TRAJECTORY PATTERN MINING

Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Martha Axiak
Marco Muscat

Introduction

- Nowadays data on the spatial and temporal location is objects is available.
- Gps, GSM towers, etc
- What can we do with this data?

Introduction

- Nowadays data on the spatial and temporal location is objects is available.
- Gps, GSM towers, etc
- What can we do with this data?
- Data mine it for patterns!! (of course)

Trajectory Pattern Mining

- A way to represent Spatial patterns.

$$Bistro\ Square \xrightarrow{10min} Appleton\ Tower \xrightarrow{2hr30min} Potterrow$$

$$Chambers\ Str \xrightarrow{15min} The\ Mound \xrightarrow{5min} Waverly\ Station$$

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Possible uses

- Prediction of movement
- Aggregate movement behaviour
- Region of interests discovery
- Discovery of traffic flow & blockages

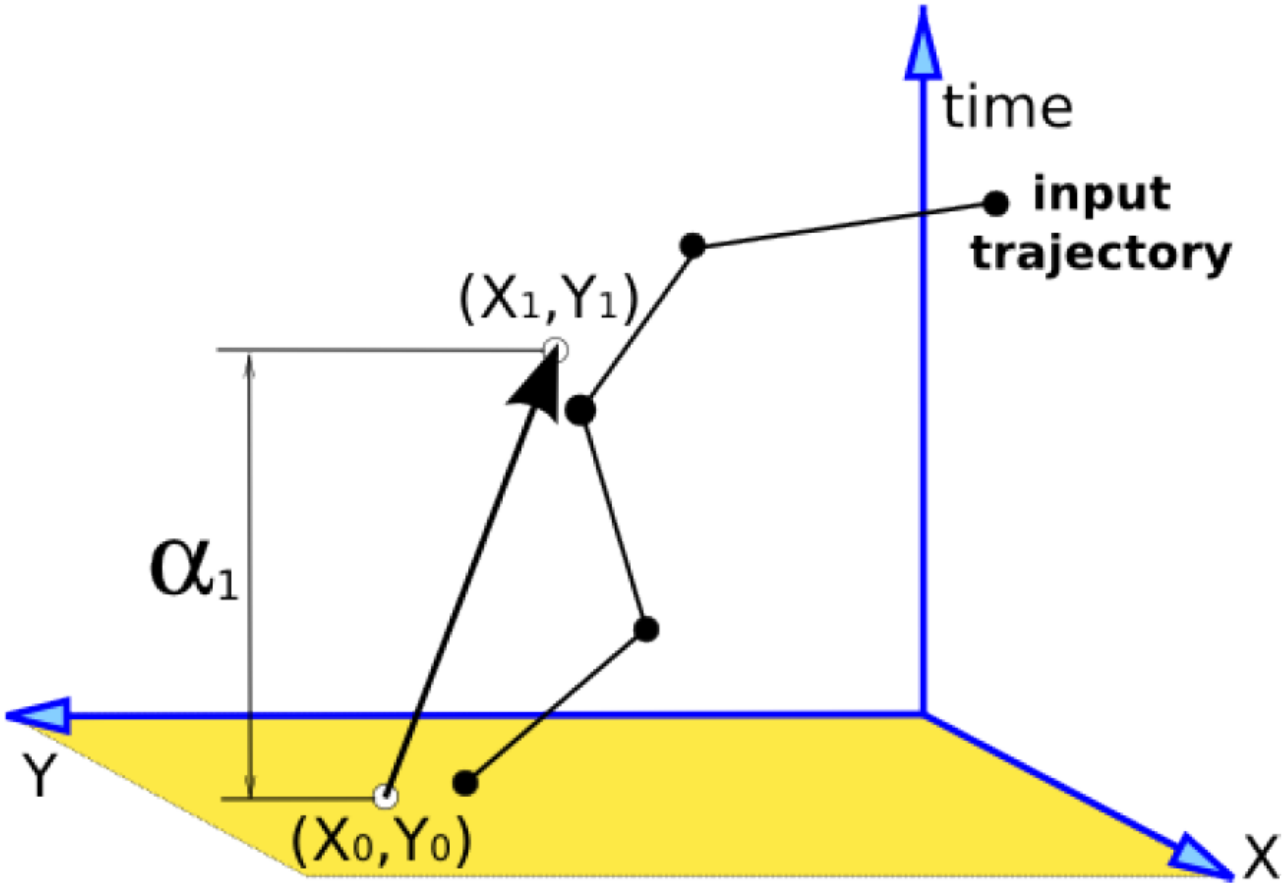
Structure of presentation

- Introduction
- Mechanics
  - Data structure
  - Regions of Interest (ROI)
  - T-Patterns using these ROI
- Results

Data structure

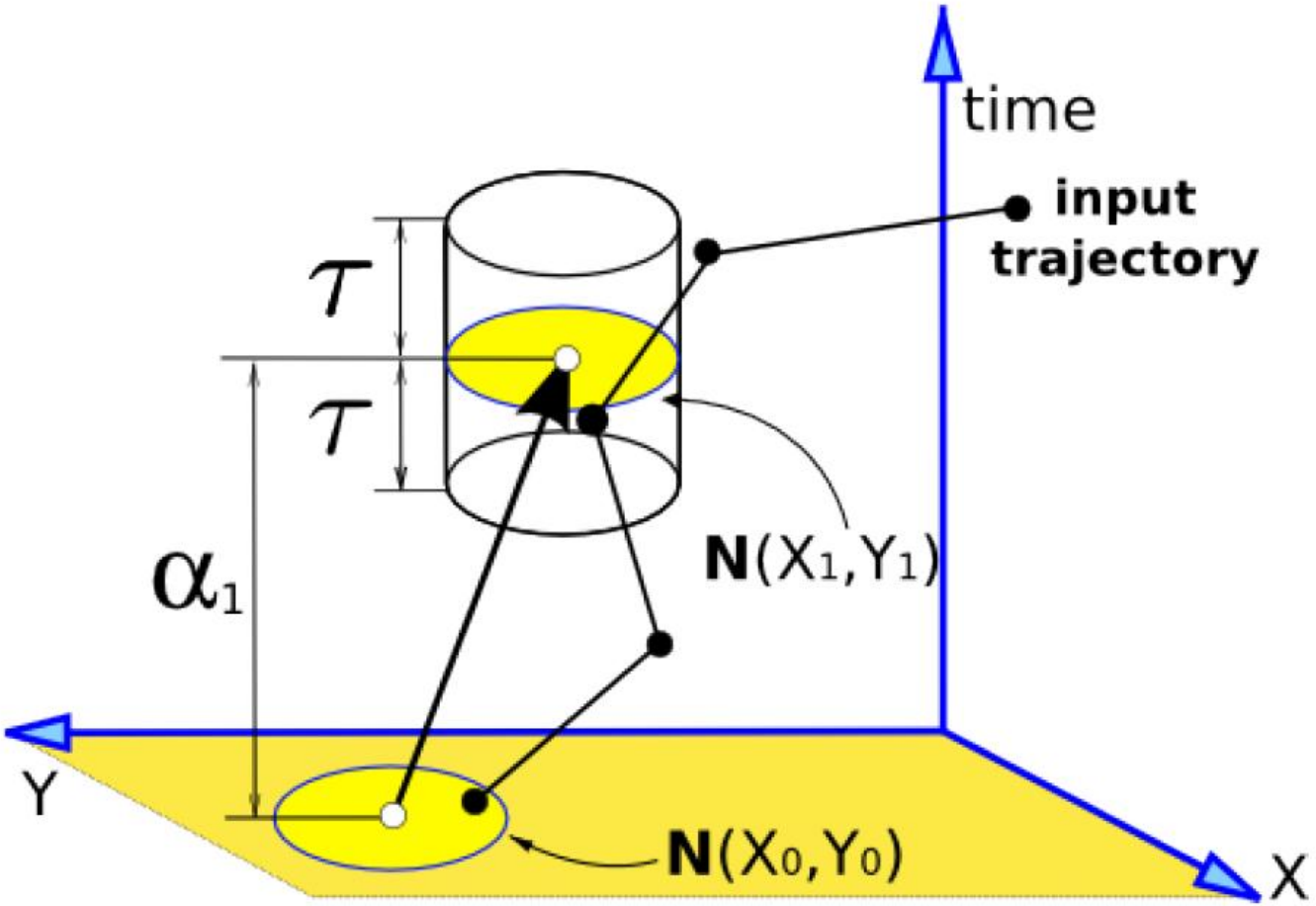
- Raw Input
  - List of $(Lat, Long, timestamp)$
  - Multiple Moving Objects
- Convert
  - $(Lat, Long, t) \xrightarrow{covert} (x, y, t) \qquad (x, y) \in R^2$
- T-Patterns
  - $T = s_0 \xrightarrow{\alpha_1} s_1 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_N} s_N$

Example: Raw input



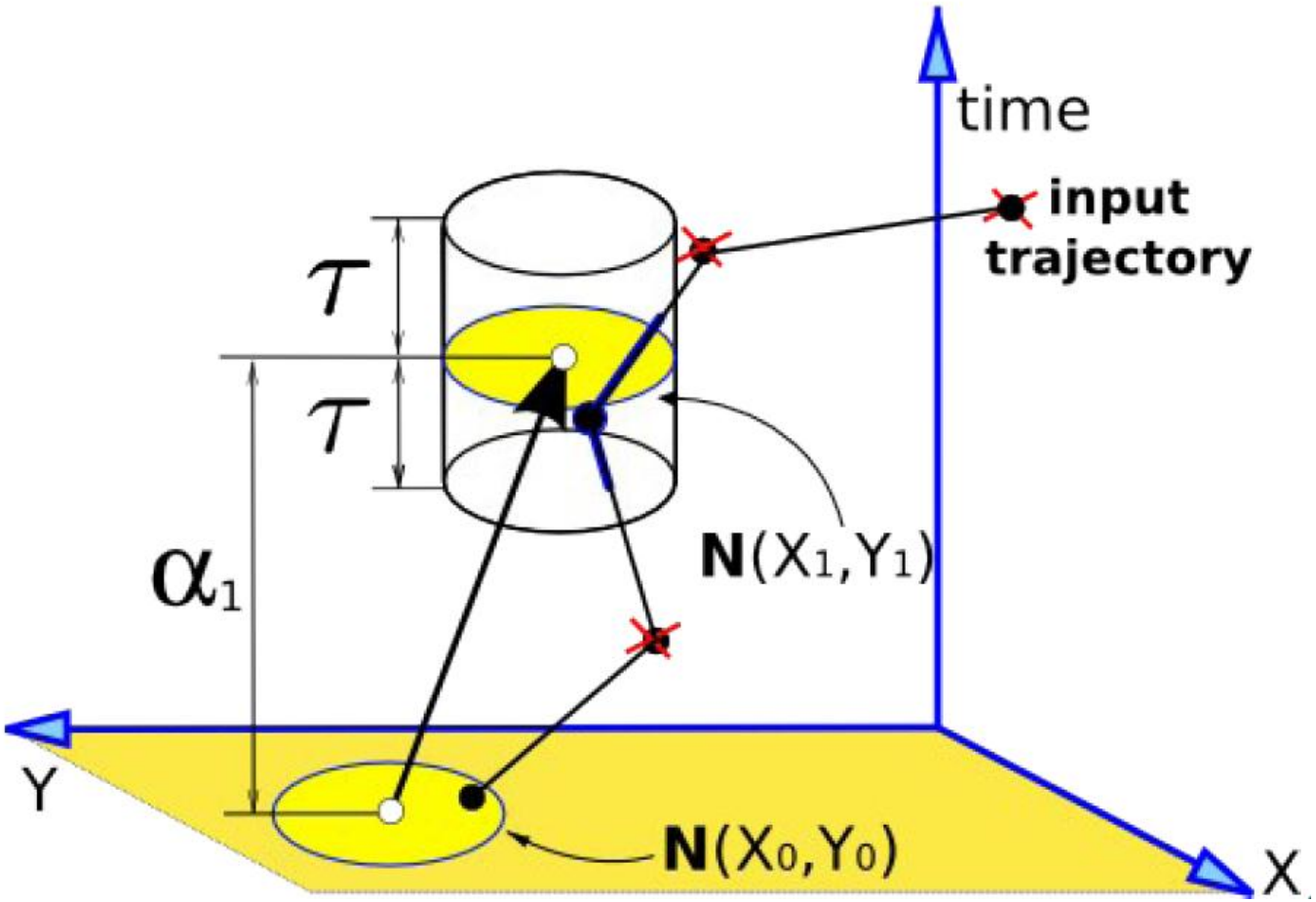
Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Example: T-Patterns



Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Example: T-Patterns



Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Support Threshold

- Support: Percentage of trajectories which are contained within a pattern.

Example: Support threshold of 0.2.

- A T-Pattern is kept only if 20% of the trajectories support it.

Curse of Dimensionality

- Data in 3 dimensional space (x, y, time).
- With Pattern length 3:

$$R^{(3points * 2D)+(2\ transition * 1D)}combinations$$

- As pattern length increases to n, search space

$$Search\ Space\ =\ R^{(n*2)+\left((n-1)*1\ \right)} = R^{3n-1}$$

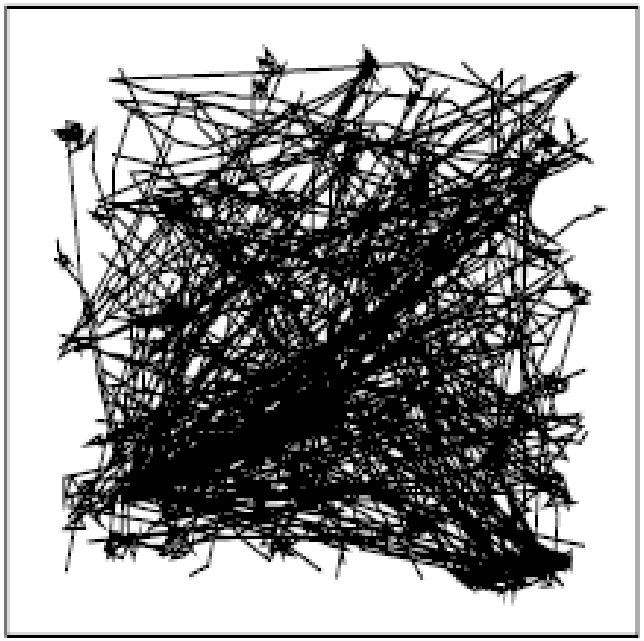
- Therefore intractable.
- Reduce dimensions through Regions of Interest.

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

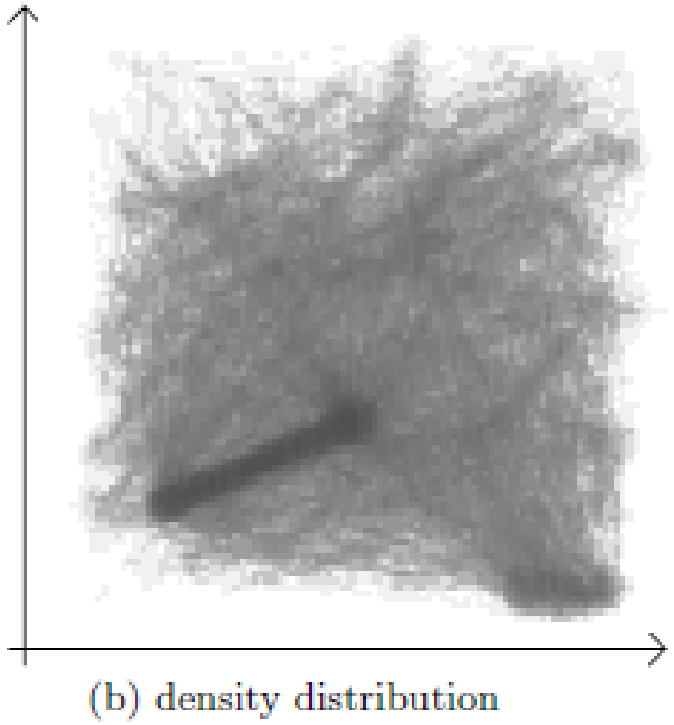
Regions of Interest

- Static pre-processed ROI
- ROI discovery
  - List of candidate places
  - Dynamic ROI
    - Popular points detection
    - ROI construction

Popular Points detection



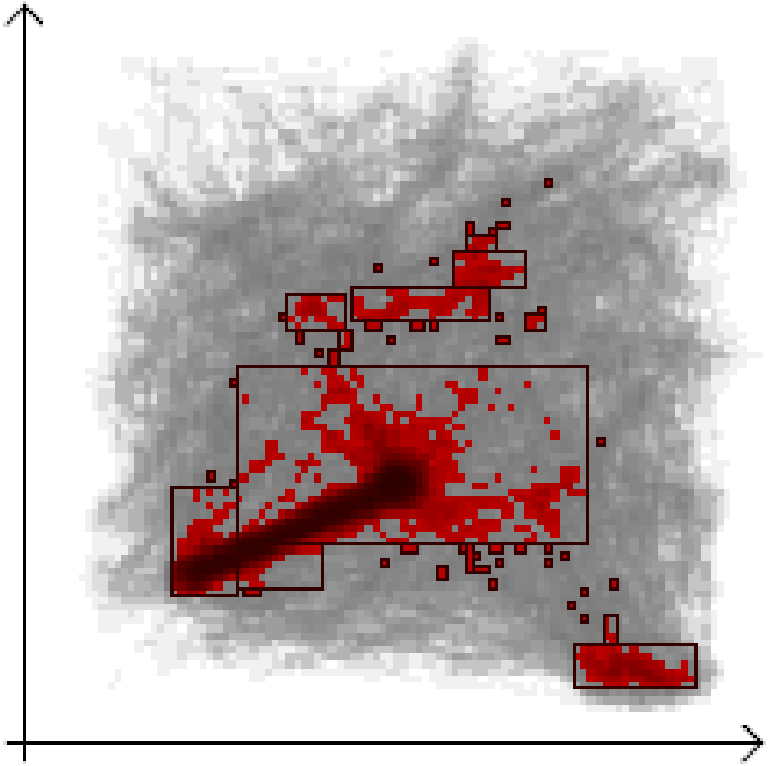
(a) input trajectories



(b) density distribution

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

ROI construction



(c) dense cells and extracted RoI

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

T-Pattern Mining (1)

- Two Approaches
  - Static
  - Dynamic
- Input: Trajectories and threshold parameters
  - Neighbourhood and time threshold
  - Support threshold
- Output: T-Patterns

T-Pattern Mining (2)

- Step-wise Heuristic
- Any frequent T-Pattern of length n+1 is the extension of some frequent T-Pattern of length n.

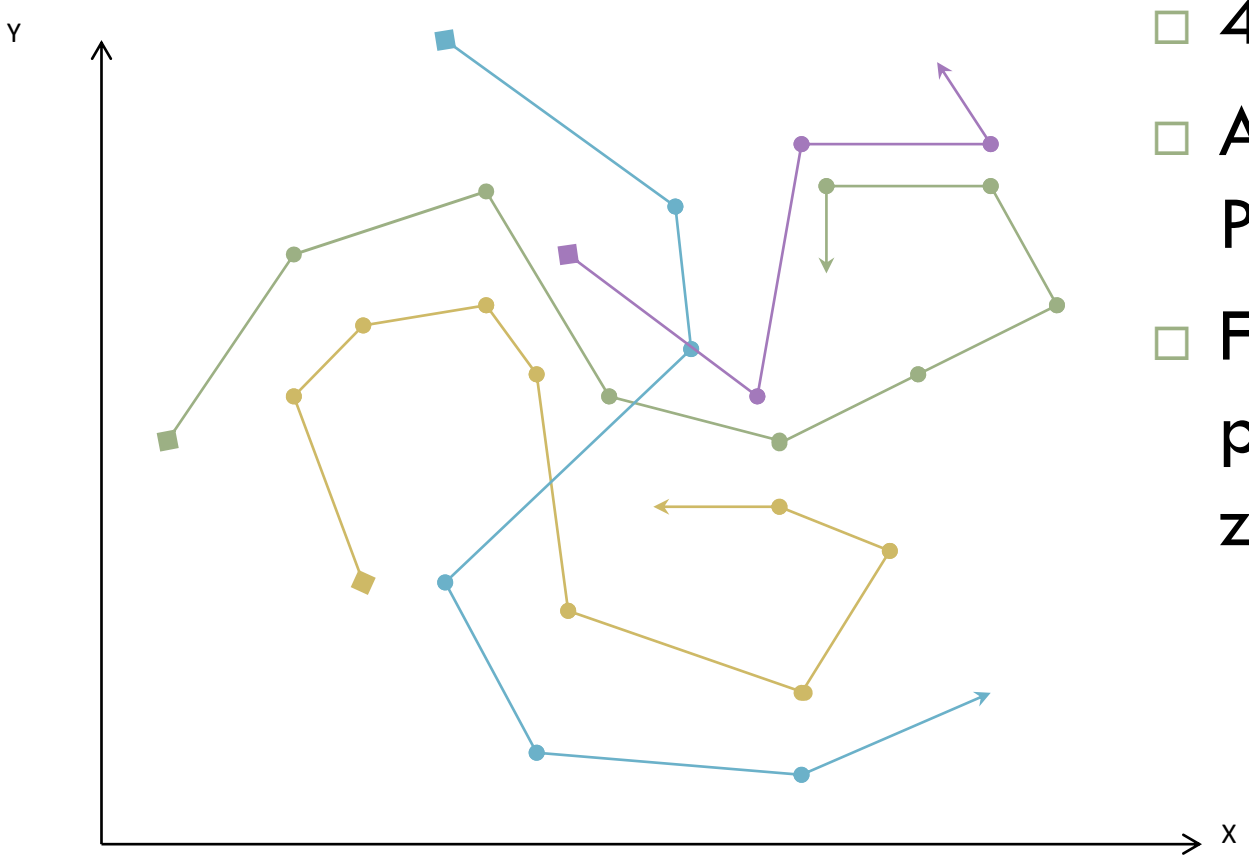
$$(x_0, y_0) \xrightarrow{\Delta t_1} \cdots \xrightarrow{\Delta t_{n+1}} (x_{n+1}, y_{n+1}) \preccurlyeq_{N,\tau} T$$

$$(x_0, y_0) \xrightarrow{\Delta t_1} \cdots \xrightarrow{\Delta t_n} (x_n, y_n) \preccurlyeq_{N,\tau} T$$

Static discovery of T-Patterns

- Compute ROIs over the whole trajectories.
- Covert (x, y, t) trajectories to (ROI, t)
- Find sequences of ROIs which satisfy support condition.

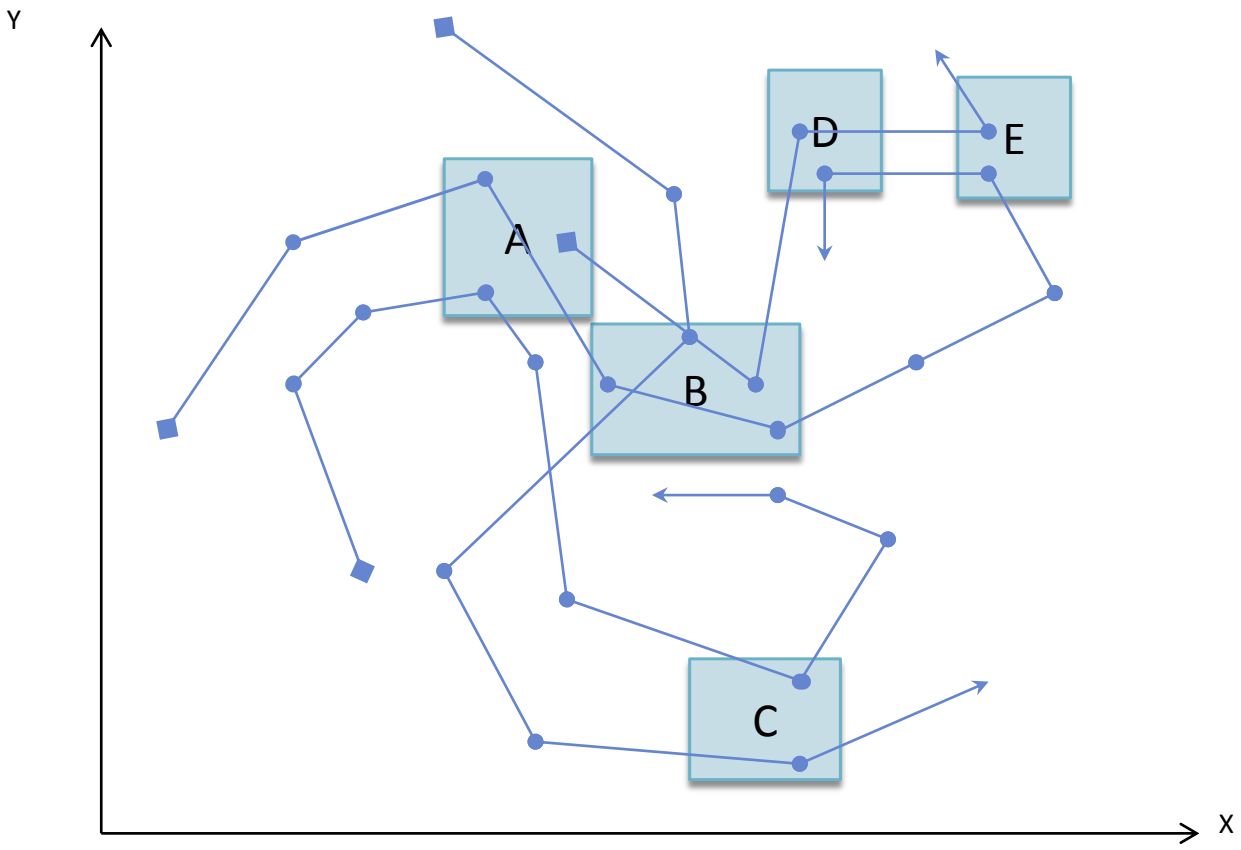
Dynamic Discovery of T-Patterns (1)



- 4 trajectories.
- Aim is to find T-Patterns.
- First iteration prefix length is zero.

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

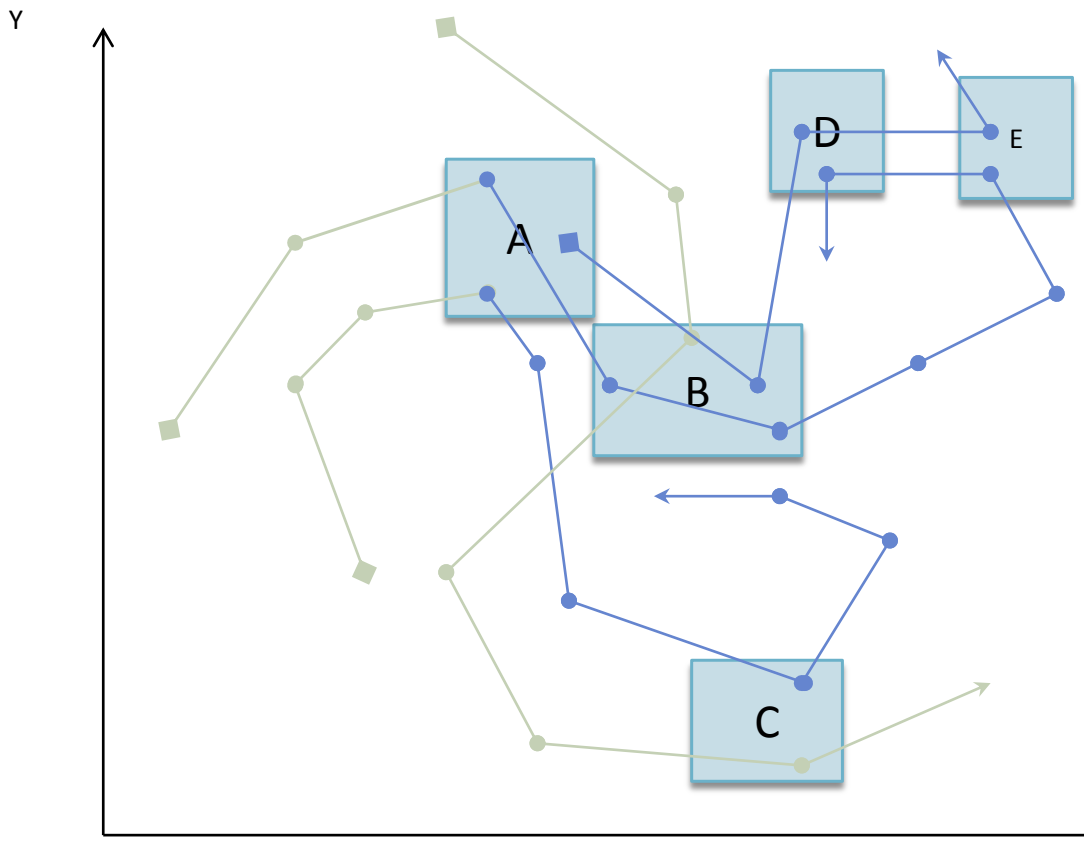
Dynamic Discovery of T-Patterns (2)



- Detect Regions of Interest
- e.g. *support(r)* ≥ 0.5
- For each region find possible projections to other regions

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (3)



- With respect to region A.
- Projections:

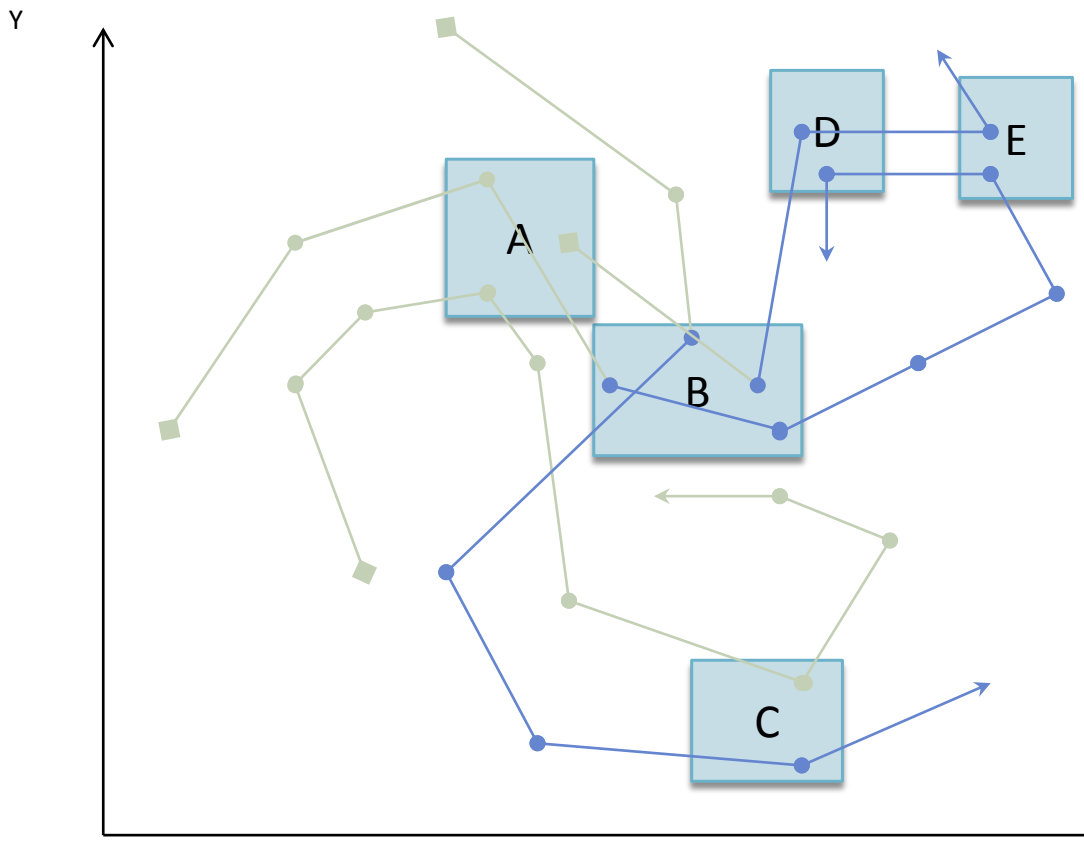
  A → B

  A → C
- Add new configuration

  (T', <A>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (4)



- With respect to region B.
- Projections:

  B → C

  B → D

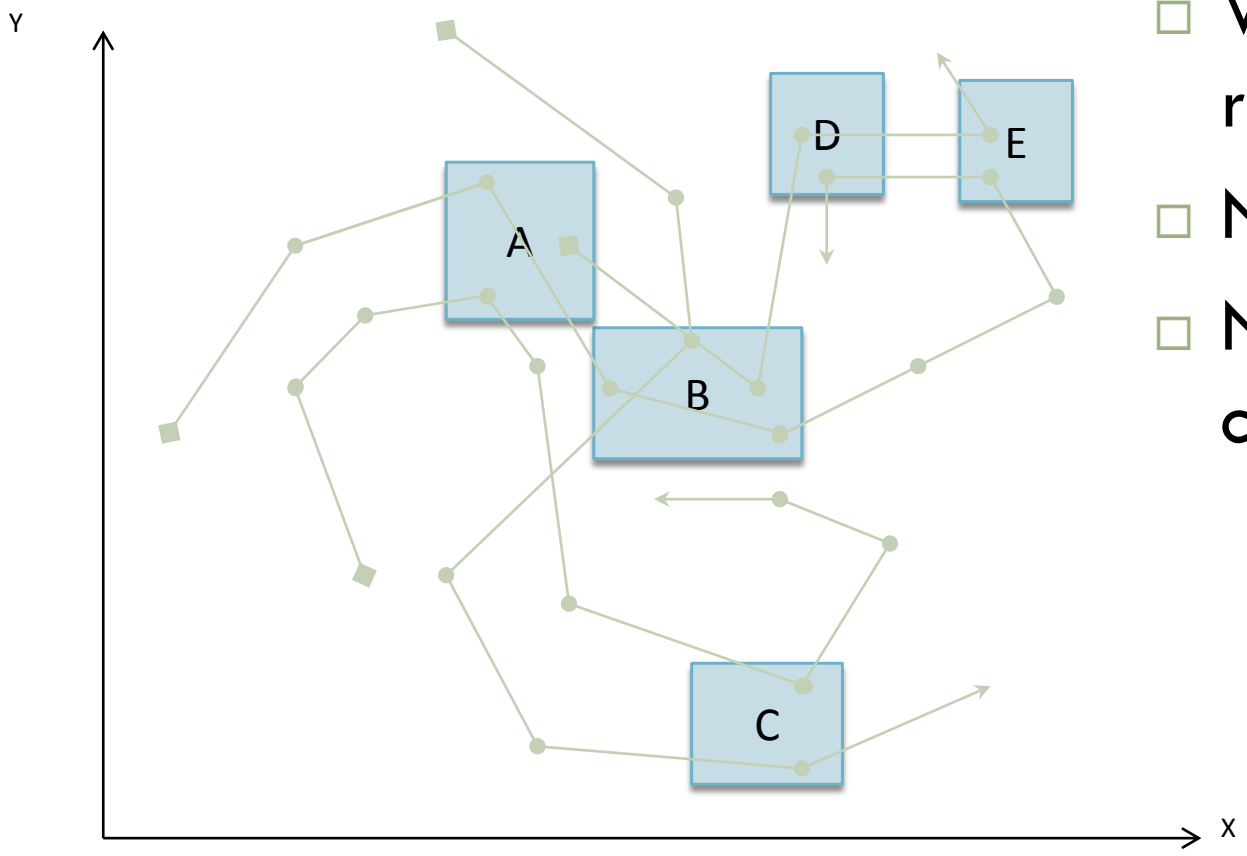
  B → E

- Add new configuration

  (T', <B>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

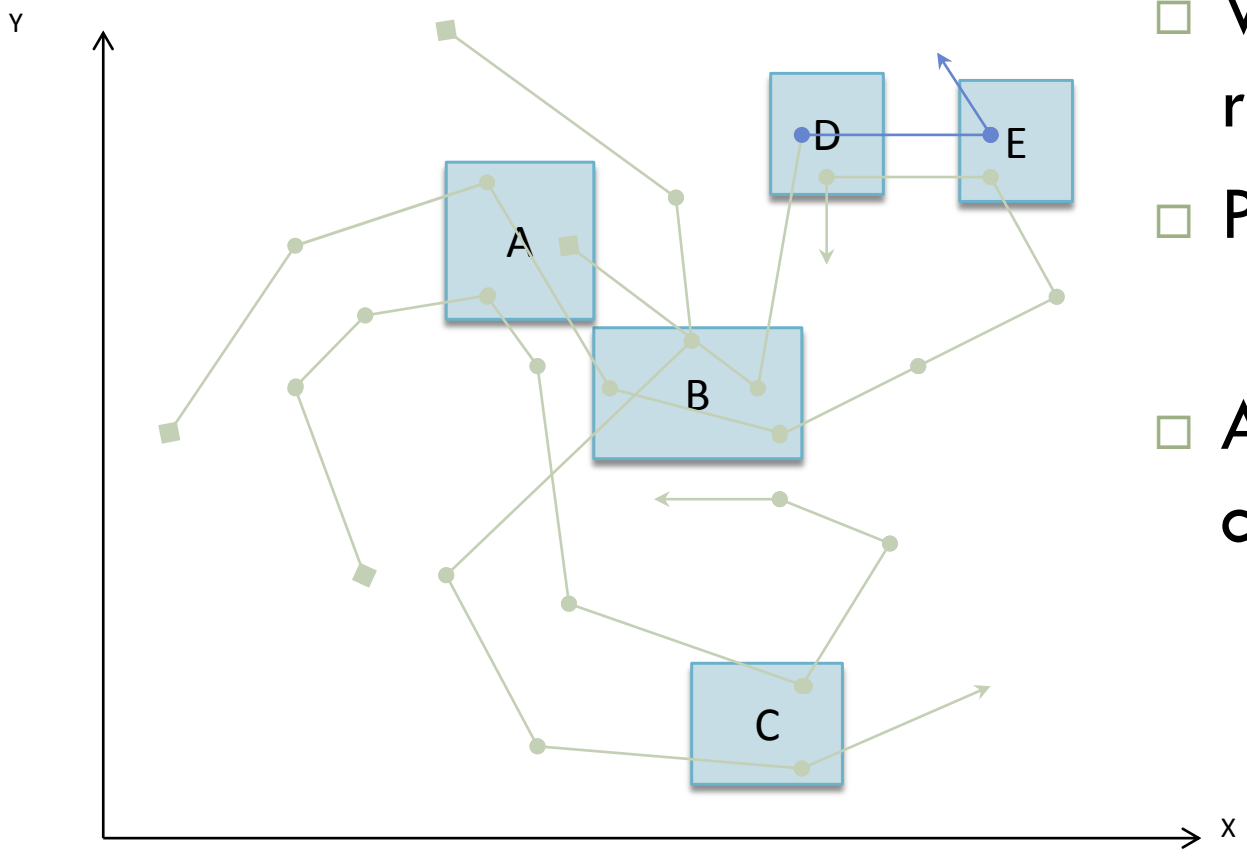
Dynamic Discovery of T-Patterns (5)



- With respect to region C.
- No Projections
- No new configuration

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (6)



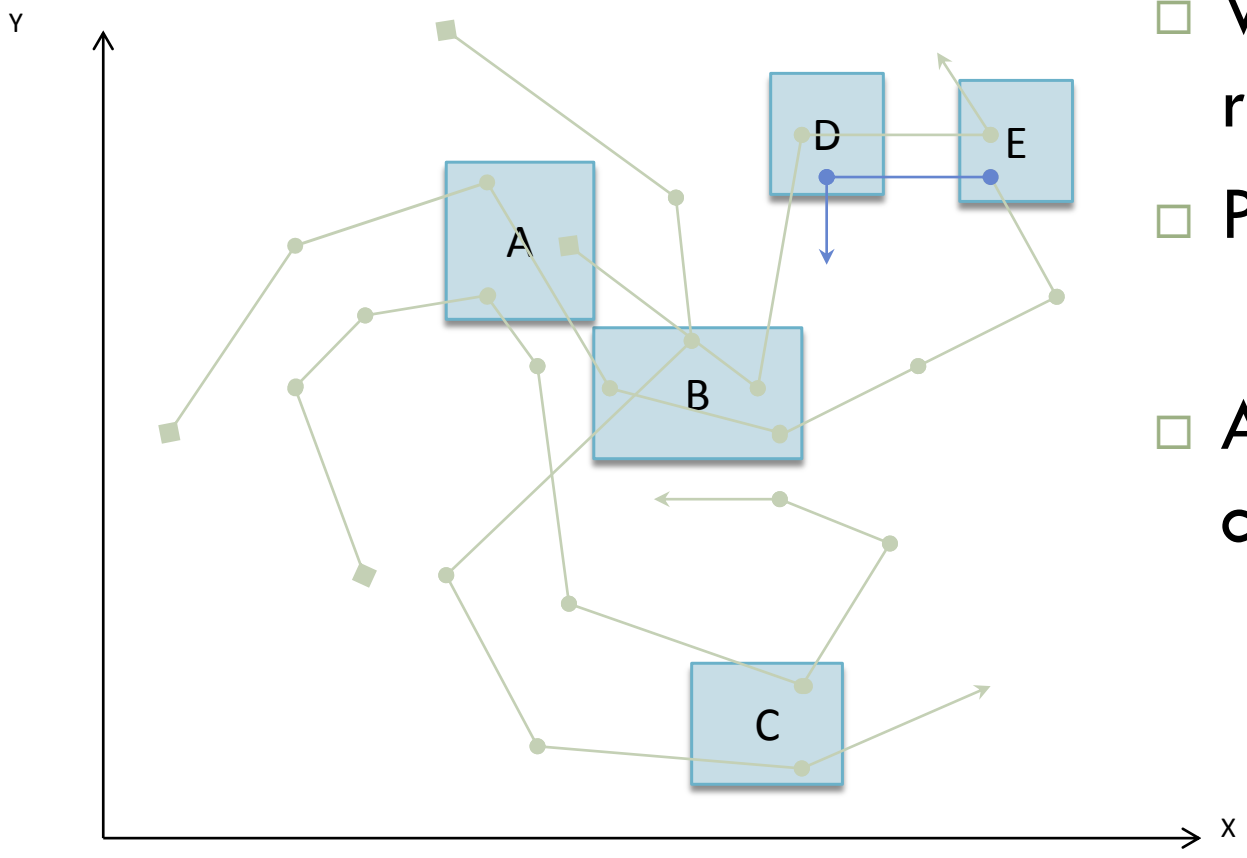
- With respect to region D.
- Projection

  D → E
- Add new configuration

  (T', <D>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (7)



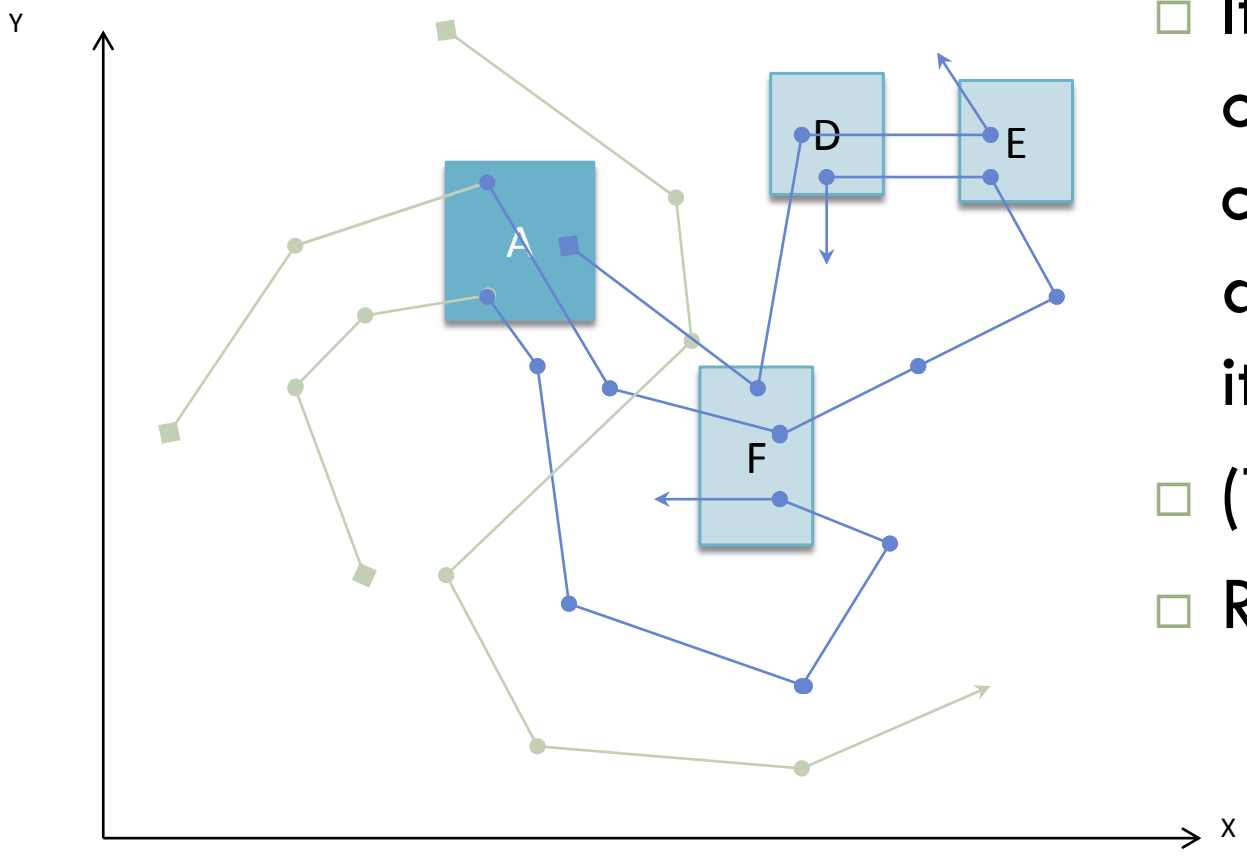
- With respect to region E.
- Projection

  E → D
- Add new configuration

  (T', <E>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

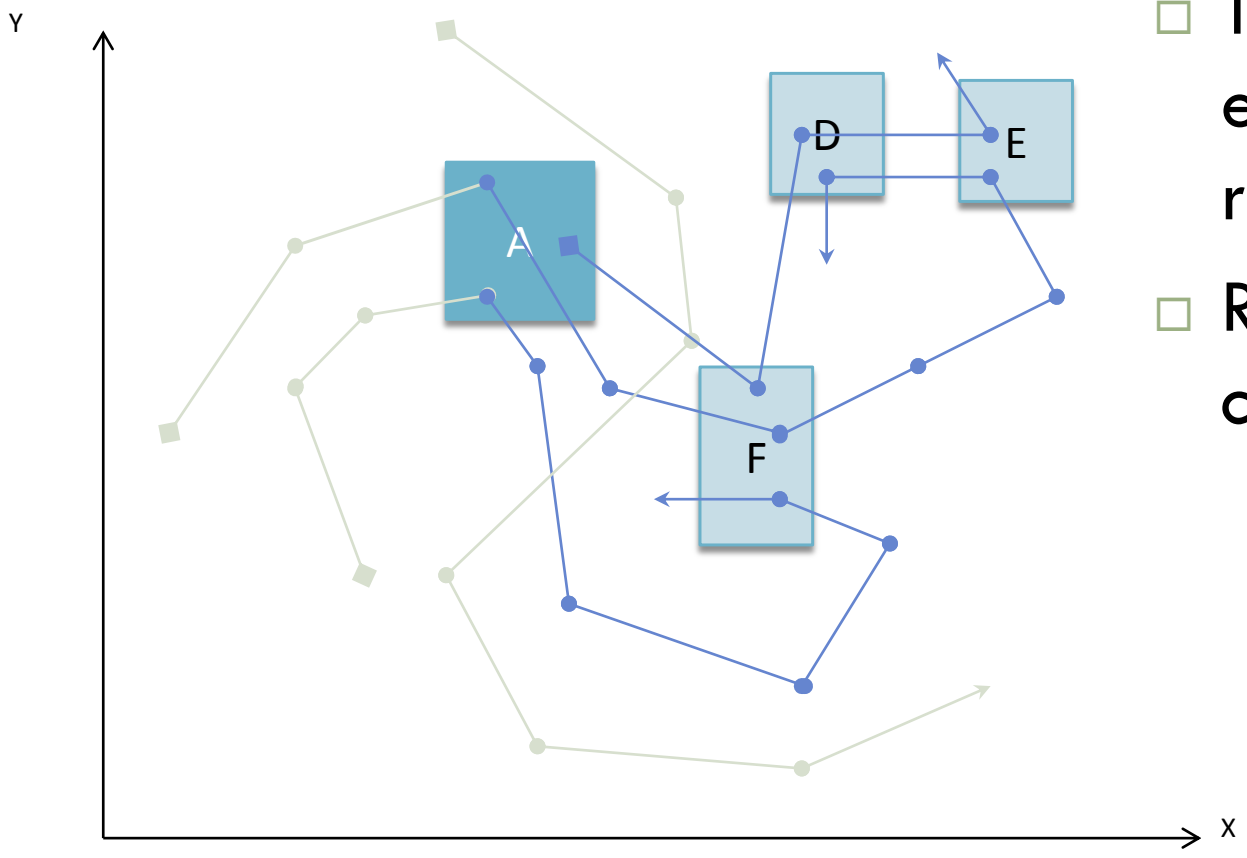
Dynamic Discovery of T-Patterns (8)



- Iteration 2 works on each configurations added in prior iteration.
- (T', <A>)
- Recompute ROIs.

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (9)

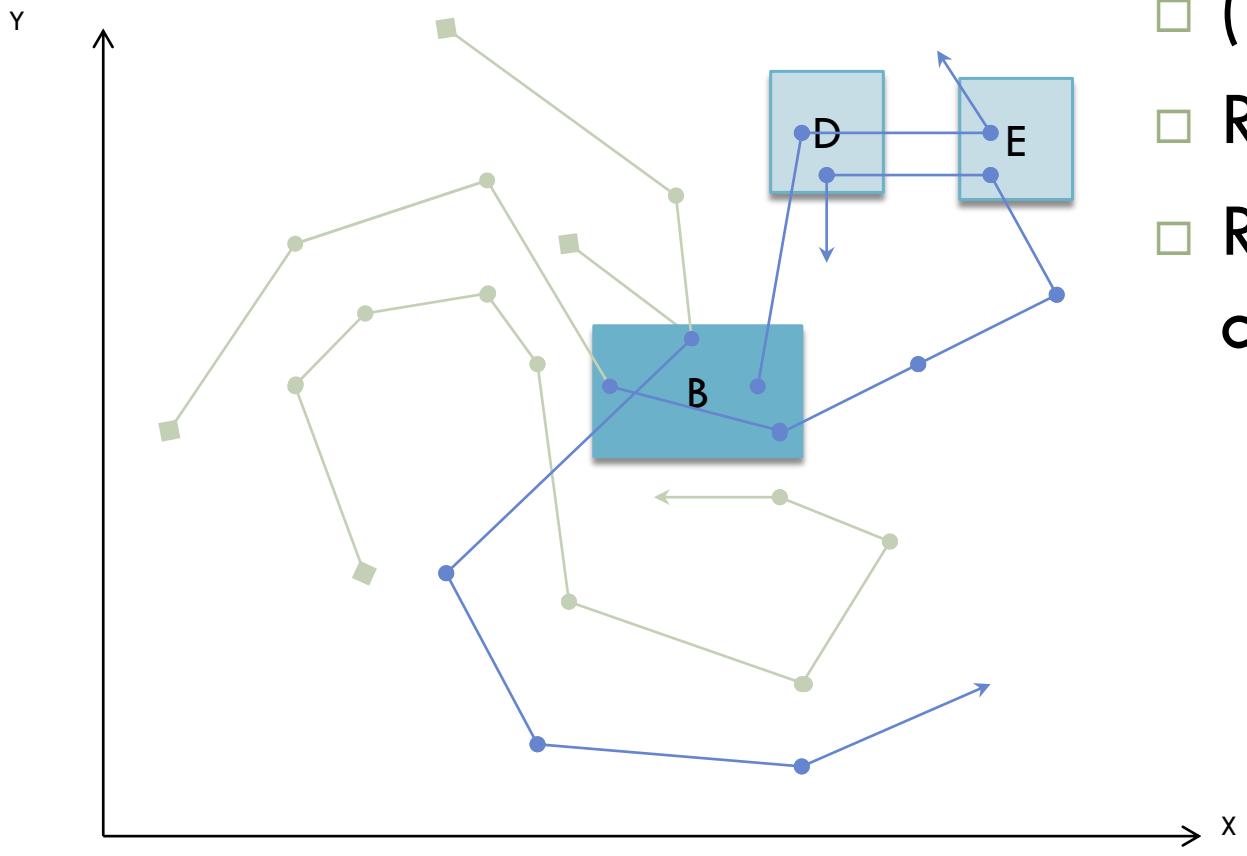


- Test projects from each of the regions.
- Results in configuration

  (T', <A→F>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (11)



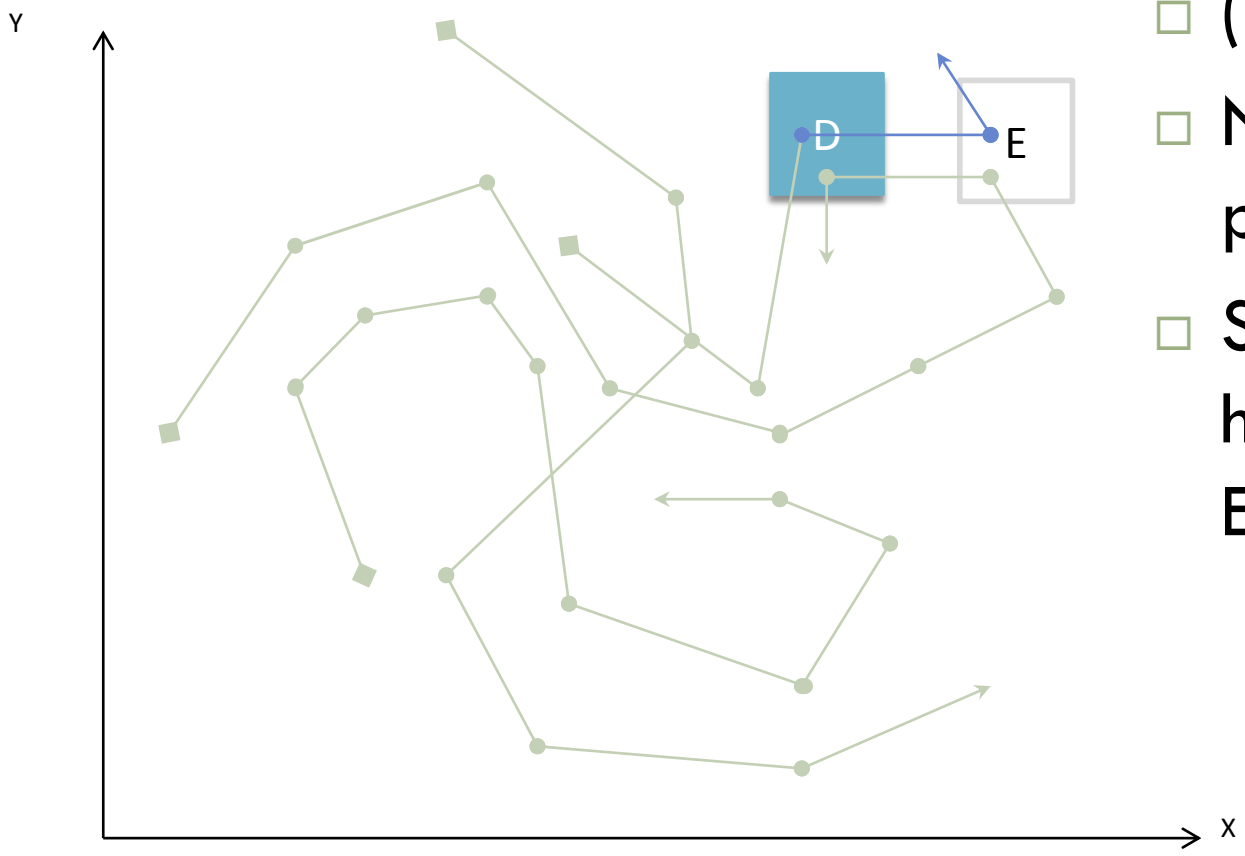
- (T', <B>)
- Recompute ROIs.
- Results in configuration

  (T', <B→D>)

  (T'', <B→E>)

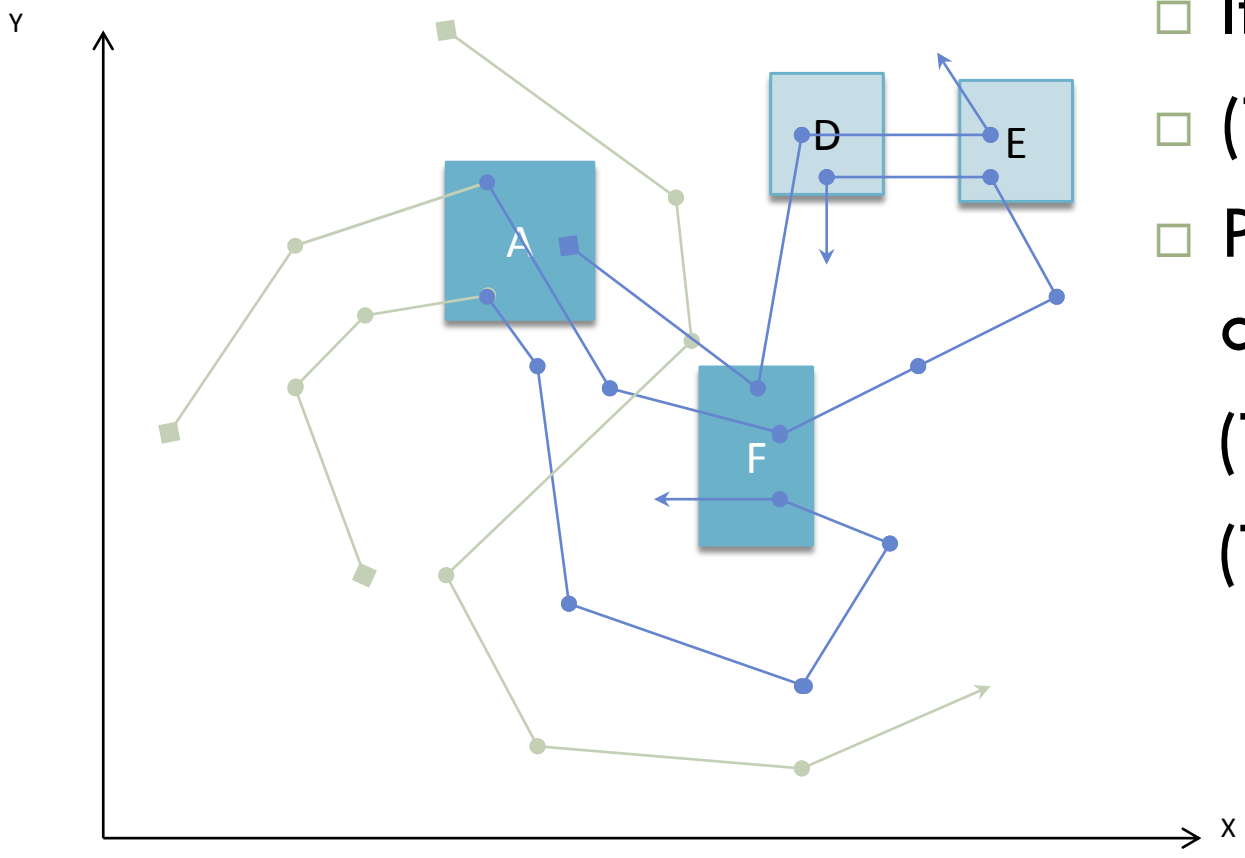
Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Dynamic Discovery of T-Patterns (12)



- (T', <D>)
- No possible projections.
- Same thing happens for region E.

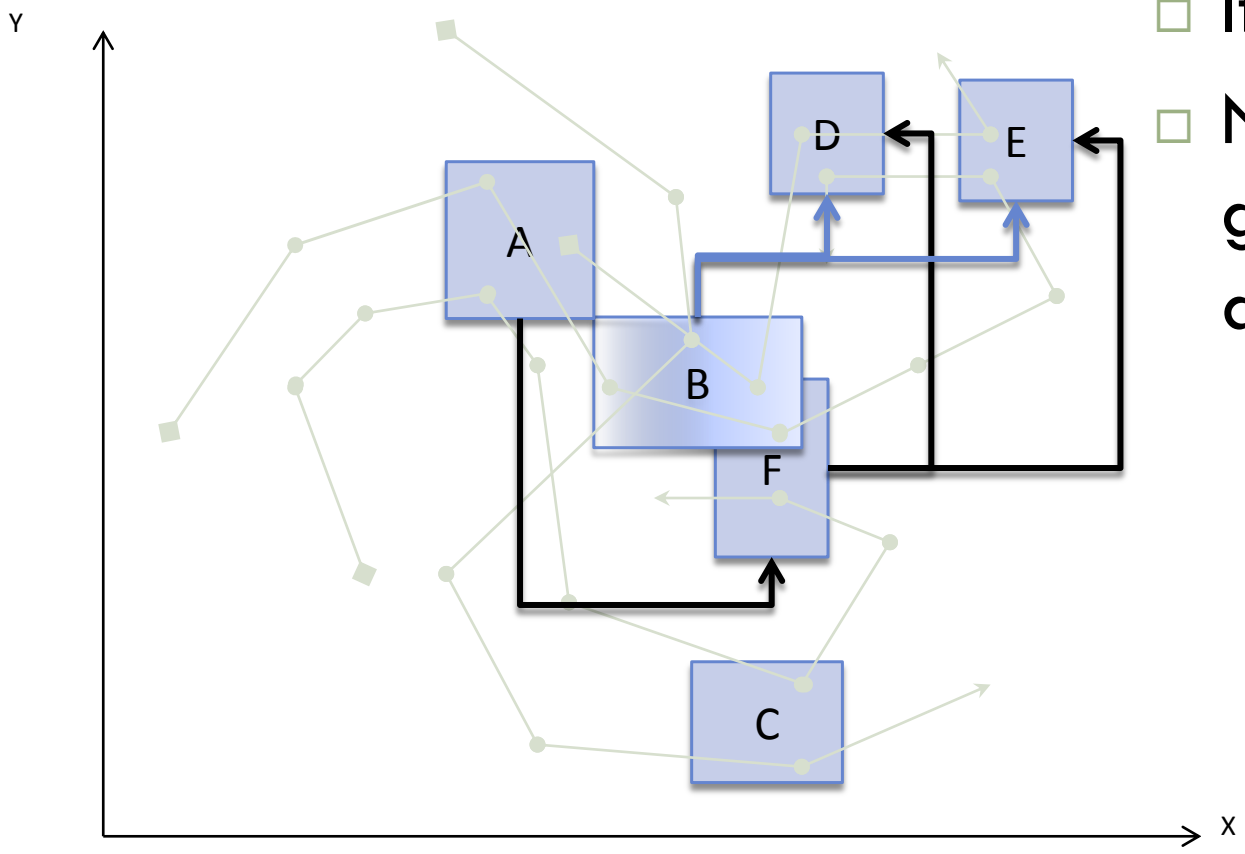
Dynamic Discovery of T-Patterns (11)



- Iteration 3
- (T', <A→F>)
- Produces configurations
  (T', <A→F→D>)
  (T', <A→F→E>)

Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

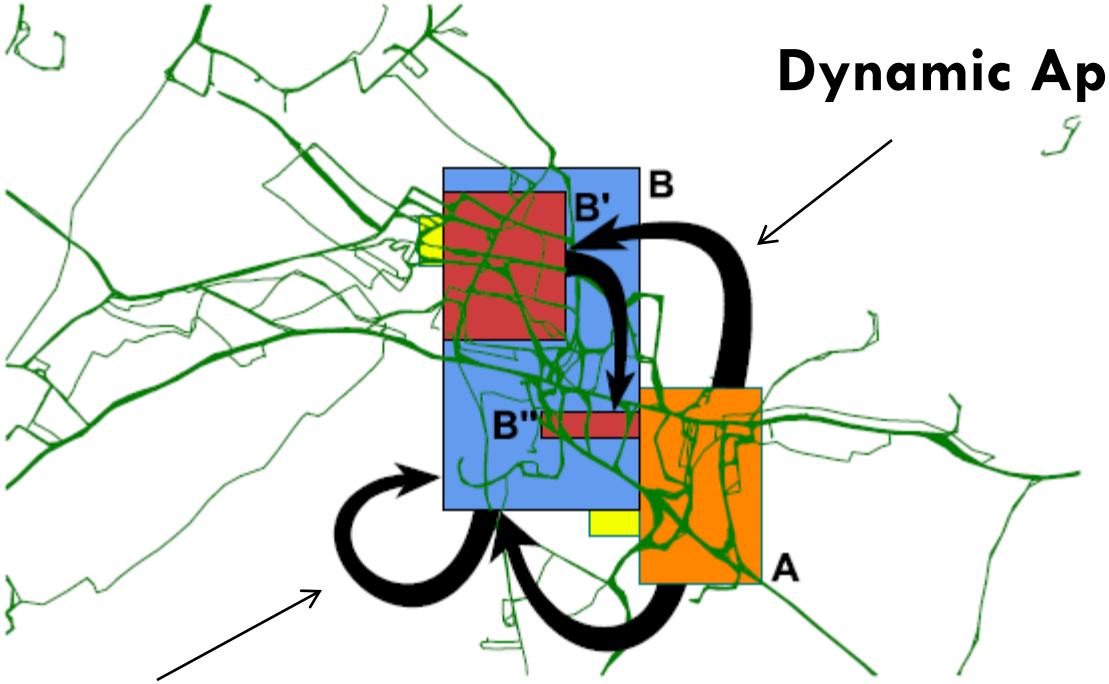
Dynamic Discovery of T-Patterns (11)



- Iteration *4*
- No more paths to go therefore stops and outputs paths:
  - A → F → D
  - A → F → E
  - B → D
  - B → E

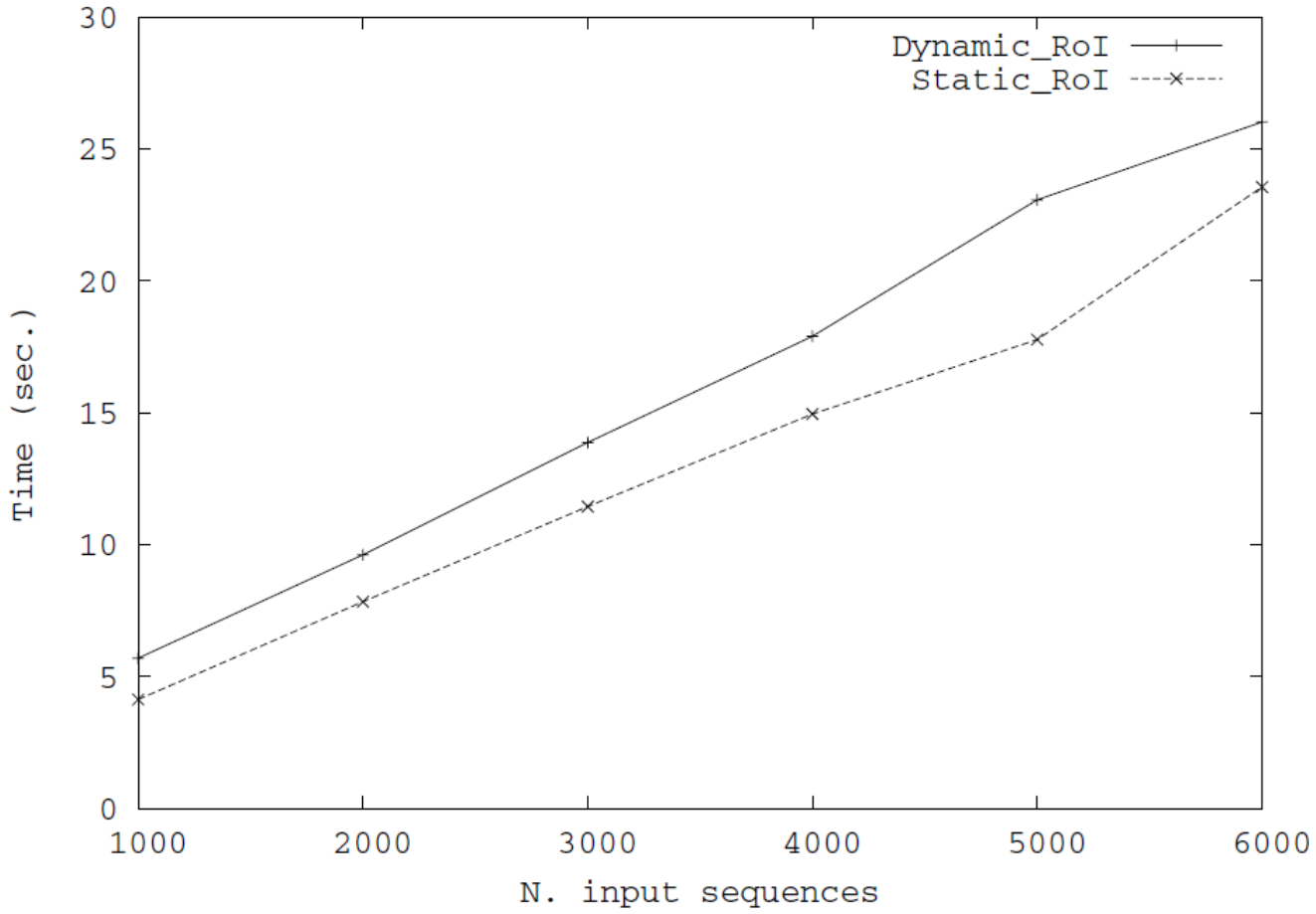
Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Results



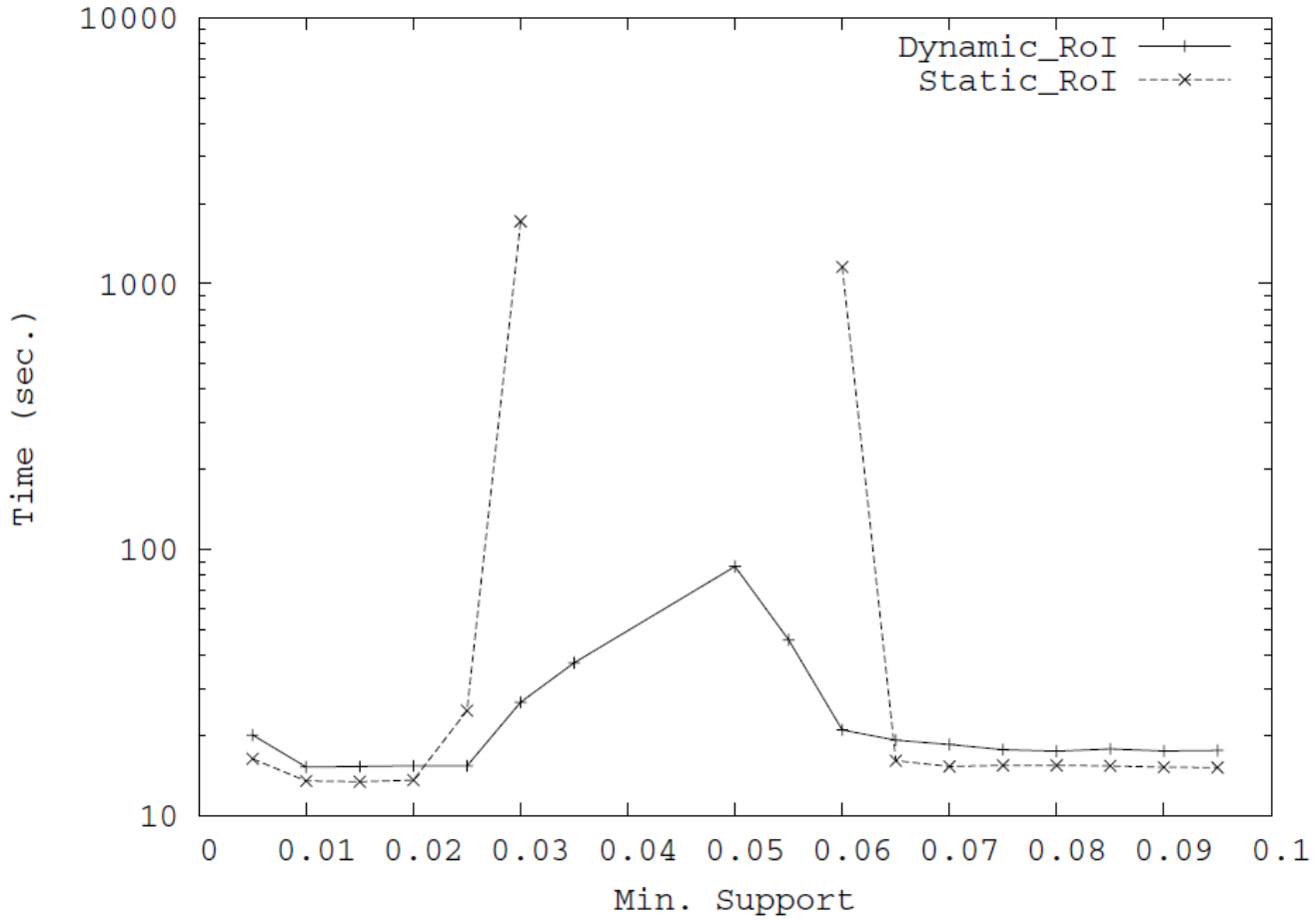
Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Results



Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

Results



Trajectory Pattern Mining: Fosca Giannotti, Micro Nanni, Dino Pedreschi, Fabio Pinelli

To conclude

- Lack of comparative results
- Idea of ROI is good for applications looking for simple and course results.
- Cited many times for ROI discovery algorithm

- *Where Next*: T-Patterns used to predict where a person will move next, given the current sequence.
- DAEDALUS: Extends SQL to enable TAS queries.

References

- Giannotti, F., Nanni, M., Pinelli, F., and Pedreschi, D. 2007. Trajectory pattern mining. http://doi.acm.org/10.1145/1281192.1281230
- F. Giannotti, M. Nanni, and D. Pedreschi. Efficient mining of sequences with temporal annotations http://www.siam.org/meetings/sdm06/proceedings/032giannottif.pdf
- Monreale, A., Pinelli, F., Trasarti, R., and Giannotti, F. 2009. WhereNext: a location predictor on trajectory pattern mining http://doi.acm.org/10.1145/1557019.1557091
- Ortale, R., Ritacco, E., Pelekis, N., Trasarti, R., Costa, G., Giannotti, F., Manco, G., Renso, C., and Theodoridis, Y. 2008. The DAEDALUS framework: progressive querying and mining of movement data. http://doi.acm.org/10.1145/1463434.1463497

Thanks!