

PFLOCK Report

Andres Calderon

University of California, Riverside

August 30, 2019

LA_10K Maximal disks finding

$$\mu = 5, \delta = 15.$$

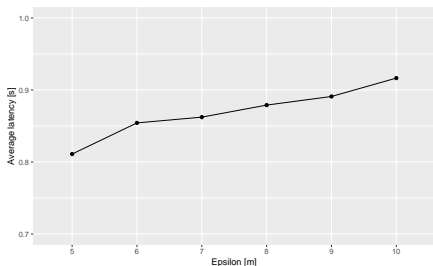


Figure: Latency.

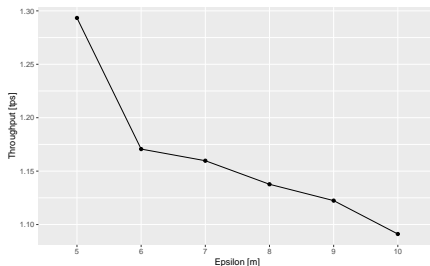


Figure: Throughput.

LA_10K Join consecutive time instants

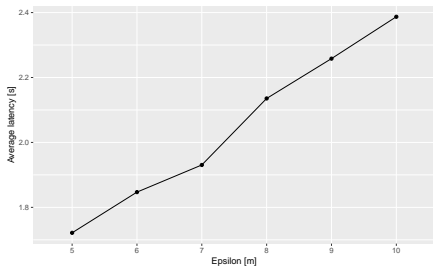


Figure: Latency.

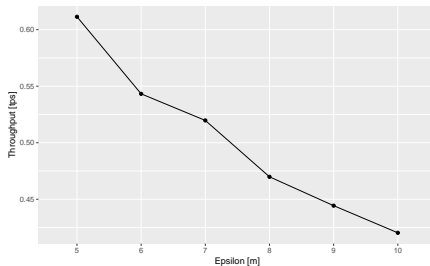


Figure: Throughput.

LA_5K Maximal disks finding

$$\mu = 5, \delta = 15.$$

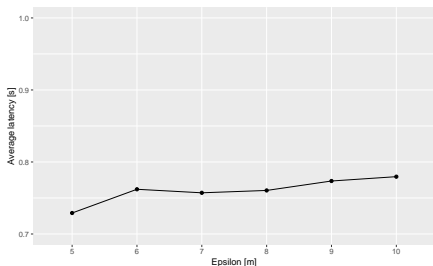


Figure: Latency.

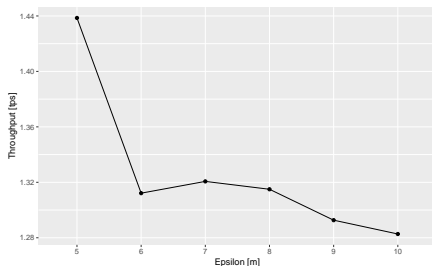


Figure: Throughput.

LA_5K Join consecutive time instants

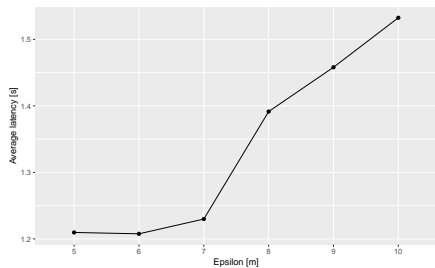


Figure: Latency.

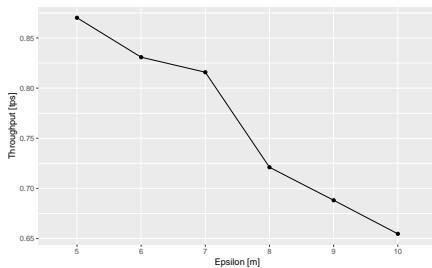


Figure: Throughput.

Some remarks

- ▶ Optimal number of partitions is small (≈ 24) due to the relatively small number of trajectories. It does not exploit all of the available cores.
- ▶ Chen et.al. (2019) exploits parallelism by mining several (η) snapshots at a time.

Issues with GeoLife dataset

- ▶ I got the same number as Table 2 in Chen et.al. (2019), but the number of snapshots (92,645) is actually the number of snapshots for the longest trajectory.
- ▶ The real life span of the dataset is huge (170'985.600 snapshots).
- ▶ If any aggregation/interpolation is done the maximum number of moving objects at the same time is only 20, Actually the maximum number of users in the dataset is 182.
- ▶ In the last email, Dr Chen confirms that the ϵ value ranges between 8.2Km and 49.5Km.