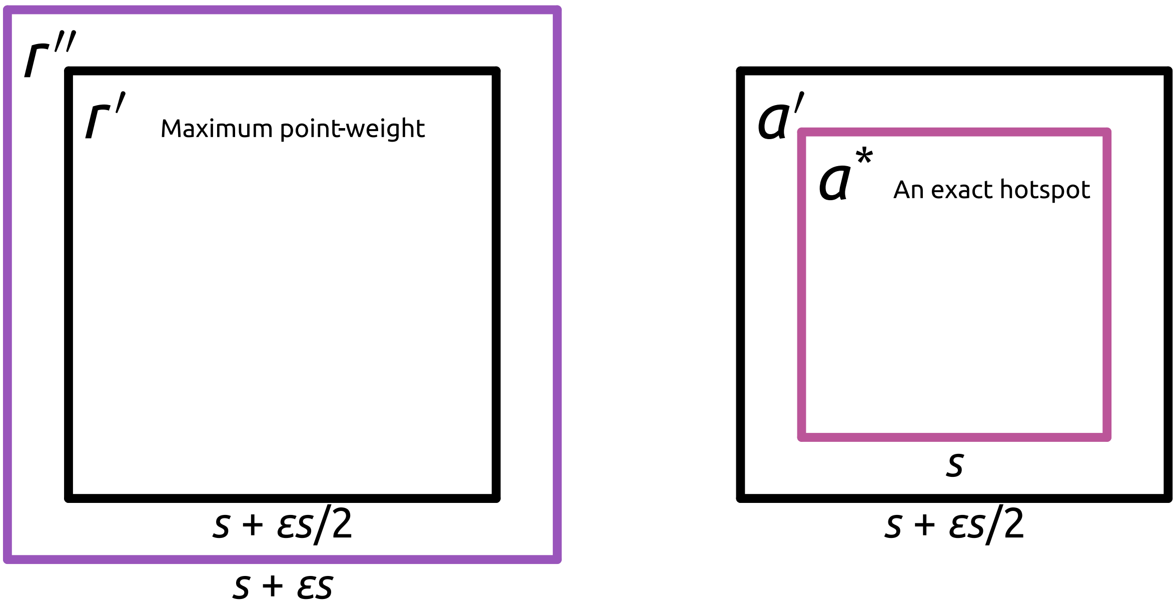
Size-Approximate Hotspot

Obtaining a $(1 + \varepsilon)$-approximate hotspot ($\varepsilon > 0$)

- Tiling (a is the total length of edges)
  - An edge of length d: at most $\left\lceil \frac{2d}{\varepsilon s} \right\rceil$ points
  - Total number of points: at most $\frac{2a}{\varepsilon s} + n$
  - $\varphi = \frac{a}{ns}$ (ratio of average edge length to s)
  - Total number of points: $O(\frac{n\varphi}{\varepsilon})$
  - Plane sweep complexity $O(\frac{n\varphi}{\varepsilon} \log \frac{n\varphi}{\varepsilon})$

Size-Approximate Hotspot

- r': square of side length $s + \varepsilon s/2$ with max point-weight
- a^*: an exact hotspot
- $w'(a') \geq w(a^*)$; Lemma 1
- $w'(r') \geq w(a')$; r' has the maximum point-weight
- $w(r'') \geq w'(r')$; Lemma 2



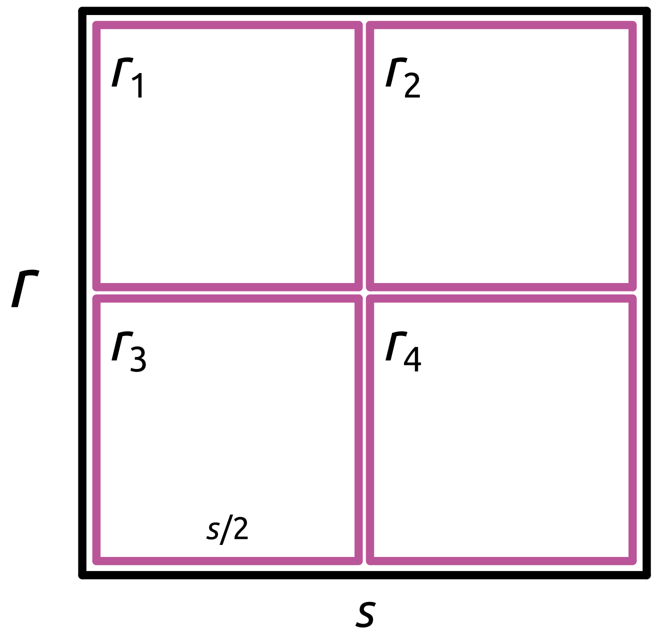
Duration-Approximate Hotspots

Obtaining 1/4-duration-approximate hotspot

- $h' \geq h/4$
  - h: the weight of an exact hotspot with side length s
  - h': the weight of an exact hotspot with side length $s/2$

Half-Sized Squares

- r: an exact hotspot
  - $w(r_1) + w(r_2) + w(r_3) + w(r_4) \geq w(r)$
- $w(r_i) \geq w(r)/4$, for some i



Obtaining 1/4-duration-approximate hotspot

- $h \leq 4h'$
- $(1 + 1)$-size-approximate hotspot for $s' = s/2$ and $\varepsilon = 1$
  - r: square with side length $2s' = s$
  - $w(r) \geq h' \geq \frac{1}{4} h$
- Time complexity $O(n\varphi \log n\varphi)$

Concluding Remarks

- $(1 + \varepsilon)$-approximate: in $O(\frac{n\varphi}{\varepsilon} \log \frac{n\varphi}{\varepsilon})$
  - φ is the ratio of average edge length to s
- 1/4-duration-approximate: in $O(n\varphi \log n\varphi)$