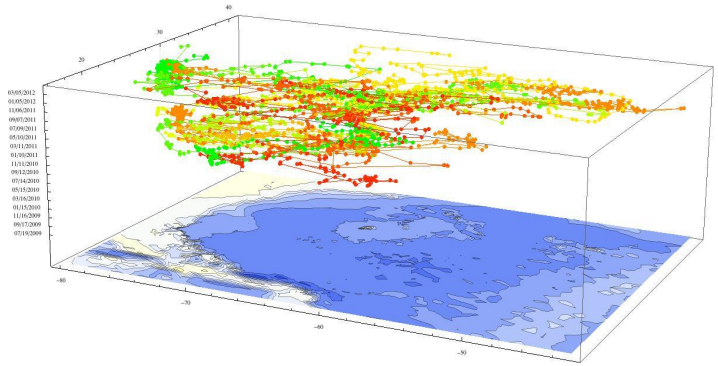
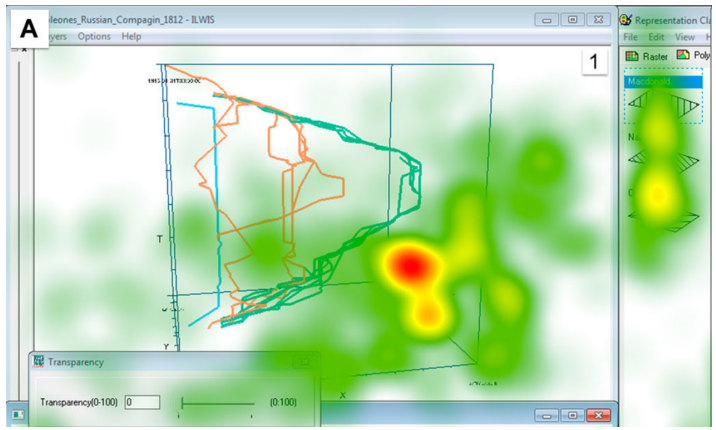
Parallel Detection of Movement Patterns in Large Spatio-temporal Datasets

February 1, 2019

Trajectory datasets

- Sensors, sensors everywhere!!!
- Anything that could move, will be tracked...
- Some applications:
  - Social behavior
  - Ecology (birds, sharks, ...)
  - Climate change (icebergs, cyclones, ...)
  - Software...





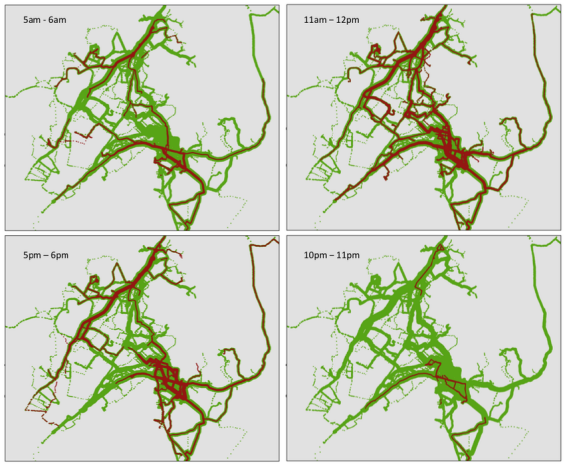
Complex movement patterns

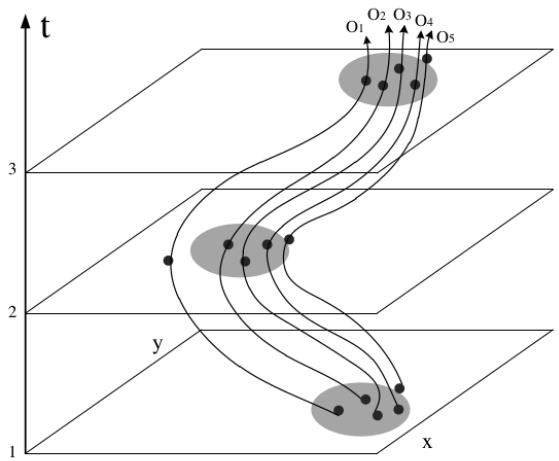
Previous works focus on traditional queries:

- Range, Nearest Neighbors, Similarity, ...

Recent works look for the aggregate behavior:

- Moving clusters, Convoys, **Flocks**, Swarms, Gatherings, ...



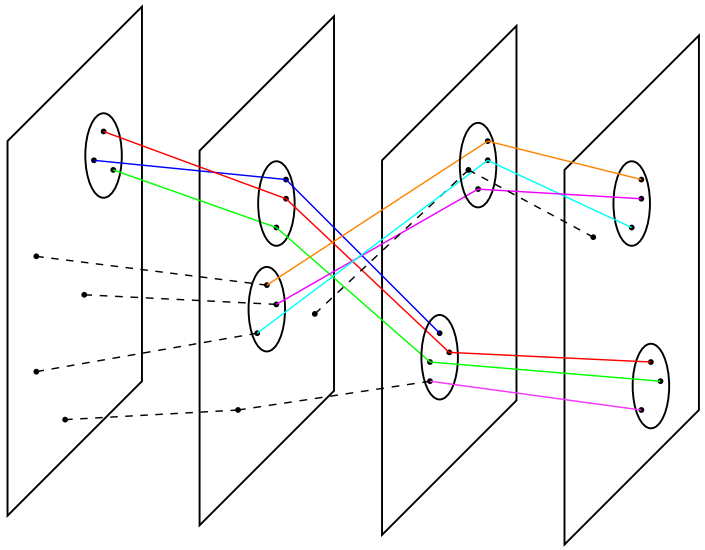


(b) Flock, convoy and swarm

What is a flock???

Definition ($(\mu, \varepsilon, \delta) - flock$)

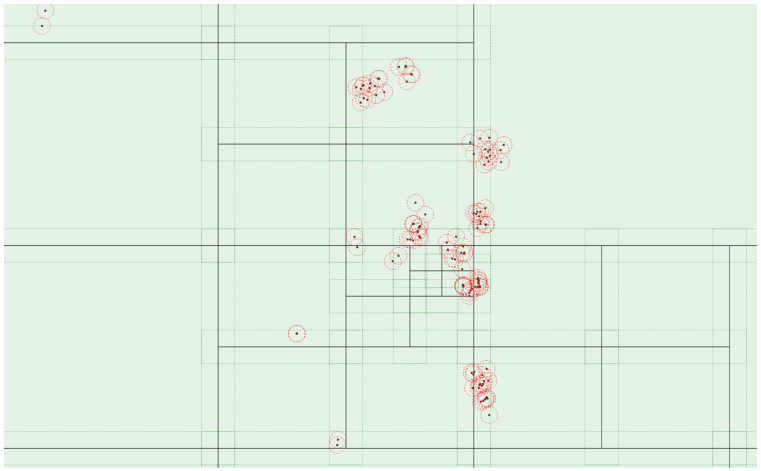
Sets of at least μ objects moving close enough (ε) for at least δ time intervals (Benkert et al, 2008).

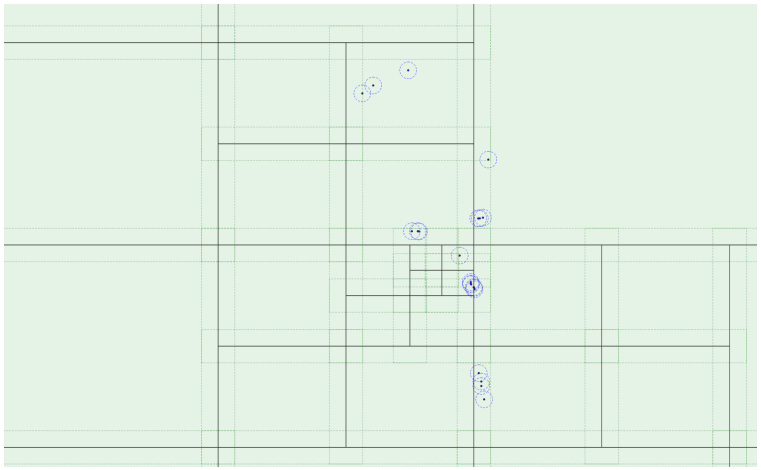


- Vieira et al. (2009) proposed BFE algorithm (first polinomial solution).
- Drawbacks:
  1. Find disks is costly. They can be at any place.
  2. Huge amount of duplicate and redundant disks.
  3. Join between time intervals was a Cartesian product.

Contributions

1. Boost the detection of disks through a parallel approach (spatial partitioning + expansions).
2. Apply a frequent pattern mining approach to improve disk filtering (local + merge approach).
3. Use parallel distance joins to improve combination between consecutive time intervals (a Distance parameter).





Preliminar results

- Implementation using GeoSpark. Synthetic datasets using SUMO (Simulation of Urban Mobility).
- Berlin network (OSM). $\approx 20K$ points per timestamp, 10 timestamps. Varying ε ($\mu = 3$ and $\delta = 3$).