Analytic Operators for Trajectories

ESR2.4 Song WU, started in September 2021

Home Supervisors (ULB): Esteban Zimányi, Mahmoud Sakr
Host Supervisor (AAU): Kristian Torp

Outline

Introduction to trajectory data

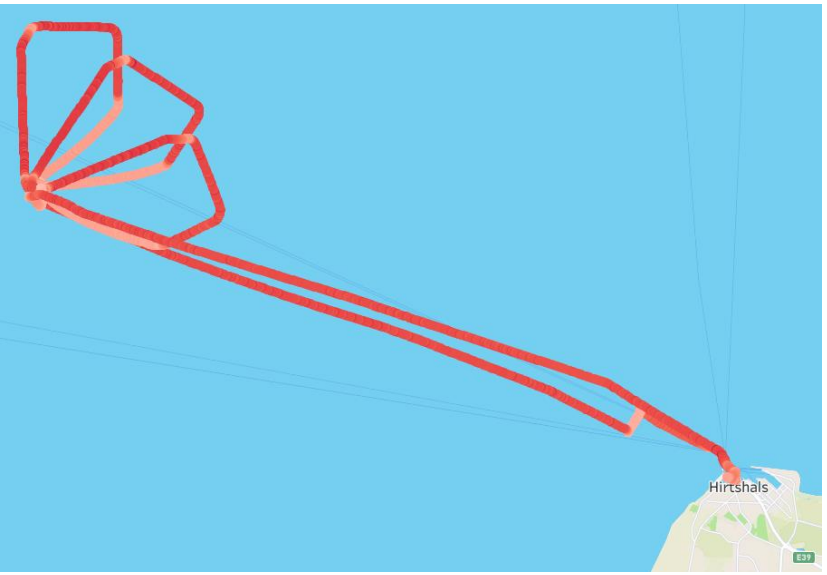
Trajectory Similarity Measures

Semantic Trajectory Segmentation

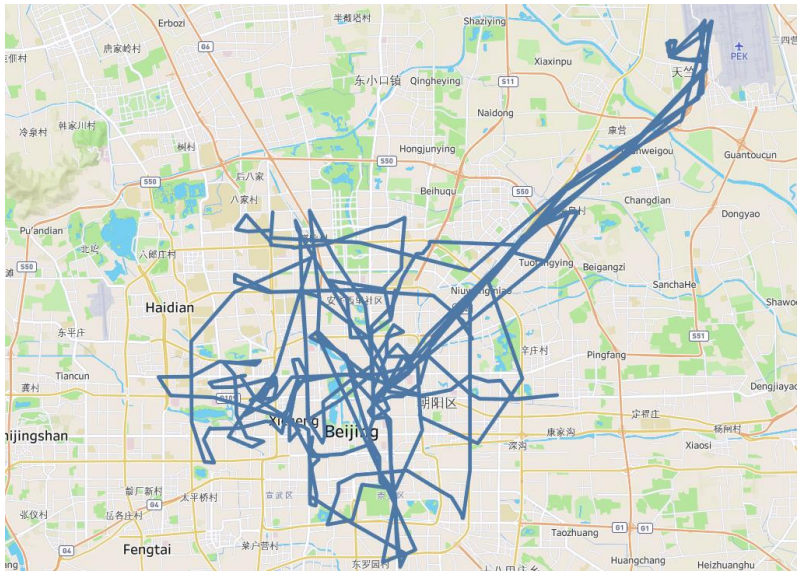
Research goals

Trajectory data is generated in many domains:

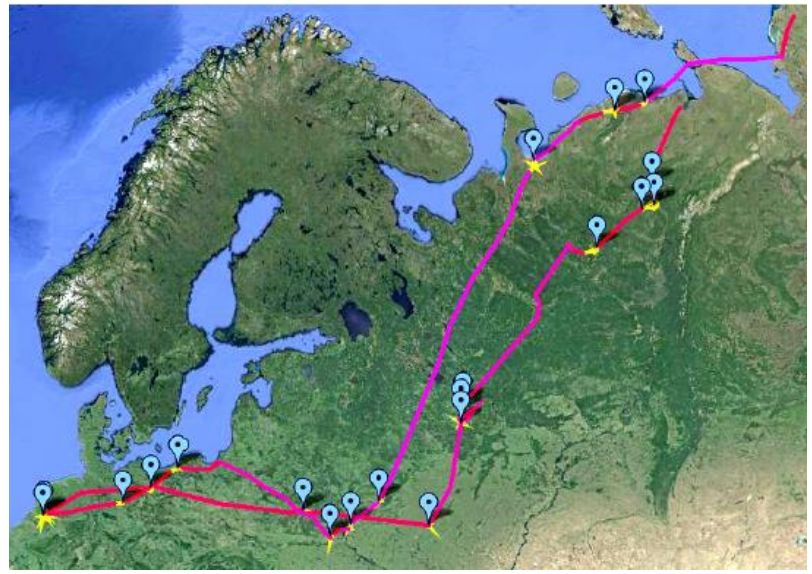
maritime, urban, animal behavior, etc.



ship trajectory



taxi trajectory

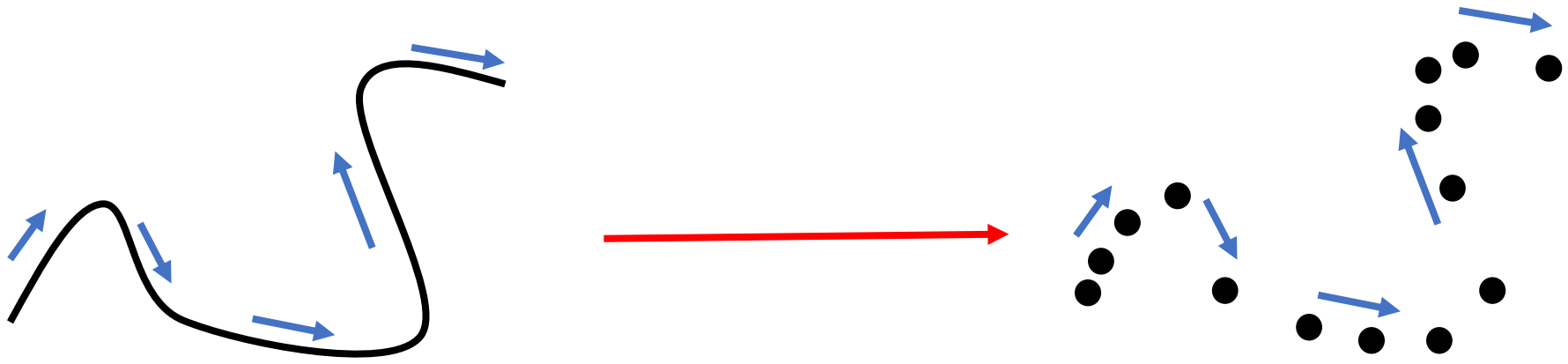


bird trajectory

How trajectory data is collected?

In the physical world,
the real movement of an object
is a continuous mapping from
time to space

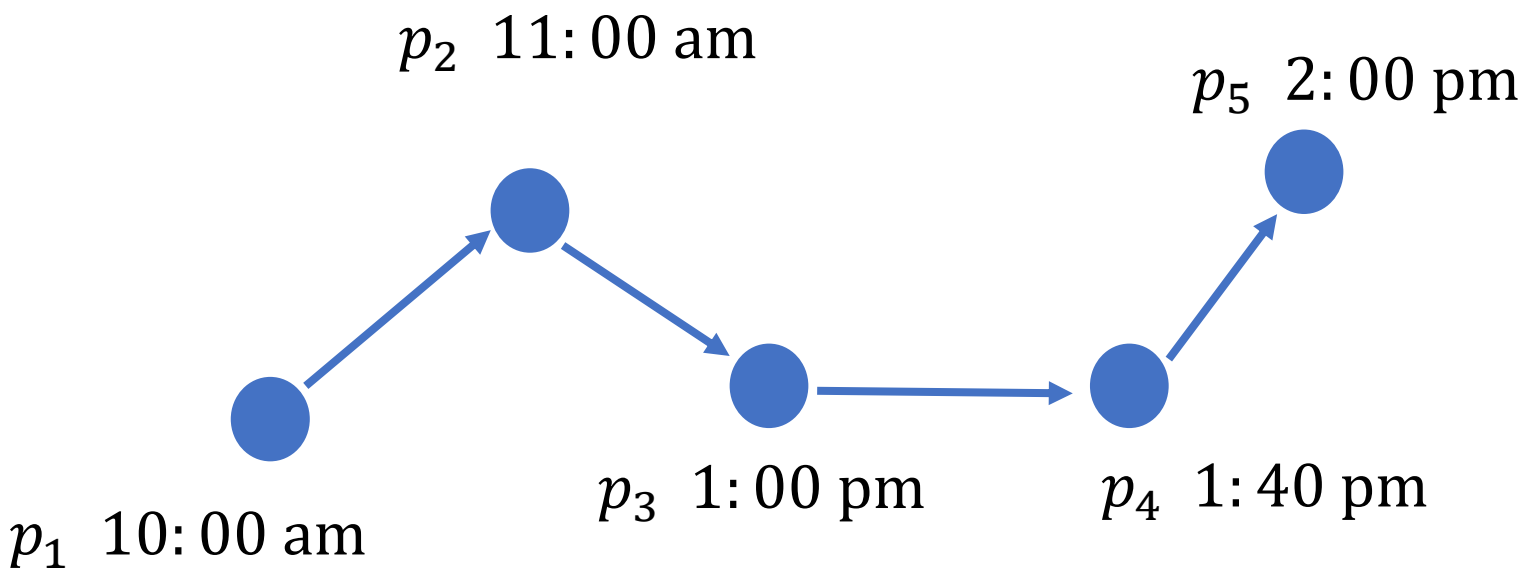
Due to technology limitations,
such a continuous curve is usually
recorded in a discrete way



Two common ways to model trajectories in the literature

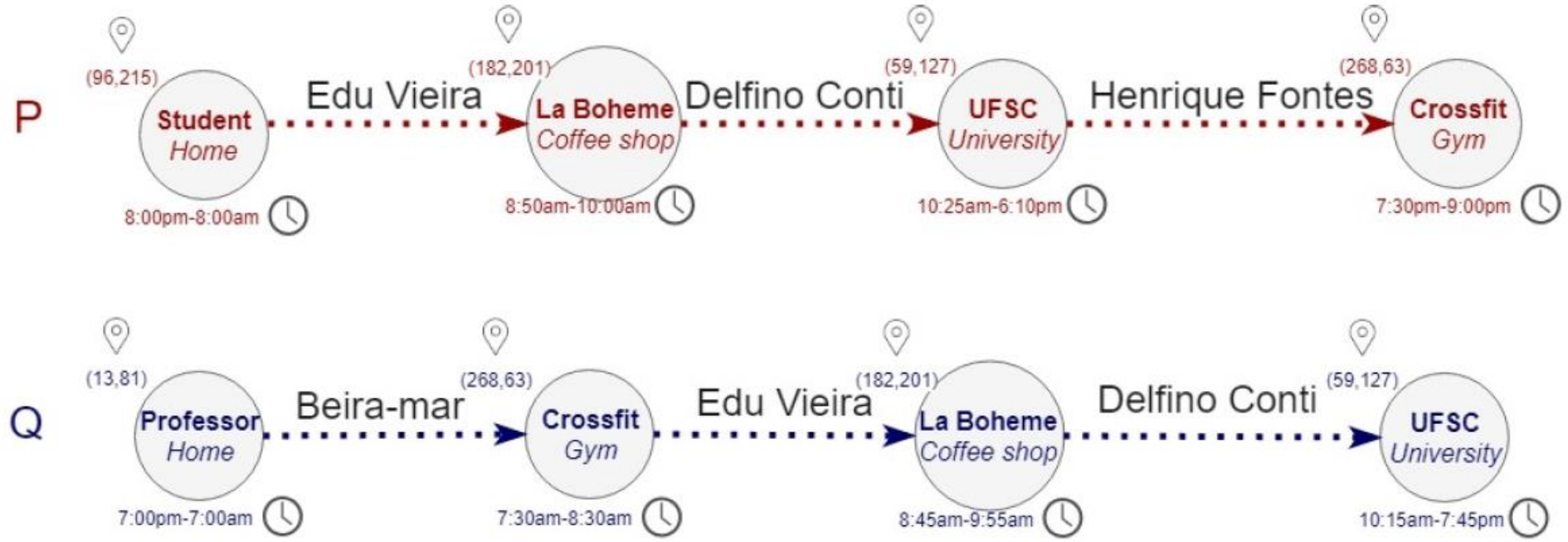
raw trajectory: an ordered sequence of time-stamped points

i.e. $T =< p_1, p_2, \ldots p_n >$, where $p_i =< t_i, x_i, y_i >$



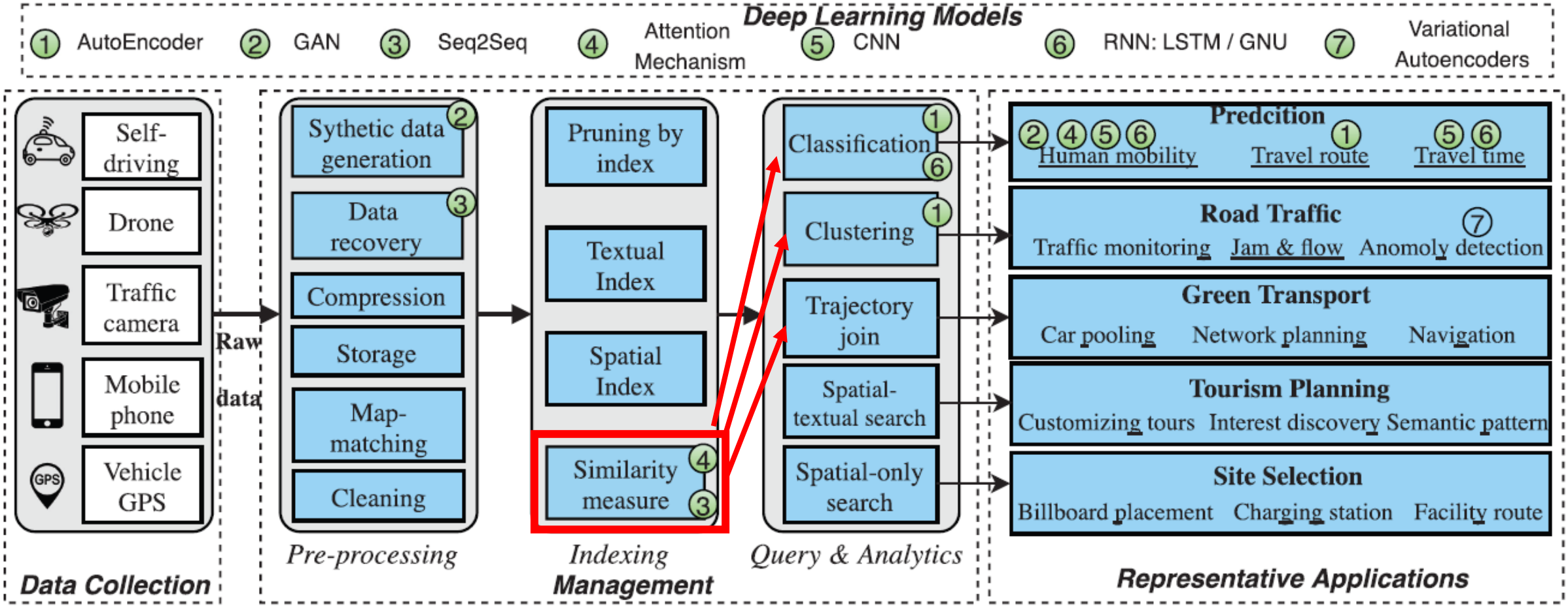
Two common ways to model trajectories in the literature

semantic trajectory: a trajectory is enriched/integrated with context/semantic information, e.g. POI category, transportation mode.



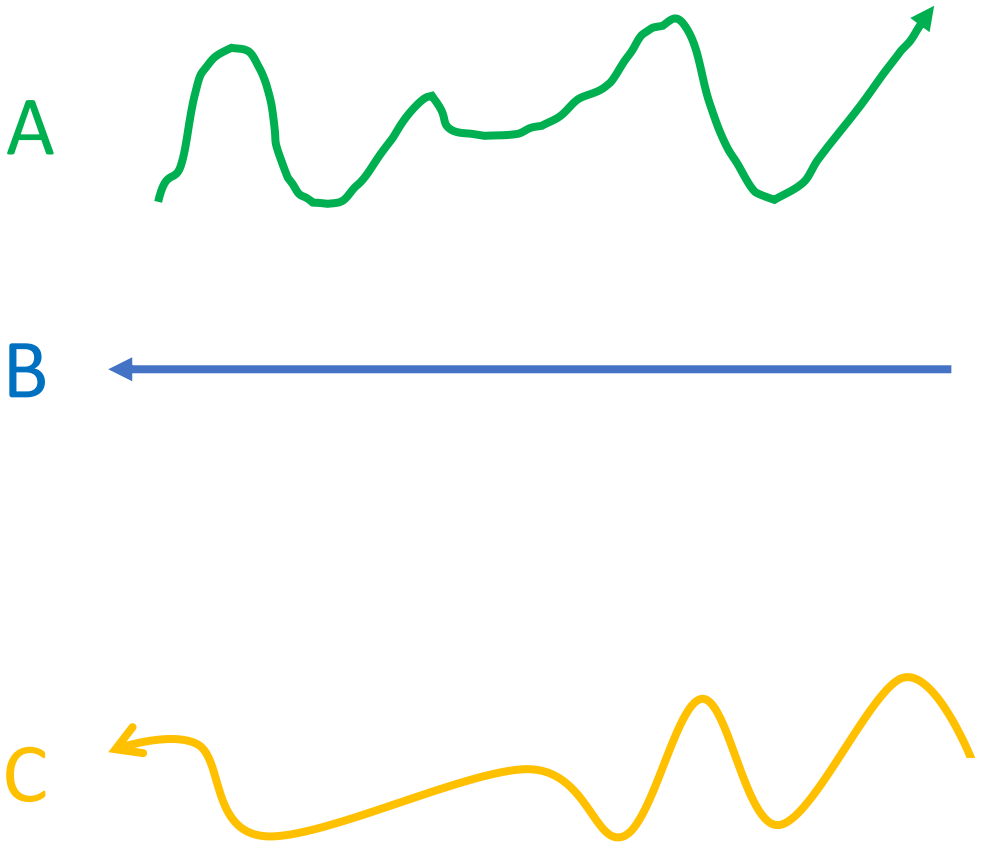
An example of semantic trajectory (taken from [6])

What are the applications of trajectory data?



Overview of typical trajectory-related tasks & applications (taken from [9])

Given two trajectories P and Q, a trajectory similarity measure S evaluates how similar P and Q are to each other.



The most similar pair of trajectories is

- A and B, in terms of spatial proximity
- A and C, in terms of trajectory shape
- B and C, in terms of moving direction

Many trajectory similarity measures have been proposed.

For raw trajectories:

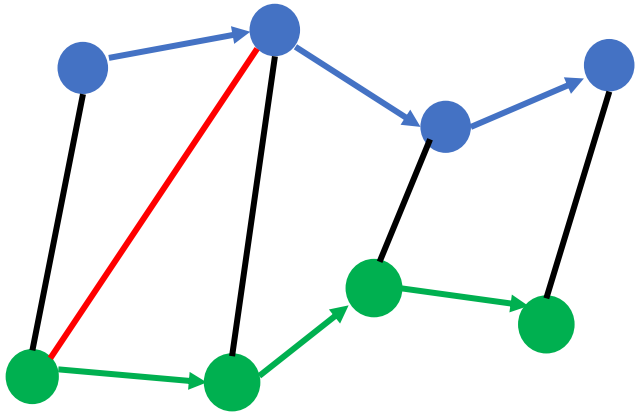
- Merge Distance [1]
- Uncertain Movement Similarity [2]
- synchronized spatial-temporal trajectory similarity [3]
- Normalized Weighted Edit Distance [4]
- Edit Distance with Projections [5]
- …

For semantic trajectories:

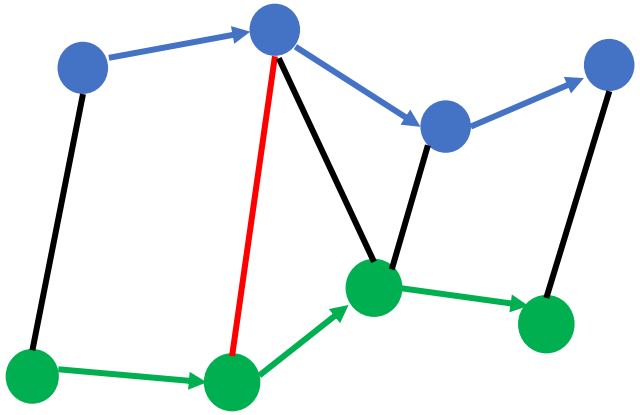
- Stops and Moves Similarity Measure [6]
- Multiple-Aspect Trajectory Similarity [7]
- Trajectory Forest Similarity [8]
- …

Each of them has its own pros and cons

Discrete Fréchet Distance: the minimum length of the leash required for both the person and the dog to move forward



The first matching

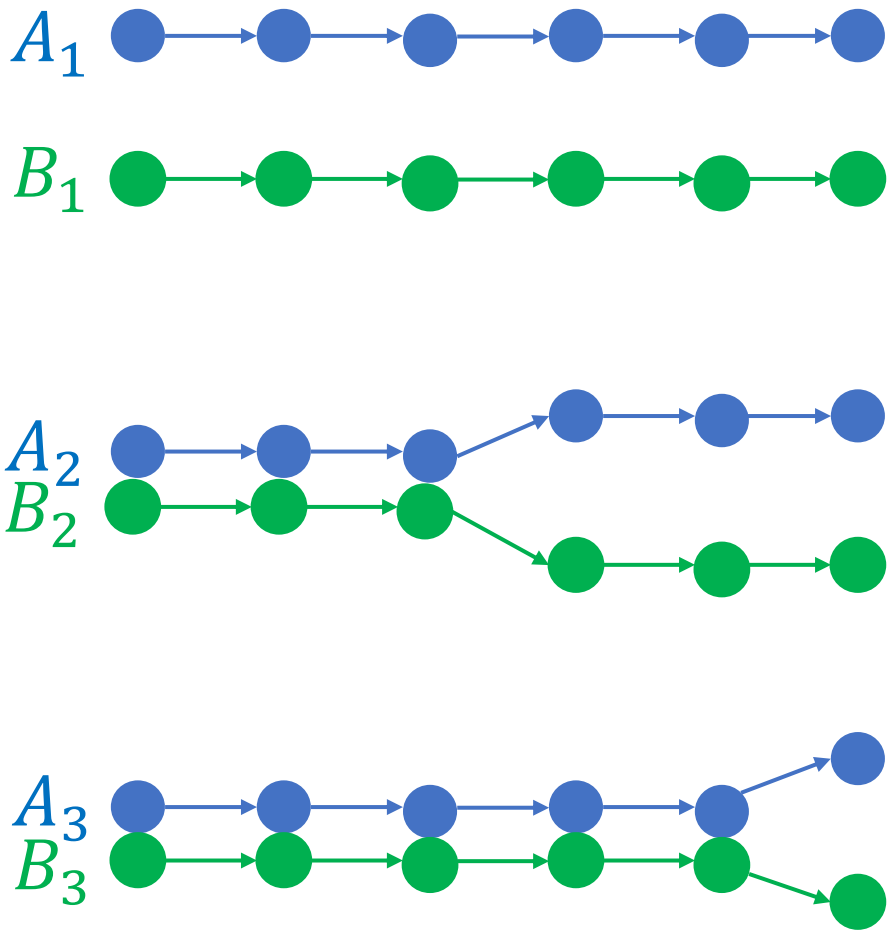


The second matching

So on…

The shortest length of the red segment among all possible matchings

One weakness of the Discrete Fréchet Distance



Under Discrete Fréchet Distance,
A_1 and B_1, A_2 and B_2, A_3 and B_3 have the same distances.

How to compare different similarity measures?

- robustness to noise / sampling rate / trajectory size
- is a metric or not
- the order of points is respected or not
- temporal information is used or not
- discrete or continuous
- matching strategy: all points need to be matched or not
- …

What are the main differences between ship trajectories and car trajectories?

No road network in the sea, so usually map-matching is not needed.

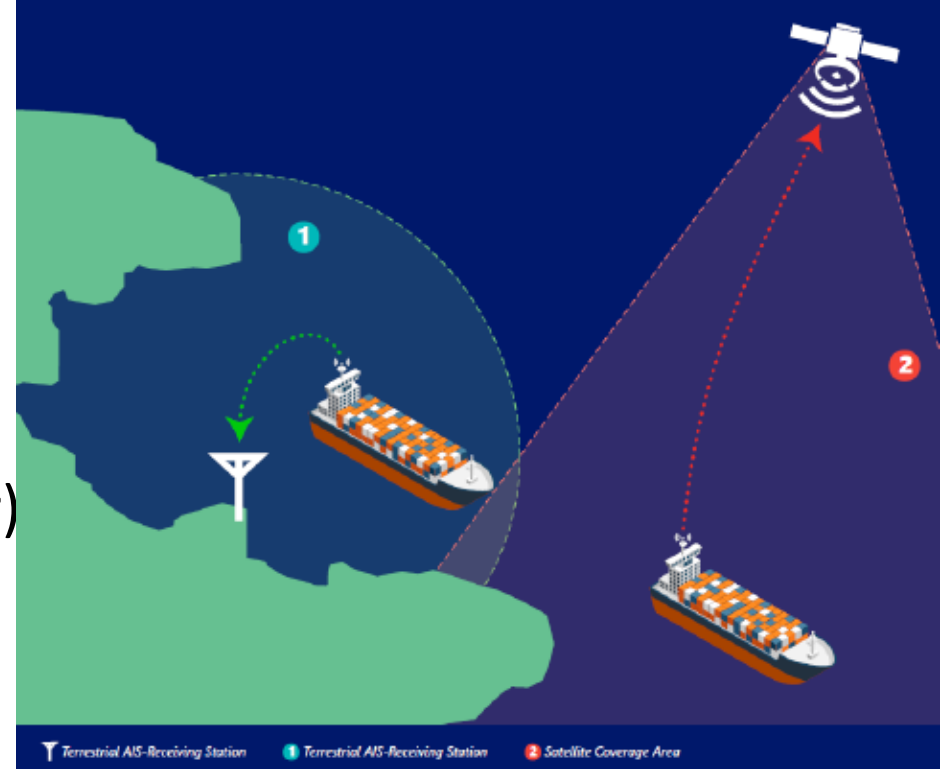
No traffic lights.

Movement of vehicles is constrained by surrounding vehicles.

Vessel trajectories usually cover a larger spatial extent.

What is the data source for ship trajectory?

- **Automatic Identification System (AIS)**



- **static and voyage-related information**

  MMSI (Maritime Mobile Service Identification number)
  ETA (Estimated Time of Arrival)
  Destination, Ship Type, Navigational Status, etc.

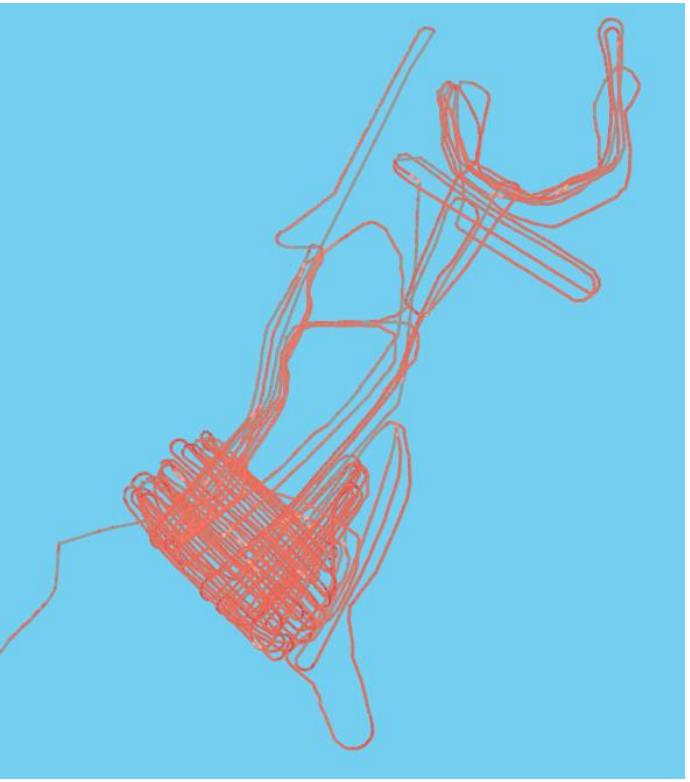
- **dynamic information**:

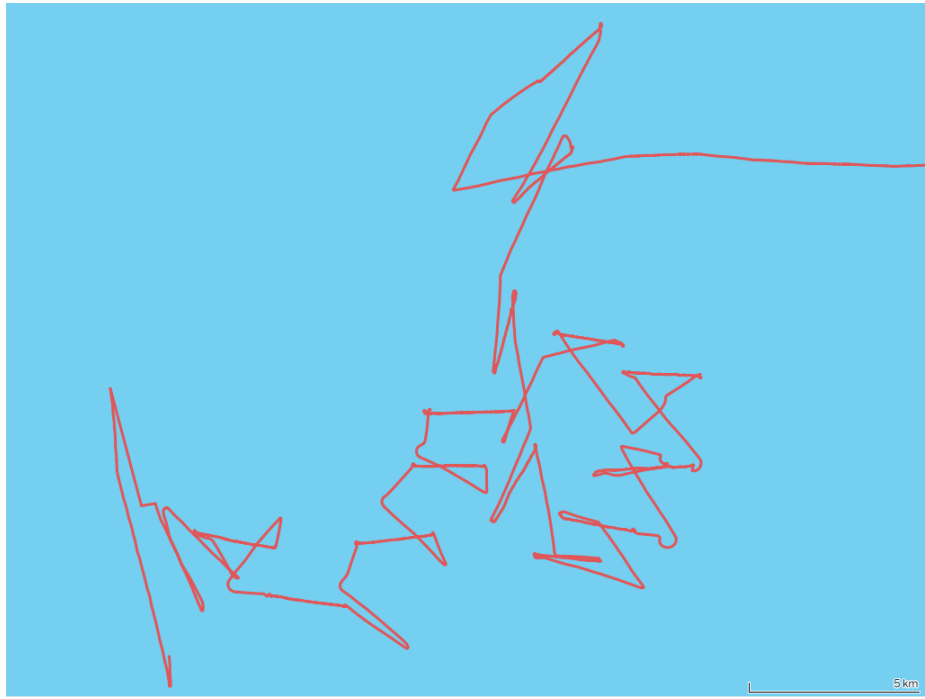
  vessel position, heading, speed, etc.

can be manipulated by humans

How can trajectory similarity measures be used in maritime domain?

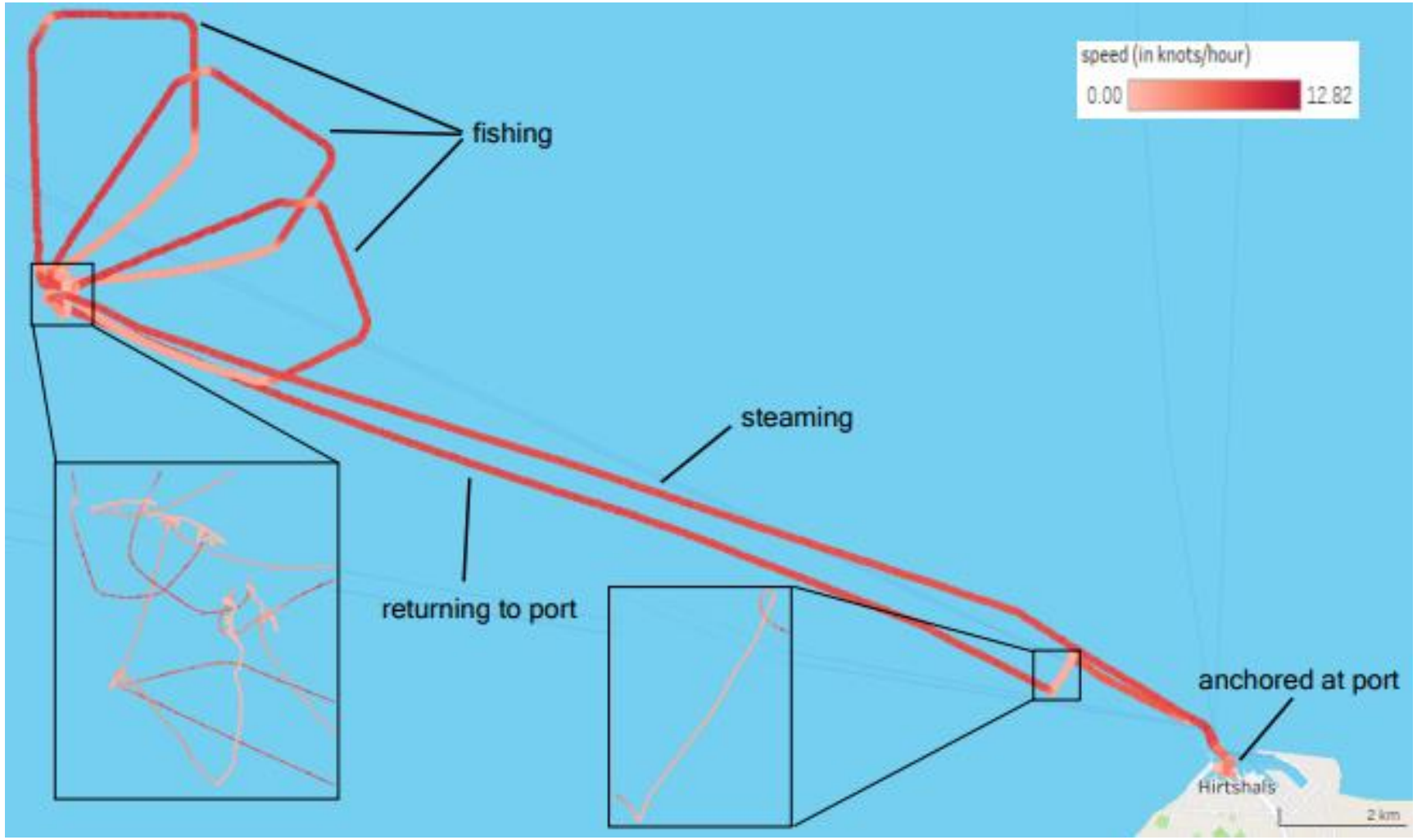
- retrieve similar trajectories in terms of fishing methods
- can help detect illegal, unregulated, and unreported fishing.



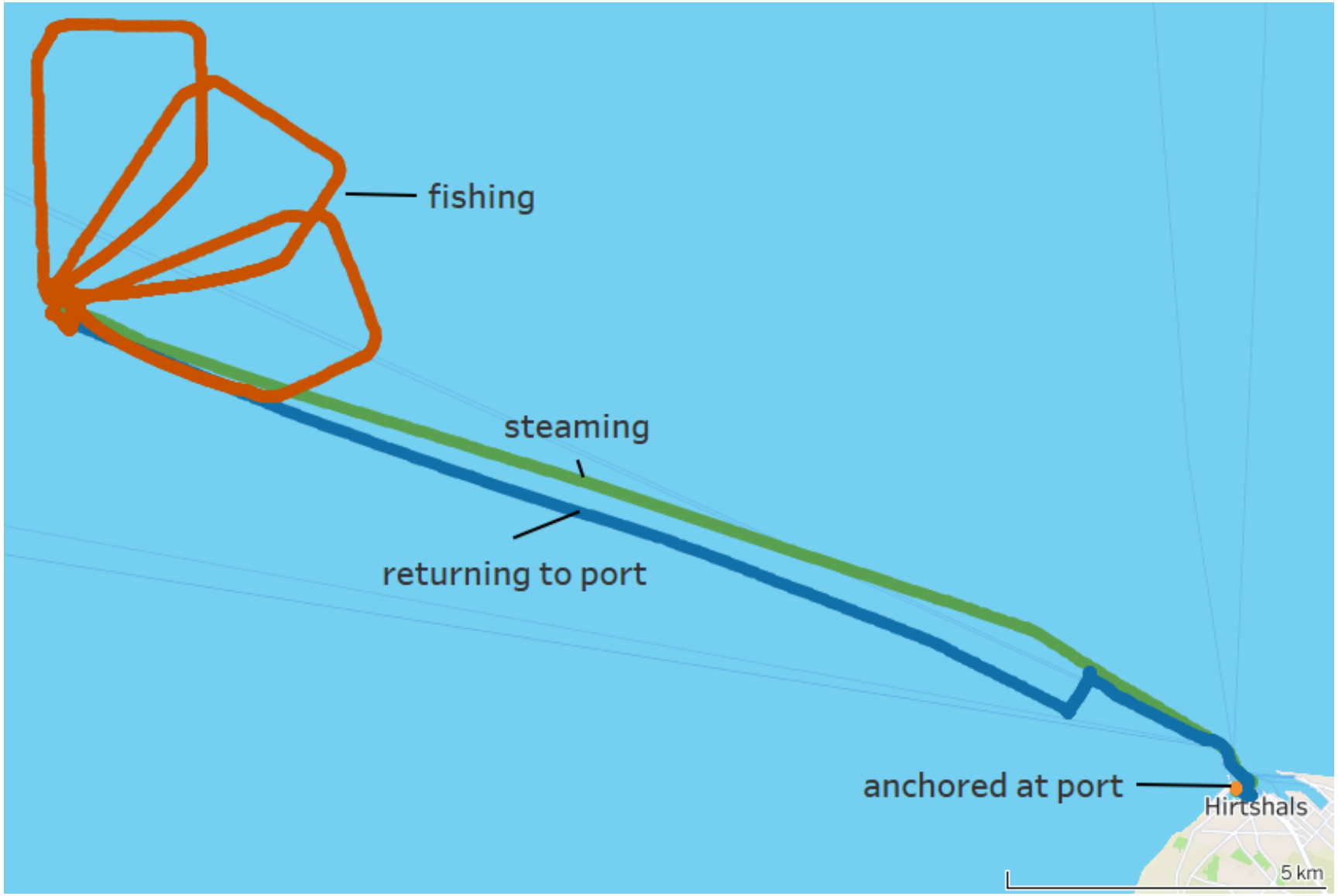


Semantic Segmentation of Fishing Trajectories

- How to split a fishing trajectory into sub-trajectories representing meaningful activities, i.e. fishing, sailing, etc.



What is the desired segmentation result?



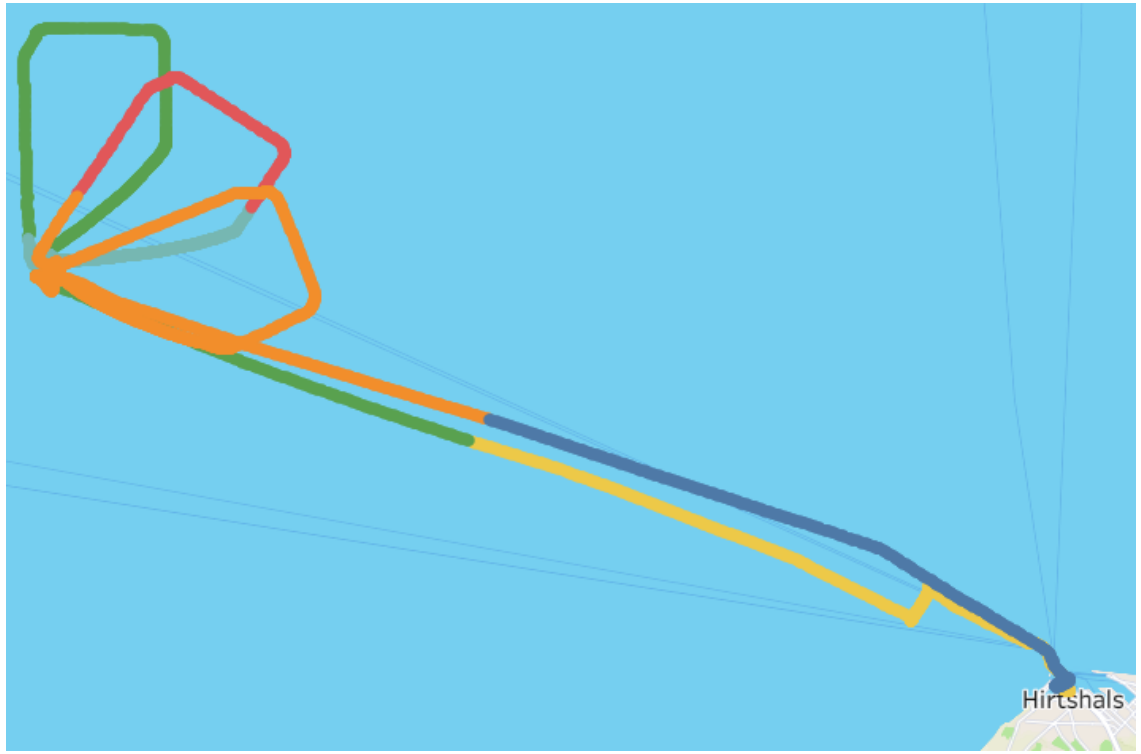
Do existing segmentation algorithms work? Not very well.

- Some aim to split trajectories into a series of stops and moves, like CB-SMoT
- most existing segmentation algorithms return segments without labels.
- one common assumption in many trajectory segmentation approaches is that the resulting segments should have high homogeneity w.r.t. some spatiotemporal criteria or features of trajectory points within segments

Segmentation results by CB-SMoT: using speed as main criteria

Segmentation results by Warped Kmeans: a variant of Kmeans

$k = 4$



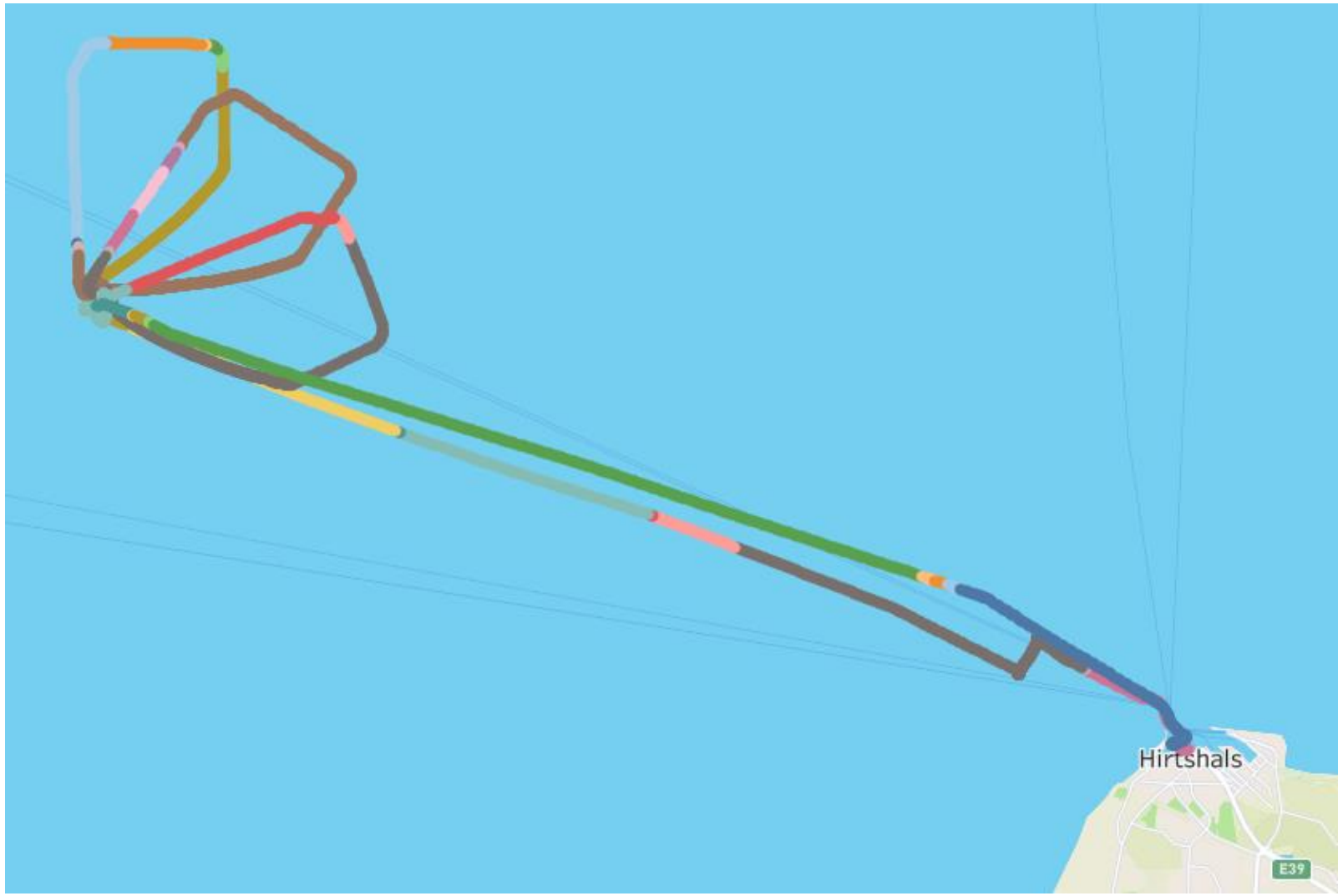
$k = 6$

Segmentation results by SWS: calculate error rate for each point

Linear Regression kernel

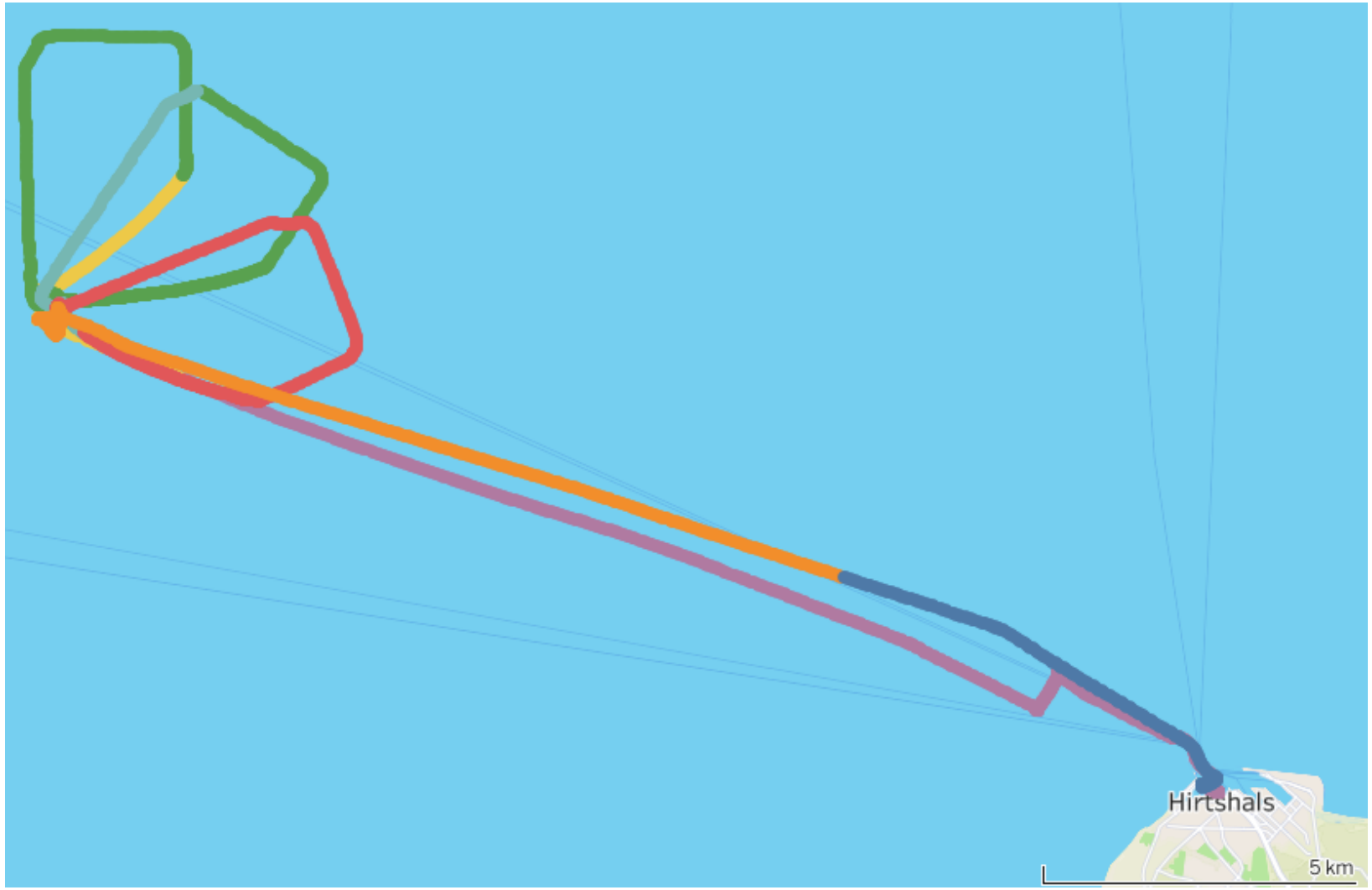
Window Size = 7

Percentile = 99.5



Segmentation results by GRASP-UTS: based on Minimum Description Principle

Minimum Segment Duration
= 2 hours



Research goals

- more meaningful trajectory analysis operators to gain insights into ship behaviors and maritime activities
- design efficient index structures to support the proposed operators
- bring proposed solutions to open source software, like MobilityDB

References

[1] Ismail, Anas, and Antoine Vigneron. "A new trajectory similarity measure for GPS data." *Proceedings of the 6th ACM SIGSPATIAL International Workshop on GeoStreaming*. 2015.

[2] Furtado, Andre Salvaro, et al. "Unveiling movement uncertainty for robust trajectory similarity analysis." *International Journal of Geographical Information Science* 32.1 (2018): 140-168.

[3] Zhao, Peng, et al. "SST: synchronized spatial-temporal trajectory similarity search." *GeoInformatica* 24.4 (2020): 777-800.

[4] Dodge, Somayeh, Patrick Laube, and Robert Weibel. "Movement similarity assessment using symbolic representatio of trajectories." *International Journal of Geographical Information Science* 26.9 (2012): 1563-1588.

[5] Ranu, Sayan, et al. "Indexing and matching trajectories under inconsistent sampling rates." *2015 IEEE 31st International Conference on Data Engineering*. IEEE, 2015.

[6] Lehmann, Andre L., Luis Otavio Alvares, and Vania Bogorny. "SMSM: a similarity measure for trajectory stops and moves." *International Journal of Geographical Information Science* 33.9 (2019): 1847-1872.

[7] Petry, Lucas May, et al. "Towards semantic-aware multiple-aspect trajectory similarity measuring." *Transactions in GIS* 23.5 (2019): 960-975.

[8] Varlamis, Iraklis, et al. "A novel similarity measure for multiple aspect trajectory clustering." *Proceedings of the 36th Annual ACM Symposium on Applied Computing*. 2021.

[9] Wang, Sheng, et al. "A survey on trajectory data management, analytics, and learning." *ACM Computing Surveys (CSUR)* 54.2 (2021): 1-36.