Fine-Tuned Compressed Representations of Vessel Trajectories

Giannis Fikioris (NCSR Demokritos),
Kostas Patroumpas (Athena Research Center),
Alexander Artikis (University of Piraeus and NCSR Demokritos),
Georgios Paliouras (NCSR Demokritos),
Manolis Pitsikalis (University of Liverpool)

Complex Event Recognition Group
cer.iit.demokritos.gr
Introduction

- AIS data precious in maritime navigation & safety.
- Data volume may be overwhelming and redundant for analysis; stakeholders “decimate” AIS data for storage.
- We apply a methodology for maintaining trajectory synopses online from AIS messages.
- Intelligently discard majority of incoming locations with minimal error in trajectory representation.
- We propose a ML approach for fine-tuned parametrization of trajectory compression.
- Method can improve quality of trajectory synopses and recognition of complex maritime events.
- Tested on real-world AIS datasets and applicable in industrial settings.
Vessel Trajectory Compression

- Find the *critical points* of a trajectory (e.g. turning and acceleration points) with which one may approximately reconstruct the original trajectory.

- Critical points rely on compression parameter values for speed, angle, etc.

- Different parameter values are necessary in different areas/datasets (e.g. sampling rates).

- Different vessel types require different parameter values (e.g. tankers vs. fishing boats).
Compression Parameters

- Original parameter values (thereafter called *default parameters*) were common for all vessels, picked using expert domain knowledge for a particular dataset.
- Tedium task, cannot be done for every ship type in every dataset.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \theta$</td>
<td>Angle Threshold ($^\circ$)</td>
</tr>
<tr>
<td>$m$</td>
<td>Buffer Size ($locations$)</td>
</tr>
<tr>
<td>$\Delta T$</td>
<td>Gap Period ($seconds$)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>Historical Timespan ($seconds$)</td>
</tr>
<tr>
<td>$v_{min}$</td>
<td>No Speed Threshold ($knots$)</td>
</tr>
<tr>
<td>$v_\theta$</td>
<td>Low Speed Threshold ($knots$)</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Speed Ratio</td>
</tr>
<tr>
<td>$D$</td>
<td>Distance Threshold ($meters$)</td>
</tr>
</tbody>
</table>
 Genetic Algorithm

- Automatically learn the best parameters per vessel type in an AIS dataset.

- **Minimize** two variables:
  1. *Compression Error*, i.e. the RMSE between the original noiseless trajectories and the ones approximately reconstructed from the synopses.
  2. *Compression Ratio*, i.e. the fraction of critical points to the noiseless points.
Optimization Function

- To minimize both the RMSE and the Ratio simultaneously, we minimize a function of the form

\[\text{Ratio} + \text{ReLU}(\text{RMSE} - \theta)\]

- ReLU is the Rectified Linear Unit function, defined \(\text{ReLU}(x) = \max(x, 0)\).

- \(\theta\) is a hyper-parameter that controls the desired RMSE:
  - If \(\text{RMSE} \leq \theta\), then the optimization function effectively becomes Ratio.
  - If \(\text{RMSE} > \theta\), then the optimization function effectively becomes RMSE (because Ratio takes small values).
Optimize incrementally, i.e. process the dataset in batches.
Optimize on 1st batch, evaluate on 2nd.
Use previous optimization results to avoid cold start when optimizing on batches 1 and 2.
Complex Event Recognition

Simple Events → Event Recognition System (RTEC) → Recognized Complex Events

Complex Events Definitions
Empirical Setup

Two real world datasets.

◊ “Brest” dataset
Covers the area of Brest, France
6 months (Oct. ‘15 to Mar. ‘16)
zenodo.org/record/1167595

◊ “Mediterranean” dataset
Covers the Mediterranean Sea
6 months (Mar. to Aug. ‘16)
Provided by MarineTraffic - marinetrack.com

A large mount of AIS positions have already been discarded from the Mediterranean dataset, to reduce its size.

github.com/GiannisFikioris/
Genetic-Algorithm-for-Synopses-Generator
Empirical Evaluation – 6-Fold Cross Validation

- Fishing Boats Brest Dataset
- Mediterranean Dataset, March
- Passenger Ships Brest Dataset
- Fishing Boats Mediterranean Dataset, June
Empirical Evaluation – Incremental Optimization
Passenger Ships – Mediterranean Dataset

Training progress

Evaluation Phase
## Composite Event Recognition

<table>
<thead>
<tr>
<th>Composite Event</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drifting Vessel</td>
<td>0.56</td>
<td>0.88</td>
<td>0.69</td>
</tr>
<tr>
<td>Search &amp; Rescue</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>Loitering</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Piloting</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Ship-to-ship Transfer</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Tugging</td>
<td>0.99</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Vessel Under Way</td>
<td>0.99</td>
<td>0.78</td>
<td>0.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input</th>
<th>Avg. Time</th>
<th>Std. Time</th>
<th>Worst Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Dataset</td>
<td>41.1 sec</td>
<td>22.1 sec</td>
<td>131.3 sec</td>
</tr>
<tr>
<td>Synopses [Prev]</td>
<td>10.1 sec</td>
<td>7.3 sec</td>
<td>55.9 sec</td>
</tr>
<tr>
<td>Synopses [GA]</td>
<td>7.2 sec</td>
<td>4.0 sec</td>
<td>24.5 sec</td>
</tr>
</tbody>
</table>
Summary and Future Work

- Genetic Algorithm that automatically learns optimal compression parameters for a ship type in a dataset (thus avoiding manual, tedious work).
- Genetic Algorithm supports incremental optimization, training in data batches.
- Synopses support the online recognition of complex maritime events, without compromising predictive accuracy.

- We are investigating ways of further improving complex maritime event recognition.