How big data enriches maritime research

– A critical review of Automatic Identification System (AIS) data applications

Dong Yang[a], Lingxiao Wu[a]*, Shuaian Wang[a], Haiying Jia[b], Kevin X. Li[c]

[a] *Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hong Kong, China*

[b] *Center for Applied Research at NHH, Norwegian School of Economics, Bergen, Norway*

[c] *Department of International Logistics, Chung-Ang University, Dongjak-Ku, South Korea*

Abstract

The information-rich vessel movement data provided by the Automatic Identification System (AIS) has gained much popularity over the past decade, during which the employment of satellite-based receivers has enabled wide coverage and improved data quality. The application of AIS data has developed from simply navigation-oriented research to now include trade flow estimation, emission accounting, and vessel performance monitoring. The AIS now provides high frequency, real-time positioning and sailing patterns for almost the whole world's commercial fleet, and therefore, in combination with supplementary databases and analyses, AIS data has arguably kickstarted the era of digitization in the shipping industry.

In this study, we conduct a comprehensive review of the literature regarding AIS applications by dividing it into three development stages, namely, basic application, extended application, and advanced application. Each stage contains two to three application fields, and in total we identified seven application fields, including (1) AIS data mining, (2) navigation safety, (3) ship behavior analysis, (4) environmental evaluation, (5) trade analysis, (6) ship and port performance, and (7) Arctic shipping. We found that the original application of AIS data to navigation safety has, with the improvement of data accessibility, evolved into diverse applications in various directions. Moreover, we summarized the major methodologies in the literature into four categories, these being (1) data processing and mining, (2) index measurement, (3) causality analysis, and (4) operational research.

Undoubtedly, the applications of AIS data will be further expanded in the foreseeable future. This will not only provide a more comprehensive understanding of voyage performance and allow researchers to examine shipping market dynamics from the micro level, but also the abundance of AIS data may also open up the rather opaque aspect of how shipping companies release information to external authorities, including the International Maritime Organization, port states, scientists and researchers. It is expected that more multi-disciplinary AIS studies will emerge in the coming years. We believe that this study will shed further light on the future development of AIS studies.

Keyword: AIS data, Data mining, Navigation safety, Ship behavior analysis, Environmental evaluation, Advanced applications of AIS data

1. Introduction

After diesel propulsion and containerization, big data technologies have been widely recognized as a significant advancement within the shipping, one that will impact most shipping operation patterns over the next one or two decades.

The AIS was developed to avoid ship collision accidents, and it has been used in maritime transportation for over two decades. Originally, AIS data was highly regionalized and originally difficult to collect, because direct ship-to-ship, ship-to-coast, or coast-to-ship AIS communication was limited to the very-high-frequency (VHF) radio wave range, which only covered 10-20 nautical miles. Since 2008, satellites equipped with AIS receivers have been able to receive AIS data transmitted by onboard AIS transceivers worldwide. On the one hand, these satellites serve as supplementary data sources for ships and coastal authorities in busy port areas where terrestrial AIS receivers may be overloaded due to the large volume of data. On the other hand, they provide an easy means for AIS data to be collected on a global scale in almost real time. At present, AIS data can be easily collected from commercial websites that provide access to AIS databases (e.g., the websites of companies such as Elane, ExactEarth, Marine Traffic, ORBCOMM, Spacequest, and Spire). With the constantly improving quality and completeness of AIS data, the applications of AIS data have expanded from navigation safety to include many other aspects too.

The rapid growth in publications, and the increasingly diverse topics covering AIS applications, have indicated the expansion of AIS in recent years. An updated AIS can in general improve traditional maritime studies in three ways. First, the global coverage and easy accessibility of AIS data extend the study scope from an individual ship or particular region to global or regional territories. For example, AIS data has been applied to mapping ship activities (Kaluza et al., 2010) and to assessing their impact on environments (Winther et al., 2014) at a global level. Second, the unparalleled high-resolution and real-time characteristics of AIS data improve the accuracy of shipping behavior analysis. For example, AIS data has been used to improve the accuracy of ship trajectory extraction and prediction (Arguedas et al., 2018) and traffic monitoring (Perera et al., 2012). Third, the discreteness and abundance of individual ship data enable researchers to look at problems from a bottom-up view rather than a top-down view. For example, the global oil trade has been forecast using AIS data by aggregating shipping volume on the sea (Adland et al., 2017); port performance (Chen et al., 2016) and the environmental effects of shipping (Gerritsen et al., 2013) are being evaluated on a per-ship basis.

The expansion of AIS applications is expected to continue and even accelerate with ongoing improvements in the quality and accessibility of AIS data. With this background, it is important for academics, practitioners, and government to understand the status of AIS applications (so that they can use AIS for the same purpose without inventing the wheel again). Therefore, in this paper, we research the following: (1) How AIS data can be applied to improving shipping industry efficiency and to solving practical problems; and (2) what the potential is for future applications of AIS. In particular, we believe that the paper's summation will help relevant academics, practitioners, and governments in two ways: (1) For researchers or practitioners who have already worked on AIS data, they will become aware of extended or related work on the further application of AIS, which will help improve their current work; and (2) for researchers and practitioners who are not familiar with AIS data, we show them the potential for improving their work using AIS data. We have also identified the limitations in existing applications of AIS data, and proposed directions in which academics, industry, and governments can do further work.

Although several articles have reviewed the methods used in AIS, they have focused on data mining and the applications of AIS data to navigational safety (Tu et al., 2017, Sidibé and Shu, 2017, and Zhao et al., 2014). To the best of our knowledge, no existing study has provided a comprehensive and up-to-date review regarding the applications of AIS data for solving a broad range of maritime problems. Academia and the industry have both taken significant steps forward with AIS. Hence, summarizing previous related works and forecasting the direction for future studies on AIS are significant. This study aims at presenting a panoramic view of state-of-the-art AIS data studies in order to stimulate research perspectives.

This rest of this paper is organized as follows: Section 2 presents an overview of AIS, describing what it is, its development, and its limitations. Section 3 introduces the method and structure used in the literature review. Section 4 reviews studies on the basic applications of AIS data, including data mining and navigational safety. Section 5 focuses on extended applications of AIS data to the analyses of ship behavior and environmental risks caused by shipping activities. Section 6 reviews more advanced applications of AIS data. Section 7 discusses the research trends of AIS data applications. Finally, Section 8 proposes the potential extensions for applications of AIS data. The abbreviations used in the paper are summarized in Table 1.

<Table 1>

2. Overview of AIS

AIS was developed in the 1990s, with the primary goal of preventing ship collisions and enhancing navigation safety. Through the use of VHF, ships equipped with AIS can broadcast and receive messages to and from other ships or coastal authorities that are also equipped with AIS. The AIS enables ships and coastal authorities to communicate with one another over a long distance. The International Maritime Organization (IMO) requires all international voyage ships with a gross tonnage above 300, and all passenger ships, to be equipped with an AIS transmitter (IALA, 2004). In addition to the IMO, governments and other authorities in different nations enforce AIS applications in ships registered with them to improve safety and security.

The AIS transceivers are of two types (Classes A and B), having different numbers of reported data fields and reporting frequencies. The information broadcast by a ship's AIS transceiver (Class A) can be grouped into 11 data fields, which can be further classified into 3 types, namely, static information, dynamic information, and voyage-related information. Dynamic information is automatically transmitted by a Class A AIS transceiver every 2-10 seconds, depending on the ship's speed while it is underway, and every 3 min while it is anchored. At the same time, a Class A AIS transceiver's interval between broadcasting static and voyage-related information is 6 minutes, regardless of navigational status. Class B transponders transmit a reduced set of data when compared with Class A transponders, omitting the IMO number, draught, destination, ETA, rate of turn, and navigational status. The reporting intervals from Class B transponders are also sparser when compared with those of Class A transponders, being a minimum of 5 seconds. Table 2 provides a detailed classification and description of these data fields.

<Table 2>

Apart from the information that can be obtained directly from AIS data, additional information can be derived by combining AIS data with data from other databases. For example, the following information can be derived:

- Port-to-port average speed, which is equal to the port-to-port voyage distance between two ports divided by voyage time, where voyage time can be calculated based on the time stamps reported at the two ports; the voyage distance can be found from the corresponding navigation distance tables.
- Cargo weight, which can be estimated based on draught and ship sizes (Jia et al. 2018).
- Technical ship specifications, including DWT, capacity, design speed, and design

draught, can be obtained from fleet databases, such as Clarkson's world fleet register, by using the IMO number.

- Port-to-port bunker consumption, which can be estimated based on the speed, distance between two ports, and technical ship specifications, such as DWT and capacity.

Although AIS provides a powerful and easy-to-access database for maritime practitioners and researchers, errors and inaccuracies may still exist in AIS data. Most errors and inaccuracies occur in data that is manually entered into the system. Such data includes static information, such as MMSI, ship's length and width, IMO number, name, type, call sign, and voyage-related information, such as the ETA, draught, and intended destination. For example, an MMSI, which ought to be unique for each ship, can be shared by different ships (Mazzarella et al., 2013). Besides, empty or inaccurate ETA reports are very frequent in AIS data (Watson et al., 2015). However, even data that is automatically generated by the sensors can be erroneous. For example, Harati-Mokhtari et al. (2007) indicated that an AIS may report erroneous positional information if the position fixing system is not working or is improperly connected to the AIS equipment. In extreme cases, it has even been reported that some ships have turned off their AIS. Therefore, one important step when dealing with AIS data is to identify and filter incorrect and inaccurate information.

3. Literature review method and structure

Since the introduction of AIS to maritime transportation, its promising practical applications have been gradually recognized in academic circles, starting with AIS data mining and its applications to navigation safety. Thereafter, with the improvement in data quality and accessibility, studies gradually expanded from navigation safety into broader and more advanced applications of AIS data. AIS data has been applied in a great variety of topics in maritime research, involving different disciplines (including Statistics, Information Techniques, Biology, and Environmental Science). In this paper, to provide a comprehensive illustration as to how AIS data has been applied in solving different problems, and to trigger new thoughts about its application across different disciplines, we conduct a systematic review of the various applications of AIS data. During the review process, the literature is searched in a detailed and thorough manner, and relevant topics, methodologies, and trends are identified, analyzed and synthesized to produce a panoramic view of AIS data studies.

We searched the databases of Scopus, Sciences Citation Index, Google Scholar, and journals, including but not limited to Accident Analysis & Prevention, Maritime Economics & Logistics, Marine Policy, Maritime Policy & Management, Marine Pollution Bulletin, Ocean Engineering, Transportation Research Part A, Transportation Research Record, and The Journal of Navigation, using AIS as the keyword. A total of 171 articles were found (including 128 journal papers, 39 conference papers, and 4 technical reports). Table 3 shows the top 14 journals that have published at least 2 papers on AIS data applications. We divided the publishing time into five periods with intervals of 3-4 years. Among the 171 studies reviewed in this paper, 104 were published in these journals, and the top 3 journals contributed to more than half of the journal publications, these being The Journal of Navigation, Marine Pollution Bulletin, and Ocean Engineering. The number of papers published in each period grew steadily from 2003 to 2014, and has undergone a sharp increase since 2015. With regard to the distribution of journals, papers in the first 2 periods from 2003 to 2008 were mainly published in journals that focus on navigation safety and the marine environment (i.e., The Journal of Navigation and Marine Pollution Bulletin). The next 2 periods have witnessed a rapid increase in the number of papers published in journals, such as Marine Policy, Ocean Engineering, and Transportation Research Record, which focus on ship behavior analysis. The distribution of journals during the last period was the most diversified, implying that research attention has widened out to cover a broader range of applications.

<Table 3>

In terms of their topics, we divided the literature into three areas: basic applications of AIS data (BA), extended applications of AIS data (EA), and advanced applications of AIS data (AA). Among the 171 articles we reviewed, 81 papers fall under BA, 69 under EA, and 21 under AA. For each area, 2 to 3 themes were identified. The overview of literature classification is given in Figure 1. The studies will be reviewed in the next sections, following the structure in this figure.

<Figure 1>

4. Basic applications of AIS data

One practical difficulty in applying AIS data is that the volume of data is extremely large, even for traffic data within a restricted geographical area covering a short period. For example, if AIS data is transmitted every 10 s, then a total of over 3 million records can be generated for a single ship in 1 year. The AIS data for 5,000 ships over 3 years includes approximately 40 billion records. Therefore, AIS data mining is the research basis of many relevant studies. AIS was initially introduced to avoid ship collisions and improve navigation safety and this serves as the most basic application of AIS data. Because this part has already been reviewed by Tu et al.

(2017), Sidibé and Shu (2017) and Zhao et al. (2014), in this section we only briefly explain the ideas and review the most up-to-date studies.

4.1 AIS data mining

Data mining is a knowledge extraction process based on raw data. Raw AIS data comprises a group of spatially and temporarily scattered points from which only limited information can be directly obtained. Therefore, data mining is significant for AIS data and provides the foundation for the majority of studies based on AIS data. The most common methods for processing and mining AIS data include trajectory extraction, trajectory clustering, and trajectory prediction.

Trajectory extraction and clustering play a vital role in AIS data applications because they form the foundation of many further studies (including trajectory prediction and analysis, anomaly detection, and collision avoidance). Trajectory extraction refers to the construction of a ship's trajectory based on the reported spatiotemporal sequence data. Trajectory clustering is based on the extraction results, which refers to the algorithms for grouping similar trajectories as a whole, thereby discovering common trajectories (Lee et al., 2007). Trajectory extraction and clustering using AIS data have elicited considerable attention in the literature, and the most recent explorations include those of Arguedas et al. (2018) and Wang et al. (2017).

Based on trajectory extraction and clustering, trajectory prediction can be carried out to predict a ship's short-term future position and trajectory. Trajectory prediction enables navigators or coastal authorities to detect possible threats and take preventive actions as early as possible. A route prediction algorithm based on the Ornstein-Uhlenbeck stochastic process was proposed by Pallotta et al. (2014), in which historical trajectories extracted from raw AIS data were used to estimate the parameters of the algorithm.

4.2 Navigation safety

A considerable number of early studies have been conducted on how to utilize AIS data to effectively avoid shipping accidents. These studies have mainly developed in two directions: collision avoidance for ships and traffic surveillance enhancement for coastal authorities.

Ship collision avoidance has been receiving attention from mariners and researchers since the 1970s. With the application of AIS data, collision avoidance has gone beyond direct ship-to-ship communication, and AIS data is now being used in state-of-the-art techniques that aim to avoid ship collisions and enhance navigation safety. These techniques include ship domain construction, collision risk assessment, and route planning.

Ship domain refers to the surrounding effective waters in which the navigator of a ship wants to keep clear of other ships or fixed objects (Goodwin, 1973). In the study of Hansen et al. (2013), AIS data was used in ship domain construction. In addition to ship domain construction, AIS data is also used in assessing ship collision risks (e.g., Li et al., 2017). Studies on risk assessment can generally be classified into five categories: (1) Calculation of the closest point of approach or time to closest point of approach among ships; (2) quantification of collision risks; (3) detection of situations where ships are under high threat of collision; (4) estimation of collision frequency in certain areas; and (5) development of precautionary systems for collision avoidance.

In navigation, route planning aims to help ship navigators identify a route with a low collision risk and detour cost. Route planning is built upon the results of ship domain construction and collision risk assessment. The most commonly used methods in route planning are heuristic algorithms (e.g., evolutionary algorithms and ant colony optimization). Such studies include those of Tsou and Hsueh (2010), and Kim et al. (2017).

AIS data can also be used to improve the surveillance capability of coastal authorities, and the two main foci of studies in this category are anomaly detection and maritime traffic monitoring.

In maritime navigation, anomaly detection is used by coastal authorities to detect illegal, suspicious, or unsafe behaviors (Kowalska and Peel, 2012). The AIS data provides a highly efficient approach for ship anomaly detection, whereby historical data is used to construct normal patterns (i.e., trajectory clustering), and ship behavior is monitored using real-time data (Zhen et al., 2017).

Another application of AIS data for traffic surveillance is to develop a maritime traffic monitoring system. Perera et al. (2012) developed a maritime traffic monitoring system based on vessel detection, tracking, state estimation, and trajectory prediction. In addition, Pan et al. (2012) introduced a visualization system for analyzing traffic situations and detecting dangerous shipping areas.

5. Extended applications of AIS data

The AIS reports ship positioning information with unparalleled resolution. Such information enables efficient analyses of ship behavior in different areas. In addition, the shipping industry is one of the largest contributors of various pollutants, and with the surge in real-time ship position data, analysis and monitoring of the environmental risks of shipping activities has become feasible. Following on from navigation safety, the application of AIS data has thus been extended to analyzing ship behavior and the impact of ships on the environment. This section provides a review of the extended applications of AIS data.

5.1 Ship behavior analysis

Ship trajectories include information about ship behavior. The following review is conducted in terms of three categories of ship behavior that we summarized from the literature, namely, fishing activities, ship behavior in open waters, and ship behavior in restricted waters.

Fishing activities

Fishing activities have elicited considerable attention in the literature due to their tremendous impact on marine ecosystems. Some studies have explored the potential of using AIS data to monitor illegal, unreported, and unregulated (IUU) fishing activities. Studies in this area have focused on IUU activity identification and estimation of the catch amount caused by IUU fishing activities. For example, a logistic regression model was proposed by Sheng et al. (2018) to classify and identify unknown types of ship. In addition to monitoring IUU fishing activities, AIS data has also been used to analyze the footprint (i.e., fishing area and intensity) of fishing activities. The footprints of fishing activities of different regions have been explored by researchers using AIS data reported by fishing ships in these regions. For instance, Vespe et al. (2016) proposed a method for mapping the fishing footprint at the European level by extracting fishing activities from AIS data.

Ship behavior in open waters

Some studies have been conducted on ship behavior in open waters where traffic flow has no boundary limits. Studies in this regard have focused on analyzing the spatial and temporal distributions of traffic. An example of these studies is the work of Breithaupt et al. (2017), who constructed the boundaries of shipping routes between ports along the Atlantic coast of the US, based on individual trajectories obtained from AIS data. They found that 95% of traffic is encompassed by such boundaries. In addition, AIS data has been used to analyze the density of global and regional shipping networks. Kaluza et al. (2010) constructed a global network of cargo ship movements using AIS data.

Ship behavior in restricted waters

Ship behavior in restricted waters focuses on shipping channels with limited widths (including straits, bays, ports, and inland waterways). Different from studies of open waters that pattern the overall traffic flow, studies of restricted waters have tended to analyze detailed navigation patterns (including ship drafts, speeds, and courses) of ships in a certain shipping channel. For example, using AIS data Mitchell and Scully (2014) assessed the influence of tide on the transit times of ships that called at deep-draft ports in the US. They also used AIS data to analyze the travel time of ships sailing in different inland waterways in the US. Zhang et al. (2017), using the historical trajectories of ships sailing in Singapore port waters, estimated the ship traffic volumes in different origin-destination pairs in a port.

It is worth mentioning that many navigation infrastructures, such as dredged channels, navigation locks, and coastal jetty and breakwater structures, are maintained within the aforesaid restricted waters. Using AIS data to analyze ship behaviors in such restricted waters also helps to evaluate the service level of the navigation infrastructures within them (e.g., Scully and Mitchell, 2017).

5.2 Environmental evaluation

Maritime transport emits around 1000 million tons of CO_2 annually, according to the estimates of the IMO, and is responsible for approximately 2.5% of global greenhouse gas (GHG) emissions. Apart from GHG emissions, shipping also contributes to other environmental threats, including oil spills, ocean noise, and seabed damage. It also poses threats to marine creatures.

Ship emission analysis

The AIS data has been widely applied in monitoring ship emissions. Most studies in this area have considered the calculation of emission inventories, and AIS data is commonly used to derive ship activities. A shipping emission inventory calculation method was considered by Winther et al. (2014). The authors proposed a method for calculating emission inventories for Arctic shipping based on AIS data and other factors related to emission. The impact of ship emissions on the air quality of sea and coastal areas has also been investigated. Kivekäs et al. (2014) introduced a method for estimating the contribution of ship traffic to the concentration of aerosol particles in the downwind of shipping routes, in which AIS data was used to identify major shipping routes.

Oil spill risk analysis

Another important application of AIS data to environmental risk analysis is to assess the possibility of oil spill accidents. Eide et al. (2007) proposed a dynamic environmental risk model for oil spills, and their model estimates the real-time environmental risk and forecasts the future environmental risk of drift grounding accidents for oil tankers. In addition to risk assessment and monitoring, some works have focused on after-accident resilience. Longépé et al. (2015) developed a method for identifying a potential polluter based on AIS data, SAR images, and the oil drift model.

Ecosystem impact analysis

The impact of shipping activities on ecosystems has also elicited considerable attention from shipping academia. Researchers have considered the impact of shipping activities, such as shipping-related noises, oil spills, and seabed damage, on marine creatures (e.g., Campana et al., 2017, Allen et al., 2018 and Gerritsen et al., 2013).

On top of this, AIS data also provides the possibility of better protecting seaborne animals by enabling better monitoring of shipping activities in critical habitat areas. For example, Reeves et al. (2007) proposed using AIS data to monitor the speeds and routes of ships sailing in the habitat areas of North Atlantic Right Whales.

Green shipping

To reduce emissions generated by shipping activities, several green shipping strategies have been proposed, and AIS data has been used to evaluate such strategies. Using AIS data, Jia et al. (2017a) investigated the impact of the virtual arrival policy, which aims to reduce the waiting time of ships at anchor for energy efficiency. A similar study was conducted by Andersson and Ivehammar (2017a), in which the authors proposed reducing the anchorage time of ships by adjusting their speed. AIS data was used to evaluate the effectiveness of the policy (i.e., the reduction of fuel consumption and emissions) for the Port of Gothenburg and ports in the Baltic Sea. In their inspiring work, Watson et al. (2015) discussed the perspective of using an integrated information system that includes data from AIS, shipping agencies and port authorities to minimize the anchorage time of ships by optimizing their sailing speed.

6. Advanced applications of AIS data

In recent years, the applications of AIS data have expanded into a broader scope with increasingly more advanced topics. We can categorize these studies into three distinct groups: trade analyses, ship and port performance evaluation, and Arctic shipping.

6.1 Trade analysis

With its increasing availability and completeness, AIS data can provide more detailed and timely trade statistics compared with traditional data sources (e.g., official customs data and port throughput data). Thus, Adland et al. (2017) adopted AIS data as an alternative statistic source for global trade analyses. An empirical study based on crude oil export statistics was conducted in their research. Jia et al. (2017b) proposed an algorithm for automatically generating seaborne transport pattern maps based on AIS data. The algorithm automatically detects major ports and zones and aggregates "real-time" trade flows among them. Prochazka and Adland (2017) used AIS data to analyze the location distribution of VLCC oil tankers on a global level.

6.2 Ship and port performance

AIS data is also used to evaluate the performance of ships. Factors involved in evaluating ship performance include ship utilization, speed, and voyage cost.

Jia et al. (2018) measured deadweight utilization using ship draught derived from AIS data. They ascertained that although AIS data alone is insufficient for tracking cargo flows on a per-ship basis, it can be used to obtain average utilization rates at ship-type levels. An empirical study to identify the impacts of different factors on ship capacity utilization was conducted by Adland et al. (2016). Adland and Jia (2016), using AIS data with a regression model, detected how ship speed is affected by other shipping factors. Andersson and Ivehammar (2017b) proposed a dynamic route planning method to optimize shipping cost in the Baltic Sea region. To show the effectiveness of the proposed method, a cost-benefit analysis was conducted based on AIS data collected from ships in this region.

AIS data has also been applied in analyzing port performance. In the work of Chen et al. (2016), ship positions obtained from an AIS were used to derive the ship traffic, container throughput, berth utilization, and terminal productivity of container ports. Jia et al. (2017c) used AIS data to evaluate the connectivity of major Norwegian ports. AIS data has also been used to assess the resilience of port operations following major disasters and other disruptive events (Farhadi et al. 2016). More recently, the U.S. Department of Transportation Bureau of Transportation Statistics has been using AIS data to measure the dwell time of ships at different ports in the US. (e.g., Bureau of Transportation Statistics, 2018).

6.3 Arctic shipping

With the recession of Arctic sea ice and the discovery of new natural resources in the Arctic, Arctic shipping has become extremely popular (Buixadé Farré et al., 2014). As the most popular ship tracking tool, AIS data has also been utilized in studying Arctic shipping.

One of the most direct applications of AIS data in Arctic shipping is to improve shipping surveillance. Aase and Jabour (2015) pointed out that knowledge of ship positioning derived from AIS data can also help improve the efficiency of polar search and rescue. Reeves et al. (2012) suggested using AIS data to improve the monitoring of ship activities in the Arctic area so as to reduce environmental damage caused by Arctic shipping.

The AIS data is also used to study the development trend of Arctic shipping. Eguíluz et al. (2016) presented a seasonal pattern of shipping activities in Arctic areas using ship densities derived from AIS data. They also analyzed the changes in Arctic shipping patterns for the period of 2010-2014. Löptien and Axell (2014) analyzed the relationship between shipping speed, as derived from AIS data, and the sea ice conditions in Arctic areas.

7. Summary

On the basis of the reviewed literature, this section first presents the trend of publications and the relationships among different research areas. Then, we demonstrate the evolution of research topics and summarize the major methods applied.

7.1 Overall trend of publications

In Figure 2 we summarize the distribution of the 171 articles on AIS studies over the three areas covering the period from 2003 until the present (March 2018). It shows that the number of studies has grown rapidly since 2010, and the topics have also become more diverse. The number of extended application papers currently exceeds the number of basic application papers, and since 2014 the number of advanced application papers is nearly equal to that of basic application papers.

<Figure 2>

7.2 Summary of data, disciplines, and relationships of AIS studies

Figure 3 presents a panoramic overview of the reviewed studies. In particular, we demonstrate the data used in these studies, the disciplines involved, and the relationships among different research themes.

As shown in the figure, AIS serves as the data source for all these studies, and is complemented by other data sources when it comes to a specific area of study. In a similar way, as far as the involved disciplines go, while Statistics forms the backbone for all research areas, knowledge from other disciplines (e.g., Computing Science, Biology, and Information Techniques) is also essential for studies with different focuses.

We also note that among all the different research areas, AIS data mining constitutes the foundation for other studies, since it converts the scattered and disordered raw data to the organized and readable information. Therefore, finding data mining methods that can handle a larger volume of data with higher processing speeds and greater accuracies is always a hot topic in AIS data studies. This explains why each year a considerable proportion of studies are focused on basic applications of AIS data (see Figure 2).

In addition to data mining, ship behavior analyses constitute another integral part of the research matrix. Based on the data provided by data mining, ship behavior analyses generate patterns of seaborne traffic in various locations. These patterns, augmented by other data, are used as the foundation for many advanced topics that look at maritime transportation from different angles (e.g., environmental analyses, trade analyses, and Arctic shipping).

<Figure 3>

7.3 Evolution of research topics

The AIS data has been applied to various applications such data mining, maritime safety enhancement, ship behavior analysis, environmental evaluation, trade analysis, ship and port performance evaluation and Arctic shipping analysis. Figure 4 presents the evolution of research topics in AIS data applications, where studies are aggregated into distinct items according to the topics and publishing times. Items displayed in the bold text represent the first appearance of the corresponding topics.

Figure 4 shows that initial topics focused on navigation safety (e.g., anomaly detection, route planning, domain construction, risk assessment, and traffic monitoring). Later, study focus shifted to ship behavior analysis (e.g., behavior analysis in open and restricted waters and fishing activities) and environmental evaluation (e.g., oil spill risk analysis, emission analysis, ecosystem impact analysis, and green shipping). With the improvement in data accessibility, the AIS data has recently spread to diverse other applications in various directions (e.g., trade analysis, Arctic shipping, and ship and port performance).

<Figure 4 >

7.4 Summary of major methods applied in the studies

This section summarizes the methods used in the reviewed literature. The methods are divided into four categories: data processing and mining, index measurement, causality analysis, and operational research. Table 4 summarizes the major methods along with their typical applications.

<Table 4>

Data mining focuses on deriving knowledge from raw AIS data. The most basic methods for AIS data mining are trajectory extraction, clustering, and prediction. Many applications, including anomaly detection (Zhen et al., 2017), maritime monitoring (Perera et al., 2012), and shipping density analyses (Kaluza et al., 2010), have been implemented based on trajectory extraction or prediction results. Visualization techniques have also been utilized in AIS data studies to improve maritime monitoring (Pan et al., 2012) or help analyze trade trends (Jia et al., 2017b).

The AIS provides data that records shipping activities with unparalleled high resolution. This data not only makes shipping activities more visible but also makes them more analyzable. Many studies have developed methods of building evaluation indices to measure the performance of shipping activities by using AIS data. These methods and applications include ship domain construction (Hansen et al., 2013), collision risk assessment (Li et al., 2017), ship emission inventory (Winther et al., 2014), oil spill risk assessment (Eide et al., 2007), evaluation of green shipping policies (Jia et al., 2017a), and Arctic shipping trend analysis (Eguíluz et al., 2016).

Causality analysis aims to identify relationships among different shipping relevant factors, as well as between them and external factors. Adland and Jia (2016) and Adland et al. (2016) investigated how certain identified factors affect ship capacity utilization rate and ship speed, respectively. Sheng et al. (2018) explored the relationship between ship types and ship traffic patterns. Löptien and Axell (2014) determined the relationship between ship speed and ice conditions in Arctic shipping. Econometric tools, such as regression, are the most used methods of solving these problems.

Operational research (OR) serves as a decision support tool. In current studies on AIS data application, route planning is the only area where OR techniques are applied. Various effective algorithms, including the ant colony optimization, the genetic algorithm, and other heuristic algorithms (Tsou and Hsueh, 2010, and Kim et al., 2017), have been proposed for route planning.

8. Future research opportunities

The applications of shipping AIS big data analysis in academic research are relatively new, but we nevertheless hope that we have provided glimpses of its potential from the studies reviewed in this paper. Undoubtedly, the applications of AIS data will be further expanded in the foreseeable future. We will hereby outline some promising future applications and directions of AIS data enriched research.

To begin with, AIS is still expanding its influence. Applications of AIS have already been expanded from sea navigation to navigation in inland waterways around the globe. More and more ships sailing in rivers are equipped with the AIS, and more and more inland AIS receivers are being built (e.g., Inland Navigation Europe, 2016). This provides a brand-new angle for industrial practitioners, governments, and researchers to look at inland water transportation.

For individual vessels and fleets, AIS data analytics will provide a more comprehensive understanding of the voyage performance, such as speed optimization, weather routing, and hull conditions. These factors are key drivers of the bunker consumption of a ship, and thus have a direct impact on both emissions by vessels and company operating costs. Currently, the operational level settings of these factors, to the degree that shipowners and operators have good control of them (apart from what are determined in shipping contracts), are largely decided based on experience gained from using limited optimization support tools. Further research can be conducted in the direction of dynamic optimization, using a combination of vessel sailing patterns and external impact factors, such as weather, to achieve vessel performance optimization.

From the viewpoint of improving fleet productivity by way of increasing capacity utilization, particularly in tramp shipping, choosing the port to which to reposition an empty ship is one of the most important decisions. This is both an operational issue faced by shipping companies and an interesting research topic. With the trend of increasing vessel sizes and orderbooks, it has been suggested that the world tanker fleet currently only has half of the productivity of that in the 1970s (Stopford, 2018). Low capacity utilization resulting from a high proportion of ballast voyages and low load factor in laden voyages, may be one of the reasons. However, due to the lack of availability of cargo size data on actual vessel voyages, such line of research hardly exists. The draft information from AIS is a valuable factor in estimating cargo sizes (Jia et al., 2018). Thus, the transparency provided by AIS data allows researchers to examine shipping market dynamics from the micro level, as opposed to relying on assumptions in the literature.

The shipping industry is pursuing to achieve zero emissions. Before reaching this goal, however, all parties involved in the ecosystem need to contribute their part. However, due to the nature of the shipping business, such as high mobility and cross-border operations, the governance of environmental regulations is very fragmented. Therefore, emission monitoring becomes an extremely complex issue. The abundance of AIS data may open up the rather opaque aspect of how shipping companies release information to external authorities, including the IMO, port states, scientists and researchers. While the literature has touched upon this research domain, as reviewed in session 5.2.1, these papers mostly focus on regional port areas. This leaves the vast proportion of deep sea ship operations intact and unobserved. The assessment of emissions in port areas is understandably important, but so also should the complete shipping operations. Future research should expand on emission accounting, based on AIS data to deep sea transportation fractions.

In addition to emission assessment, as Watson et al. (2015) pointed out, traffic authorities should investigate the possibility of implementing coordinating mechanisms that synchronize AIS data with information from other sources. Such mechanisms will help AIS users to validate information broadcast by the system and to overcome the challenges brought about by errors and inaccuracies in AIS data (as we illustrated in Section 2). More importantly, doing this will contribute to a smarter, greener, and more efficient shipping industry, by enabling the many entities involved in the shipping industry to cooperate with each other at an unprecedented level. Some pioneering efforts have already been made in this direction, e.g., the European Union's MONALISA Project and STM Validation Project (interested readers may refer to Sea Traffic Management, 2018), and we believe that more studies are required in this area.

Moreover, we expect that even more multi-disciplinary studies will emerge in the coming years. For instance, in areas of marine insurance and piracy, knowledge can be developed through machine learning based on AIS information so as to identify anomaly behaviors and sailing patterns. Digitization in port management, berth allocation, and pilot employment can benefit from enhanced utilization of AIS data. Improved supply chain management with autonomous operations (vessels, vehicles and companies) can also benefit from AIS analytics.

Akin to the role of AIS in the maritime field, the development of new technologies such as E-Commerce, Internet of Things, and Smart Cities will create masses of data for researchers from relevant disciplines. With the enrichment of these micro individual data, we are then able to solve macro problems based upon a bottom-up approach, which better reflects the mechanism of complex matters. We firmly believe that applications of AIS data will inspire and stimulate even more data-related studies from various areas.

References

Aase, J. G., & Jabour, J. (2015). Can monitoring maritime activities in the European High Arctic by satellite-based Automatic Identification System enhance polar search and rescue? *The Polar Journal*, *5*(2), 386-402.

Adland, R., & Jia, H. (2016). Dynamic speed choice in bulk shipping. *Maritime Economics & Logistics*, 1-14.

Adland, R., Jia, H., & Strandenes, S. P. (2016). The determinants of vessel capacity utilization: The case of Brazilian iron ore exports. *Transportation Research Part A: Policy and Practice, 110*, 191-201.

Adland, R., Jia, H., & Strandenes, S. P. (2017). Are AIS-based trade volume estimates reliable? The case of crude oil exports. *Maritime Policy & Management, 44*(5), 657-665.

Allen, A. S., Yurk, H., Vagle, S., Pilkington, J., & Canessa, R. (2018). The underwater acoustic environment at SGaan Kinghlas-Bowie Seamount Marine Protected Area: Characterizing vessel traffic and associated noise using satellite AIS and acoustic datasets. *Marine Pollution Bulletin, 128*, 82-88.

Andersson, P., & Ivehammar, P. (2017a). Green approaches at sea – The benefits of adjusting speed instead of anchoring. *Transportation Research Part D: Transport and Environment*, *51*, 240-249.

Andersson, P., & Ivehammar, P. (2017b). Dynamic route planning in the Baltic Sea Region – A cost-benefit analysis based on AIS data. *Maritime Economics & Logistics, 19*(4), 631-649.

Arguedas, V. F., Pallotta, G., & Vespe, M. (2018). Maritime traffic networks: From historical positioning data to unsupervised maritime traffic monitoring. *IEEE Transactions on Intelligent Transportation Systems, 19*(3), 722-732.

Breithaupt, S. A., Copping, A., Tagestad, J., & Whiting, J. (2017). Maritime route delineation using AIS data from the atlantic coast of the US. *The Journal of Navigation, 70*(2), 379-394.

Buixadé Farré, A., Stephenson, S. R., Chen, L., Czub, M., Dai, Y., Demchev, D., Efimov, Y., Graczyk, P., Grythe, H., Keil, K., Kivekäs, N., Kumar, N., Liu, N., Matelenok I., Myksvoll, M., O'Leary, D., Olsen, J., Pavithran.A.P., S., Petersen, E., Raspotnik, A., Ryzhov, I., Solski, J., Suo L., Troein, C., Valeeva, V., van Rijckevorsel, J. & Wighting, J. (2014). Commercial Arctic shipping through the Northeast Passage: routes, resources, governance, technology, and infrastructure. *Polar Geography, 37*(4), 298-324.

Bureau of Transportation Statistics (2018). Port Performance Freight Statistics Annual Report to Congress 2018. Retrieved January 27, 2019 from https://www.bts.gov/sites/bts.dot.gov/files/-docs/browse-statistical-products-and-data/port-performance/224751/ppfsp-annual-report-2018. pdf.

Campana, I., Angeletti, D., Crosti, R., Luperini, C., Ruvolo, A., Alessandrini, A., & Arcangeli, A. (2017). Seasonal characterisation of maritime traffic and the relationship with cetacean presence in the Western Mediterranean Sea. *Marine Pollution Bulletin, 115*(1-2), 282-291.

Chen, L., Zhang, D., Ma, X., Wang, L., Li, S., Wu, Z., & Pan, G. (2016). Container port performance measurement and comparison leveraging ship GPS traces and maritime open data. *IEEE Transactions on Intelligent Transportation Systems*, *17*(5), 1227-1242.

Eguíluz, V. M., Fernández-Gracia, J., Irigoien, X., & Duarte, C. M. (2016). A quantitative assessment of Arctic shipping in 2010-2014. *Scientific Reports, 6*, 30682.

Eide, M. S., Endresen, Ø., Breivik, Ø., Brude, O. W., Ellingsen, I. H., Røang, K., Hauge, J., & Brett, P. O. (2007). Prevention of oil spill from shipping by modelling of dynamic risk. *Marine Pollution Bulletin, 54*(10), 1619-1633.

Farhadi, N., Parr, S. A., Mitchell, K. N., & Wolshon, B. (2016). Use of Nationwide Automatic Identification System Data to Quantify Resiliency of Marine Transportation Systems. *Transportation Research Record: Journal of the Transportation Research Board*, (2549), 9-18.

Gerritsen, H. D., Minto, C., & Lordan, C. (2013). How much of the seabed is impacted by mobile fishing gear? Absolute estimates from Vessel Monitoring System (VMS) point data. *ICES Journal of Marine Science, 70*(3), 523-531.

Goodwin, E. M. (1973). A statistical study of ship domains. *The Journal of Navigation, 26*(1), 130-130.

Hansen, M. G., Jensen, T. K., Lehn-Schiøler, T., Melchild, K., Rasmussen, F. M., & Ennemark, F. (2013). Empirical ship domain based on AIS data. *The Journal of Navigation, 66*(6), 931-940.

Harati-Mokhtari, A., Wall, A., Brooks, P., & Wang, J. (2007). Automatic Identification System (AIS): Data reliability and human error implications. *The Journal of Navigation, 60*(3), 373-389.

IALA (2004). IALA Guideline No. 1028 On the Automatic Identification System (AIS) Volume 1, Part I Operational Issues Edition 1.3. Retrieved April 2, 2018 from https://www.e-navigation.nl/sites/default/files/universal-automatic-identification-ais-volume-1-part-1-operational-issues-1028.pdf.

Inland Navigation Europe (2016). Inland Navigation Europe 2016 Annual Report. Retrieved January 27, 2019 from http://www.inland navigation.eu/media/76057/ INE_AR _2016 _final-_spreads_links_small.pdf.

Jia, H., Adland, R., Prakash, V., & Smith, T. (2017a). Energy efficiency with the application of Virtual Arrival policy. *Transportation Research Part D: Transport and Environment, 54*, 50-60.

Jia, H., Lampe, O. D., Solteszova, V., & Strandenes, S. P. (2017b). An automatic algorithm for generating seaborne transport pattern maps based on AIS. *Maritime Economics & Logistics, 19*(4), 619-630.

Jia, H., Lampe, O. D., Solteszova, V., & Strandenes, S. P. (2017c). Norwegian port connectivity and its policy implications. *Maritime Policy & Management*, *44*(8), 956-966.

Jia, H., Prakash, V., & Smith, T. (2018). Estimating vessel payloads in bulk shipping using AIS data. *International Journal of Shipping and Transport Logistics*. In press.

Kaluza, P., Kölzsch, A., Gastner, M. T., & Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface, 7*(48), 1093-1103.

Kim, D., Hirayama, K., & Okimoto, T. (2017). Distributed Stochastic Search Algorithm for multi-ship encounter situations. *The Journal of Navigation, 70*(4), 699-718.

Kivekäs, N., Massling, A., Grythe, H., Lange, R., Rusnak, V., Carreno, S., Skov, H., Swietlicki, E., Nguyen, Q. T., Glasius, M., & Kristensson, A. (2014). Contribution of ship traffic to aerosol particle concentrations downwind of a major shipping lane. *Atmospheric Chemistry and Physics, 14*(16), 8255-8267.

Kowalska, K., & Peel, L. (2012). Maritime anomaly detection using Gaussian Process active learning. In *Information Fusion (FUSION), 2012 15th International Conference on* (pp. 1164-1171). IEEE.

Lee, J. G., Han, J., & Whang, K. Y. (2007). Trajectory clustering: A partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data* (pp. 593-604). ACM.

Li, L., Lu, W., Niu, J., Liu, J., & Liu, D. (2017). AIS data-based decision model for navigation risk in sea areas. *The Journal of Navigation*, 1-15.

Longépé, N., Mouche, A. A., Goacolou, M., Granier, N., Carrere, L., Lebras, J. Y., Lozach, P., & Besnard, S. (2015). Polluter identification with spaceborne radar imagery, AIS and forward drift modeling. *Marine Pollution Bulletin, 101*(2), 826-833.

Löptien, U., & Axell, L. (2014). Ice and AIS: Ship speed data and sea ice forecasts in the Baltic Sea. *The Cryosphere, 8*(6), 2409-2418.

Mazzarella, F., Alessandrini, A., Greidanus, H., Alvarez, M., Argentieri, P., Nappo, D., & Ziemba, L. (2013). Data fusion for wide-area maritime surveillance. In *Proceedings of the COST MOVE Workshop on Moving Objects at Sea, Brest, France* (pp. 27-28).

Mitchell, K. N., & Scully, B. (2014). Waterway performance monitoring with automatic identification system data. *Transportation Research Record: Journal of the Transportation Research Board,* (2426), 20-26.

Perera, L. P., Oliveira, P., & Soares, C. G. (2012). Maritime traffic monitoring based on vessel detection, tracking, state estimation, and trajectory prediction. *IEEE Transactions on Intelligent Transportation Systems, 13(3)*, 1188-1200.

Pallotta, G., Horn, S., Braca, P., & Bryan, K. (2014). Context-enhanced vessel prediction based on Ornstein-Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results. In *Information Fusion (Fusion), 2015 18th International Conference on* (pp. 1152-1159). IEEE.

Pan, J., Jiang, Q., Hu J., & Shao Z. (2012). An AIS data visualization model for assessing maritime traffic situation and its applications. *Procedia Engineering, 29*, 365-369.

Prochazka, V., & Adland, R. (2017). Insight into oil-tanker market behaviour gained from AIS data. Paper presented at IAME 2017 Conference.

Reeves, R. R., Read, A. J., Lowry, L., Katona, S. K., & Boness, D. J. (2007). Report of the North Atlantic right whale program review. Marine Mammal Commission, Bethesda, Maryland.

Reeves, R., Rosa, C., George, J. C., Sheffield, G., & Moore, M. (2012). Implications of Arctic industrial growth and strategies to mitigate future vessel and fishing gear impacts on bowhead whales. *Marine Policy, 36*(2), 454-462.

Scully, B. M., & Mitchell, K. N. (2017). Underkeel clearance reliability model for dredged navigation channels. *Transportation Research Record: Journal of the Transportation Research Board*, (2611), 41-49.

Sea Traffic Management (2019), Sea Traffic Management Projects, Retrieved January 27, 2019 from https://www.stmvalidation.eu/projects/.

Sheng, K., Liu, Z., Zhou, D., He, A., & Feng, C. (2018). Research on ship classification based on trajectory features. *The Journal of Navigation*, *71*(1), 100-116.

Sidibé, A., & Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. *The Journal of Navigation*, *70*(4), 847-858.

Stopford, M. (2018). Leading change into the future for shipping. Leadership 2018 Conference, Bergen, Norway.

Tsou, M. C., & Hsueh, C. K. (2010). The study of ship collision avoidance route planning by ant colony algorithm. *Journal of Marine Science and Technology, 18*(5), 746-756.

Tu, E., Zhang, G., Rachmawati, L., Rajabally, E., & Huang, G. B. (2017). Exploiting AIS data for intelligent maritime navigation: A comprehensive survey from data to methodology. *IEEE Transactions on Intelligent Transportation Systems*. Doi: 10.1109/TITS.2017.2724551.

Vespe, M., Gibin, M., Alessandrini, A., Natale, F., Mazzarella, F., & Osio, G. C. (2016). Mapping EU fishing activities using ship tracking data. *Journal of Maps*, *12* (sup1), 520-525.

Wang, J., Zhu, C., Zhou, Y., & Zhang, W. (2017). Vessel spatio-temporal knowledge discovery with AIS trajectories using co-clustering. *The Journal of Navigation*, *70*(6), 1383-1400.

Watson R.T., Holm H., & Lind M. (2015). Green Steaming: A Methodology for Estimating Carbon Emissions Avoided. Thirty Sixth International Conference on Information Systems, Fort Worth, The U.S.A.

Winther, M., Christensen, J. H., Plejdrup, M. S., Ravn, E. S., Eriksson, Ó. F., & Kristensen, H. O. (2014). Emission inventories for ships in the arctic based on satellite sampled AIS data. *Atmospheric Environment, 91*, 1-14.

Zhang, L., Meng, Q., & Fwa, T. F. (2017). Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E: Logistics and Transportation Review*. Doi: 10.1016/j.tre.2017.07.011.

Zhao, Z., Ji, K., Xing, X., Zou, H., & Zhou, S. (2014). Ship surveillance by integration of space-borne SAR and AIS – review of current research. *The Journal of Navigation*, *67*(1), 177-189.

Zhen, R., Jin, Y., Hu, Q., Shao, Z., & Nikitakos, N. (2017). Maritime anomaly detection within coastal waters based on vessel trajectory clustering and Naïve Bayes Classifier. *The Journal of Navigation, 70*(3), 648-670.