RESEARCH NOTE

ON

AIS DATA ANALYSIS & RETRIEVAL

-by

Atharv Anirudha Kurdukar

Final Year B.Tech. in Information Technology

Dr. Babasaheb Ambedkar Technological University, Lonere

BACKGROUND:

As per UNCTAD Review of Maritime Transport 2018, there are around 1 million vessels present at sea at any moment in time. It's around $829 billion industry, shipping around 10.7 billion tons of cargo from one point to another every year. Thus, it is mandatory to have a monitoring system for tracking these vessels around the globe. Automated identification System (AIS) provides a way to track these vessels in (almost) real-time.

But, the AIS data sent by the vessel is prone to errors. So, it is necessary to filter this data in order to get appropriate results. Also, the collective data of all the vessels also provides a way to get some additional information regarding the vessel traffic. All of this requires analysis of available data in multiple ways.

AIS data analysis has multiple aspects. It can be used for defence applications, safety & security, maritime study, environmental protection, etc. This research note describes the role of data analysis in these aspects while stating the research conducted on the same.

Also, the AIS data is readily available for exploring, but collecting specific data for research purposes is not so easy. This research note elaborates the methods available for retrieval of data from various sources.

CONCEPTS:

Automatic Identification System (AIS)

The Automatic Identification System (AIS) is an automated, autonomous tracking system which is extensively used in the maritime world for the exchange of navigational information between AIS-equipped terminals.[1] It is a tool for monitoring and identifying maritime traffic which uses transponders to send messages through Very High Frequency (VHF) marine bandwidth [2].

AIS was primarily designed to operate in one of the following modes:

- In a ship-to-ship mode for collision avoidance.
- As a means for coastal states to obtain information about a ship and its cargo.
- As a traffic management tool when integrated with a Vessel Traffic System (VTS) [3]

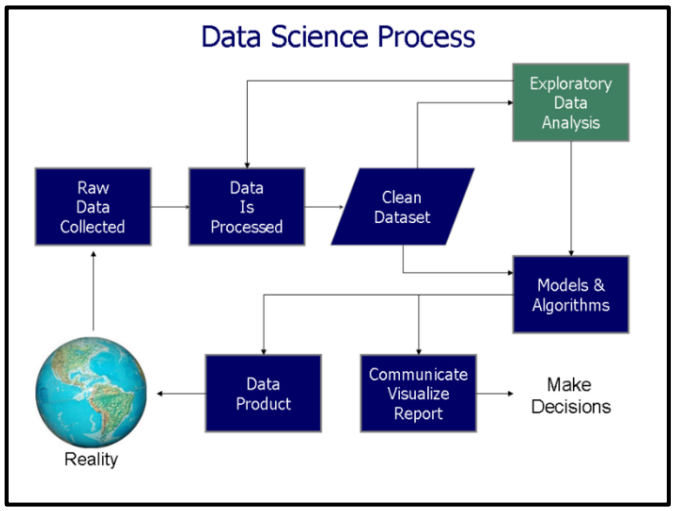
Originally, AIS was used terrestrially, meaning the signal was sent from the boat to land, and had a range of roughly 40 nautical miles (74 km). As ships began sailing further and further away from land, with the advent of Satellite AIS (S-AIS), they began sending the signal to low orbit satellites, which then relayed information back to land.

The AIS data that are exchanged are divided into three different types [3]:

1. **Static data** (e.g., vessel name, the dimensions of the vessel, IMO Number).
2. **Dynamic data** (e.g., vessel position, course over ground and heading, timestamp in UTC).
3. **Voyage-related data** (e.g., ship's draft, hazardous cargo type, destination & ETA)

Data Analysis

Data analysis is a process of **inspecting, cleansing, transforming** and **modeling data** with the goal of discovering useful information, informing conclusions and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, and is used in different business, science, and social science domains. In today's business world, data analysis plays a role in making decisions more scientific and helping businesses operate more effectively.[7]



Data analysis is a process for obtaining raw data and converting it into information useful for decision-making by users. Data are collected and analyzed to answer questions, test hypotheses or disprove theories.[8]

There are several distinguished phases of data analysis, as shown in fig. These phases are iterative and may provide feedback to the real world as well.

Data Retrieval

Data retrieval means obtaining data from a database management system in a well-defined format. In this case, it is considered that data is represented in a structured way, and there is no ambiguity in data. The database management system is used to select demanded data from the database.

Reports and queries are the two primary forms of the retrieved data from a database. A query language, such as **Structured Query Language (SQL)**, is used to prepare the queries.

DISCUSSION:

History

AIS was developed in the 1990s as a high intensity, short-range identification and tracking network. The 2002 IMO SOLAS Agreement included a mandate that required most vessels over 300GT on international voyages to fit a Class A type AIS transceiver. This was the first mandate for the use of AIS equipment and affected approximately 100,000 vessels. In 2006, the AIS standards committee published the Class B type AIS transceiver specification, designed to enable a simpler and lower-cost AIS device. The AIS transmission was firstly meant for terrestrial communication only.

Since 2005, various entities have been experimenting with detecting AIS transmissions using satellite-based receivers and, since 2008, companies such as **exactEarth**, **ORBCOMM**, **Spacequest**, **Spire**, and also many government programs have deployed AIS receivers on satellites.

Need for AIS Data Analysis

The AIS system was originally designed for radar augmentation and vessel traffic services (VTS), but AIS data can be used to collect information about traffic in a given area [4], which can be exploited in research, i.e. AIS becomes a source of big data. For more than a decade, researchers have increasingly used AIS data to analyse maritime traffic in a number of issues such as ship surveillance, tracking, security, collisions, shipping noise levels, and vessel emissions [5]. However, since AIS was not designed for research, there are several challenges in making use of the data for research, for example, the limited coverage of land-based receivers and problems storing and analysing the data [6].

Challenges in Tracking Vessels using AIS

There are certain challenges associated with tracking vessels using AIS data. The major problem lies due to incorrect data. As per a report, around 85% of the times, the data received from AIS vessel has at least one of the parameters which is incorrect. The problem of incorrect data may include many factors such as:

1. **Problems in equipment:** If the AIS transceiver gets a problem, it may transmit the default MMSI & IMO number instead of the number associated with the device. This results in an anomaly and may cause trouble while tracking.
2. **Errors in entering data:** As the voyage-related data need to be entered manually, it has a potential of human error. This error may be due to typing mistakes or due to the inability of the user to understand the situation.
3. **Falsification of data:** Sometimes, the data may be entered incorrectly by the user intentionally. This may be due to the malicious activities being carried out.
4. **Spoofing of data by outsiders:** As the data is being transmitted over the air, it is susceptible to intervention. A malicious nearby ship may capture the data, modify it and retransmit it for its own profit.

Another challenge is the inaccuracy in locating vessels via satellite (S-AIS). S-AIS uses Low Earth Orbiting (LEO) satellites. Thus, it is difficult to get precise position of the vessel while in deep sea due to changes in position of the satellite.

Also, it is not possible to mandate AIS in vessels below 300GT. Thus, it is very difficult to track smaller vessels. The AIS transceiver can also be turned off, leaving the ship unable to track.

Examples of AIS Data Analysis

1. Collision Avoidance

- **Simulate traffic rules:** In Kujala et al.2009 [11], they used AIS data to simulate traffic routes including the number of vessels, departure times, vessel dimensions, and sailing speed to estimate the frequency, time, and the location of collision incidents in the Gulf of Finland.
- **Calculate collision risk on route:** Son et al. (2012) [12] suggested a fuzzy algorithm to estimate collision risk among multiple ships in Korea using real-time AIS data, while Idiri and Napoli (2012) [13] proposed a rule-based method applied to the movements of ships under changing sea conditions which would give an identification of the risks in real-time and potentially trigger an alarm to help prioritize interventions.
- **Identify danger zones:** Jiacai et al.2012 [14] used the rate of turn, the ship acceleration, and ship encounters extracted from the AIS data, in order to develop an index of dangerous areas.

2. Fleet & Cargo Tracking

- **Identify vessel status as "sailing" or "stopped":** One of the notable studies was conducted in 2015 by Cazzanti and Pallotta [15] to identify and characterize main stationary areas for vessels. They created an algorithm to detect vessels' status as "sailing" or "stopped" using AIS data.
- **Analyse of fishing locations & fishing behaviour:** In the same study, If the identified stop area is not a port or harbour and it is far from coastline, it is asserted to be a fishing area for fishing vessels.

3. Underwater Data Analysis

- **Ocean ambient noise mapping:** The AIS can play an important role in allowing us to estimate the extent, quantity, and impacts of the shipping radiated noise. Many notable efforts have been made to model/map the radiated noise of the ships indicated through AIS by generating shipping radiated noise through the length and speed parameters derived from AIS data and applying channel models on them in order to supplement marine spatial planning, acoustic capacity building and many more such applications.

- **Ocean current estimation:** Thomas D Jakub [27] in his work has proposed estimation of ocean surface current based on the Ship Drift using the location parameters from AIS Data.

AIS Data Retrieval

Basically, there are three ways for retrieval of AIS data:

1. **Websites:** Most of the online vessel traffic systems (VTS) are available for free to explore. It includes websites such as **FleetMon Explorer**, **MarineTraffic**, **VesselFinder**, etc. These websites provide basic information received from the coastal stations for free, but to get additional information, the user has to pay a membership fee. Also, dome websites restrict the S-AIS data for premium membership users. The data is made available in almost real-time and is presented in user-friendly manner.
2. **Web Servers:** Some websites such as ____ have a web API, which provides raw as well as processed data in JSON format. This data is mostly available via paid membership plans. The data is provided whenever the API call is made. Some websites also provide historical data over a geographical area, path of a vessel fleet, etc.
3. **Government Bodies:** Many countries provide AIS data related to their country for free via government portals. The data is historical as opposed to the real-time data provided by private entities. This data is available for free for research purposes. Some countries do not have a public government portal, but provide the data on-demand for research.
4. **Commercial Applications:** There are certain commercial Vessel Traffic Management Systems (VTMS) which are specifically built and / or customized for ports, ship owners, ship agents, and defence. They provide highly accurate real-time data with possibly AI prediction for vessel routes and other advanced features.

WAY AHEAD:

Opportunities for AIS Data Analysis

1. **Data Quality Improvement:** Reducing data error rate is the most significant challenge. This can be reduced to some extent by reducing human errors. Some of the solutions could be:
    - Training may be provided to the crew in order to reduce the errors.
    - AI based solutions can be used to correct the typing mistakes.
    - Algorithms can be developed to identify a ship using multiple parameters instead of relying solely on MMSI or IMO number.
    - The developed algorithms should be implemented on coastal stations, which will correct the data and broadcast the corrected data rather than repeating the same ambiguous data.
2. **Marine Spatial Planning:** In India, marine spatial planning is not yet done. The research has been conducted by different authorities on their own, but there is no proper integration as of now. Thus, the AIS data analysis can be used to identify the major routes and plan the future missions accordingly for better economical benefits.

Opportunities for AIS Data Retrieval

Data Availability

Currently, the data is very much restricted and is not freely available for research purposes. Some governments such as US & UK provide raw historical data for free via their portals, but this need to be expanded for other governmental agencies as well. This will generate more opportunities for further analysis.

REFERENCES:

[1] ***"What is the Automatic Identification System (AIS)?"***, MarineTraffic

[2] Iphar, Clément & Napoli, Aldo & Ray, Cyril & Alincourt, Erwan & Brosset, David. (2016), ***"Risk Analysis of falsified Automatic Identification System for the improvement of maritime traffic safety"***.

[3] Leica Geosystems Inc., ***"The Complete Guide to Automatic Identification System"***

[4] K. Gunnar Aarsæther and T. Moan, ***"Estimating Navigation Patterns from AIS"***, Journal of Navigation, vol. 62, no. 4, pp. 587–607, 2009.

[5] Richard L. Shelmerdine, ***"Teasing out the detail: How our understanding of marine AIS data can better inform industries, developments, and planning"***, Marine Policy, Volume 54, 2015,

[6] M. Robards, G. Silber, J. Adams, J. Arroyo, D. Lorenzini, K. Schwehr, J. Amos, ***"Conservation science and policy applications of the marine vessel Automatic Identification System (AIS)—a review"***, Bull. Mar. Sci., 92 (1) (2016), pp. 75-103

[7] Xia, B. S., & Gong, P. (2015), ***"Review of business intelligence through data analysis"***, Benchmarking, 21(2), 300-311. doi:10.1108/BIJ-08-2012-0050

[8] Judd, Charles and, McCleland, Gary (1989), ***"Data Analysis"***, Harcourt Brace Jovanovich. ISBN 0-15-516765-0.

[9] Shridhar Prabhurman, Dr. (Cdr) Arnab Das, ***"Automatic Identification System (AIS) Data Analysis"***, Maritime Research Center, Pune

[10] Shreeya Nelekar, Dr (Cdr) Arnab Das, ***"Application of Machine Learning and Artificial Intelligence for Automatic Identification System (AIS) Database"***, Maritime Research Center, Pune

[11] Kujala P, Hänninen M, Arola T, Ylitalo J (2009) ***"Analysis of the marine traffic safety in the Gulf of Finland"***. Reliab Eng Syst Saf 94(8):1349–1357

[12] Son NS, Kim SY, Lee CS (2012) ***"The monitoring system of collision risks among multiple ships"***. Annual of Navigation 19(1):149–163

[13] Idiri B., Napoli, A. (2012). ***"The automatic identification system of maritime accident risk using rule-based reasoning"***. In Proc.of the 7th International Conference on System of Systems Engineering - IEEE SOSE. pp. 125-130

[14] Jiacai P, Qingshan J, Jinxing H et al (2012) ***"An AIS data visualization model for assessing maritime traffic situation and its applications"***. Procedia Eng Elsevier:365–369

[15] Cazzanti L, & Pallotta G. (2015) ***"Mining maritime vessel traffic: promises, challenges, techniques"***. In OCEANS 2015-Genova. 1-6. IEEE

[16] Jakub, Thomas Dominic, ***"Estimating Ocean Currents from Automatic Identification System Based Ship Drift Measurements"*** (2013). Aerospace Engineering Sciences Graduate Theses & Dissertations. 55.