

Decentralized System Intelligence in Data Driven Networks for Shipping Industrial Applications: Digital Models to Blockchain Technologies

Lokukaluge P. Perera
Karen Czachorowski

Accepted version of article in
OCEANS 2019 – Marseille

Original version:

DOI: [10.1109/OCEANSE.2019.8867045](https://doi.org/10.1109/OCEANSE.2019.8867045)

Decentralized System Intelligence in Data Driven Networks for Shipping Industrial Applications: Digital Models to Blockchain Technologies

Lokukaluge P. Perera
UiT The Arctic University of Norway
Tromsø, Norway
Prasad.Perera@uit.no

Karen Czachorowski
USN University of South-Eastern Norway
Vestfold, Norway
Karen.Czachorowski@usn.no

Abstract—Data driven networks applicable for shipping industrial applications to create decentralized system intelligence are considered in this study. Such system intelligence can facilitate to improve the respective operational efficiency in local (i.e. vessel operations) and global (i.e. logistics operations) scales in shipping as the main advantage. The main features of these data driven networks are summarized in the first part of this study. Two applications of digital models and blockchain technologies are discussed and compared with their features to illustrate their similarities and differences in the second part of this study. A digital model represents a vector based mathematical structure derived from ship performance and navigation data sets and has categorized as a low-level information model. It is also believed that the respective data sets from industrial IoT (internet of things) should go through such low-level models to improve their quality. These data driven networks can be used to quantify ship performance and navigation conditions, where the outcome can also be used to improve vessel energy efficiency and reduce engine emissions in a local scale. A blockchain represents a decentralized, distributed and digital ledger system in a public domain and can handle and record transactions executed by many users. That has categorized as a high-level information model due the high quality data sets from industrial processes that these networks are handling. Such data driven networks can be used to formulate various logistics operations in shipping and optimize their operational conditions in a global scale. The outcomes of these data driven networks can be used to improve operational efficiency and reduce the respective costs in the shipping industry.

Keywords—*Digital Models, Blockchain Technology, Data Driven Networks, System Intelligence, Shipping Industry.*

I. INTRODUCTION

The shipping industry, with modern vessels and shore based operational centers, collects large scale data sets, so called big data, in which should be utilized to improve various sectors of its operations. However, the utilization of such data sets may not be an easy task due to the respective data handling challenges [1]: data volume, velocity, variety and veracity. The volume and velocity issues addressed under data management-based applications relate to the scale (i.e. the size) and speed (i.e. the processing speed) of data sets, respectively. The variety and veracity issues addressed under data analytics-based applications relate to the forms (i.e. the types of measured parameters and their formats) and uncertainty (i.e. the quality) of data sets, respectively.

Since the shipping industry collects various forms of data sets from different sources (i.e. sensor data, industrial

processes data, etc.), their quality can vary. The quality can be an issue for the data sets coming from industrial IoT (internet of things). However, the data sets coming from industrial processes may not have such data quality issues. However, the data quality can eventually effect on the outcome of the analysis, therefore appropriate tools and techniques to improve the same should be considered. The quality improved data sets can be used to support various industrial applications.

Since conventional data handling approaches often fail to address data handling challenges that relate to variety and veracity issues, data drive networks are considered by various industrial applications. Machine learning (ML) and artificial intelligence (AI) based approaches can be accommodated into such networks consisting of various advanced analytics. It is also expected that such networks have the capabilities to improve the data quality by accommodating appropriate advanced analytics. This study also investigates analytics-based data driven networks developed to support massive data sets from the shipping industry. Hence, that can also provide appropriate solutions to the real-time data handling challenges. However, these challenges and solutions can often relate to the respective industrial applications. One should note that real-time data sets can assign the property of discreteness into these networks in these situations.

Advanced data analytics can also introduce additional system intelligence, i.e. due to the respective information extracted from the data sets, into these networks. A given set of rules should be used to develop these data driven networks as well as advanced analytics. Hence, this can also be seen as a decentralized system intelligence approach. It is expected that decentralized system intelligence can provide flexible operational solutions, especially in a local level, to the respective industrial application. That can also be a better solution to highly complex distributed industrial applications depend on massive data sets. Hence, this study illustrates such intelligent networks in shipping supported by various data sets from industrial sensors and processes. The main features of such data driven networks are discussed in the following section.

II. DATA DRIVEN NETWORKS

A. Network Architecture

In general, a network can be defined as a group of nodes that are interconnected in a pre-defined distribution. Such network architecture can be series, parallel or a combination

of both and relates to the respective system or application requirements. Since this study is interested in data driven networks, these nodes can often be categorized as data/information clusters. One should note that the respective information in each node should be derived from the data sets. There are several steps that should be taken to transform data into information [2], where advanced analytics can be used. The information on each node can often be updated with real-time data from industrial sensors and processes. These information nodes introduce the respective discreteness in the data driven networks, therefore that can be categorized as digital networks. Furthermore, these nodes have the capability of extracting, updating, receiving or transmitting information over the network through incoming and outgoing data sets. That can also be a part of a larger data management framework. Such activities in the network nodes can create distributed system intelligence supported by decentralized information handling processes. Hence, these data driven networks can support to overcome modern industrial challenges in the shipping industry by utilizing the respective data sets.

B. Node Classification

There are several main features that can be observed under the nodes of such data driven networks. A general representation of a data driven network with several nodes is considered to explain its features in this study. However, these features can vary with respect to the industrial application. Each network node should consist of a quantifiable information structure supported by the respective data sets. In general, these nodes can classify into two main categories: honest and not-so-honest nodes. It is expected that honest nodes, with higher computational power and information, create the main network (or the chain) and that network preserves its information integrity. The honesty in these nodes relate to the trust of the respective information derived from the real-time data sets. Not-so-honest nodes can also play an important role in the respective data driven networks in some situations. That role may depend on the respective industrial application. However, such nodes should also be identified and either be accommodated or eliminated (i.e. isolated) from the main network to preserve its information integrity. The inclusion or exclusion of such nodes depends on the respective industrial application, as mentioned before.

Each honest node consists of a representative information structure, in a higher level, and supported by a data cluster, in a lower level. Since these clusters can be updated through real-time data, the respective information structure, in each node, can also be updated. This information structure that relates to network intelligence in each node can have varying shapes throughout the network. One should note that the structural shape of each node relates to its respective information. The information structure in each node should be time-stamped, properly to preserve its integrity [3]. Since these networks update with real-time data, that can introduce possible structural changes. The structural changes with respect to the time line in data driven networks should be preserved to trace their evolutions. Hence, the network evolution can provide extensive information on the past and present conditions of the industrial application. Furthermore, the same network information can also be used to predict future conditions of the industrial application in some situations.

These data driven networks can accommodate additional features of self-learning, self-cleaning, and self-compression & expansion. In general, a given set of rules should be assigned into these networks to grow from its initial node, as mentioned before. The network can add or remove honest nodes, in accordance with its rules, i.e. the self-learning feature. However, not-so-honest nodes can also be added or removed during this process. The network should have enough system intelligence to identify such nodes, and isolate or remove from the respective network, as required. That feature represents self-cleaning capabilities of such data driven network. If the network structure is well-defined, then that can be used to scale, i.e. compress and recover, the respective information in its nodes. That feature can introduce self-compression & expansion capabilities into these networks and improve the network operational efficiency. In general, any intelligent network should have these important features to interact with the respective data sets.

A network with distributed system intelligence may not have a deterministic leader node to make all its decisions. However, there are some situations that these networks may require to have a leader to make the final decision. A leader node can be appointed under the respective network rules in such situations. Furthermore, these nodes can also communicate each other, i.e. exchange information, to formulate group decisions under various logical operations. Therefore, the network decisions represent its system intelligence and support the respective industrial applications. However, the domain knowledge from the respective industry should also be integrated into such data driven networks to make smarter decisions. This domain knowledge can be incorporated by either network rules or logical operations. Two types of data driven networks with possible shipping industrial applications are discussed in this study: digital models and blockchain technologies.

III. DIGITAL MODELS

A. Network Structure

Digital Models are another form of data driven networks developed to quantify ship performance and navigation conditions. Ship performance and navigation parameters collected by onboard IoT as data sets are considered to derive these networks [4]. A general representation of a digital model is presented in Figure 1. As presented in the figure, the network consists of several nodes as data clusters, i.e. C_1, C_2, C_3 , etc., that relate to localized vessel navigational and ship system operational conditions. Each data cluster can be represented by a vector that locates its mean value, i.e. μ_1, μ_2, μ_3 , etc.. Hence, the entire digital model can be represented by a group of mean vectors in a higher level. Furthermore, each data cluster can have several structural vectors, i.e. $Z_{1,1}, Z_{1,2}, Z_{1,3}$, etc., in relation to its mean vector in a lower level. These structural vectors have also been categorized as singular vectors and that can be the building blocks for digital models. It is expected that future mathematical models for electrical and mechanical systems can be presented by such digital models, i.e. data driven networks, derived by the system data sets. One should note that this concept is a slight deviation from the digital twin. These data driven networks are application oriented models, i.e. digital representative systems, and may not capture the total complexity in the system behavior.

A given set of rules can be deployed to create such digital models, i.e. network structures, from ship performance and navigation data sets. The rules should formulate the number of data clusters by the mean vectors and the structural vectors of each data cluster in relation to the same mean vectors. One should note that these structural vectors represent the respective correlations among ship performance and navigation parameters. Hence, the combination of mean and structural vectors represents the structure of the same data set. The data structure becomes the mathematical model for vessel navigational and ship system operational conditions. Since these models can be application oriented, the respective ship performance and navigation parameters should be selected in accordance with the application requirements.

This is a digital representation of vessel navigational and ship systems operational conditions. Hence, the representative vessel behavior can move from one data cluster (or node) to another due to navigational and operational changes. Furthermore, the information stored in each node can vary due to the updates from real-time data sets. These data clusters can be classified as honest and not-so-honest nodes under the same data driven network. A not-so-honest node is also presented in Figure 1 as a data anomaly cluster. Furthermore, some data points beyond the data clusters are also presented as outliers in the same figure and those outliers may relate to the data quality issues. The respective sensors data from industrial IoT can often introduce such data anomalies and outliers into the digital models.

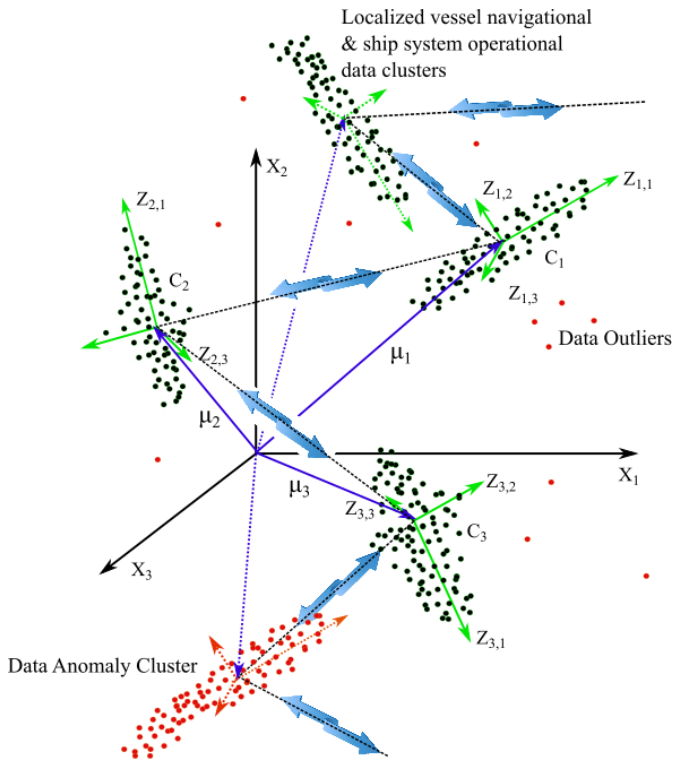


Fig. 1. Digital model formation

B. Engine Data

The initial concepts of digital models are developed by investigating ship performance and navigation data sets, where the respective data clusters, i.e. nodes, that relate to

vessel navigational and ship system operational conditions are observed. A data density map from a marine engine of a selected vessel is presented with respect to engine power, shaft speed and fuel consumption values in this section to illustrate the above claim. An example of a data density map is presented in Figure 2 and that represents marine engine speed-power data for the same vessel. The respective data clusters are represented as kernel density estimation, i.e. data density maps, to improve the information visibility. One should note that this data set has three data clusters, i.e. three network nodes that relate to the respective engine operational modes. Hence, this engine has three operational modes represented by three data clusters.

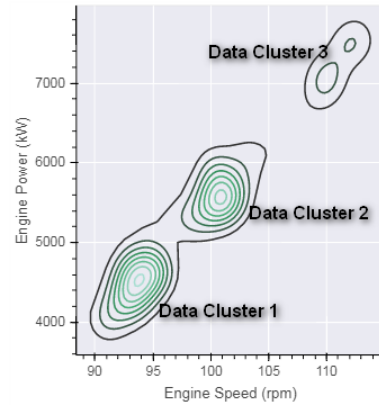


Fig. 2. Marine engine speed-power Data

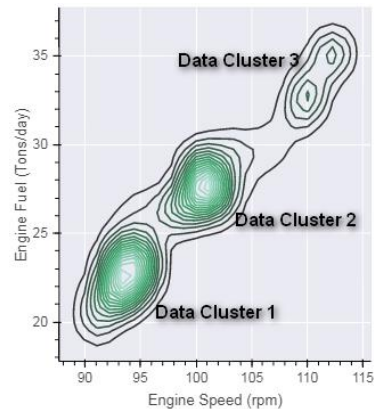


Fig. 3. Marine engine speed-fuel consumption data

The respective engine speed-fuel consumption data for the same marine engine are presented in Figure 3. These data clusters also relate the engine modes in the previous figure (as denoted in both figures). Since these engine speed-power-fuel consumption clusters relate to each other, that create the respective digital model with three data clusters. e.g. this data driven network consists of three marine engine related network nodes. Hence, a digital model can exist in a high dimensional data space and the space dimensionality relates to the number of system parameters measured by industrial IoT. That can also introduce complex geometrical shapes into the digital models.

However, the visualization of such digital models can be a difficult task due to its space dimensionality. Hence, these data density maps can be used to visualize a projected view of such data driven networks to support the model

development process. However, these digital models can further be expanded by considering other ship performance and navigation parameters. e.g. the respective trim and draft conditions can introduce additional data clusters (see Figure 4 & 5). Therefore, other ship performance and navigation parameters can reorganize the digital models along the existing clusters or create additional clusters to support such data driven networks.

C. Data Anomalies

There are several additional benefits that can be obtained through the structural shapes of digital models. Firstly, the respective data anomalies can be identified and isolated from the digital models. As discussed previously, these data driven networks can consist of not-so-honest nodes to represent data anomaly regions. Since ship performance and navigation data come from various IoT, the quality of such data sets may not be higher. Data anomalies can relate to system abnormal events and sensor faults of vessel navigational and ship system operational conditions. The respective data anomalies can be identified from the same data structure by observing its abnormal behavior. That can be done in two levels. Firstly, the domain knowledge can be used to identify data anomaly clusters that often relate to system abnormal events. Secondly, the outlier of the data structure can be used to identify possible data anomalies [5] that often relate to sensor faults. Hence, the respective network structures can also be used to detect, isolate and recover various data anomalies, i.e. a good solution to the data veracity issue.

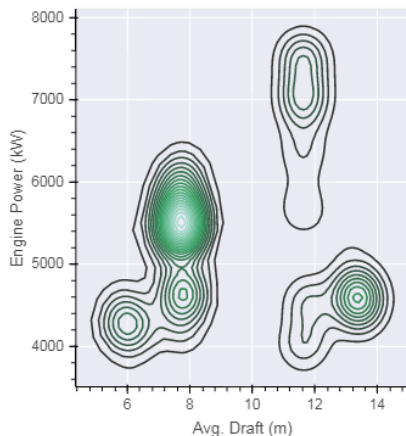


Fig. 4. Marine engine average draft-power data

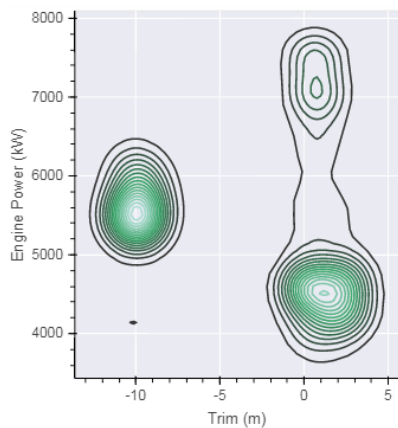


Fig. 5. Marine engine trim-power data

The information on system abnormal events is important to improve the operational efficiency of ocean going vessels. Such information can be used to arrange appropriate maintenance for the vessel and ship systems. The information on sensor faults can also be used to identify the respective faulty sensors and replace or arrange maintenance for the same [6]. The sensor faults can be determined from not-so-honest nodes or outliers of the network, where these faults can be isolated or recovered by the same. That can be seen as a process, where the respective data sets from industrial IoT go through a low-level digital model to improve their quality.

Secondly, the structural shapes of digital models can be used to support various shipping industrial applications, e.g. ship energy efficiency, emission control and system reliability [4]. Since these models are derived from ship performance and navigation data sets, that can be a representative model for vessel and ship system behavior. Furthermore, such models can be a good solution to the data variety issue, since various forms of data sets can be combined through data driven networks. Such approaches have not been used by the shipping industry to the authors' knowledge. Furthermore, these models can be placed onboard vessels and updated with real-time ship performance and navigation data sets to enhance the respective applications.

IV. BLOCKCHAIN TECHNOLOGY

A. Network Structure

Blockchain technology is another form of data driven networks that consist of decentralized transactional processes with additional data management technologies [7]. That can be a higher level model with compared to the digital models. Since these networks can have enough system intelligence to execute the respective transactions by themselves, any third-party institutions in the middle to support the same may not require. One should note that such third-party institutions often preserve the transactional integrity, security and privacy in such situations. That can be the main advantage of having blockchain technology in industrial applications, where the respective paper work can be eliminated. Furthermore, the transactional costs can be reduced, transparency can be improved, and speed can be increased due to the same reasons. This technology has been highlighted by the recent developments in Bitcoin cryptocurrency [8]. According to the world economic forum, about 10% of the global GDP will be stored in blockchain type technologies by 2027 [9].

A general overview of a blockchain formation is presented in Figure 6. The nodes of this data drive network represent the respective blocks. As presented in the figure, the main chain consists of honest nodes of the network. It is expected that the main chain with adequate system intelligence controls the entire network. Furthermore, the identical blocks with the transactional information are shared and stored in the main chain. One should note that the information stored in every node is similar, due to the same reason. That feature is a deviation from digital models. The genesis block represents the first block, i.e. the birth, of the blockchain (see Figure 6). The orphaned blocks can be categorized as not-so-honest nodes that may not be a part of the main chain. However, those can be either valid blocks

created due to the time delay in the networks or invalid blocks created by cyber-attacks with enough computational power to interfere with the respective transactions.

Since these networks may not be controlled by a single entity, i.e. distributed system intelligence, that can have extensive robustness, i.e. no single point failures. It is also noted that such networks are impossible to corrupt, i.e. introduce any data anomalies, by influencing any node, since the respective information has been shared among many nodes. Therefore, such networks can be smarter, and the computational power requirement to introduce data anomalies can be higher. On the other hand, these networks can detect, isolate and recover such data anomalies, faster. These anomalies represent not-so-honest nodes, i.e. erroneous nodes, and can be a part of the network. That may also reduce the respective network integrity in some situations.

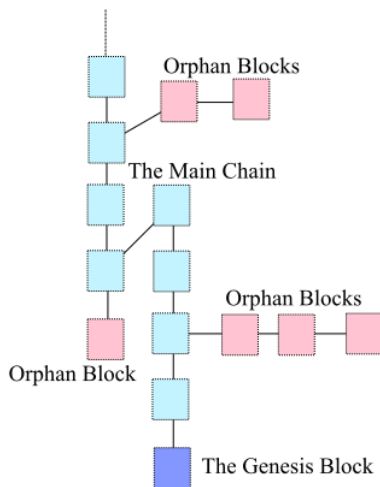


Fig. 6. Blockchain formation

To overcome such efforts, the honest nodes, i.e. the main chain, should be supported by adequate computational power. While these networks are getting expended, the integrity of the blockchain can grow due to the same reason. In general, the blockchain can have a higher number of honest nodes and a fewer number of not-so-honest nodes due to its maturity. Therefore, the age can improve the respective network integrity of blockchain technology. A shipping industrial application of managing a logistics network in relation to block technology is discussed in the following section to further elaborate these data driven networks.

B. Logistics Operations in Shipping

Shipping logistics is a complex network of peers that raises and transacts vast amounts of data at very different levels, being most of these data sets not tracked or used for any relevant purpose. Within the operational layers, there are several possible gates for data collection, which can be used to raise events to trigger sequential actions. The smart contract functionality in blockchain technology can add an appropriate logic into the network, allowing event and data-driven triggers to optimize shipping operations. i.e. a simple “if-then-else” logic can be triggered by a GPS signal when a vessel is at a certain port distance to share needed documentation to the nodes in the network, allowing the vessel berthing and unloading through a simpler process. At the same time, the structure of the blockchain, i.e.

decentralized information storing and sharing, can track, maintain and distribute the events raised by the respective data in the network, allowing sequential triggers to be started from each event in one or more nodes. The main blockchain protocols also allow setting channels and similar privacy handling in the respective information, which can contribute as an intelligent network. Channels, private data collection and private smart contracts are “sub-blockchain” networks within pre-defined parties inside the main blockchain network [11, 12, 13]. They allow routing data from the right source to the right parties with security, without the need for additional activities or institutes.

An in-depth case study of a real blockchain application is currently in a development phase at Aker BP, a Norwegian Oil and Gas operator in the Norwegian Continental Shelf. The respective experiences gained through the same development phase are summarized in this section. The company’s project called exChain is developing a modular software that applies blockchain for data sharing by creating a protected while efficient environment for information handling in a multiple parties’ network. A general framework of exChain is presented in Figure 7. The exChain project aims to optimize the business processes in offshore logistics and supply chain by developing digital solutions and environments to enable autonomous operations through trusted data exchanges.

The project is developing a blockchain consortium to securely share data from the operator and its outside stakeholders involved in a logistics and supply chain, such as inland and offshore transporters, suppliers, terminal bases/warehouses and etc. The blockchain network was tested in a prototype completed in the first quarter of 2019. It was built on Hyperledger Fabric, i.e. a blockchain framework implementation, with channels to give the privacy needed. After testing some scenarios, the project decided to test Ethereum Quorum, which will lead to a ledger comparison to define the best protocol for the company’s and industrial goals. An important note is that the system design in the development phase allows blockchain technology to work as a sharing network and give a sense to the shared data.

The blockchain network is not the central figure in the project, as it relies on the blocks for data sharing with simple logic, leaving the business logic to be handled and orchestrated through an interoperability framework that is the responsible for orchestrating data for consumption as needed. This framework sets the rules for the data flow, security and software integration through APIs (Application programming interfaces) to allow seamless direct and indirect integration between exChain, the company’s different business units, blockchain and the network participants’ own software. The software structure consists of focused modules and with REST (Representational state transfer) based APIs for communication between client and server, constituting in a typical three tier application structure. A SQL/no-SQL database is used for the data storage and web API constitutes in the middle tiers processing. This connects to a gateway for integrating the blockchain, data lakes and/or IoT platforms, chosen for security purposes [14]. The preliminary results from testing the technology applications include increased data visibility, traceability and lower barrier to technological innovation [11].

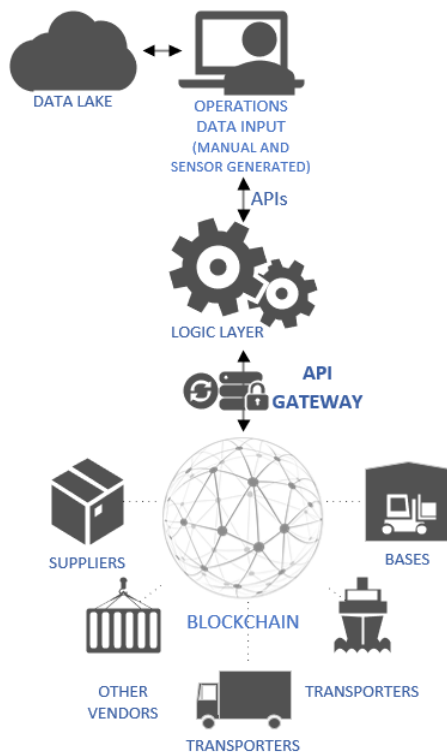


Fig. 7. exChain general system architecture

V. CONCLUSIONS

An introduction of decentralized data driven networks towards system intelligence in shipping industrial applications is presented in this study. Furthermore, the representative networks of digital models and blockchain technologies are also discussed. Digital models represent a network structure that can be derived from ship performance and navigation data sets. The outcome of digital models can be used to quantify ship performance and navigation conditions, i.e. to improve energy efficiency and reduce engine emissions from ocean going vessels [14, 15]. Blockchain represents a decentralized, distributed and digital ledger in a public domain and the network can handle and record transactions based on operational data from various sources. One should note that the blockchain is a high-level cleaner data driven network than the digital model due to the quality of the data sets.

These networks are governed by a given set of rules and that introduce the respective distributed system intelligence. However, there are several challenges have been observed with respect to these networks. The network structure can be highly complex, therefore the required computational power to update such networks can be higher. That can also compromise the network integrity, where not-so-honest nodes, i.e. data anomalies or cyber-attacks, with higher computational power may succeed to challenge the respective operations. Since these networks can be updated in real-time data, that can further increase the required computational resources. However, the scaling of the networks or having enough computational power and speed can be two possible solutions to overcome such situations. These data driven network technologies can be used in various ship operations to optimize and improve operational

efficiency and reduce the respective costs both on local and global scales.

ACKNOWLEDGMENT

This work is supported by the MARKOM2020 project, a development project for maritime competence established by the Norwegian Ministry of Education and Research in cooperation with the Norwegian Ministry of Trade and Industry.

REFERENCES

- [1] L.P. Perera and B. Mo, "Machine Intelligence based Data Handling Framework for Ship Energy Efficiency," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 10, 2017, pp. 8659-8666.
- [2] L.P. Perera, "Industrial IoT to Predictive Analytics: A Reverse Engineering Approach from Shipping," In *Proceedings of the 3rd Norwegian Big Data Symposium (NOBIDS)*, Trondheim, Norway, November, 2017, pp. 12-17.
- [3] S. Haber, and W. S. Stornetta, "How to time-stamp a digital document", *Journal of Cryptology*, vol. 3, no. 2, 1991, pp. 99-111.
- [4] L.P. Perera, and B. Mo, "Ship Performance and Navigation Information under High Dimensional Digital Models," *Journal of Marine Science and Technology*, (DOI: 10.1007/s00773-019-00632-5).
- [5] L.P. Perera and B. Mo, "An overview of Data Veracity Issues in Ship Performance and Navigation Monitoring," In *Proceedings of the 37th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2018)*, Madrid, Spain, June, 2018 (OMAE2018-77669).
- [6] L.P. Perera "Marine Engine Centered Localized Models for Sensor Fault Detection under Ship Performance Monitoring," In *Proceedings of the 3rd IFAC Workshop on Advanced Maintenance Engineering, Service and Technology*, (AMEST'16), Biarritz, France, vol. 49, no. 28, 2016, pp. 91-96.
- [7] J. Yli-Huumo, D. Ko, S. Choi, S. Park, K. Smolander, "Where Is Current Research on Blockchain Technology?—A Systematic Review," *PLoS ONE* 11(10), 2016, e0163477. (A1)
- [8] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," URL: <http://bitcoin.org/bitcoin.pdf>. (A2)
- [9] V. Dieterich, M. Ivanovic, T. Meier, S. Zapfel, M. Utz and P. Snader, "Application of Blockchain Technology in the Manufacturing Industry," FSBC working paper, Frankfurt School Blockchain Center, 2017.
- [10] K. Czachorowski, M. Solesvik, & Y. Kondratenko, "The Application of Blockchain Technology in the Maritime Industry," In *Green IT Engineering: Social, Business and Industrial Applications* (pp. 561-577). Springer, Cham, 2019.
- [11] Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. In *2017 IEEE International Congress on Big Data (BigData Congress)* (pp. 557-564). Honolulu, HI, USA: IEEE. <https://doi.org/10.1109/BigDataCongress.2017.85>
- [12] Buterin, V. (2018). Consortium Chain Development. Retrieved from <https://github.com/ethereum/wiki/wiki/Consortium-Chain-Development>
- [13] <https://github.com/jpmorganchase/quorum-examples/tree/master/examples/7nodes>
- [14] K. Czachorowski, A. H. Gausdal, "Towards Data Interoperability and Sharing through Blockchain: A Use-Case in Autonomous Supply Chain & Logistics Operations in the Norwegian Continental Shelf Oil and Gas Industry". Forthcoming, 2019.
- [15] K.Q. Bui and L.P. Perera, "The Compliance Challenges in Emissions Control Regulations to Reduce Air Pollution from Shipping," In *Proceedings of the MTS/IEEE OCEANS 19*, Marseille, France, June, 2019 (Accepted).
- [16] B. Mo, C. Steinebach, L.P. Perera, P. Dehli, and T. F. Lim, "An Automated System for Fleet Benchmarking and Assessment of Technical Condition," In *Proceedings of the 36th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2017)*, Trondheim, Norway, June, 2017 (OMAE2017-61219).