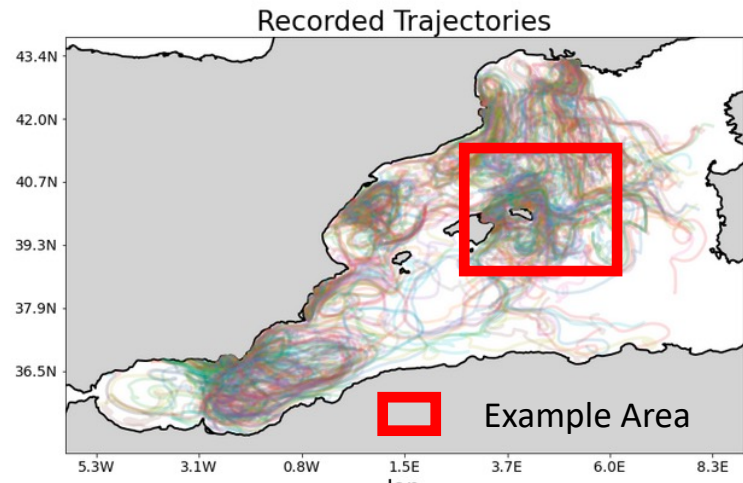


Efficient Pathway Identification from Geospatial Trajectories

Carola Trahms^{1,2}, Patricia Handmann¹, Willi Rath¹, Matthias Renz² and Martin Visbeck^{1,2}

¹ GEOMAR Helmholtz Centre for Ocean Research Kiel, ² Kiel University Kiel, Germany

- Lagrangian analysis is widely used in the earth-sciences
 - Used to analyze the output of e.g. ocean circulation models
 - Large sets of virtual particles can be integrated within the three-dimensional, time-evolving velocity fields.

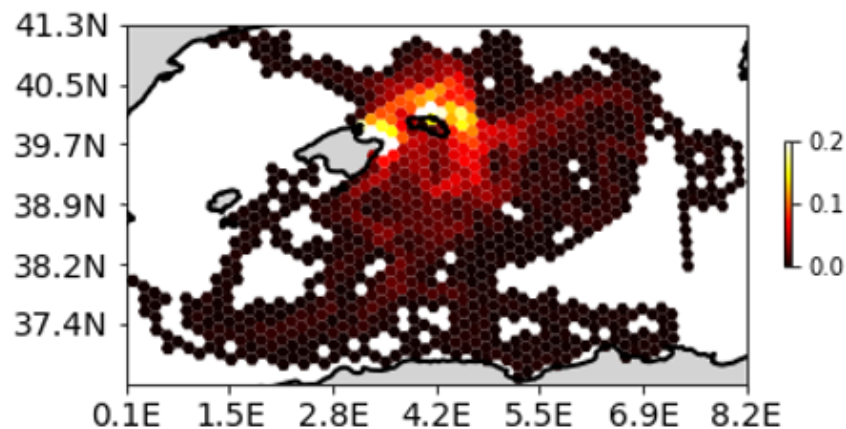


Challenge:

Fast analysis of huge amounts ($\sim 10^6$) of Lagrangian trajectory data

State-of-the-Art approach in ocean sciences

- Binning of particles [*Seville et al. 2017*]
 - Subdivide space into cells along a fixed longitude-latitude grid and count number of particles in each cell
 - resulting probability density maps
 - *Transit-time distributions*
 - Results are predominantly displayed on regular grids
 - Qualitatively and visually identify main pathways
- Loose direction information
- Extracting exact pathways remains a challenge
- Loss of direction of movement
- Disregards temporal dependencies



State-of-the-Art approach in computer sciences

- Trajectory clustering (TraCLUS) [Lee 2007]:
 - Segments trajectories and uses DBScan [Ester et al. 1996] to cluster the trajectory segments
 - Segmentation preprocessing step is very expensive and does not scale well on bigger amounts of trajectories
 - DBScan requires optimization of two parameters (epsilon and min number of neighbours) -> takes a long time on huge trajectory data sets or automated optimizing for the clusters needs to be defined.
 - Long runtimes

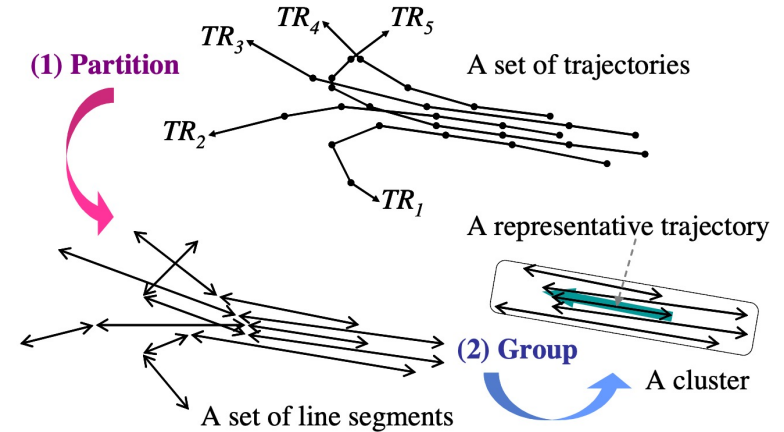
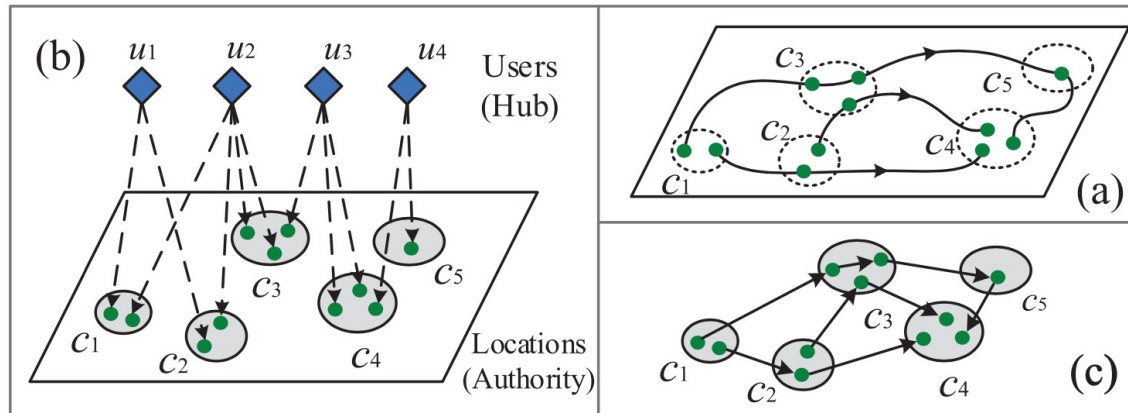


Figure 3: An example of trajectory clustering in the partition-and-group framework. [Lee 2007]

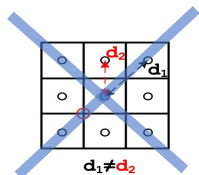
State-of-the-Art approach in computer sciences

- Transition graph networks for organic trajectories [Zheng 2015]:
 - Identifying key locations from the trajectories by, e.g., identifying stay points where particles lingered for a longer time (nodes)
 - Connections btw. key locations (edges) indicate transitions
- Computationally expensive to identify key locations.
- not yet established

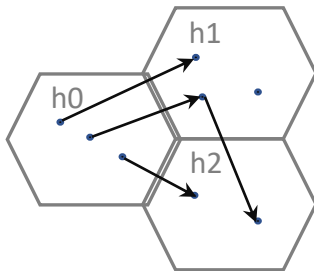


[Zheng 2015]

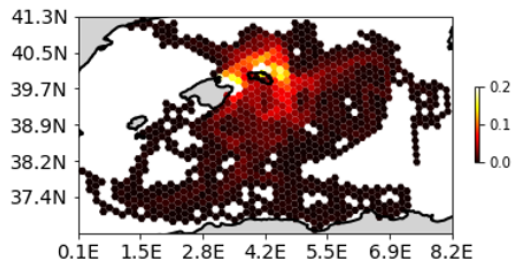
Mixed approach marine sciences and computer sciences



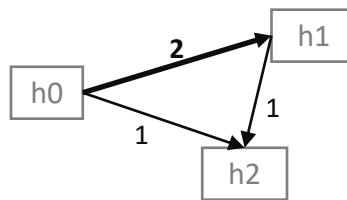
1) Hexagons (H3)



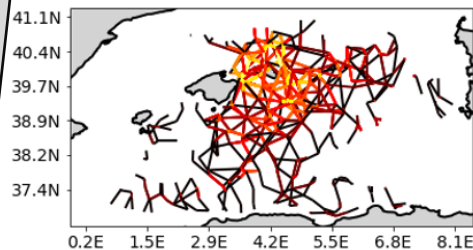
- Binning with established and fast H3 algorithm
 - Hexagon centers are all equidistant
 - 6 moving directions



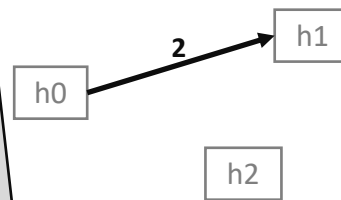
2) Transition graph network



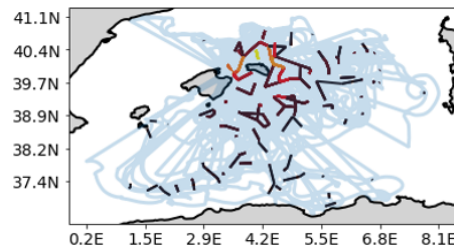
- Count number or transitions from one box (node) to another (edges)



3) Graph Algorithm:

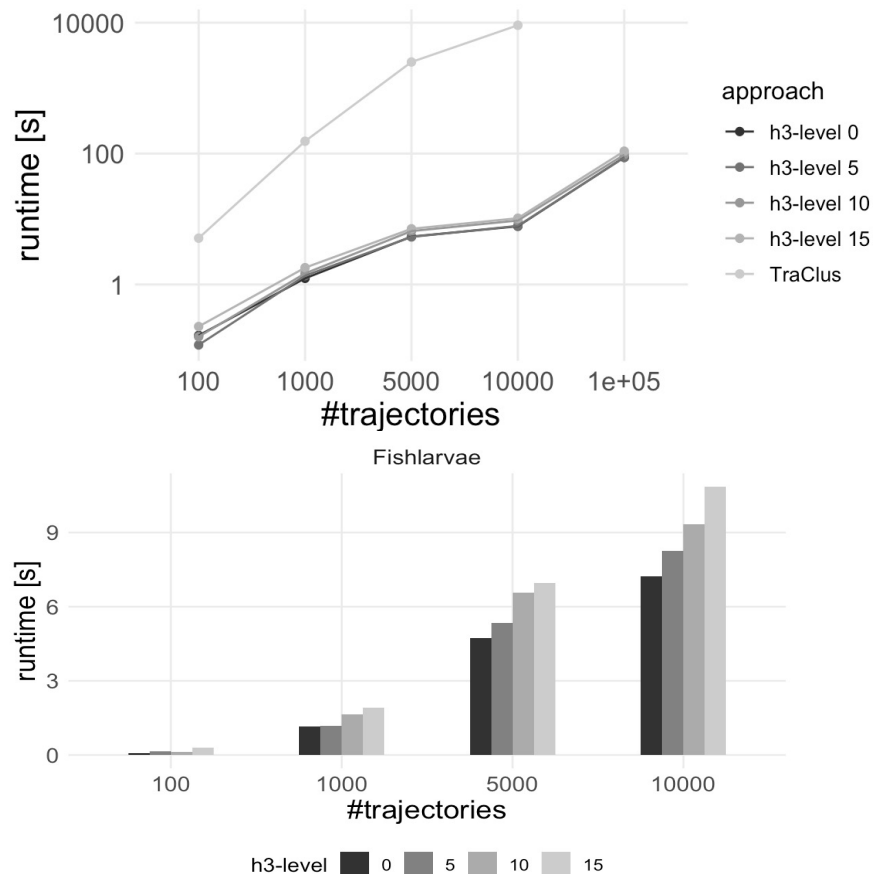


- Filter edges: First degree neighborhood filtering
- Includes directional information
- Shows major pathways



Mixed approach marine sciences and computer sciences

- Scales well with amount of particles
- Very fast compared with TraClus
- Parallelizing approach could leverage speed even more



Literature/Cites

- H3: Uber's Hexagonal Hierarchical Spatial Index (<https://github.com/uber/H3>)
- *Lee 2007* : Jiawei Han Jae-gil Lee. 2007. Trajectory Clustering: A Partition-and-Group Framework.
- Ester et al. 1996 : Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A Density- Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (Portland, Oregon) (KDD'96). AAAI Press, 226–231.
- Zheng 2015: YuZheng.2015.TrajectoryDataMining.ACMTransactionsonIntelligentSystems and Technology 6, 3 (2015), 1–41. <https://doi.org/10.1145/2743025>