International Journal of Computer and Information Technology (ISSN: 2279 – 0764)
Volume 02– Issue 01 January 2013

Trajectory Analysis using Automatic Identification Systems in New Zealand Waters

Premalatha Sampath
Whitireia New Zealand
450 Queens Street
Auckland
Premalatha.Sampath@whitiriea.ac.nz

David Parry
School of Computing and Mathematical Sciences and AURA Lab
Auckland University of Technology
Auckland, New Zealand
Dave.Parry@aut.ac.nz

Abstract— **Trajectory analysis is one of the actively researched areas of spatio-temporal databases. Exploring and analysing large datasets of movement data has become a vital part of research in many disciplines and decision-making fields. The major challenge in analysing process of trajectory data is to visualize, understand and extract meaningful patterns out of millions of locations collected from Automatic Identification Systems (AIS) points. AIS are used in maritime environments to assist in tracking and monitoring vessel movements. AIS datasets are real movement datasets recorded from dynamic vessels and consisting of voluminous raw data. To analyse such datasets required a systematic and methodical process. The first phase focused on development of a decoder to extract significant information from raw data. The extracted information was then utilized to perform knowledge discovery on movement data from dynamic objects. A spatio-temporal approach was applied to perform trajectory analysis on decoded datasets. The paper focuses on optimization of information to discover hidden knowledge from raw datasets. The purpose of optimizing information is to conduct trajectory analysis in order to identify the characteristics of vessels in New Zealand waterways. The discovered knowledge can also be applied in other fields such as safety, security and additional navigational aids. The study used real movement datasets of maritime domain provided by Kordia New Zealand recorded between March 2011 and May 2011. The experimental results indicate that the proposed methods could be successfully applied to perform trajectory analysis of vessel movement**.

Keywords-component; AIS, Trajectory, spatio-temporal, movement data

I. INTRODUCTION

Modern positioning and identification technologies enable the collection of large amounts of data from moving objects. Data about movements of these moving objects are often collected as 'trajectories' in space and time. The movements are recorded as a series of geographic locations with respective time stamps. Exploring and analysing large datasets of movement data has become a vital part of research in many disciplines. There seems to be increased interest in many disciplines in the study of trajectory analysis. Such as behavioral sciences [5], ecology [1], transportation modeling [2], cartographic planning [3], terrestrial change [4], urban transportation [5], marine traffic [6] and vessel movements [7].

Auckland University has conducted a research study on homing pigeons using global positioning devices datasets [10]. A joint study of vegetation communities conducted by the University of Otago and CSIR–NRE, Earth Observation Research Group, South Africa examined ground-based permanent photo-point images [8]. The former is a spatio-temporal analysis approach to trajectories whereas the latter is a spatial analysis (object oriented approach) using large-scale aerial photographs or satellite images. No studies have been conducted on New Zealand waters to date in regards to trajectory analysis on vessel movements. This is the motivation and rationale for this investigational study.

A. Methodology

The study used the Design Science methodology approach described by [9]. The goal of design science research is to produce a suitable artifact in the form of a construct, model, method or installation. The raw data is extracted from large data files collected as part of the Automatic Identification System (AIS).

1) The AIS data

AIS provide a means for vessels to transmit data electronically including vessel identification, position, speed and course with Vessel Traffic Services (VTS). The data is shared with other vessels. AIS utilize Global Positioning Systems (GPS) technology in conjunction with shipboard sensors and VHF radio communication equipment. This helps automatic exchange of navigation information electronically.
The use of AIS in maritime monitoring started in 2004 [6]. Global navigation satellite systems (GNSS) are extensively used in large ships of over 300 tons and passenger vessels. AIS are mainly used for collision avoidance and navigation aid as well as search and rescue operations.

AIS allow vessels to broadcast ship data electronically at regular intervals. These data include vessel identification, position, course speed and other crucial information of the vessels. Each vessel is identified with its own unique identification number called Maritime Mobile Service Identity (MMSI). The vessel transmits continuously, providing comprehensive and detailed data set for individual vessels. The received data from vessels are in the National Marine Electronics Association (NMEA) format.

These raw data are comprised of critical information that could be applied to a range of vital applications, such as vessel monitoring, maritime surveillance [10], security, search and rescue, traffic management, collision avoidance [12]. Additional value-added services are also traffic density plot, monitoring vessel location and movement, area of interest, historical track analysis or similar trajectory analysis. Such information in essential in investigating and identifying oil-spills from suspect vessels [11], verifying that vessel traffic complied with navigation rules or evaluating the need for additional navigational aids.

An artifact to decode vital information from raw data has been designed using VB.net. Standard procedures and methods were followed according to the IALA guidelines on AIS [12]. The analysis was conducted on trajectories of passenger boats, ferries and high speed craft based on velocity trends of speed over time.

To perform trajectory analysis requires the extraction of knowledge about these moving objects from massive datasets [13]. The analysis of raw data was conducted with a two-phased methodological process. Phase 1 primarily focused on handling the large dataset and developing a decoder. Given a large dataset (2GB - 30 million rows) of spatio-temporal movement observations the goal is to perform segmentation into sets of data from different periods. The objective of segmentation is to handle datasets of manageable size in order to overcome the main memory issue. The purpose of the decoder is to decode the raw data to read the information transmitted from the vessel. The decoder has been developed in VB.net and it can extract information such as **Maritime Mobile Service Identity** (**MMSI**), latitude, longitude, speed over ground (SOG) and time. The decoder was designed to intake raw data in text files and output the decoded dataset in text files. As the decoded dataset was thus made available it was imported to an Excel file. The most important challenge was centered on cleaning the decoded data of noise, outliers and duplications. Subsequently, the dataset was well prepared for the next phase: to conduct trajectory analysis in Geographical Information Systems (GIS). Phase 2, trajectory analysis was conducted using a spatio-temporal approach that included a series of methods such as projecting trajectories on ArcGIS, plotting velocity trend graphs [3] and clusters based on stop or start movements.

The above-mentioned applications provide the opportunity for refined vessel movement analysis and detailed data sets also provide challenges in data management and analysis. This paper also discusses challenges involved in optimizing information from NMEA format raw data, which are decoded into meaningful AIS messages.

II. AIS DATA

A. How AIS Works

Vessel identifiers such as the vessel name and VHF call sign are made available during initial equipment installation Number equations consecutively. Equation program. This information is included in the transmission along with location information from the vessel's global navigation satellite system receiver and gyrocompass [6]. AIS are used by marine vessels in coordination with VTS to monitor vessel location and movement. AIS transmitters normally send data every 2 to 10 seconds depending on moving vessel's speed while underway and every 3 minutes while the vessel is at anchor. The format is described in table 1.

Table 1 AIS Data elements

Maritime Mobile Service Identity (MMSI) - a unique 9-digit identification number
Navigation status - "at anchor", "underway using engine(s)", or "not under command"
Rate of turn - right or left, degrees per minute
Speed over ground - knots (normally 0.1 knot resolution)
Position Accuracy
Longitude and Latitude - 1/10,000 minute
Course over ground - relative to north to 0.1 degree
Time stamp - Coordinated Universal Time (UTC) accurate to nearest second when the data originated or generated.

B. Classification of AIS Information

The data is received by AIS receiver and relayed to computer - based navigation system, encoded in NMEA sentences (64-bit plain text). The information received can be classified into 2 categories: ship information and voyage information [14].

Table 2: List of information contained in AIS string

Ship information	***Voyage information***
Ship name	Current latitude
Type of vessel	Current longitude
Call sign	Speed over ground
MMSI number	Course over ground
IMO number	Heading
Registration	Rate of turn
Length	Navigational status
Beam	Destination
Draft	ETA

According to [15] ship information can be further divided into static and dynamic information,

- Static Information - vessel's name, IMO number, MMSI number, dimensions.
- Dynamic Information - vessel's position, speed, current status, course and rate of turn.
- Voyage-specific information - destination, ETA and draught.

The original messages decoded are NMEA Type 1 2 or 3, an example message is shown below

```
!AIVDM,1,1,,B,177KQJ5000G?tO`K>RA1wUbN0TKH,0*5C
```

. Each character is used to code 6 bits of data and binary (2's complement) coding is used for the latitude and longitude. A full description of the coding can be found at (http://gpsd.berlios.de/AIVDM.html).

III.. PROCESS OF KNOWLEDGE DISCOVERY (PHASE 1)

This is the first stage of the study to open the original dataset from MSSQL server 2008 and then decode the raw movement data. This stage or phase also focused on handling the large dataset and developing a decoder. Given a large dataset (2GB - 30 Million rows) of spatio-temporal movement observations the goal is to perform segmentation into sets of certain period of data. The objective of segmentation is to handle the datasets in manageable size to overcome the main memory issue.

Step1: - ***Decoding Raw Data:-*** The original dataset is provided by Kordia New Zealand as raw data in NMEA format. The provided raw data is exported to MS SQL Server 2008. The datasets are raw AIS information in NMEA format and are time stamped along with TalkerID and SentenceID. These datasets might provide inconsistent data, so a SQL query is used to select valid rows.

Step 2: The selected raw sentence is saved as a text file. The next step involved is to decode raw data using the developed decoder. The decoded dataset is collected as OUTPUT text file. The purpose of the decoder is to decode the raw data to read the information transmitted from the vessel. The decoder was developed in VB.net and can extract information such as **Maritime Mobile Service Identity** (**MMSI**), latitude, longitude, speed over ground (SOG) and time.

Step 3: The OUTPUT text file of each database has to be imported to an Excel sheet in order to perform next level tasks.

Step 4:- ***Data Cleaning***: One of the major challenges involved in handling raw datasets is cleaning them to ensure the removal of noise and inconsistent data.

Step 5: ***Data Selection:*** A subset of the decoded dataset is selected to perform analysis and imported as an Excel file of manageable size in order to avoid exceptional memory errors. The original data was 2GB, and such huge datasets could not be handled as a single file because the system encountered exceptional memory errors and was not able to respond to queries.

Step 6: This step exports the cleaned decoded dataset to the ArcGIS tool to conduct trajectory analysis. There was a particular focus on the speed parameter and other space (Longitude and Latitude) and time attributes. Another important attribute used to identify particular vessels is MMSI. Other attributes may be utilized in future studies. The decoded datasets were thus obtained and taken to next level to conduct trajectory analysis.

IV.TRAJECTORY ANALYSES (PHASE 2)

Trajectory analysis was conducted using a GIS tool (ArcGIS) to study and understand trajectory of pattern of different type of vessels. First the decoded dataset was projected on the GIS tool. Then a 2D graph of speed over time was plotted against to study their speed parameter. Following a detailed study of a frequency distribution of speed, a particular vessel type was selected and thoroughly investigated on speed profile. Also a spatio-temporal approach using selected clustering technique was applied to study the vessel's anchored positions and provide insights into the vessels particular characteristics. The objective of applying various methods such as plotting trajectories on GIS tool and studying speed profile over time and anchored positions was to rigorously investigate and identify characteristic features of different type of vessel movements.

A. Types of vessels

There are more than 18 classes of vessels defined in our data set, but we have concentrated on three, High Speed Craft, Passenger Vessels and Ferry Boats. In normal speech these terms may be interchangeable but they have specific meaning according to to the International Maritime Organization (IMO).

1. *High Speed Craft*

High-speed craft (SOLAS X/1.2):- A high speed water vessel for civilian use, also popularly called Fastcraft or Fast ferry. These vessels are a special category because of the potential risk they pose to other vessels because of their speed. Mostly these craft serve as passenger ferries; at times the large craft also known as catamarans also carry cars and other vehicles. High speed crafts are capable of maximum speed of at least 3.7 metres per second (m/s).

High speed craft travel 3 times faster than normal boats or ferry and may be used for personal use or carrying passengers. To investigate its speed characteristics and type of usage, a particular high speed vessel was selected. The study showed its average speed was 30 knots and spatio-temporal analysis revealed that it made frequent return trips between Auckland and Beachlands.

2. *Passenger Vessels*

A passenger ship has the capacity of carrying more than 12 passengers (SOLAS I/2). Passenger Vessel is a basic category because of the implications for rescue etc. One passenger boat's trajectory showed its path between Mahau sound and Pelorus sound. This trajectory suggested that this particular boat was involved in a sight-seeing activity or field trip tour-based role. A similar role was played by another ship. The analysis revealed that the average speed of these two passenger boats varied between 8 and 13 knots. Likewise, velocity graphs revealed more details of the roles they played.

3. Ferries

A ferry is primarily used to carry passengers and sometimes vehicles and cargo. AIS messages can use self-description of the type of vessel, so there is not a clear differentiation of ferries from other vessels. Most ferries operate on services on regular, frequent and return services.

Ferry boats were frequently observed between Wellington Harbour and Picton Harbour. The selected sample dataset used for investigation revealed that these ferries travelled via a Cook Strait route and passed Arapawa Island to reach Picton Harbour. In addition, it was also noticed that this route was frequently used by various types of boats, ferries and ships.

A. Example data

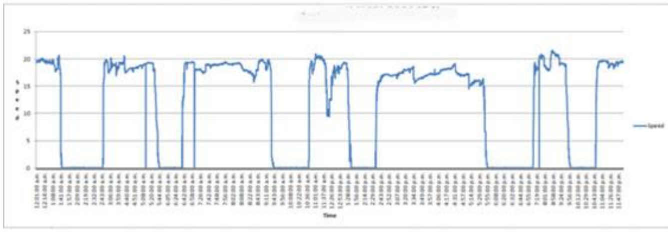

Figure 1: Velocity Profile of a ferry

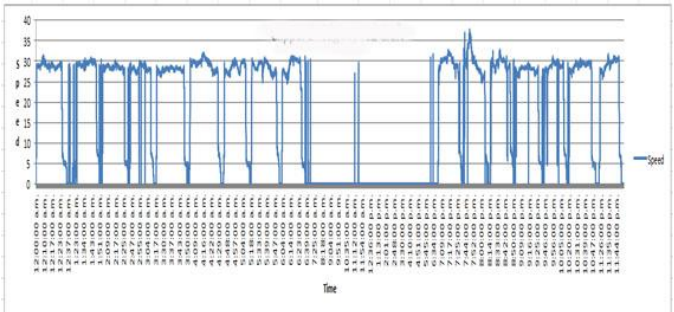

Figure 2: Velocity Profile of a High Speed Craft

The two examples (Figure 1 and Figure 2) demonstrate that the velocity profiles of different vessels can be easily distinguished by eye.

Frequency distribution of velocity could also be calculated Figure 3.

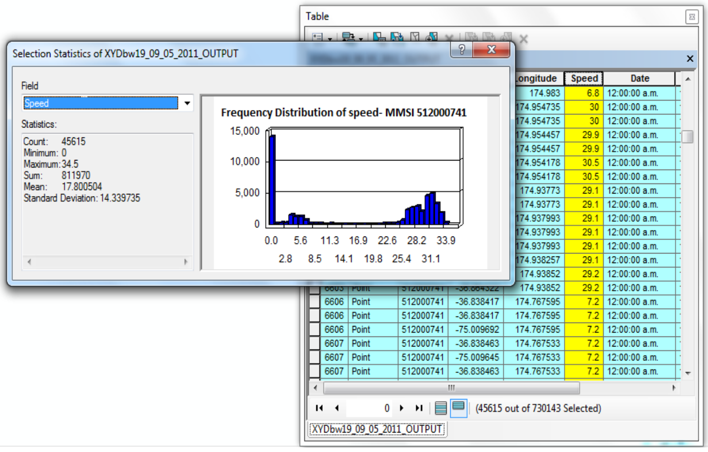

Figure 3 : Frequency distribution of velocity

Other means of displaying the data include mapping via arcGIS or another software package. Figure 4 demonstrates the concentration of vessel tracks near a port.

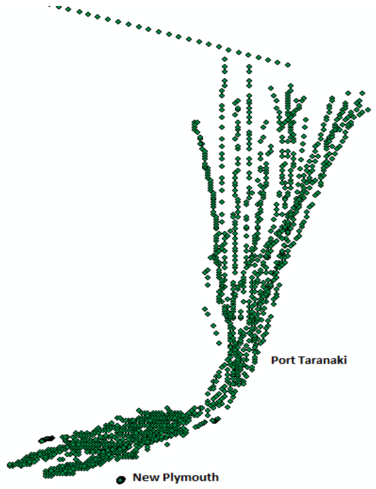

Figure 4 : Multi-vessel tracking

VI. CONCLUSIONS

This was an early attempt to study trajectories exclusively focused on New Zealand waters. The motivation was to identify patterns over time made by moving objects and their type and characteristics.

Movement parameters were identified in terms of location, speed and time. Movement parameters were categorized into spatial (location-oriented), temporal (time-based) dimensions. Selecting this approach had three purposes: to observe the time each type of vessel spent at particular locations, to obtain insight into movement data of each vessel type, and to the characteristics of each vessel type. Various trajectory patterns were identified during this analytical study, such as meeting, concentration, constancy in direction, constancy in speed and moving in clusters. The trajectories were found to be smooth and continuous in respect to time.

18 different types of vessels were identified in the dataset. Other types of vessels will be considered for trajectory analysis in future studies. The two-phased methodological study can be applied to identify the characteristics of any type of vessel.

Overall, conducting trajectory analysis utilizing real raw movement data was challenging in terms of handling, managing and performing analysis. On the other hand, the analysis revealed the hidden knowledge in movement dataset of AIS. Clustering-based stops and moves of trajectories [16] could be incorporated with the spatio- temporal approach in the future, and machine learning techniques may be useful to identify abnormal or potentially hazardous trajectories.

ACKNOWLEDGEMENTS

Our thanks to Kordia New Zealand for providing us with the raw AIS datasets between March 2011 and May 2011.

REFERENCES

[1] K. M. Kubach, M. C. Scott, and J. S. Bulak, "Recovery of a temperate riverine fish assemblage from a major diesel oil spill," *Wiley Online Library -Freshwater Biology,* vol. 56, pp. 503–518, 2011.

[2] H. Wang and R. Zimmermann, "Processing of continuous location-based range queries on moving objects in road networks," *IEEE Transactions on Knowledge and Data Engineering,* vol. 1, pp. 1-15, 2010.

[3] Y. Ping, T. Xinming, and W. Shengxiao, "Dynamic Cartographic Representation of Spatio-temporal data," *International Society for Photogrammetry and Remote Sensing. ,* vol. 37, pp. 7-12, 2008.

[4] C. Plumejeaud, H. Mathian, J. Gensel, and C. Grasland, "Spatio-temporal analysis of territorial changes from a multi-scale perspective," *International Journal of Geographical Information Science,* vol. 25, pp. 1597–1612, 2011.

[5] O. Wolfson and B. Xu, "Spatio-temporal databases in urban transportation," *IEEE Data Engineering,* vol. 33, pp. 1-8, 2010.

[6] P. Jiacaia, J. Qingshan, H. Jinxing, and S. Zheping, "An AIS data Visualization Model for Assessing Maritime Traffic Situation and its Applications," presented at the 2012 International Workshop on Information and Electronics Engineering, 2012.

[7] N. Willems, H. v. d. Wetering, and J. J. v. Wijk, "Evaluation of the Visibility of Vessel Movement Features in Trajectory Visualizations," *Wiley Online* vol. 30, pp. 801-810, 2011.

[8] P. Michel, R. Mathieu, and A. F. Mark, "Spatial analysis of oblique photo-point images for quantifying spatio-temporal changes in plant communities," *Applied Vegetation Science,* vol. 13, pp. 173–182, 2009.

[9] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Systems Research," *MIS Quaterly,* vol. 28, pp. 75-105, 2004.

[10] Z. Feixiang, "Mining ship spatial trajectory patterns from AIS database for maritime surveillance," presented at the 2nd IEEE International Conference on Emergency Management and Management Sciences (ICEMMS), Beijing 2011.

[11] M. S. Eide, O. Endresen, P. O. Brett, J. L. Ervik, and K. Røang, "Intelligent ship traffic monitoring for oil spill prevention: risk based decision support building on AIS," *Marine Pollution Bulletin,* vol. 54, pp. 145-148, 2007.

[12] IALA, "IALA Technical clarifications on recommendation ITU-R M. 1371-1," International Association of Marine Aids to Navigation and Lighthouse Authorities2004.

[13] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. D. Macedo, F. Porto, and C. Vangenot, "A Conceptual view on trajectories," *ScienceDirect,* vol. 65, pp. 126-146, 2008.

[14] NAVCEN, "Automatic Identification System Overview. ," U.S. Coast Guard Navigation Center2010.

[15] IMO, "International Maritime Organization," 2011.

[16] A. T. Palma, V. Bogorny, B. Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," in *2008 ACM symposium on Applied computing,* 2008, pp. 863-868.