Semantic Segmentation of AIS Trajectories for Detecting Complete Fishing Activities

Song Wu[1], Esteban Zimányi[1], Mahmoud Sakr[1], Kristian Torp[2]

[1]Université Libre de Bruxelles, Belgium; [2]Aalborg University, Denmark

1. Motivation

Illegal, unreported and unregulated (IUU) fishing does harm to

- marine environment
- sustainable use of marine resources, etc.

So it is important to know when&where a ship may have conducted fishing activities.

2. Problem Definition

Input: a trajectory T represented as a sequence of timestamped points $(p_1, t_1), \cdots, (p_n, t_n)$

Output: a sequence of labelled segments $<(S_1,l_1), \ldots, (S_k,l_k)>$, where

- $\bigcup S_i = T$
- $l_i \in$ {fishing, non-fishing} is the label for the segment S_i

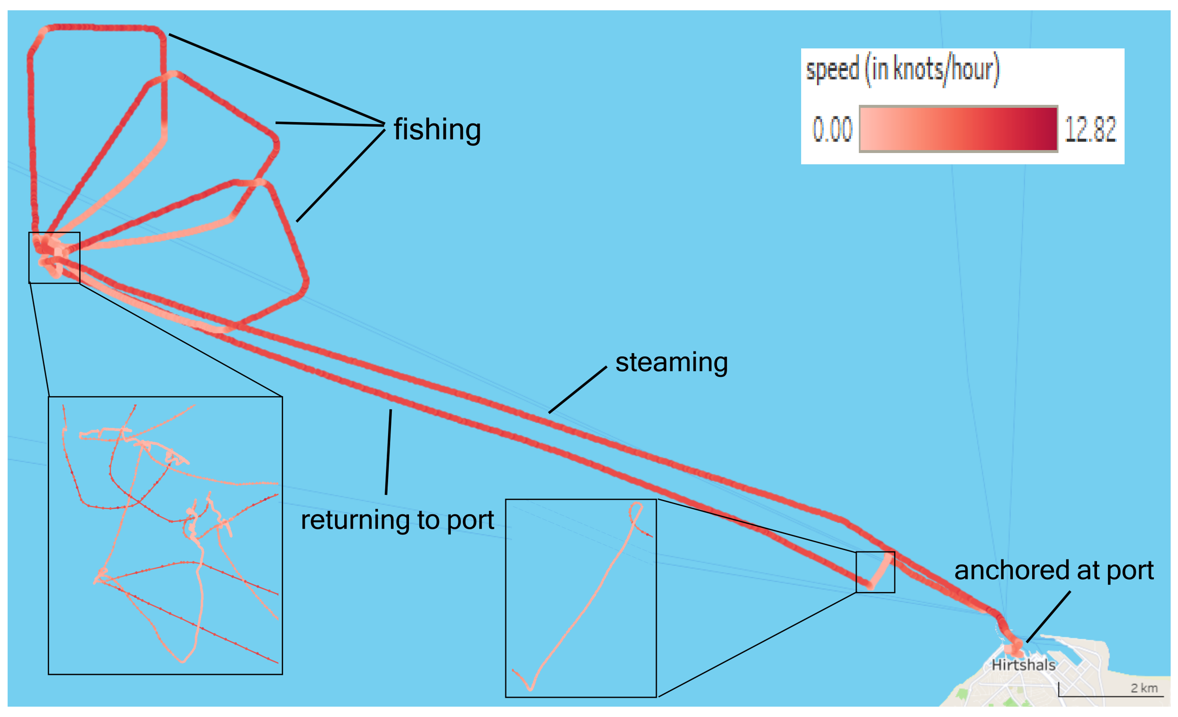
3. Related Work

Limitations of Existing Trajectory Segmentation Algorithms:

- some work simply treat trajectories as a sequence of stops and moves, such as CB-SMOT and DB-SMOT.
- segments are returned without labels, such as GRASP-UTS, WKmeans, SWS, WS-II.
- Many studies assume that that returned segments should have high homogeneity w.r.t. some spatiotemporal criteria or features of points, such as GRASP-UTS.

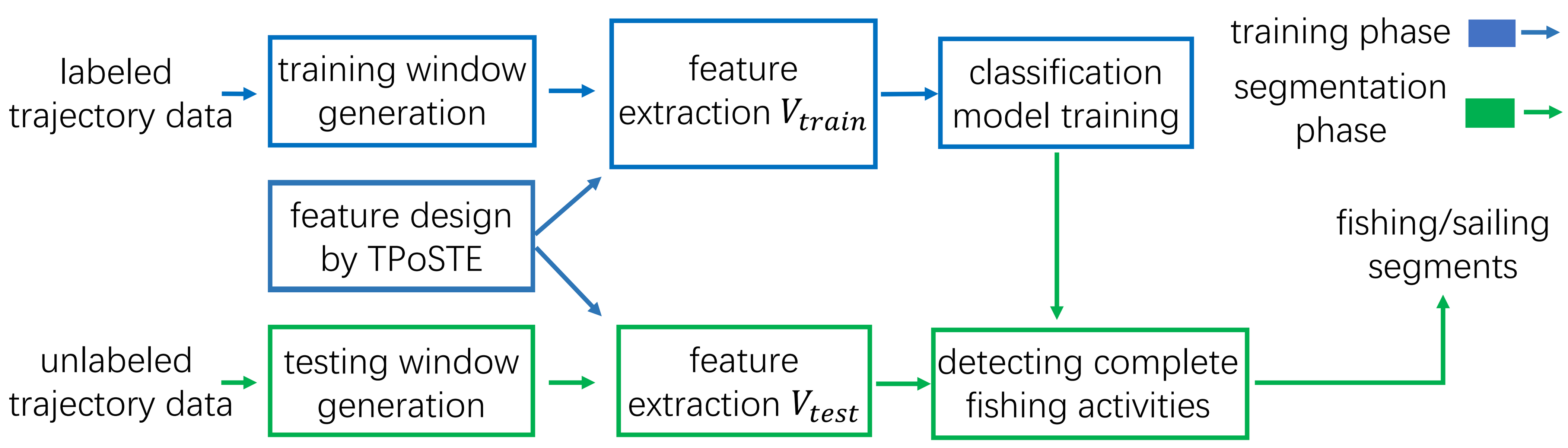
Two Observations:

- Movement patterns during fishing can be complex depending on the gear type and the situation on the spot.
- It is difficult, if not impossible, to design effective spatiotemporal criteria for the detection of fishing activities.

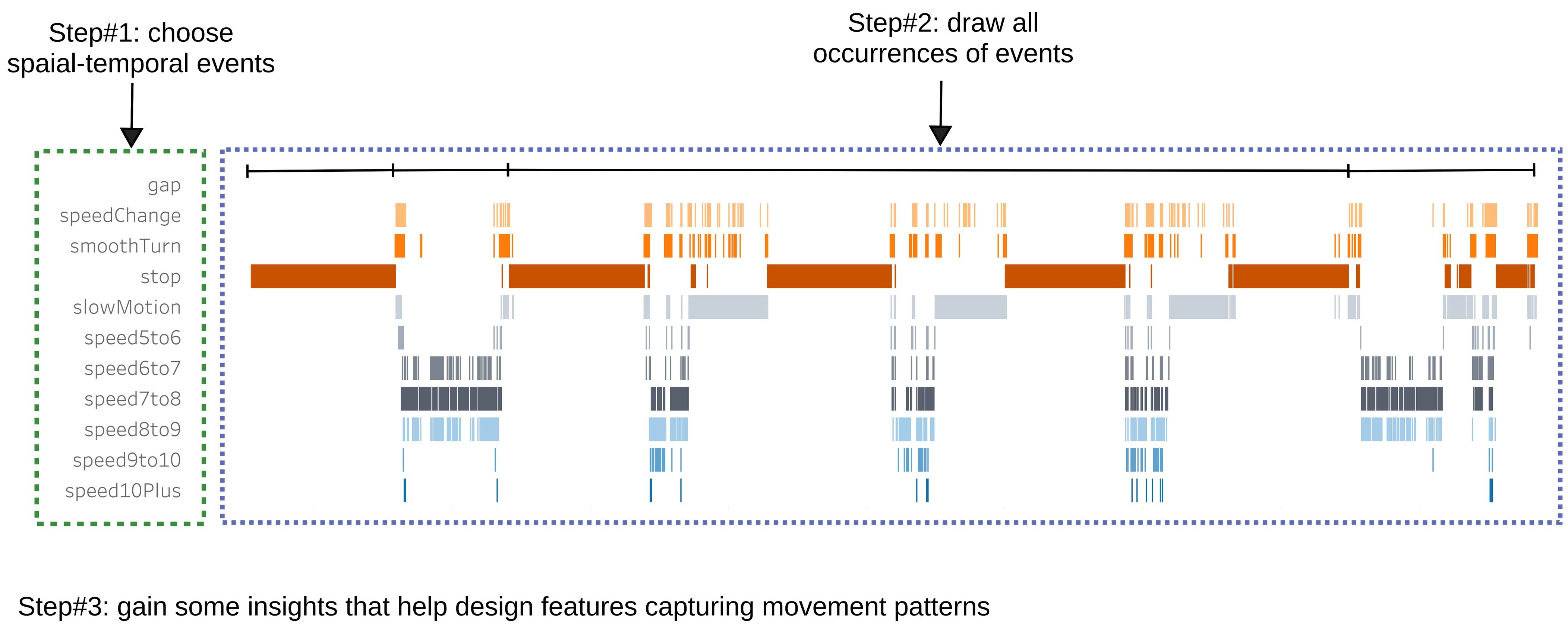


4. Methodology

4.1 Overall Workflow of the Proposed Methodology



4.2 Feature Design by TPoSTE



4.3 Window-Based Trajectory Segmentation using Run-Length Encoding (WBS-RLE)

Window Generation Strategy:

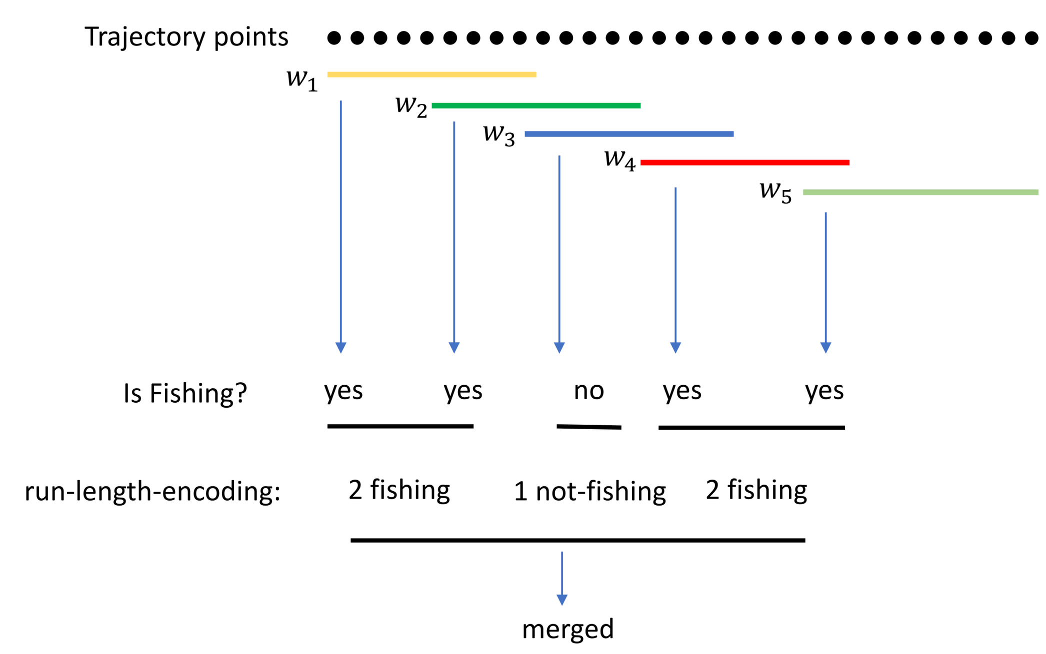
- a window is required to contain at least $size_w$ points and its duration is larger than a time threshold t_w.
- two adjacent windows have some overlap indicated by $ratio$.

Run-Length Encoding technique

- an alternating sequence of counts $\ldots, a_{fishing}, b_{sailing}, c_{fishing}, \ldots$ is obtained from the labeled windows.

A complete fishing activity A is a maximal subsequence of counts that:

- A starts and ends with fishing counts.
- each triplet $<a_{fishing}, b_{sailing}, c_{fishing}>$ in A fulfills $a \geq b$ and $b \leq c$ to correct occasional classification errors.

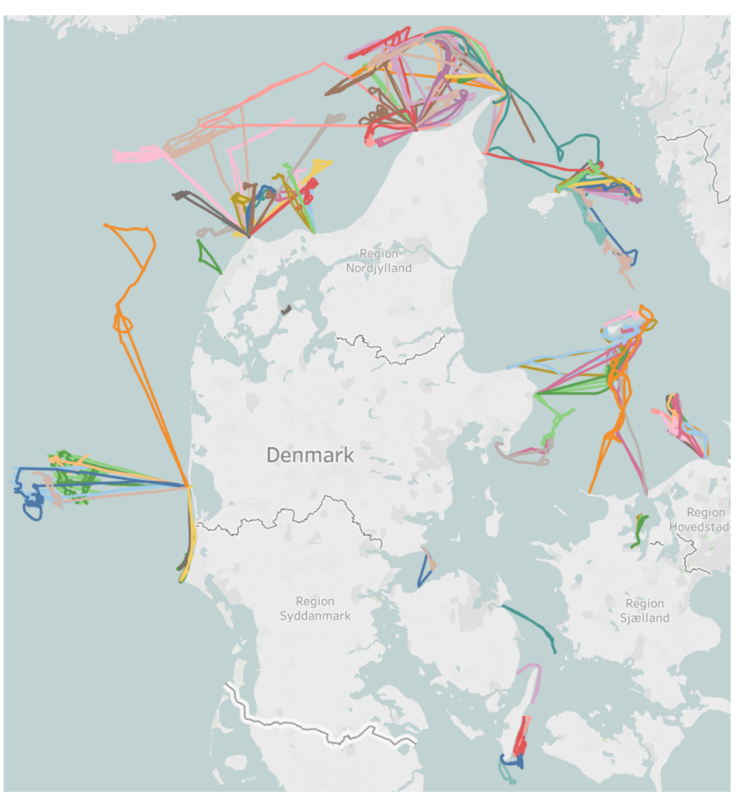


5. Experimental Results

5.1 Dataset

We manually labeled 128 trajectories between Nov 14, 2021 and Nov 20, 2021.

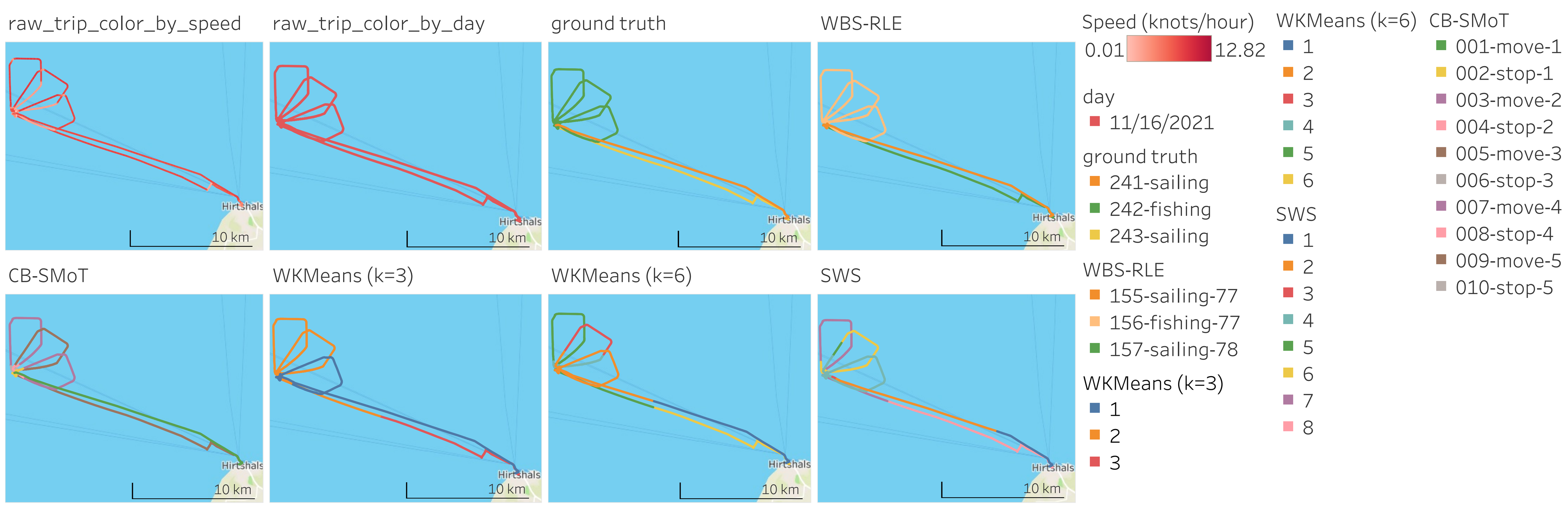
- publicly available from *Danish Maritime Authority*
- average sampling gap: 10.63 seconds
- # of points: 1,080,220
- 31 trajectories used for training



5.2 Average Performance on the 97 Testing Trajectories

method	purity	coverage	harmonic mean	# of segments
WBS-RLE	0.890	0.974	0.927	2.670
CB-SMoT	0.859	0.885	0.859	5
WKMeans (k=3)	0.878	0.840	0.855	3
WKMeans (k=6)	0.932	0.619	0.741	6
SWS	0.954	0.759	0.837	9.855

5.3 Segmentation Result for the Trajectory #220051000-2



5.4 All Segmentation Results