MARITIME INFORMATICS TECHNOLOGY

Michalis Michaelides[1], Department of Electrical Eng., Computer Eng. and Informatics, Cyprus University of Technology, 30 Arch. Kyprianos Str., 3036 Limassol, Cyprus, michalis.michaelides@cut.ac.cy

Herodotos Herodotou, Department of Electrical Eng., Computer Eng. and Informatics, Cyprus University of Technology, 30 Arch. Kyprianos Str., 3036 Limassol, Cyprus, herodotos.herodotou@cut.ac.cy

Patrick Schmitt, BearingPoint GmbH, Erika-Mann-Straße 9, 80636 München, Germany, patrick.schmitt8194@web.de

Mathias Karlson, Department Mobility and systems, Digital Systems, RISE Research Institutes of Sweden, Lindholmspiren 3A, SE-417 56 Gothenburg, Sweden, mathias.karlsson@ri.se

Ioannis Kyriakides, University of Nicosia Research Foundation and Cyprus Marine and Maritime Institute, Cyprus, kyriakides.i@unic.ac.cy

Daniel Hayes, Cyprus Subsea Consulting and Services C.S.C.S. Ltd., Cyprus, hayesdan@cyprus-subsea.com

Xiuju Fu, Institute of High Performance Computing, A*Star, Singapore, fuxj@ihpc.a-star.edu.sg

ABSTRACT

Smart sensor technologies and digitisation of information are transforming today’s world, including the maritime sector. Ships are currently equipped with hundreds of sensors for monitoring various parameters of interest related to the physical environment in which a vessel is operating (i.e., ocean data), the characteristics and state of the vessel, and the physiological and mental condition of the crew. Ports are also equipped with advanced monitoring systems and tracking technologies, using various sensors, such as inertial sensors, ultrasonic sensors, eddy current sensors, radar, LiDAR, imaging sensors, and RFID readers and tags, which allow ports to provide essential services in a faster and more efficient manner. The collected data from these various sensors include both spatial and temporal information. They can be linked both with a geographical location and the time of occurrence of a specific event. To extract useful information from the data, we need to have appropriate techniques for data acquisition, management, analysis, and visualisation. Such intelligent algorithms will empower human users to ‘make sense’ of the spatiotemporal data and provide enhanced decision support. This paper provides more details on each of these important dimensions of dealing with spatiotemporal data.

Keywords: *Maritime Informatics, Digitisation, Information Technology in Shipping and Ports, Spatiotemporal Data Analysis*

[1] Author for correspondence and presenting the paper

1. INTRODUCTION

Smart sensor technologies and the digitalisation of data are transforming today's world, including the maritime sector (Aslam et al., 2020). Ships are currently becoming equipped with hundreds of sensors for monitoring various parameters of interest related to the physical environment in which a vessel is operating (weather, temperature, sea conditions, etc.) and the vessel's characteristics and state. On the other hand, wearable devices that can indicate the crew's physiological and mental condition are possibilities, too.

Ports are also becoming equipped with advanced monitoring systems and tracking technologies which allow port authorities to provide essential services in a faster and more efficient manner using various sensors, such as inertial sensors, ultrasonic sensors, eddy current sensors, radar, LiDAR, imaging sensors, and RFID readers and tags. The "Ocean of Things" project at the United States (US) Defence Advanced Research Projects Agency (DARPA) aims to likewise wire up the high seas with swarms of floating, connected sensors (Anderson, 2020). The collected data from these various sensors include both spatial and temporal information, as they can be linked both with a geographical location and the time of occurrence of a specific event. To extract useful information from the data, we need to have in place appropriate techniques for *data acquisition*, *management*, *analysis*, and *visualisation*. Such intelligent algorithms will empower human users to "*make sense*" of the *spatiotemporal* data and provide enhanced decision support.

This paper synthesises the material from the *third* section of a new book, Lind, M., Michaelides, M. P., Ward, R., and Watson, R. T. (Eds.). (2020). *Maritime Informatics*, Springer. Maritime Informatics concerns the application of information systems to increasing the efficiency, safety, and ecological sustainability of the world's shipping industry and the third section of the book covers the important dimensions of dealing with spatiotemporal data in chapters 19-23.

The chapters outline what is considered state of the art in each of these areas concerning the underlying technology and how it can be tailored to the maritime world's needs.

In particular, chapter 19: *Big Maritime Data Management* (Herodotou et al., 2020) provides an extensive overview of the maritime data value chain and discusses state-of-the-art technological solutions for managing and processing maritime data in efficient and effective ways. Next, chapter 20: *Spatiotemporal Data Analytics for the Maritime Industry* (Schmitt et al., 2020*)* presents the state-of-the-art in spatiotemporal analytics and provides an overview of its practical applications in the maritime industry. Chapter 21: *Data Visualisation Tools for Enhanced Situational Awareness in Maritime Operations* (Karlsson et al., 2020) elaborates on creating meaningful data visualisations providing decision-support to enable distributed coordination in self-organised ecosystems involving multiple actors.

Chapter 22: *Intelligent Maritime Information Acquisition and Representation for Decision Support* (Kyriakides et al., 2020) reviews technological advancements in intelligent information acquisition and representation to aid human users and organisations make sense of the complex maritime environment. Finally, chapter 23: *AIS Data Analytics for Intelligent Maritime Surveillance Systems* (Fu et al., 2020) reviews maritime traffic surveillance systems for spatiotemporal data collection. Then a computational framework is presented to efficiently compress, transfer, and acquire the necessary information for the further analysis of large-scale automatic identification system (AIS) data. The rest of the paper is organised as follows. A more extensive description outlining the most important results for each of the chapters is provided in Section 2. The paper concludes with Section 3.

2. APPLICATIONS OF MARITIME INFORMATICS TECHNOLOGY

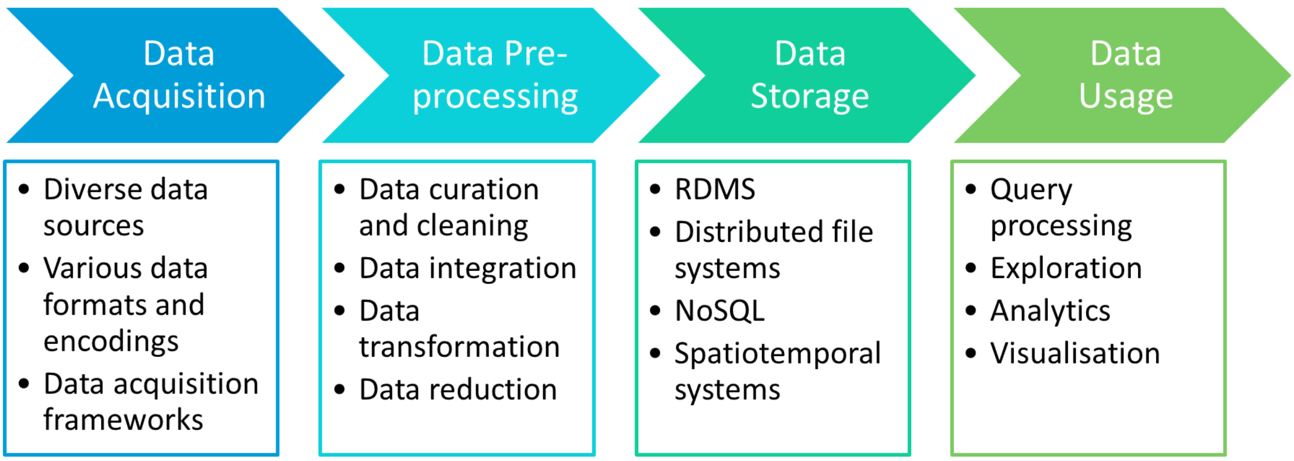
2.1. Big Maritime Data Management

Massive amounts of heterogeneous data are now continuously collected by modern maritime transport and logistics companies, ship owners, ship agents, and port authorities due to the digitalisation of the field and the increasing use of smart sensing devices (Lytra et al., 2017). For example, MarineTraffic, an AIS ship tracking platform, collects over 520 million AIS messages a day, containing the position, course, and speed of vessels travelling the oceans worldwide (Perobelli, 2016). Besides, port stakeholders and authorities are amassing various port call data associated with the arrival, berthing, loading/unloading, and departure of vessels to/from the ports (Michaelides et al., 2019). Simultaneously, at least 5GB of oceanographic, environmental, and meteorological data are recorded daily by several sensors deployed at sea (Lytra et al., 2017).

As more data is collected and stored, maritime stakeholders are investing in new technological solutions for managing and processing maritime data in efficient and effective ways to extract deep insights from the data, which will automate various decision-making processes. Optimising port operational efficiency (Yang et al. 2018), tracking cargo in real-time (Yeoh et al., 2011), improving fuel consumption (Besikçi et al., 2016), and preventing accidents (Zhao et al., 2014) are only a few examples of enabling key application scenarios that can have a substantial impact in the maritime industry.

The *maritime data value chain* defines the series of activities needed to manage maritime data properly during the entire data life cycle and extract value from data (Ferreira et al., 2017). The European Commission considers the data value chain to be 'the centre of the future knowledge economy, bringing the digital developments opportunities to the more traditional sectors (e.g., transport, financial services, health, manufacturing, retail)' (DGConnect, 2013, p. 4). The four key activities identified are: (1) *data acquisition* for gathering data across different and geo-distributed data sources; (2) *data pre-processing* for cleaning, integrating, transforming, and linking data; (3) *data storage* for storing data in a persistent and scalable way; and (4) *data usage* for processing the data and extracting value. Figure 1 illustrates the main activities associated with the maritime data value chain.

Figure 1: The maritime data value chain



Source: Herodotou et al. (2020)

Data acquisition in the maritime domain entails the process of collecting data about port visits, vessel routes, sensory equipment information, and ocean/weather conditions (among others) from a wide variety of sources such as port community systems, automatic identification systems (AIS) and weather stations, and Internet of Things (IoT) devices (Lytra et al., 2017). The collected data come in many formats and encodings (e.g., GeoJSON, GML, KML, RDF), can be structured, semi-structured, or unstructured in nature, and often contain both a geographical position and a time component (i.e., are spatiotemporal data). Some typical examples of maritime spatiotemporal data are a moving vessel whose location continuously changes over time, or oceanic measurements collected by sensors on-board a drifting buoy. Specialized data acquisition frameworks are employed for gathering the data from each source and delivering it to the downstream data storage systems (Curry, 2016). Apache Kafka and Apache Flume are two open-source distributed systems that are commonly used for collecting and sharing data due to their performance and fault-tolerant characteristics.

Data pre-processing involves a set of methods for cleaning, integrating, transforming, and linking the data to ensure that the stored information satisfies the desired data quality requirements for its effective use (Yablonsky, 2018). Several data faults are possible, depending on the source and communication link characteristics, such as data noise, outliers, missing values, data corruptions, and bias. The process of data curation and cleaning provides technological and methodological data management support for removing faulty data, replacing missing data, smoothing out noisy data, and correcting inconsistent data (Han et al., 2011). After cleaning the data from each source individually, data integration techniques (e.g., data consolidation, federation, and propagation) are used for combining data based on a common timeline and frame of reference into a single, organised, and coherent view (Wang et al., 2016). Next, data transformation techniques convert data formats and structures into other forms that are more suitable for data storage or analysis. The multitude of data formats in the maritime domain has led to utilising a wide spectrum of data transformation techniques such as numeric normalisation, aggregation, generalisation, and attribute construction (Han et al., 2011). Finally, data integration techniques can improve data interoperability via linking existing repositories of relevant data with raw maritime data coming from various sources (Lytra et al., 2017).

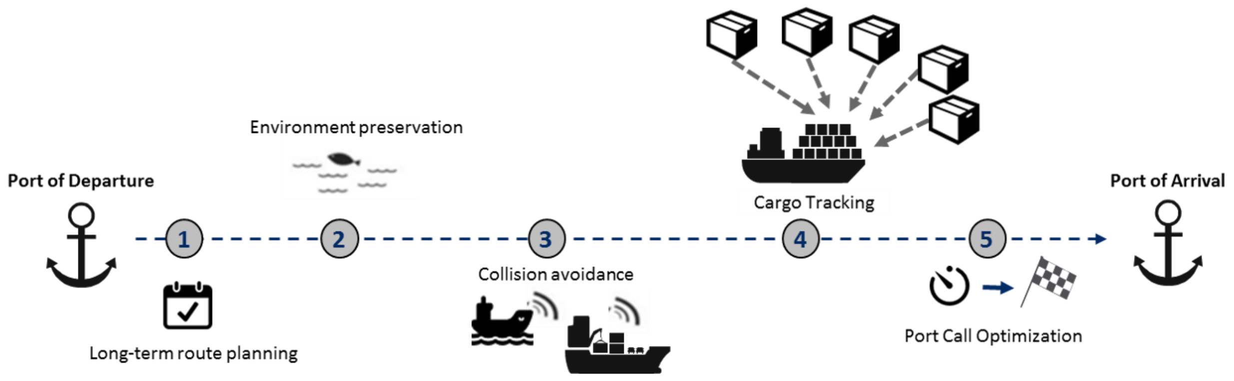
Data storage refers to methods and systems for storing data in a way that ensures data persistence, availability, and scalability (Ferreira et al., 2017). These storage technologies often employ data partitioning, distributed replication, compression, and indexing to provide higher-level applications with fast and easy access to the underlying data. While traditional data warehouses and relational database management systems are widely used to store maritime data, the semi/unstructured and spatiotemporal nature of data has shifted attention towards newer storage solutions. In particular, NoSQL data stores such as MongoDB, HBase, and Cassandra, are gaining popularity as they support more flexible data models and store data in more scalable ways (Wang et al., 2016). At the same time, specialised spatial or spatiotemporal systems such as PostGIS and GeoMesa have emerged for dealing with entities/objects containing space and time information (Xiong et al., 2017). Unlike other data storage solutions, spatiotemporal ones can manage the dynamic properties of objects (e.g., a vessel moving in space and time) efficiently and effectively.

Data usage involves query processing, analytics, and visualisation systems and tools for analysing the collected maritime data and extracting value for various data-driven business activities. The appropriate use of these tools can significantly improve a maritime enterprise's competitiveness by reducing operational costs, optimising existing business processes, and providing better services to end-users (Cavanillas et al., 2016). Several query processing systems have been developed over the years, enabling browsing, searching, reporting, finding correlations, identifying patterns, and predicting relations across maritime data (Curry, 2016), while offering different query interfaces and functionalities for processing data. For instance, Hive and SparkSQL provide SQL-like functionality to analyse and generate reports for structured data such as port management, transport, and logistics data. On the other hand, Presto and Impala specialise in interactive data processing and provide query processing with low latency. Finally, GeoMesa and GeoSpark provide spatiotemporal querying and analytics with specialised features for geometric and geographic processing, topogeometry functions and topologies, as well as raster processing and analysis.

2.2. Spatiotemporal Data Analytics for the Maritime Industry

Having the right information at the right time is a key ingredient to creating value for any business, and the maritime industry is no different. Data analytics makes a difference for various maritime industry stakeholders from optimising vessels routes and preparing ports for efficient operations to reduce pollution and save the environment. Some key application scenarios are (1) long-term route planning, (2) environment preservation, (3) collision avoidance, (4) cargo tracking and (5) port-call optimisation. All those scenarios emerge along a voyage (see Figure 2) and are providing use cases for spatiotemporal analytics.

Figure 2: Application areas for spatiotemporal analytics along a voyage



Source: Schmitt et al. (2020).

2.2.1. General Characteristics of the Outlined Application Areas

The five outlined cases have a common understanding of information across all its involved actors to derive valuable information out of a large set of spatiotemporal data. As operations along a voyage process usually involve various actors, messaging standards are needed to ensure the correct interpretation of exchanged information and achieve efficient collaborative decision making. Understanding the maritime ecosystem is crucial for the definition of efficient messaging standards and collaborative infrastructures. Given the high interconnectedness of actors, the elaborated examples are underlined by the application of graph data analysis. It has been proven to be an efficient way to deal with spatiotemporal data analytics (George and Shekhar, 2008; Gunturi and Shekhar, 2017). Finally, each example is linked to the maritime transportation industry's three primary goals as defined by the European Commission—safety, sustainability, and efficiency (European Commission, 2016).

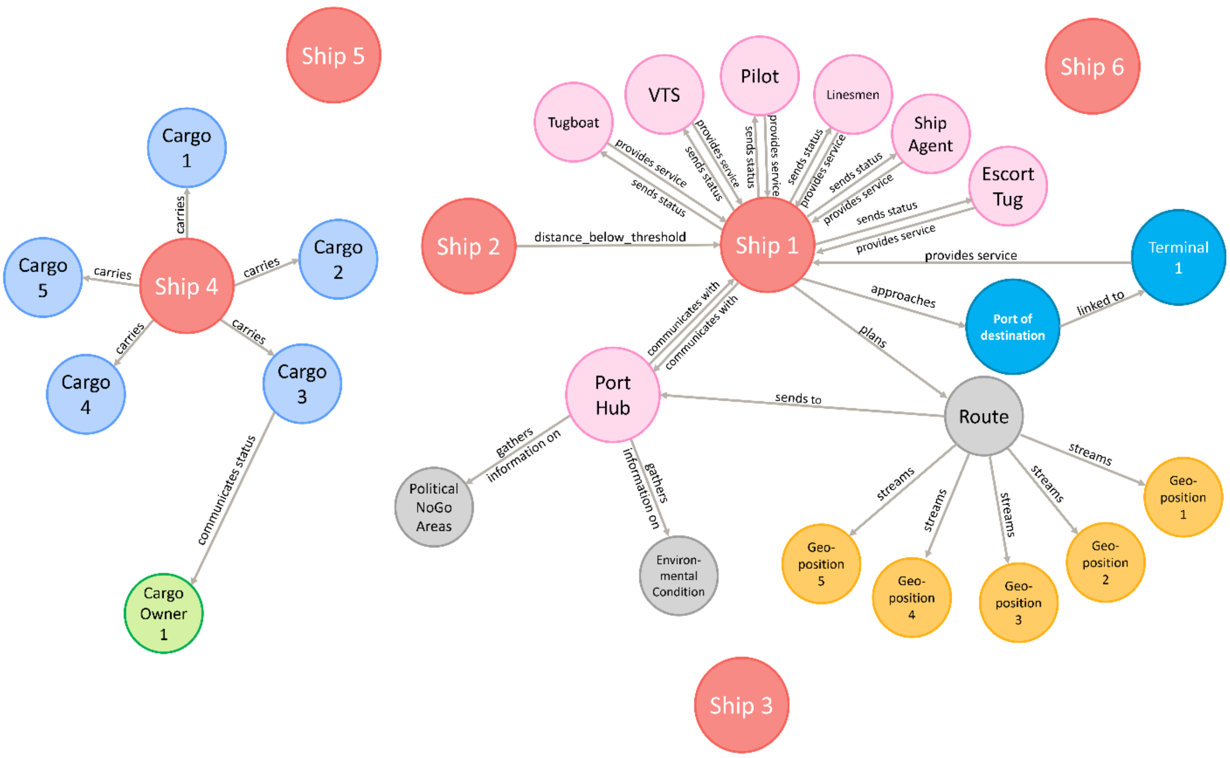
2.2.2. Individual Assessment of each Application Area

While the goal is to create a holistic framework sketching exemplarily the important key actors and their relationships with each other along a voyage, a particular focus needs to be given to each application area in future research. Due to different circumstances, actors and overall targets, there is no one-size-fits-all approach across those use cases to achieve the defined goals. The main stakeholders in each of the application areas shown in Figure 2 can usually be assigned to the groups of port authorities, ship operators or cargo owners – while the specific actors, their extent of involvement and their actions differ largely dependent on the application area. For instance, while the long-term route planning (application area 1) and the short-term route adjustment process (application area 3) is not of primary interest to the cargo owner, the cargo owner is more interested in monitoring the condition of the transported goods throughout the shipment processes (application area 4).

2.2.3. The Big Picture: Using a Holistic Framework to Reflect the Entire Ecosystem

Since all involved actors around a voyage process usually work on achieving the common goal of transporting goods safely, timely, and in an environmentally sound manner, achieving goals in one application area in this complex ecosystem can lead to positive spillover effects. For instance, a more efficient port call process (application area 5) does not only have direct positive impacts on port efficiency (lower idle time) and sustainability (less fuel usage) but also benefits the cargo owner because it enables better planning capability of subsequent operations. This interconnectedness of events and actions is sketched in Figure 3.

Figure 3: Holistic framework of application areas for spatiotemporal data analytics along a vessel's voyage



Source: Schmitt et al. (2020).

Although such a framework is, to some extent, always a simplification of reality, by looking at the big picture of such a complex ecosystem, the relationships across various actors can be understood better. Acquiring a profound understanding of the various actors involved along a voyage and their relationships with each other is the basis for efficient collaborative decision making in the maritime industry.

2.3. Data Visualisation Tools for Enhanced Situational Awareness in Maritime Operations

Much of the data benefits from being visualised for specific needs, such as coordinating port call operations where the multitude of actors, constituting a self-organised ecosystem, sharing data amongst themselves to create situational awareness based on secure digital data streams. This means presenting shared data in real-time in a standardised format to generate value for the maritime industry. As port call operations are complex and engage multiple actors, different visualisation tools are useful to address the many concerns raised by the different actors.

Visualisation is not a new phenomenon. When our ancestors created cave paintings, they were attempting to communicate via images. In modern times, visualisations can be animations, images, videos or diagrams such as information about performance management, decision support or a business intelligence system. Data visualisation is the graphical representation of data to help the viewer meaningfully interpret, analyse, and understand large complex sets of data to support informed decisions. When it comes to visualisation concerning support for maritime decision making and planning, we are at an early stage. Hence, Maritime Informatics researchers must develop and deploy new forms of visual data representation as they gain access to more data about voyage and port operations and learn how various visualisation forms improve decision-making.

2.3.1. Spatiotemporal Data Visualisation

Spatial data refers to the location (or space), such as a specific berth in a port; or information providing directions to different locations. Temporal data refers to time, such as when a ship berthed. By combining spatial data with temporal data, the so-called spatiotemporal data refers to data containing both the time and the location of an object or its status. This makes it possible to reference according to a certain location and time. For example, in the information statement: “Mare Liberum was first published in 1609 in the Netherlands” (Wikipedia 2020), “1609” is temporal data and “Netherlands” is spatial data. Spatiotemporal data can also be used to give directions, for example, “the ship arrived at berth 519 in the Port of Gothenburg at 11:15 am”, or “Flight SK144 departures from gate 19A at 07:00 am”, as shown in Figure 4.

Figure 4: Spatiotemporal data showing flight information

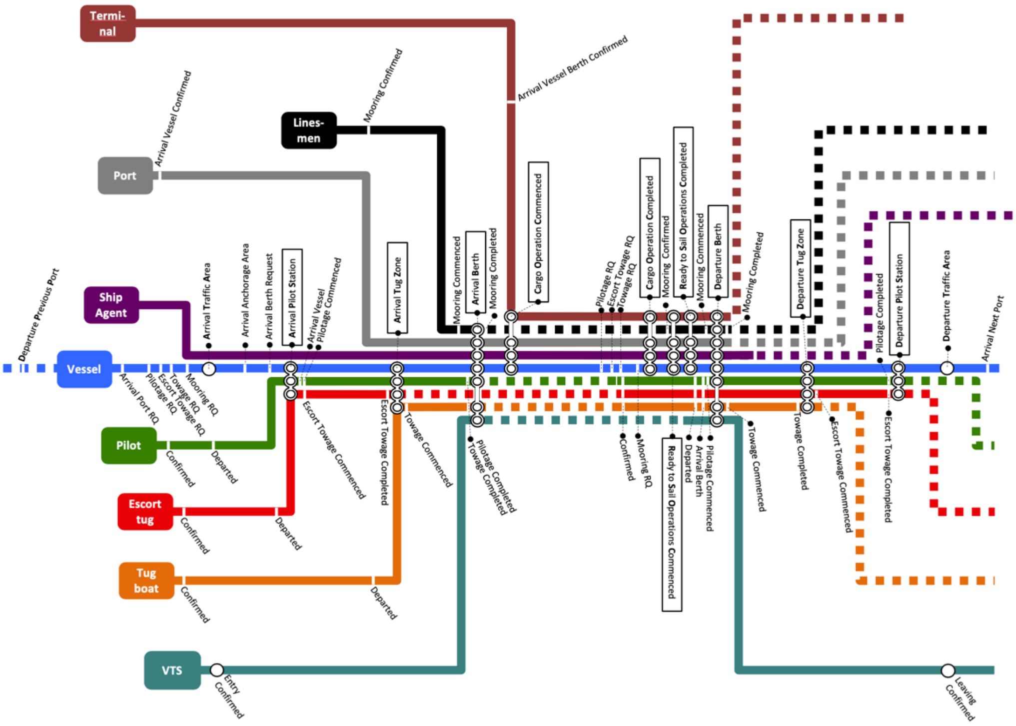
Source: Karlsson et al., (2020).

Since 2005, most ships engaged on international voyages, together with many others engaged in the coastal passage, must be equipped with AIS (International Maritime Organisation, 1974). The AIS laid the ground for the digitalisation of shipping (see Watson et al., 2020), as it can provide spatiotemporal data related to a ship’s movements in different time intervals, such as the identity of the ship, its position, speed, and heading. The presentation of AIS data on an Electronic Chart Display and Information System (ECDIS) display enables officers on a ship to visualise the presence of other ships in their vicinity in real-time, thereby increasing navigation safety. Tracking AIS information could be useful for actors within a port who want to see the present position of a ship that they are anticipating to calculate when they can expect her to arrive, especially when they have doubts about the accuracy of data previously received.

2.3.2. Visualisation of a Port Call Using a Metro Map

In order to optimise a port call, the different required actions need to be synchronised, and the involved actors need to be able to trust that the prediction and achievement of each episodic tight coupling of events are shared (Lind et al., 2015, Lind et al., 2016). In order to visualise the dependencies of these episodes a *"Metro Map"* (see Figure 5) visualisation has been developed to show the complexity and the interdependencies involved in a port visit. The metro map shows, for example, that in order for a towage operation to start, a ship and a tug need to be organised and synchronised to be at the same location at the same time.

Figure 5: Visualisation of a port call through a Metro Map



Source: Karlsson et al. (2020)

The Metro Map provides an easy to understand presentation of the important coordination points in a port call process. A *station* (a white round circle) on the Metro Map is where there is an episodic tight coupling. To reach these coordination points, information needs to be exchanged before reaching a station; this is illustrated as lines on connecting small grey ringed white circles, such as the coordination between a tugboat operator and a ship's master. As port call operations are complex and engage multiple actors, different visualisation tools are useful to address the many concerns raised by the different actors, which requires that visualisations are optimised to suit each actor involved in a port call, e.g. a pilot's view or a terminal's view. These visualisations, however, need to use the same data feeds to ensure the different perspectives are consistent.

2.4. Intelligent Maritime Information Acquisition and Representation for Decision Support

Activities including coastal tourism, fishing, shipping, and the development and operation of off-shore infrastructure increase intensity, conflicting interests, and impact on the natural environment. Information gaps and complex dependencies between marine environmental parameters and human maritime activities hinder the decision making processes and increase the risk of environmental pollution and accidents. The need to observe and characterise the marine and maritime activity coupled with the constraints imposed by limited information acquisition, processing, and communication resources calls for developing novel agile methods of efficient and effective data collection and interpretation. Intelligent information collection and representation will provide human users with a common operational picture for decision support that adapts to specific scenarios via efficient coordination of human and cyber-physical resources allowing decision-makers to make sense of human and environmental processes. Effective information collection and representation is provided by the smart management and agile response capabilities of cognitive systems. Cognitive systems (Haykin, 2006) are equipped with perception, reasoning, learning, action, and intelligent resource allocation capabilities. These intelligent systems actively seek missing information to enhance situational awareness under resource constraints via reconfiguration of human and cyber-physical resources. Technological advancements towards the development of cognitive systems are briefly reviewed next.

Effective information acquisition requires the use of agile data acquisition systems. Autonomous systems are increasingly used to collect marine and maritime data, often in synergy with larger research vessels that collect water samples and carry Remotely Operated Vehicles. Commercially available autonomous maritime platforms that can host sensor nodes include Autonomous Stationary Systems (ASSs), Autonomous Underwater Vehicles (AUVs) such as those in Figure 6 including gliders (D. Hayes, 2016), Autonomous Surface Vehicles (ASVs), Deep Profiling Floats (DPFs), and Autonomous Aerial Vehicles (AAVs). Autonomous Vehicles (AxVs) are capable of long endurance and offer diverse capabilities in terms of motion and can operate as swarms to improve coverage in ocean areas. ASS and AxVs, classified as maritime Internet of Things (IoT), can host edge processing and communication modules at the point of sensing. However, data sets collected by diverse remote sensing systems under the limited connectivity ocean environment are often stored in different formats and include missing or erroneous data.

Figure 6: URRready4OS project experiment: Underwater Robotics Ready for Oil Spills

Source: Kyriakides et al., (2020). Credit: Universidad Politécnica de Cartagena (UPCT)

Processing of data for standardisation and quality improvement needs to take place before information extraction for decision-making. This type of data processing ranges from automated real-time checks to delayed-mode human-driven controls. Data points are marked with quality flags, and estimates of accuracy and precision are calculated. Additionally, estimates of data quality and accuracy are used to evaluate weights for the data points and accommodate inconsistencies between data, estimates, and predictions. Data needs to include both observations and metadata about the observations. Metadata include information such as quality control procedures and flags, the type of sensor and platform that has acquired the data, and the time and place of data acquisition. The lack of metadata drastically reduces data reusability and longevity. A standard format for data and metadata is essential to be commonly used to allow sharing and interoperability. Organisations can still secure exclusive scientific analysis and publication by the initial owners during an embargo period. After this period, data can be shared with regional or global data assembly centres to foster collaboration and allow the reproducibility of scientific results. Data management that meets the principles of findability, accessibility, interoperability, and reusability (FAIR) is now encouraged by many funding agencies, governments, and international bodies.

Intelligent interpretation and representation of information is a necessary step that follows agile sensing, data standardisation, and quality assessment. Even if data are easily accessible and reusable, the data represent very different aspects of the observed natural or human processes. The data need to be fused into a common operational picture that is understandable to human users. Besides, the diverse capabilities of heterogeneous human and cyber-physical assets need to be optimised to increase information collection. Methods such as Cognitive Fusion (Kyriakides, 2019) can fuse the diverse data and reconfigure heterogeneous cognitive sensing systems that synergise observing the marine and maritime environment using an efficient allocation of resources.

In conclusion, the commercialisation of technologies, including maritime IoT with edge-processing and communications capabilities, allows the development of heterogeneous cognitive systems for intelligent data acquisition. The intelligent data acquisition and interpretation of information promote efficient data collection and the integration of human and cyber-physical capabilities. This integration of capabilities combines the benefits of fast response and accuracy by cyber-physical systems for making tactical decisions while allowing human users to *make sense* of the complex maritime environment and empowering organisations to make strategic decisions.

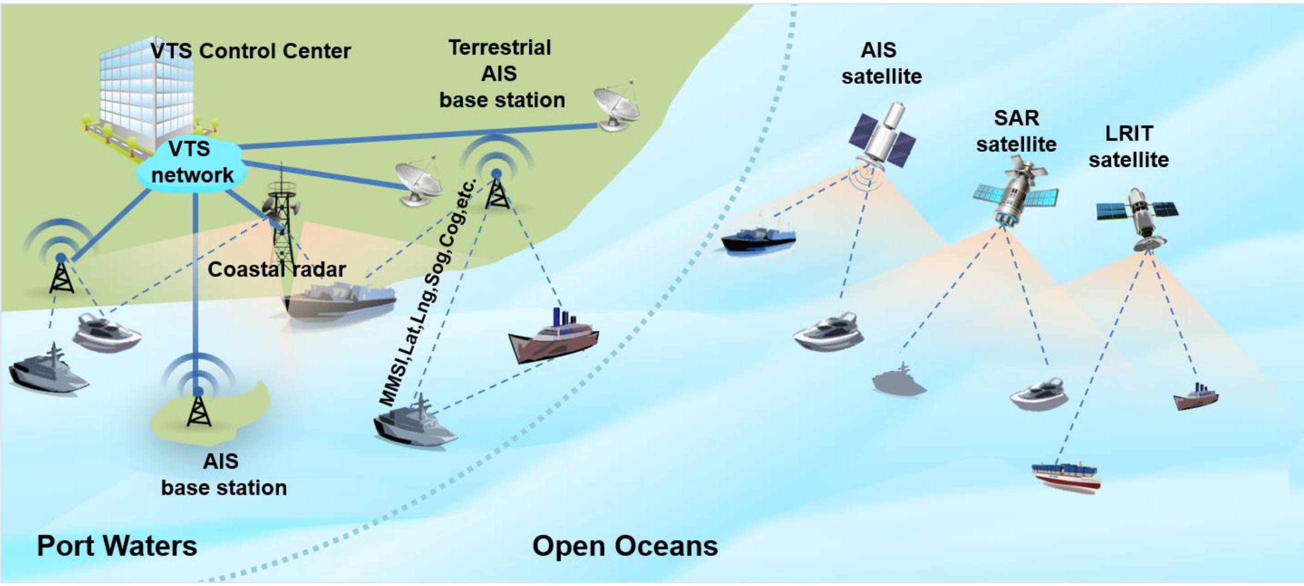
2.5. AIS Data Analytics for Intelligent Maritime Surveillance Systems

AIS, long-range identification and tracking (LRIT) system, coastal radar and satellite-borne synthetic aperture radar (SAR) are main maritime surveillance systems, which are widely used in maritime surveillance networks (see Figure 7) and provide key information for maritime traffic and operations. Terrestrial AIS and coastal radars support vessel traffic surveillance in port waters whilst LRIT, S-AIS and SAR are primary instruments used for open oceans. The corresponding traffic or sensor data containing spatiotemporal information is collected and plays key roles in maritime systems for ship tracking, monitoring, port operation visibility, maritime traffic surveillance, and safety management.

With the increasing size, the received spatiotemporal data from surveillance networks contain multiple sets of information for all the vessel traffic in the target zone. Thus data formation is often required to structure the traffic data to be vessel-based (or other traffic feature-based) as a preparatory step before further processing and analytics for intelligent actions. AIS is a typical maritime data that requires much more pre-analysis before further applications. Some common processing includes data cleaning, noise and outlier removal, among other issues. In many applications, interpolation is also needed to make data structured in a reasonable and unified time frame to facilitate the visualisation and processing. While the pre-processing algorithm ensures better data quality, it also introduces data redundancies which can further increase the volume of the already sizable AIS data.

Xu et al. (2019) presented a framework that enables the speedy access and transfer of AIS data across networks and systems while generating aggregated data for easier access, further analysis by practical applications. The lossless compression algorithm (Xu et al. 2019) is proposed, converting the AIS data into compressed archives in binary form. The output is a collection of compressed archives that are independent of each other and properly indexed. Hence, each archive can be individually decompressed on demand in order to reduce the use of disk memory, the decompressed contents are cached to favour fast subsequent access, and parallel operation can also be applied to speed up the processing. This algorithm provides significant performance gain and benefits the typical workflows, especially for AIS data which contains lots of strings and duplicate information. In the second part of the framework, each binary-file was decomposed and aggregated to grid-based data with key information retained and/or derived, which further compressed the data. Through this method, the key information of vessel movement in the data was maintained that can be easily accessed by common users, without compromising the data analytics and intelligence generated out of the data. The grid and key information were customised based on the application use case.

Figure 7: Overview of maritime surveillance networks, information systems



Source: Fu et al., (2020)

To prove the proposed method and framework, we demonstrated in a use case about how our algorithm can be used efficiently to generate easily accessible AIS data for vessels patterns at the Singapore port. The proposed framework was evaluated considering the saving of storage space and processing duration and the retaining of key information. Besides achieving more than 400 times of reducing data size and reducing processing data, we also obtained high accuracy about retaining key operation information. This was shown by examining the performance of detected vessel port operation events through the processed data. We compared the detected vessel voyage event details (who, when, where and duration) against available actual port operation data for a few vessels within Singapore and observed that 95% of the berths and 99.5% of berth stay duration were correctly identified for 88 journeys in 2017.

The emergence of big spatiotemporal data in maritime and many other sectors helps generate new insights and understanding of traditional operations and system management. It also sheds light for applying digital technology to provide intelligence for transforming the traditional operations and enhancing safety and operation efficiency from diversified angles. The proposed data pre-processing approach can be applied to efficiently store, retrieve, and analyse big AIS data with reduced computational cost. It can then be fused with other sources of digitised data to facilitate big data intelligence development towards innovation processes and applications, not only in the maritime but also in other domains.

3. CONCLUSIONS

Smart sensor technologies and digitisation of information are transforming today's world, including the maritime sector. Ships, ports, and the ocean itself are currently being equipped with hundreds of sensors for monitoring various parameters of interest. The collected maritime data from these various sensors include both spatial and temporal information, as they can be linked both with a geographical location and the time of occurrence of a specific event.

Maritime Informatics is a new discipline that involves applying information systems to increase the efficiency, safety, and ecological sustainability of the world's shipping industry. This paper synthesises the material from the *third* section of the book, Lind, M., Michaelides, M. P., Ward, R., and Watson, R. T. (Eds.). (2020). *Maritime Informatics*, Springer. In this section, entitled "*Maritime Informatics Technology,*" important dimensions of dealing with spatiotemporal data are covered, including *data acquisition*, *management*, *analysis*, and *visualisation*.

In particular, an extensive overview of the maritime data value chain is first provided, and state-of-the-art technological solutions are discussed for managing and processing maritime data in efficient and effective ways. Second, a holistic framework is described for performing spatiotemporal analytics along a ship's voyage including some practical applications like long-term route planning, environment preservation, collision avoidance, cargo tracking, and port-call optimisation. Next, creating meaningful data visualisations is presented, which can provide enhanced situational awareness in maritime operations and enable the distributed coordination of the multiple actors involved. Following this, technological advancements in intelligent information acquisition and representation are reviewed, to aid human users and organisations make sense of the complex maritime environment. Finally, maritime traffic surveillance systems for spatiotemporal data collection are reviewed and a computational framework is presented to compress efficiently, transfer, and acquire the necessary information for the further analysis of large scale automatic identification system (AIS) data.

As the ones presented in this paper, such intelligent algorithms and technologies can empower human users, both academics and practitioners alike, to "*make sense*" of the complex, spatiotemporal data associated with the maritime ecosystem, and provide enhanced decision support tools and solutions for the maritime stakeholders.

ACKNOWLEDGMENTS

This work was co-funded by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (STEAM Project: INTEGRATED/0916/0063 and MARI-Sense Project: INTEGRATED/0918/0032).

REFERENCES

Aslam S, Michaelides M., and Herodotou H., "Internet of Ships: A Survey on Architectures, Emerging Applications, and Challenges", *IEEE Internet of Things Journal* (*Special Issue on IoT for Smart Ocean)*, vol. 7, no. 10, pp. 9714-9727, Oct. 2020, https://doi.org/10.1109/JIOT.2020.2993411.

Anderson C., “American researchers want to fill the oceans with sensors”, *The Economist*, Mar 14, 2020, https://www.economist.com/science-and-technology/2020/03/14/american-researchers-want-to-fill-the-oceans-with-sensors

Besikçi, E. B., Arslan, O., Turan, O., & Ölçer, A. (2016). An Artificial Neural Network based Decision Support System for Energy Efficient Ship Operations. *Computers & Operations Research*, 66, pp. 393–401.

Cavanillas, J. M., Curry, E., & Wahlster, W. (2016). *New Horizons for a Datadriven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*. Springer.

Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In J. M. Cavanillas, E. Curry,&W.Wahlster (Eds.), *New Horizons for a Data-driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*, pp. 29–37. Springer International Publishing.

DGConnect. (2013). *A European Strategy on the Data Value Chain* (Tech. Rep.). Brussels, Belgium: European Commission. <URL: https://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3488>. [Accessed: July 29, 2020]

European Commission. (2016). *Mobility and transport – maritime* [WWW Document]. Maritime Mobility and Transport. Accessed April 12, 2019, from https://ec.europa.eu/transport/modes/maritime_en.

Ferreira, J.,Agostinho, C., Lopes, R., Chatzikokolakis, K., Zissis, D.,Vidal, M.-E.,& Mouzakitis, S. (2017). Maritime Data Technology Landscape andValue Chain Exploiting Oceans of Data for Maritime Applications. *Proceedings of the 2017 International Conference on Engineering, Technology and Innovation* (ICE/ITMC), pp. 1113–1122.

Fu X., Xiao Z., Xu H., Jayaraman V., Othman N., Chua C. P., Lind M. (2020), AIS Data Analytics for Intelligent Maritime Surveillance Systems, in M. Lind, M. Michaelides, R. Ward, R. T. Watson (Eds.), Maritime Informatics (chapter 23), Springer.

George, B., & Shekhar, S. (2008). Time-aggregated graphs for modeling spatio-temporal networks. In S. Spaccapietra, J. Z. Pan, P. Thiran, T. Halpin, S. Staab, V. Svatek, P. Shvaiko, & J. Roddick (Eds.), *Journal on Data Semantics XI, Lecture Notes in Computer Science* (pp. 191–212). Berlin: Springer. https://doi.org/10.1007/978-3-540-92148-6_7.

Han, J., Pei, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Elsevier.

Hayes D. (2016). Physical oceanography from Seaglider mission CNCY201114910 (Tech. Rep.). Oceanography Center.

Haykin S. (2006). Cognitive radar: a way of the future. *IEEE Signal Processing Magazine*, *23*(1), 30–40. Retrieved from http://ieeexplore.ieee.org/document/1593335/ doi: 10.1109/MSP.2006.1593335

Herodotou H., Aslam S., Holm H., Theodossiou S. (2020), Big Maritime Data Management, in M. Lind, M. Michaelides, R. Ward, R. T. Watson (Eds.), *Maritime Informatics* (chapter 19), Springer

IMO. (1974, November 1). *International convention for the safety of life at sea (SOLAS)* (Vol. 1184, p. 278). UN Treaty Series, London.

Karlsson M., Haraldson S., Lind M., Olsson E., Andersen T., Tichavska M. (2020), Data Visualisation Tools for Enhanced Situational Awareness in Maritime Operations, in M. Lind, M. Michaelides, R. Ward, R. T. Watson (Eds.), Maritime Informatics (chapter 21), Springer.

Kyriakides I., Hayes D., Tsiantis P. (2020), Intelligent Maritime Information Acquisition and Representation for Decision Support, in M. Lind, M. Michaelides, R. Ward, R. T. Watson (Eds.), Maritime Informatics (chapter 22), Springer.

Kyriakides I. (2019). *Cognitive Fusion for Target Tracking* (Synthesis Lectures ed.; A. Spanias, Ed.). Morgan and Claypool.

Lind, M., Haraldson, S., Karlsson, M., & Watson, R. T. (2015). *Port collaborative decision making–closing the loop in sea traffic management.* 14th International Conference on Computer Applications and Information Technology in the Maritime Industries, Ulrichshusen, Germany.

Lind, M., Haraldson, S., Karlsson, M., & Watson, R. (2016). *Overcoming the inability to predict- a PortCDM future*. 10th International Harbor Masters’ Association Congress–Global Port & Marine Operations.

Lytra, I., Vidal, M.-E., Orlandi, F., & Attard, J. (2017). A Big Data Architecture for Managing Oceans of Data and Maritime Applications. *Proceedings of the 2017 International Conference on Engineering, Technology and Innovation* (ICE/ITMC), pp. 1216–1226.

Michaelides, M. P., Herodotou, H., Lind, M., & Watson, R. T. (2019). Port-2-Port Communication Enhancing Short Sea Shipping Performance: The Case Study of Cyprus and the Eastern Mediterranean. *Sustainability*, 11(7), pp. 1912-34.

Perobelli, N. (2016, June). *MarineTraffic – A Day in Numbers*. [WWW] <URL: https://www.marinetraffic.com/blog/a-day-in-numbers/> [Accessed: July 29, 2020]

Schmitt P., Bartosiak M., Rydbergh T. (2020), Spatiotemporal Data Analytics for the Maritime Industry, in M. Lind, M. Michaelides, R. Ward, R. T. Watson (Eds.), Maritime Informatics (chapter 20), Springer.

Wang, H., Zhuge, X., Strazdins, G., Wei, Z., Li, G., & Zhang, H. (2016). Data Integration and Visualisation for Demanding Marine Operations. *Proceedings of the MTS/IEEE OCEANS 2016 Conference*, pp. 1–7.

Watson, R. T., Lind, M., Delmeire, N., & Liesa, F. (2020). Shipping: A self-organising ecosystem (Chapter 2). In M. Lind, M. Michaelides, R. Ward, & R. T. Watson (Eds.), *Maritime informatics*. Heidelberg: Springer.

Wikipedia Contributors. (2020, January 24). *Mare Liberum*. Wikipedia, The Free Encyclopedia. Retrieved February 15, 2020, from https://en.wikipedia.org/w/index.php?title=Mare_Liberum&oldid=937348672.

Xiao, Z., L. Ponnambalam, X. Fu, and W. Zhang. 2017. "Maritime Traffic Probabilistic Forecasting Based on Vessels' Waterway Patterns and Motion Behaviors." IEEE Transactions on Intelligent Transportation Systems 18 (11): 3122–34. https://doi.org/10.1109/TITS.2017.2681810.

Xiong, X., Mokbel, M. F., & Aref, W. G. (2017). Spatiotemporal Database. *Encyclopedia of GIS*, Springer Science & Business Media.

Xu, Hai-Yan, Vasundhara Jayaraman, Xiuju Fu, Nasri Bin Othman, Wanbin Zhang, Xiao Feng Yin, Deqing Zhai, and Rick Siow Mong Goh. 2019. "Efficient Compression and Pre-processing for Facilitating Large Scale Spatiotemporal Data Mining - A Case Study Based on Automatic Identification System Data." In 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 354–58. https://doi.org/10.1109/IEEM44572.2019.8978767.

Yablonsky, S. (2018). Innovation Platforms: Data and Analytics Platforms. In *Multi-Sided Platforms (MSPs) and Sharing Strategies in the Digital Economy: Emerging Research and Opportunities.* IGI Global, pp. 72–95.

Yang, Y., Zhong, M., Yao, H., Yu, F., Fu, X., & Postolache, O. (2018). Internet of Things for Smart Ports: Technologies and Challenges. *IEEE Instrumentation & Measurement Magazine*, 21(1), pp. 34–43.

Yeoh, C.-M., Chai, B.-L., Lim, H., Kwon, T.-H., Yi, K.-O., Kim, T.-H., Kwark, G.-H. (2011). Ubiquitous Containerized Cargo Monitoring System Development based on Wireless Sensor Network Technology. *International Journal of Computers Communications & Control*, 6(4), pp. 779–793.

Zhao, Y.-x., Li, W., Feng, S., Ochieng, W. Y., & Schuster, W. (2014). An Improved Differential Evolution Algorithm for Maritime Collision Avoidance Route Planning. In *Abstract and Applied Analysis*. Vol. 2014.

AUTHOR(S) BIONOTE(S)

Michalis Michaelides is an Assistant Professor with the Department of Electrical Engineering, Computer Engineering and Informatics at the Cyprus University of Technology. He has held a tenure-track position since January 2012. Michalis research interests include communication systems, wireless sensor networks, event detection and localisation, fault detection and diagnosis, fault tolerance, collaborative signal and information processing, computational intelligence with applications to environmental monitoring, intelligent systems and maritime informatics. Michalis has been involved as a principal investigator in many research projects, both local (CUT, RIF) and European (EU, FP7) including the Sea Traffic Management Validation project (EU, 2016-2019). He is currently the coordinator of the STEAM Project (RIF, INTEGRATED/0916/0063, 2019-2021). In 2014, he received the Elsevier Building and Environment Journal Best Paper Award.

Herodotos Herodotou is an Assistant Professor in the Department of Electrical Engineering, Computer Engineering and Informatics at the Cyprus University of Technology, where he is leading the Data Intensive Computing Research Lab. He received his Ph.D. in Computer Science from Duke University in May 2012. His Ph.D. dissertation work received the ACM SIGMOD Jim Gray Doctoral Dissertation Award Honourable Mention. Before joining CUT, he held research positions at Microsoft Research, Yahoo! Labs, and Aster Data. His research interests are in large-scale data processing systems, database systems, and cloud computing. In particular, his work focuses on ease-of-use, manageability, and automated tuning of both centralised and distributed data-intensive computing systems. In addition, he is interested in applying database techniques in other areas like maritime informatics, scientific computing, bioinformatics, and social computing. His research work to date has been published in several top scientific conferences and journals, two books, and two book chapters. He was/is actively participating in four large EU-funded projects, namely STMV, ENCASE, NOTRE, and ENGINITE, as well as two nationally funded projects, namely STEAM and

MARI-Sense. He is acting as the Scientific Coordinator of the STEAM project and was the Technical Coordinator of the Cyprus team in STMV.

Patrick Schmitt practices consultancy in the area of supply chain management and carbon accounting. Within the maritime supply chain industry, he has specific experience in utilising graph databases to create a more efficient and collaborative architecture across multiple actors.

Mathias Karlsson is a researcher at RISE Research Institutes of Sweden at the department of mobility and systems since 2011. He has a B.Sc., in Shipping and Logistics from Chalmers University of Technology, and a M.Sc., with a major in Informatics from University of Gothenburg. Mathias has from the beginning been involved with the development and validation of Port Collaborative Decision Making (PortCDM) which was one of the enabling concepts within the EU funded project Sea Traffic Management validation project. During his years at RISE he has been working in several different EU-funded project such as SUNSET, EMPOWER and Monalisa 2.0 and is currently working in the EU-funded project FEDeRATED focusing on supporting interoperability in supply and logistics chains where it will be demonstrated how the federative platform as proposed by the EU Digital Transport and Logistics Forum (DTLF) can work. Furthermore Mathias is also involved in several Swedish initiatives such as Digitalization, automation and electrification in small and medium sized ports in Sweden and different CDM initiatives and testbeds for the transport system in Sweden. Prior to joining RISE eight years ago, he worked for several years within the maritime sector, both as a ship agent and in other port operations in the Energy Harbour in the port of Gothenburg.

Ioannis Kyriakides received his B.S. degree in Electrical Engineering in 2003 from Texas A&M University. He received his M.S. and Ph.D. degrees in 2005 and 2008 respectively from Arizona State University. He held a research associate position throughout his graduate studies, funded by the Integrated Sensing and Processing program and the Multidisciplinary University Research Initiative of the USA Department of Defence. In the final year of his Ph.D. work, he received the University Graduate Fellowship of the Arizona State University. His research interests include Bayesian target tracking, sequential Monte Carlo methods, and cognitive fusion for heterogeneous data fusion and heterogeneous sensing node configuration. His research work includes localisation of multiple RF sources, tracking surface vehicles using passive acoustic sensing, tracking multiple targets with constraints in motion, tracking multiple targets using heterogeneous cognitive sensors. Dr. Kyriakides is currently an Associate Professor at the Engineering Department at the University of Nicosia, a researcher at the University of Nicosia Research Foundation, and a Senior Scientist at the Cyprus Marine and Maritime Institute. He was the coordinator of a Cyprus Research and Innovation Foundation project "Automated Sonar-Thermal Monitoring for Sea Border Security" with budget € 180 K. He is currently the Principal Investigator of the Research and Innovation Foundation funded Project: "MARI-Sense: Maritime Cognitive Decision Support System" with budget € 1 M (www.marisenseproject.net) and "OS Aqua: Open Sea Aquaculture in the Eastern Mediterranean" with budget € 0.8 M (www.osaqua.blue).

Daniel Hayes is the Managing Director and Oceanographer at Cyprus Subsea Consulting and Services (CSCS). He has participated in several national projects on ocean observations and modeling and was the coordinator of a national infrastructure project to set up and operate an ocean glider facility to study the seas of Cyprus using underwater gliders. Recently he has led the drafting and implementation as leader of work packages in the EU FP7 project: GROOM, the EU H2020 project BRIDGES, and the U.S. Office of Naval Research project: CINEL. He specialises in the analysis of data collected by ships, buoys and autonomous vehicles to investigate small and meso-scale physical exchange processes. He has been working on the development of the real time observing and forecasting system for the Cyprus region, including collection and analysis of data and development of numerical circulation models including the assimilation of in situ glider data. Other work consists of environmental modelling and data collection and analysis for baseline and environmental impact assessments for commercial and regulatory stakeholders. He is a member of the EuroGOOS Glider Task Team and the Steering Committee of OceanGliders, the glider component of the GOOS.

Xiuju Fu is a Senior Scientist in the Department of Systems Science at Institute of High Performance Computing, A-STAR. She received her Ph.D. in Information System, School of EEE, Nanyang Technological University, 2003. Her research areas are in intelligent modelling and simulation based on data analysis, modeling/simulation and optimisation techniques, with the applications in maritime port and transportation, logistics and supply chain, and public health. She had led multiple maritime projects in maritime traffic safety and operation efficiency for developing big data intelligence driven solutions. She had co-authored a book "Data Mining with Computational Intelligence". Her team had been awarded the Outstanding Paper Award recently for the research in logistics and supply chain risk management in 2015 IEEE International Conference on Industrial Engineering and Engineering Management.

WITH THE SUPPORT OF

CIMOSM
Centro de Investigação em Modelação e Optimização de Sistemas Multifuncionais

ECONOMIA AZUL