Maritime Trajectory Analysis

Karna Bryan
15 June, 2011
Cooperative Anti-Submarine Warfare

Autonomous Naval Mine Countermeasures

Environmental Knowledge and Operational Effectiveness

Maritime Security
Maritime Situational Awareness

An enabling capability which seeks to deliver the required **Information Superiority in the maritime environment**, to achieve a **common understanding** of the maritime situation, in order to increase effectiveness in the planning and conduct of operations

- NATO EXTAC 790

Effective understanding of anything associated with the global maritime environment impacting economies, the environment, safety and security

- Office of Global Maritime Situational Awareness, USA
“AS DATA become more abundant, the main problem is no longer finding the information as such but laying one’s hands on the relevant bits easily and quickly. What is needed is information about information.”
Satellite AIS

Not enough Data → Too much Data
Open source data sources

Information freely available on the internet is quickly growing
Task Objective

• To understand how historical maritime data can be used to better understand the maritime environment and assess what “information of information” is useful for exchange
  – Products for NATO tools (e.g. BRITE “Routes”)
  – Prototypes/techniques for NATO nations
  – Advancing information sharing between various stakeholders
M-Atlas Workbench

- Use of M-Atlas Workbench developed by University of Pisa Knowledge Discovery in Databases Lab and Consiglio Nazionale di Ricerca (CNR)
  - Originally designed for automobile traffic data
“Typical” AIS Data

Gaps in coverage, Multiple data sources
Two Main Challenges

• Data Intermittency

• Defining where one trajectory stops and another one begins
Straight of Gibraltar, 2010
Clustering on one year of data
Comparison between Cargo

- Cargo and Tanker
  - Transit only or stop at major ports

- Passenger and HSC
  - Capture Ferry behavior
Cargo Clustering

Two Largest Clusters identify transit only:
Passenger Clustering
What is Anomalous behaviour?
• Origination-Destination Matrix with efficient calculation
  – Custom areas identified to quantify movements between areas
World Port Index

3716 Ports in World Port Index
Port Detection

- World Port Index does not include all ports
- Simple distance buffers do not always work

Approach: Use clustering techniques on starts and Voronoi partitions
Role of Human Error & Data Clean-up

- Incorrect and Incomplete Data

### Table 1: Vessel Information

<table>
<thead>
<tr>
<th>ID</th>
<th>MMSI</th>
<th>IMONumber</th>
<th>CallSign</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>152709</td>
<td>242962000</td>
<td>&quot;EG(#77)</td>
<td>TANGER JET II</td>
</tr>
<tr>
<td>2</td>
<td>227840</td>
<td>242962000</td>
<td>9278919</td>
<td>TANGER 6[6 II</td>
</tr>
<tr>
<td>3</td>
<td>160747</td>
<td>242962000</td>
<td>CNA4377</td>
<td>TANGER JERAO^)Y&gt;0 H</td>
</tr>
<tr>
<td>4</td>
<td>152508</td>
<td>242962000</td>
<td>9279264</td>
<td>TANGER JET II</td>
</tr>
<tr>
<td>5</td>
<td>16131</td>
<td>242962000</td>
<td>CNA4377</td>
<td>TANGER JET II</td>
</tr>
<tr>
<td>6</td>
<td>64519</td>
<td>242962000</td>
<td>CNA4377</td>
<td>TANGER JET II ,JLH</td>
</tr>
<tr>
<td>7</td>
<td>146820</td>
<td>242962000</td>
<td>CNA4377</td>
<td>TANGER JET II B+M</td>
</tr>
</tbody>
</table>

### Table 2: Journey Details

<table>
<thead>
<tr>
<th>VesselID</th>
<th>Destination</th>
<th>StartTime</th>
<th>DurationInHrs</th>
<th>ETAMonth</th>
<th>ETADay</th>
<th>ETAHour</th>
<th>ETAMinute</th>
<th>TypeID</th>
<th>NumReports</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:10:10.000</td>
<td>0.100000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:22:12.000</td>
<td>0.200000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:40:13.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:46:14.000</td>
<td>0.100555</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:58:16.000</td>
<td>0.100000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 01:10:17.000</td>
<td>0.001368</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 02:04:23.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 02:10:25.000</td>
<td>1.602222</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:52:34.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 03:58:35.000</td>
<td>0.601111</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 04:40:39.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 04:46:40.000</td>
<td>2.204166</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>13</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 07:04:54.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 07:10:55.000</td>
<td>0.712500</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>TANGER, TARIFA, TANGER</td>
<td>2011-04-01 07:59:41.000</td>
<td>0.000000</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>60</td>
<td>14</td>
<td>1</td>
</tr>
</tbody>
</table>
Speed over Ground versus Hour

Speed Over Ground, High Speed Craft, approx. 2 months of data
Course Over Ground versus Hour

Departs Tangier after 7, 11, 15, 19

Departs Tarifa after 9, 13, 17, 21
Suez Canal Example

Trajectories + Start Points (yellow)

Longest 100 trajectories

Trajectories with 50+ points
Distributed and collaborative fusion

- Next generation data fusion algorithms
  - Lessons learned from NURC distributed multi-hypothesis tracker (DMHT)
  - Handling out-of-sequence data (near real-time application)
  - Fusion “Nodes” share hypothesis-level information
  - Generic framework for both high-level and low-level fusion

- Novel storage of hypothesis trees using distributed content addressable databases
  - Enables full versioning and parallelization of tree
  - Cryptographically signed objects
• Estimation of Recognized Maritime Picture (RMP) completeness without knowledge of sea truth
  – Integrate results into fusion algorithms
  – Asset Management for mobile sensors
  – Quantify value of multi-sensor tracks

Looking Forward

• Reconstruction of trajectories which better deals with gaps in coverage and multiple data sources
• Further incorporation of static data, and confidence measures in data
• Multi-trajectories (connecting multiple “trips”)
• Investigating more advanced statistical models for trajectory clustering
• Tie in with Semantic Logic (BRITE, ROLEX, NLD POSIDEON)
• We would be open to other application areas
Conclusions

• Maritime data will soon transition from scarce to superabundant
• Automation and ability to assimilate information from data is key
• Moving point/trajectory analysis shows promise, but must be tied to tracking and fusion
• Maritime traffic is complex
  – Data is incomplete and sometimes inaccurate
  – Routes are relatively unconstrained compared to many other applications
• Nonetheless, patterns can be derived from historical data