

SCALABLE PROCESSING OF MOVING FLOCK PATTERNS

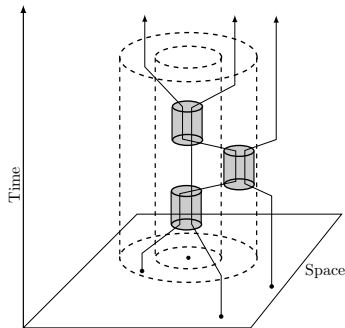
Andres Calderon · acald013@ucr.edu

University of California, Riverside

October 30, 2024

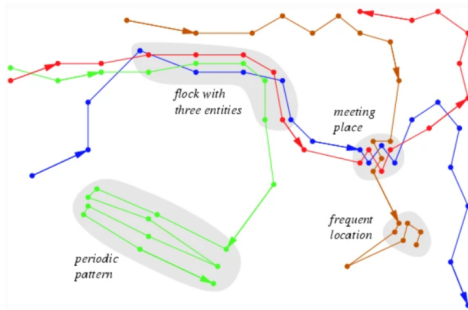
LARGE TRAJECTORY DATABASES

- A spatial trajectory is a trace in time generated by a moving entity in a geographical space.
- i.e. $p_1 \rightarrow p_2 \rightarrow \dots \rightarrow p_n$
- A trajectory is stored as a set of points, $p_i = (x, y, t)$ (spatial coordinate + time stamp).



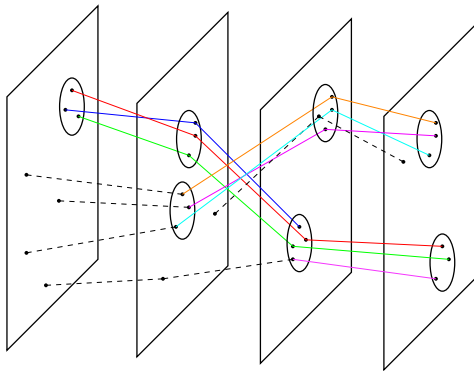
(Shoval, 2017)

MOVEMENT PATTERNS



(Gudmundsson, et al. 2008)

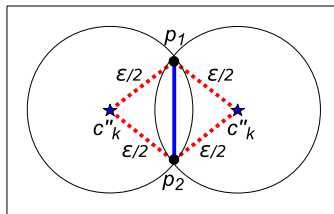
- i.e. convoys, moving clusters, swarms, gatherings, **flocks**, ...



- ε : Maximum distance between objects.
- μ : Minimum number of objects.
- δ : Minimum time the objects keep 'together'.

BASIC FLOCK EVALUATION ALGORITHM

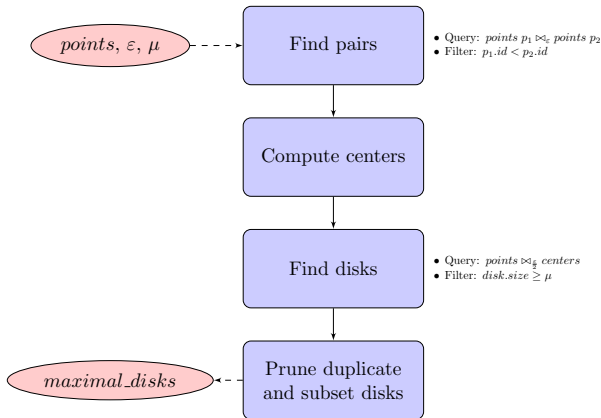
- Vieira, et al. 2009.
- The first polynomial-time solution for determining disk locations.
- Under fixed time duration it has polynomial time complexity $O(\delta|\tau|^{(2\delta)+1})$



BASIC FLOCK EVALUATION ALGORITHM

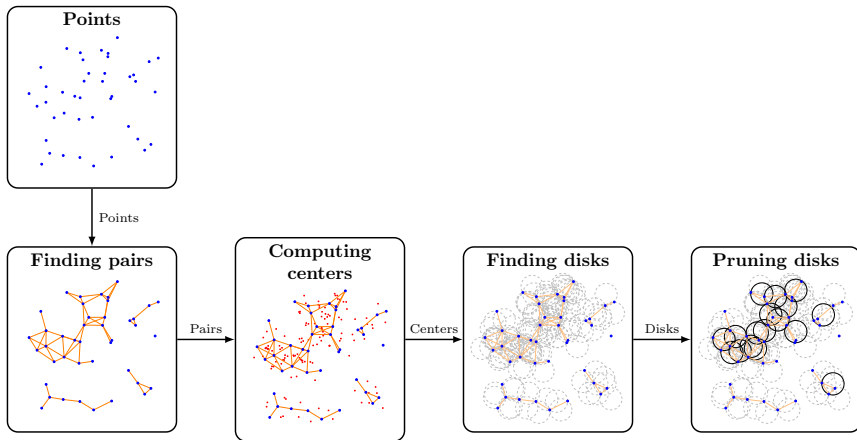
- Two main parts:
 - ▶ In the spatial domain it finds maximal disks at each time stamp.
 - ▶ In the temporal domain it joins consecutive times to match set of maximal disks.

■ BFE overview...

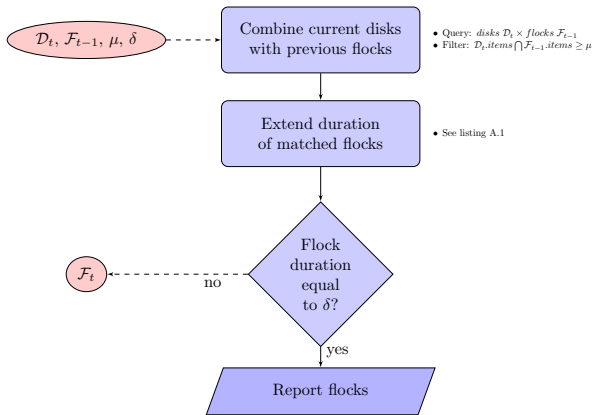


ON THE SPATIAL DOMAIN

■ BFE overview...

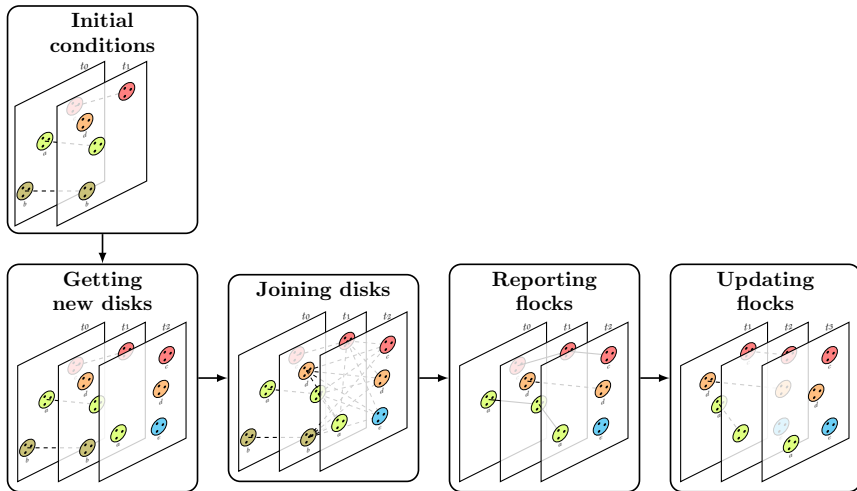


■ BFE overview...

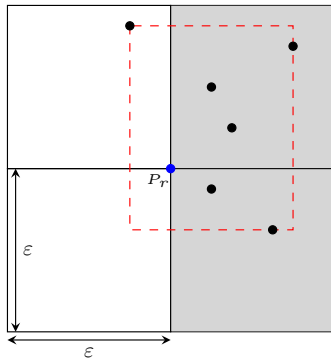
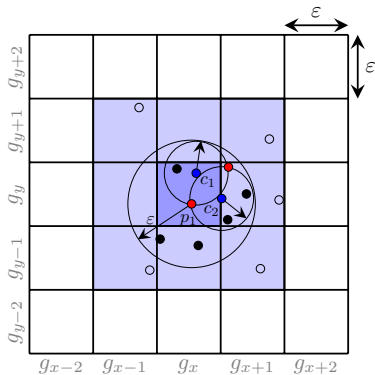


ON THE TEMPORAL DOMAIN

■ BFE overview...

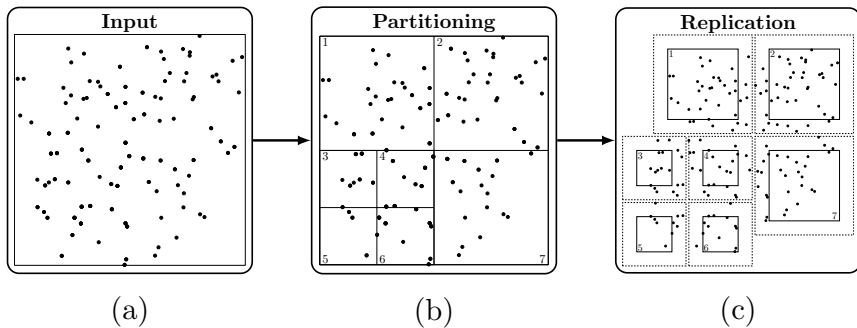


PSI ALGORITHM

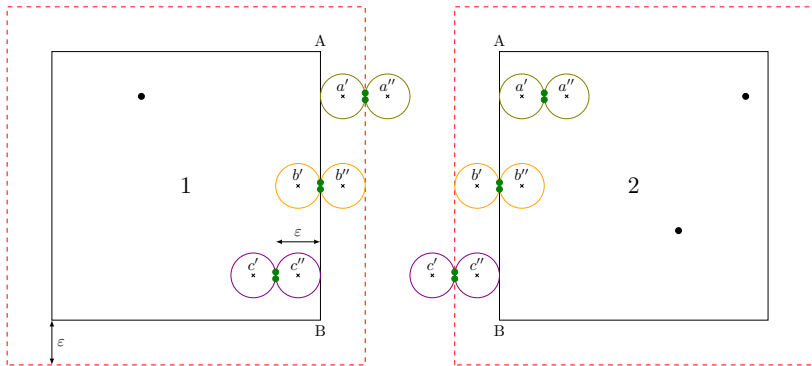


- Due to high complexity it does not scale well.
- In databases with a large number of moving entities per time stamp it has a direct impact.
- Just sequential implementation yet.
- We propose a parallel solution in both domains.

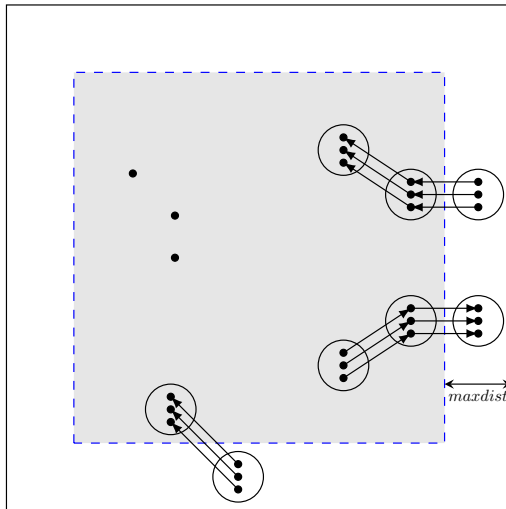
■ Parallel overview...



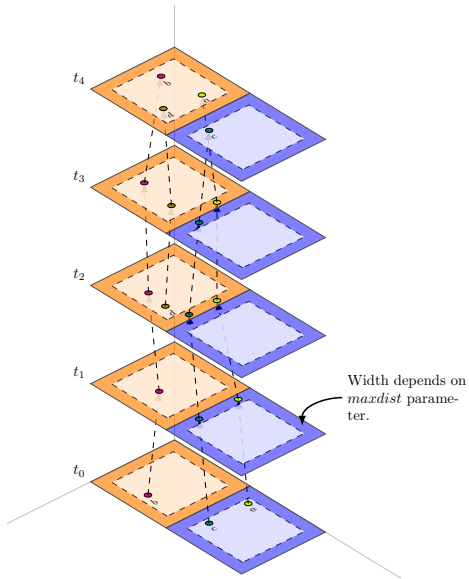
■ Parallel overview...



ON THE TEMPORAL DOMAIN

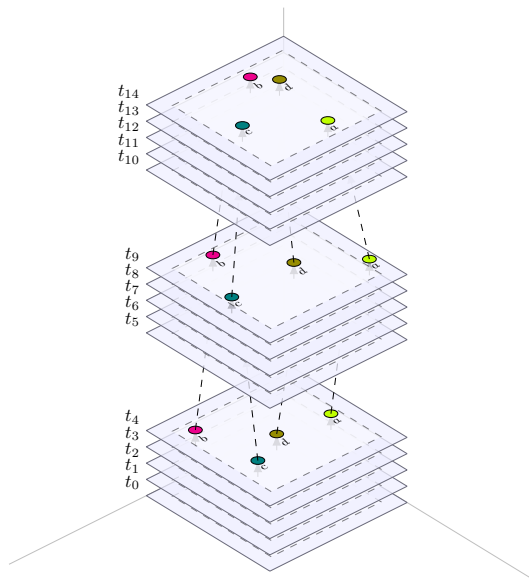


ON THE TEMPORAL DOMAIN



* a, b, c and d are flocks moving along time.

ON THE TEMPORAL DOMAIN



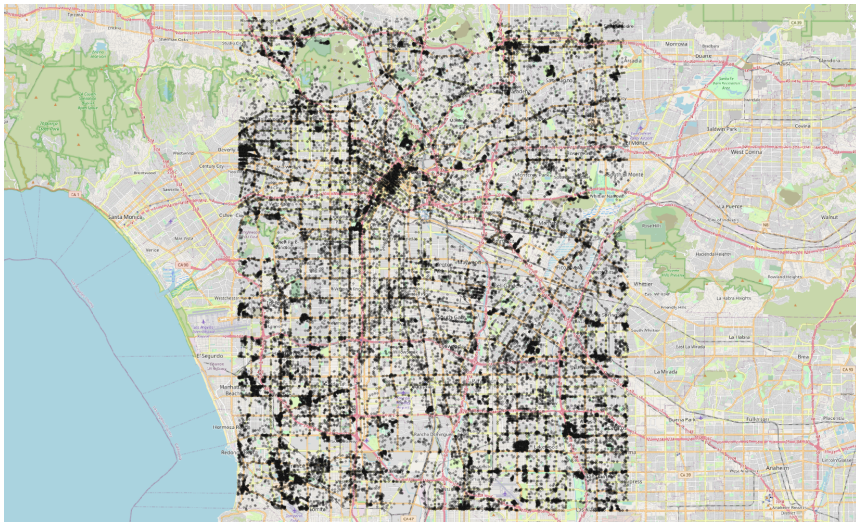
*a,b,c and d are flocks moving along time.

DATASETS

Dataset	Number of Trajectories	Total number of points	Maximum Duration (min)
Berlin10K	10000	97526	10
LA25K	25000	1495637	30
LA50K	50000	2993517	60

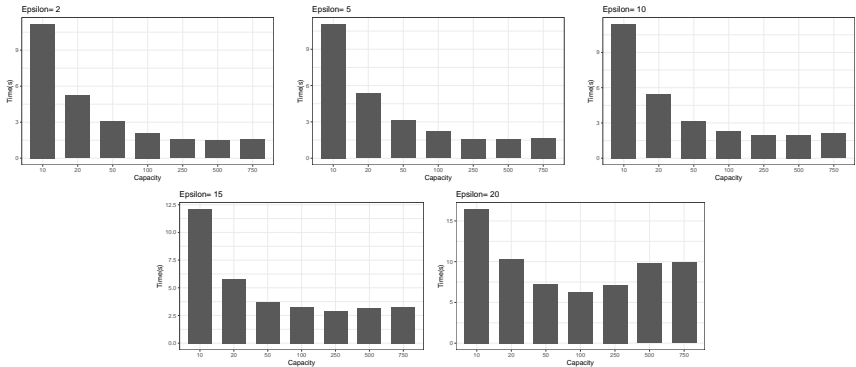
DATASETS

■ Synthetic dataset [LA: 50K objects]



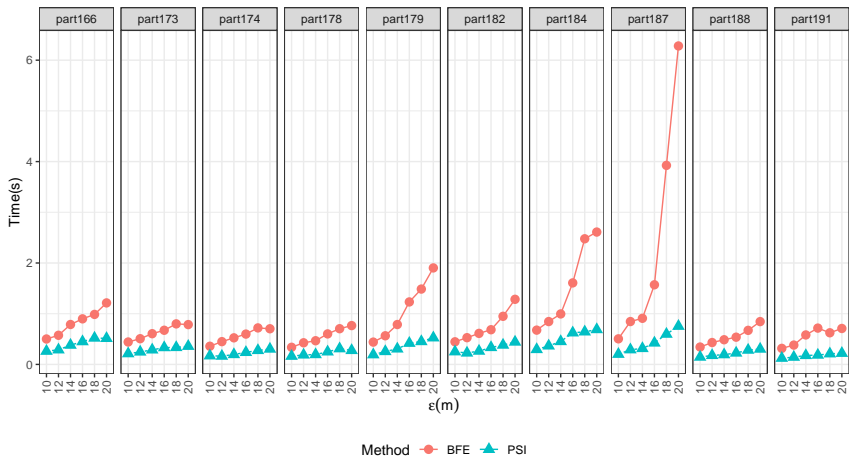
EXPERIMENTS

■ Optimizing the number of partitions for Phase 1.

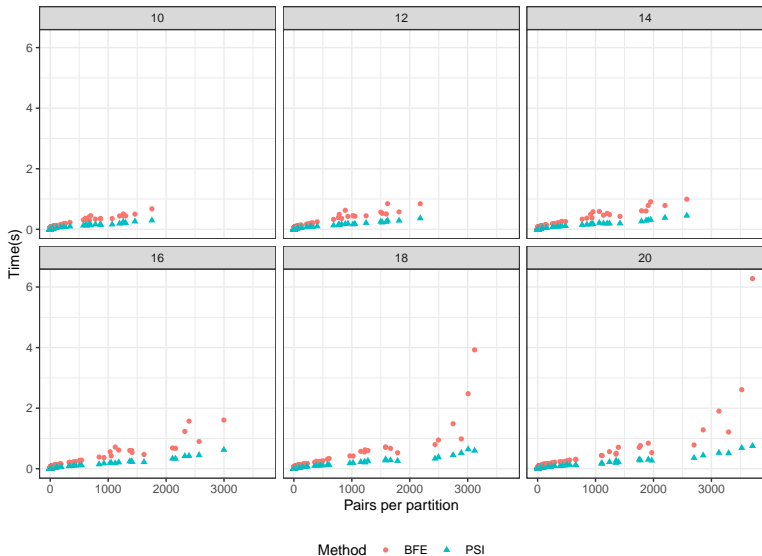


EXPERIMENTS

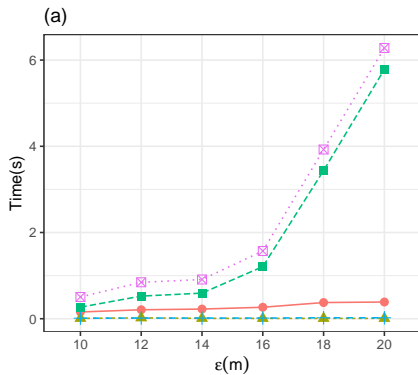
■ Analyzing most costly partitions.



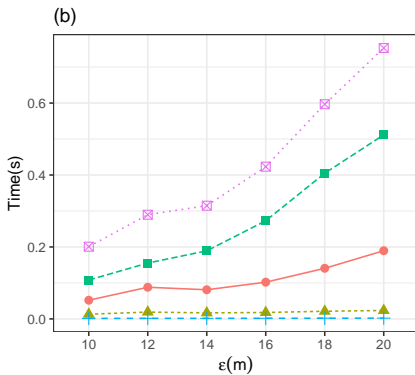
EXPERIMENTS



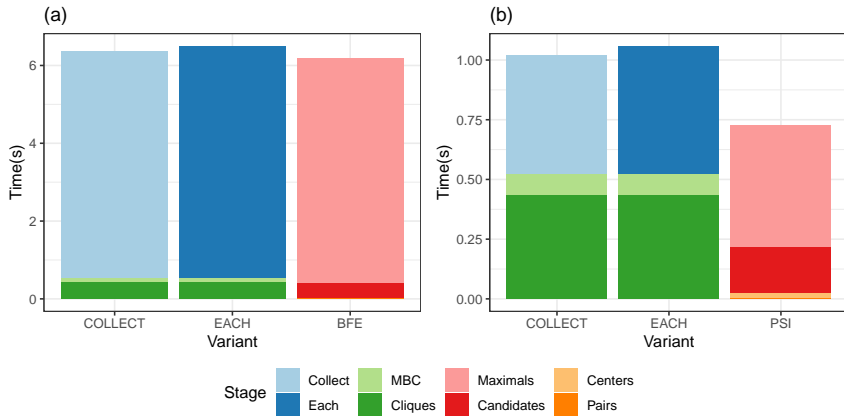
EXPERIMENTS



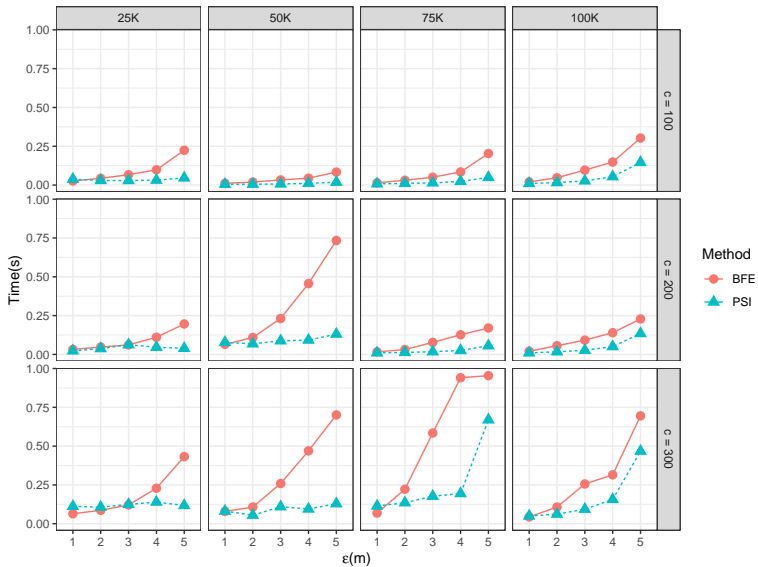
Stage —●— Candidates —▲— Centers —■— Maximals —+— Pairs —×— Total



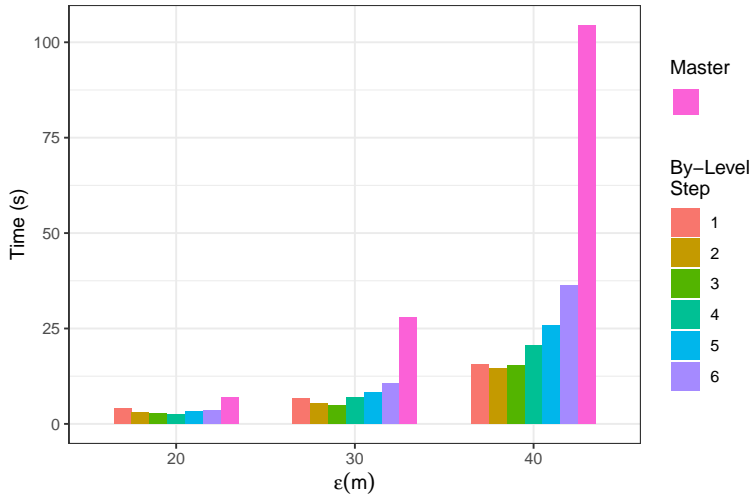
EXPERIMENTS



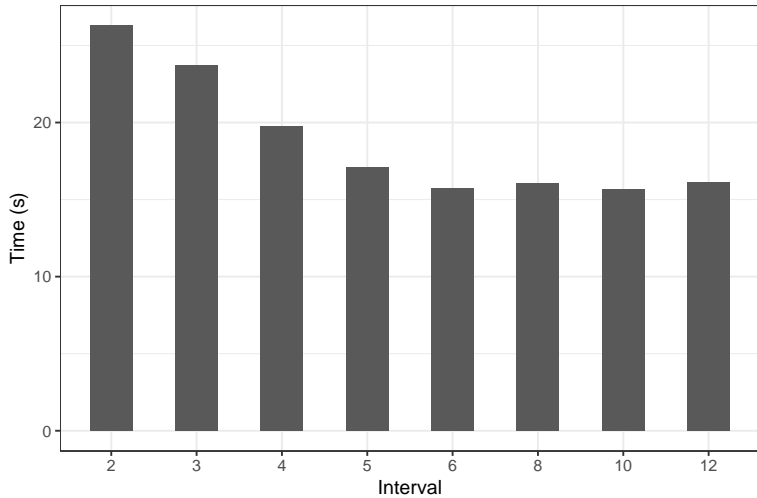
EXPERIMENTS



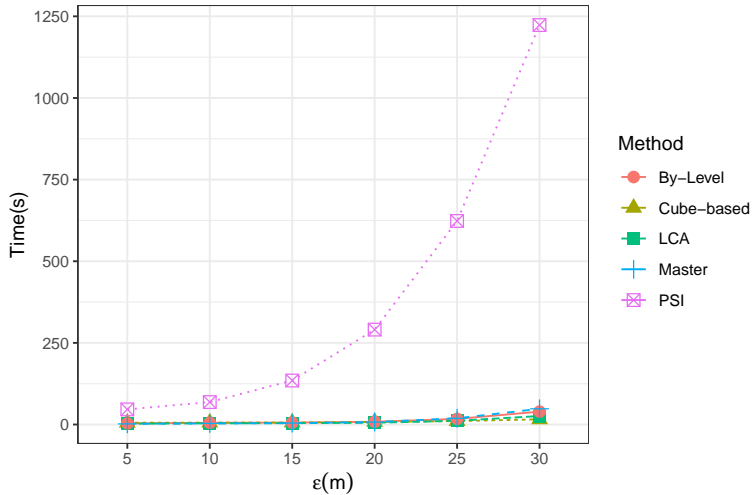
EXPERIMENTS



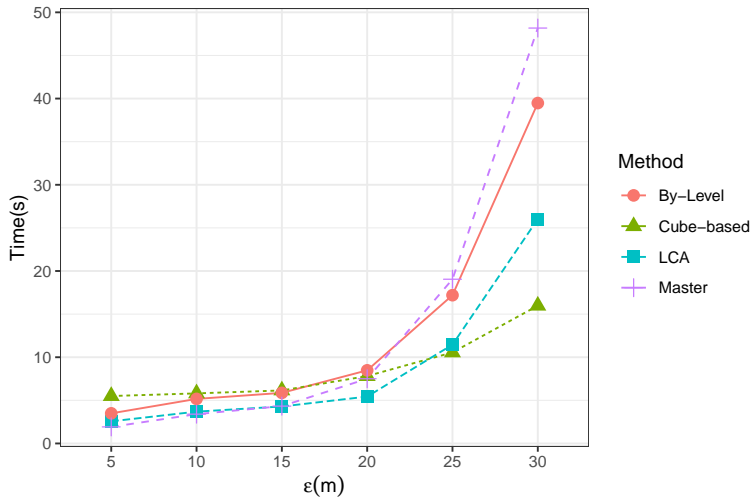
EXPERIMENTS



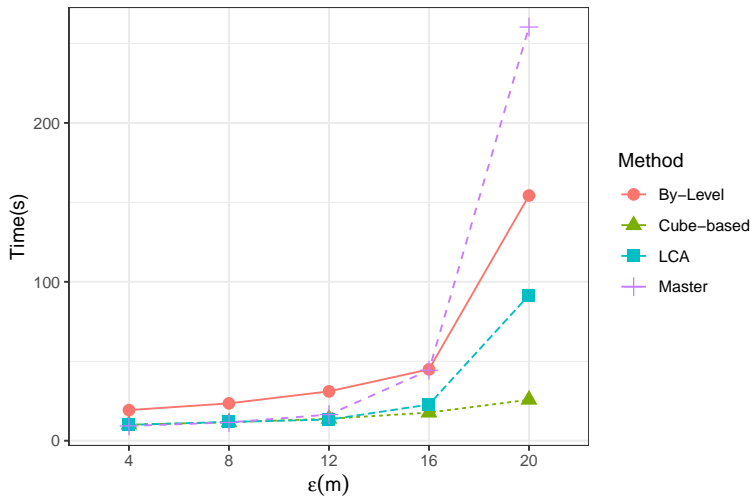
EXPERIMENTS



EXPERIMENTS



EXPERIMENTS



Thank you!