Big data for Maritime Domain Awareness: An AIS case study

Waldo Kleynhans, IMIS Global
• The world’s oceans is of critical importance to humanity as it is key to fisheries, shipping as well as the environment.

• As maritime activities increase globally, there exist a greater dependency on technology in the monitoring, control and surveillance of vessel activities.

• One of the most prominent systems for monitoring vessel activity is Automatic Identification System (AIS).

• There are however some challenges with the use of AIS data.
AIS data fidelity:

• Information could be manipulated / corrupted

• AIS receivers are not controlled in the same manner as AIS transmitters

Significant volume increase of AIS messages:

• Due to the global increase in vessels fitted with AIS transmitters as well as the proliferation of satellite and terrestrial receiving stations.

• While increased AIS data volumes is beneficial, the processing and storage of these large data volumes can become problematic.
• The study area covers a significant portion of the Globe.

• Snapshot of all AIS messages received over 24 hours for 17 March 2018 was considered.

• AIS data from this area is provided by more than a hundred AIS data sources which includes terrestrial receivers, base stations as well as satellites.

• The total number of vessels that was received during the day was 27834.

• These vessels transmitted a combined total of approximately 26 Million AIS messages during the 24 hour period.
# of reporting MMSI's for a one day period over AOI: 27834

# of AIS messages received for a one day period over AOI: 25975463
• >70% of vessel's transmitted AIS messages were received only 0-100 times

• Approx. 15% of vessel's total received AIS messages for the day > 1000

• 10% of the vessels generated 90% of the AIS messages received through the network of receivers in the AOI.
• An on-line lossy track compression methodology was used in the filtering process of the positional AIS data

• Method tracked the AIS message stream originating from each ship in an on-line fashion

• Spatio-temporal information contained in each newly presented AIS message compared to current vessel track history using a spatio-temporal distance metric

• A decision was made to include the newly presented positional information or discard it based on the maximum allowable distance metric

• The goal was to be able to do historic vessel track reconstruction as a function of the maximum allowable track distance error

• It was shown that a compression ratio of approximately 90% could be achieved when the error metric of between 60m and 70m is selected.
• Volume of AIS data is increasing significantly 40% increase over the last 4 years.

• Maritime industry needs to able to deal with with these large data volumes.

• Data pre-processing and compression is essential to enable effective downstream processing.

• A simple framework was presented to showcase an example of AIS on-line data compression.

• Initiatives such as datAcron is crucial for the practical implementation of operational systems able to effectively deal with AIS data of these orders of magnitude.