Study on route division for ship energy efficiency optimization based on big environment data

Wang, Kai; Yan, Xinping; Yuan, Yupeng; Jiang, Xiaoli; Lodewijks, Gabri; Negenborn, Rudy

DOI
10.1109/ICTIS.2017.8047752

Publication date
2017

Document Version
Accepted author manuscript

Published in
Proceedings of the 4th International Conference on Transportation Information and Safety (ICTIS 2017)

Study on Route Division for Ship Energy Efficiency Optimization Based on Big Environment Data

Kai WANG
Reliability Engineering Institute, School of Energy and Power Engineering
Wuhan University of Technology
Wuhan, China
Email: K.Wang-2@tudelft.nl

Xinping YAN
National Engineering Research Center for Water Transport Safety, (WTSC)
MOST
Wuhan, China
Email: xpyan@whut.edu.cn

Yupeng YUAN
Reliability Engineering Institute, School of Energy and Power Engineering
Wuhan University of Technology
Wuhan, China
Email: canicula2000@whut.edu.cn

Xiaoli JIANG
Maritime and Transport Technology
Faculty of Mechanical, Maritime and Materials Engineering
Delft University of Technology
Delft, The Netherlands
Email: X.Jiang@tudelft.nl

Gabriel Lodewijks
Maritime and Transport Technology
Faculty of Mechanical, Maritime and Materials Engineering
Delft University of Technology
Delft, The Netherlands
Email: G.Lodewijks@tudelft.nl

Rudy R. Negenborn
Maritime and Transport Technology
Faculty of Mechanical, Maritime and Materials Engineering
Delft University of Technology
Delft, The Netherlands
Email: R.R.Negenborn@tudelft.nl

***Abstract* —In the case of the global energy crisis and the higher sound of energy saving and emission reduction, how to take effective management measures of ship energy efficiency to achieve the goal of energy saving and emission reduction, put forward a new challenge for the development of shipping technology. The application of big data technology provides a new idea for the research of ship energy efficiency optimization management. The energy efficiency management level of the operating ship can be improved by the analysis and mining of the big data. In this paper, a big data analysis platform for ship energy efficiency management based on the widely used Hadoop platform architecture is designed. Afterward, due to the huge amount of involved data on the energy efficiency management which has exceeded the processing ability of traditional solutions, the big data analysis method is used to achieve the route division according to environmental factors, thus to lay the foundation for speed optimization in different segments of a route. Finally, a simple decision-making method of optimal engine speed based on the result of route division is proposed, which could improve ship energy efficiency and hence reduce CO_2 emission.**

Keywords—ship energy efficiency optimization; big data; energy saving and emission reduction; Hadoop;

I. INTRODUCTION

Recently, big data has rapidly become a hot topic in the field of science and technology, the business and even the governments around the world. 'Nature' and 'Science' have published monographs illustrating the opportunities and challenges that big data brought [1-2]. Data has penetrated into the domains of every industry and business and has become an important influencing factor in production. The application of big data indicates the arrival of a new wave of productivity growth and consumer surplus [3]. The United States government believes that big data is the new oil in the near future. The size of the owned data and use ability of them will become a significant part of the overall national strength for a country. All in all, Big data has become the focus of attention of the whole communities and the era of big data has come [4].

Nowadays, the rapid development of the Internet, Internet of Things and cloud computing promotes explosive growth of data in almost every industrial and commercial areas [5]. Big data has attracted wide attention from academia, industry, and government from all over the world. Kambatla et al. [6] discussed the development trend of big data, including system hardware and the achievement of data analysis technology. Moreover, Jin et al. [7] described the worldwide representative researchers based on big data and the relative great challenges, including the complexity of data and computation as well as a system. In the field of ship and ocean engineering, with the development of data acquisition technology, the big data sets of the ocean have been formed in the last several years. Huang

Supported by National Key Technology Support Program (No. 2013BAG25B03) and the Hubei Provincial Leading High Talent Training Program Funded Project (File No. HBSTD [2012]86) and the China Scholarship Council under Grant 201606950063.

et al. [8] analyzed the management structure of marine big data and introduced the relative management models and methods. He pointed out that there are potentially useful values in the large amount of data, such as tsunami and red tide warning, prevention and prediction, disaster inversion, visual hazard modeling and other functions. Meanwhile, he also analyzed the new challenges brought by the ocean big data management, such as data acquisition, storage, analysis, application as well as data quality control and security.

Ship operation data obtained from shipping information platform and monitoring platform mainly include ship dynamic data (such as ship position, speed, draft, and trim), navigation environment factors (such as the wind, wave, water depth and speed) and ship fuel consumption data. In addition, we can also obtain fuel filling, fresh water storage and supply information, dynamic information of import and export, the loading and unloading of the goods and anti-typhoon information. These data can constitute a relatively complete set of ship operation data which contain a lot of potentially important information. Through the analysis and mining of these massive data, a lot of valuable information and rules can be found. They can be used to the statistical analysis of the navigation environment and to analyze the impact of them on the ship running status and energy efficiency, thus to promote much safer and more efficient navigation of ships. Moreover, it is greatly helpful for shipping companies to achieve the optimization management and deployment of the fleet.

Among them, the research and application of big data analysis in ship navigation optimization have gained extensive attentions in the shipping industry. Ship energy efficiency is not only related to shipping sailing speed [9-10] but also influenced by the environmental factors during her voyages, such as wind speed, wind direction and water depth [11]. A real ship study by IMO shows that only 55% of the propulsive force is used to overcome the hull friction and 17% to overcome the wave resistance, while as much as 28% to overcome the increased resistance caused by the wind, current and other environmental factors [12]. On the other hand, the high degree of randomness and complexity of the environmental factors also have an important impact on working conditions and operating states of the main engine and the whole power system, and it hence influences the optimal sailing condition of ships [13]. Therefore, identifying the optimal sailing speed in different situations considering comprehensively the environmental factors and sailing state of the ship has become the focus of ship operation energy efficiency optimization and enhancement research, as it plays a significant role in facilitating the ship's economic and environmentally-friendly sailing.

Navigation environment is an important factor that influences the energy efficiency of ships. However, the knowledge on the effect of different environmental factors on ship energy efficiency is dispersed and complicated, which needs to be studied furtherly. In addition, due to the high degree of randomness of these factors, the decision of best engine speed for the optimal ship energy efficiency is subject to obvious ambiguity and uncertainty. In these regards, the feature extraction of environmental factors during a route based on statistical analysis is very important. The route division based on these statistical features can promote the achievement of speed optimization in different segments of a route. However, the amount of involved data is so huge making it exceeded the processing ability of traditional solutions. The method of big data analysis provides a new way for ship navigation optimization considering environmental factors. In this paper, based on the concept of space-time division, the distributed parallel clustering algorithm is used to divide a route into some different sections according to the environmental characteristics extracted from the obtained large number of navigation environment data of the whole route. By using the established ship energy efficiency model considering environmental factors, the optimal engine speed of different segments of the route can be obtained according to the statistical features of the environmental factors and the energy efficiency model, thus to obtain overall optimal engine speed during the whole route. It is greatly crucial to improving the level of ship energy efficiency through the speed optimization under different environmental conditions.

II. Design of Big Data Analysis Platform for Ship Energy Efficiency Management

The analysis and mining technology of big data is greatly significant to promote the optimization management of ships and the improvement of ship energy efficiency. However, It is very hard to analyze the data due to the characteristics of multi-source heterogeneous, huge scale and fast changing etc. Therefore, the design of a rational analyzing and mining platform of big data is the basis to fully explore the potential application value of big data.

A. Introduction of Hadoop-based big data analysis platform

In 2006, the project 'Hadoop' achieved two powerful open-source products: 'HDFS' and 'MapReduce'. Nowadays Hadoop has become one of the typical architectures of batch processing of big data. 'HDFS' is responsible for the storage of static data, and 'MapReduce' is used to calculate and discover the potential application value by assigning the logic calculation to each data node.

'HDFS' has the typical master-slave structure. It exposes a file system namespace which allows users to store data. A HDFS cluster includes a NameNode and several DataNodes. A NameNode is responsible for the file namespace of the management system and the access to the file by coordinating terminals. It mainly carries out the file system operation and also determines the DataNode in which the data blocks will be stored. DataNode is responsible for responding to the requests of reading and writing from terminals and achieve the function of creating, deleting and copying from the NameNode.

MapReduce programming model has been used widely in these years due to the advantages of cost-efficient performance, scalability, simplify and easy to use [14]. Its framework mainly includes 3 levels of contents: 1) Distributed file system; 2) Parallel programming models; 3) Parallel execution engines. Among them, the MapReduce parallel programming model is divided into two main stages: Map and Reduce.

B. Design of big data analysis platform for energy efficiency management

The design of big data analysis platform for energy efficiency management should make full use of the distributed computing resources and take fully the optimal allocation of physical computing resources into account, thus to achieve the most efficient parallel processing and to meet the requirement of system's adaptive scalability. The platform adopts the master-slave mode, including the physical host node, distributed computing task node, data storage node, and other relative functions, which are shown in Fig. 1.

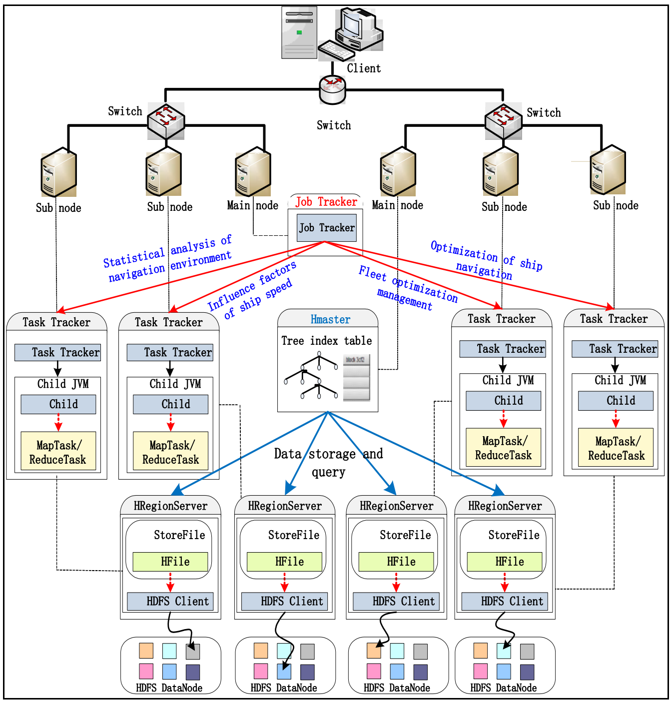

Fig. 1. Design of big data analysis platform for ship energy efficiency management

It can achieve the load balancing of computational task assignment of ship energy efficiency optimization management and the dynamic parallel capacity expansion of the data, which can improve the efficiency of data processing. The calculation tasks mainly include the statistical analysis of the navigation environment, the influence of the wind, wave, and current on the ship speed, the optimization deployment of a fleet and the navigation optimization of ships. The corresponding data storage and query tasks mainly include the data of the environment, speed influence factors, fleet optimization deployment and ship navigation optimization. Tasks can be divided and distributed according to the load balancing strategy between different nodes, and the results are finally combined to realize the distributed parallel computing model.

C. Operation principle of the system

The system is mainly divided into four functional levels, named big data aggregation layer, computing layer, optimization management and information providing layer.

The big data aggregation layer is responsible for the information collection and pretreatment. After heterogeneous conversion and distribution by the data gateway, the data will be transported to the computing layer, which is responsible for the construction of the data warehouse of ship energy efficiency management, and the achievement of data distributed processing, information retrieval and query based on the cloud architecture.

The optimization management layer is responsible for analyzing the ship optimization management scheme under different conditions, which is mainly based on the data mining and optimization decision-making models. Moreover, the information providing layer can realize the automatic release of the ship optimization management scheme according to the different needs of managers, ship owners and operators on the ship operation management.

III. APPLICATION OF CLUSTERING ALGORITHM IN THE ROUTE DIVISION

Environmental factors are main factors which influence ship energy efficiency. The statistical characteristics of these factors in different segments of a route are very different. As a result, the corresponding optimal engine speeds at different sections of the route are also different. Based on the obtained large number of data, the segments with similar environmental factors can be clustered into the same category by using the distributed parallel clustering algorithm. After which, the optimal engine speed in different segments can be obtained through the established model of ship energy efficiency considering environmental factors. Therefore, dividing the sailing route into several segments is the basis and very important for optimizing the management and control of ship energy efficiency.

Due to the large variety and huge amount of environment data, it is hard to meet the requirements of big data calculation and analysis for traditional clustering algorithm. Therefore, the application of distributed parallel clustering algorithm based on MapReduce is important to improve computing power and reduce running time. In this paper, the K-Means clustering algorithm based on MapReduce is studied and achieved. The followings focus on the clustering analysis of the environmental factors and the route division based on the K-Means algorithm.

A. Big data acquisition and preprocessing

In this paper, we take the route between Jiujiang and Nanjing on the Yangtze River as the study case. By installing the GPS receiver, shaft power meter, fuel consumption meter, wind speed and direction sensor, water depth and water speed sensor on the ship, the relative data including ship speed, shaft speed, oil consumption, wind speed, wind direction, water depth and water speed is collected. The sampling frequency is once per second, which means that about 60 thousands data sets can be obtained per voyage. Through the data acquisition of many voyages, a large number of ship navigation environment data and the corresponding energy efficiency data are obtained.

After collecting a large number of ship navigation environment and energy efficiency data, the method of data preprocessing is adopted to clean the big data, such as removing the obvious abnormal data, and interpolating the missing data. Finally, we get a lot of valid data for analysis.

B. Implementation of K-Means Based on MapReduce

The K-Means algorithm based on MapReduce mainly includes two components. One is the external loop, in which the cycle will continue until the algorithm's criterion function meets the demand. Another one is the main part of the loop and also the main calculating process. The schematic diagram of the algorithm is shown in Fig. 2.

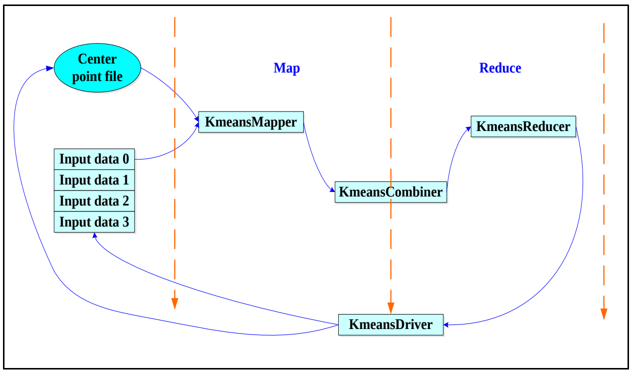

Fig. 2. Schematic diagram of the K-Means algorithm

KmeansMapper reads the input data by the setup function, and then places these inputs in the nearest cluster center according to the user-defined distance calculation method. The output ‘Key’ is the label of the category and ‘Value’ is the representation value of the category. Following which, the KmeansCombiner integrate the outputs obtained from Mapper and obtain the total output. KmeansReducer, sets a Reducer to calculate the outputs getting from the Combiner, and then integrate and output the representation value of the category with the same ‘Key’ value. KmeansDriver calculates the output through multiple cycles by continuously using the input data and the input center point, until the error reaches the threshold or it gets to the given maximum number of cycles, the algorithm will stop and output the final clustering results.

C. Division of the route based on the distributed parallel clustering algorithm

Based on the designed big data platform for ship energy efficiency management and the deployment of Hadoop cluster, the obtained navigation environment data between Jiujiang and Nanjing, including wind speed, wind direction, water speed and water depth are uploaded to the HDFS.

Due to the requirement for the sequence file of the MapReduce, we should first convert the original data to the sequence file with the type of Key is the Text, and the type of Value is the VectorWritable, through the edited conversion program of sequence file by using JAVA language,

Afterward, edit and compile the source code of the K-Means clustering algorithm based on MapReduce by JAVA language, and run the K-Means algorithm to cluster the input navigational environmental factors.

Finally, the clustering result can be obtained.

According to the clustering result, the sectors of the route with the same category of navigation environment can be divided into the same segments, thus to achieve the division of the navigation route based on the environmental factors. The obtained total five types segments can be denoted by 1-5. The five types of segments are distributed in different positions of the route. Fig. 3 shows the categories, to which the environmental characteristics in different locations belong, by the way of latitude and longitude positioning. It can be seen that the entire route can be divided into 97 intervals.

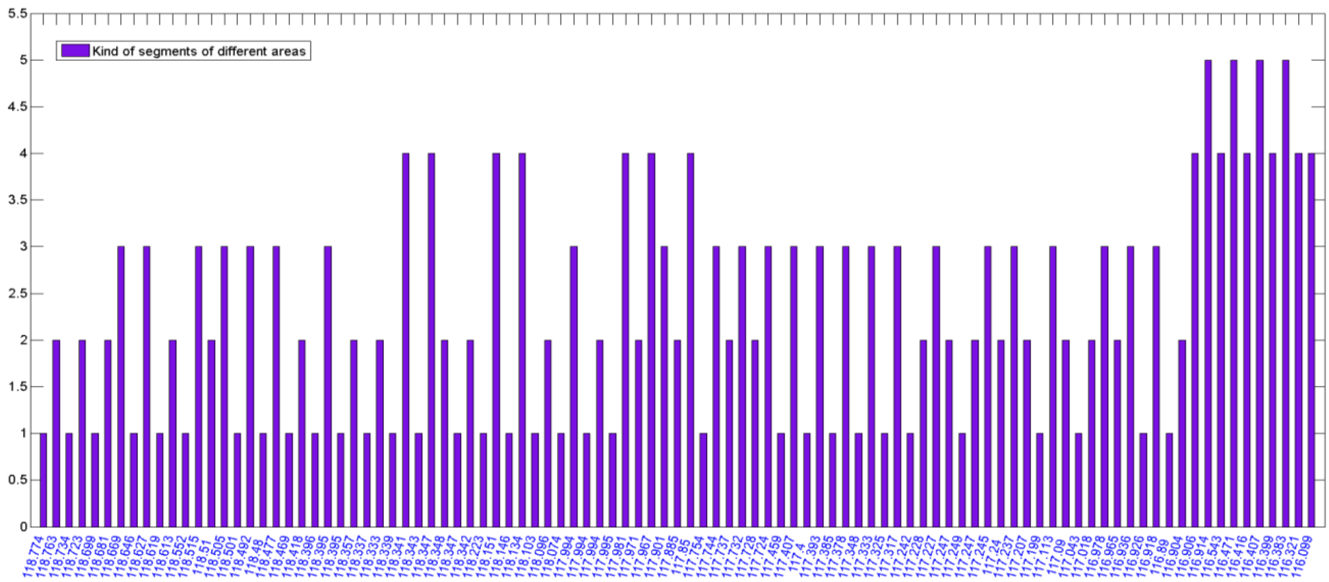

Fig. 3. Categories of navigation environment in different locations

IV. Optimization of Main Engine Speed in Different Navigational Segments

According to the characteristics of navigation environment in different segments, the optimization method can be adopted to determine the optimal speed of the main engine in each category of the navigation environment. By using the established ship energy efficiency model in [15], the fuel consumption corresponding to different engine speed in different navigational environments can be obtained, which lays the foundation for the optimization of ship main engine speed.

A. Optimization model of engine speed in different categories of segments

In this paper, the total fuel consumption of the ship during the whole route is taken as the optimization target, and the sailing time as the constraint condition, as shown in (1).

$$\begin{cases} \min \quad y = f_{oil}(x_1, x_2, x_3, x_4, x_5) \\ s.t \quad f_{time}(x_1, x_2, x_3, x_4, x_5) < T_{\text{limit}} \\ \quad\quad n_{\min} < x_i < n_{\max} \\ \quad\quad f_{oil}(x_i) > 0 \\ \quad\quad f_{speed}(x_i) > 0 \end{cases} \tag{1}$$

where, y is the total fuel consumption of the whole route (g); x_i is the engine speed of the i th category of segment (r/min) and is also the optimization object; f_{time} denotes the total sailing time for the entire route (h); $f_{\text{oil}}(x_i)$ is the fuel consumption of i th category of segment (g), and $f_{\text{speed}}(x_i)$ is the sailing speed of i th category of segment (m/s).

B. Optimization results of engine speed in different categories of segments

The best engine speed corresponding to optimal ship energy efficiency in different kinds of segments is calculated by using the non-linear optimization function 'fmincon' in the optimization toolbox of Matlab. After 14 steps, the algorithm converges to obtain the optimal solution, as showed in Fig. 4. The obtained optimization results of engine speed are shown in Fig. 5.

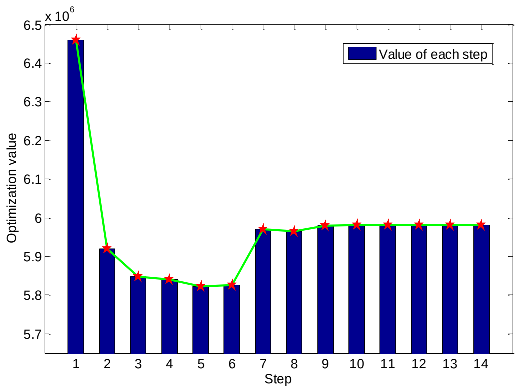

Fig. 4. Optimization value of different steps

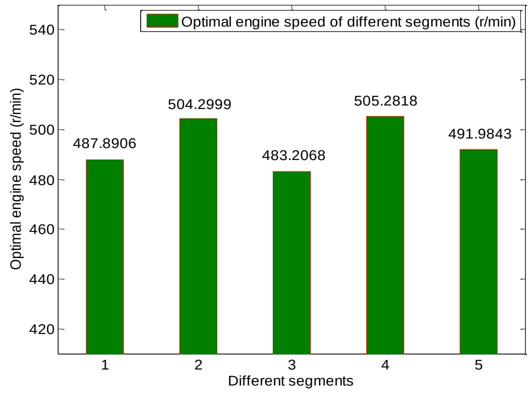

Fig. 5. Optimal engine speed of different segments

C. Analysis of ship energy efficiency optimization results

The original data of fuel consumption, main engine speed, sailing speed, shaft power and CO_2 emission in different categories of segments collected by real-ship experiment is shown in Fig. 6. Similarly, Fig. 7 shows the fuel consumption, engine speed, sailing speed, shaft power and CO_2 emission under the optimal engine speed in the same categories of segments.

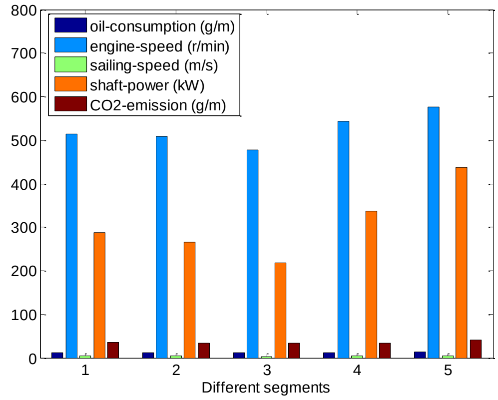

Fig. 6. Original data collected by real-ship experiment

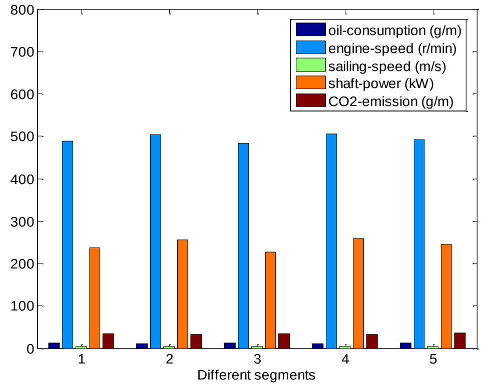

Fig. 7. Optimized data obtained by the optimization method

Comparing the results, we can see that using this optimization method can save fuel consumption by 3.09%, and also reduce CO_2 emission by about 3.09%. The total amount of the saved fuel consumption is nearly 190.68 Kg and the CO_2 emission reduction is approximately 439.71 Kg in the entire route per single voyage.

V. Conclusion

In this paper, an attempt has been made to optimize ship energy efficiency through big data analysis in the shipping industry. Based on the widely used Hadoop platform architecture, a big data analysis platform for ship energy efficiency optimization management is constructed, mainly using the Hadoop's core design of 'HDFS' and 'MapReduce'. On such a basis, the ship navigation optimization method based on the division of navigational route by the big data analysis algorithm is investigated. The results show that this method can reduce the fuel consumption of the ships effectively, thus improving the ship energy efficiency and reducing CO_2 emission.

References

[1] C. Lynch, "Big data: How do your data grow?," Nature, vol. 455, no. 7209, pp. 28-29, 2008.

[2] K. Wren, "Big data, big questions," Science, vol. 344, no. 6187, pp. 982-983, 2014.

[3] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, et al., "Big data: the next frontier for innovation, competition, and productivity," Analytics, 2011.

[4] S. Yang, K. Guo, j. Li, Y. Zhong, R. Liu, Z. Feng, "Framework formation of financial data classification standard in the era of the big data," Procedia Computer Science, vol. 30, pp. 88-96, 2014.

[5] M. Rathore, A. Paul, A. Ahmad, S. Rho, "Urban planning and building smart cities based on the Internet of Things using Big Data analytics," Computer Networks, vol. 101, pp. 63-80, 2016.

[6] K. Kambatla, G. Kollias, V. Kumar, A. Grama, "Trends in big data analytics," Journal of Parallel and Distributed Computing, vol. 74, no. 7, pp. 2561-2573, 2014.

[7] X. Jin, W. Benjamin, X. Cheng, Y. Wang, "Significance and challenges of big data research," Big Data Research, vol. 2, no. 2, pp. 59-64, 2015.

[8] D. Huang, D. Zhao, L. Wei, Z. Wang, Y. Du, "Modeling and analysis in marine big data: advances and challenges," Mathematical Problems in Engineering, vol. 2015, no. 2, 2015.

[9] J. J. Corbett, H. Wang, J. J. Winebrake, "The effectiveness and costs of speed reductions on international shipping," Transportation Research Part D: Transport and Environment, vol. 14, pp. 593-598, 2009.

[10] K. Wang, X. Yan, Y. Yuan, "Fuzzy logic method for ship energy efficiency decision-making model to determine optimal engine speed," Sea Technology, vol. 11, pp. 45-46, 2015.

[11] R. Vettor, C. G. Soares, "Development of a ship weather routing system," Ocean Engineering, vol. 123, pp. 1-14, 2016.

[12] Marine Environment Protection Committee (MEPC). Prevention of air pollution from ships (Second IMO GHG Study 2009), 2009, London.

[13] E. M. Bitner-Gregersen, C. G. Soares, M. Vantorre, "Adverse Weather Conditions for Ship Manoeuvrability," Transportation Research Procedia, vol. 14, pp. 1631-1640, 2016.

[14] M. Vaidya, "MapReduce: a flexible data processing tool," Communications of the Acm, vol. 53, no. 1, pp. 72-77, 2010.

[15] K. Wang, X. Yan, Y. Yuan, F. Li, "Real-time optimization of ship energy efficiency based on the prediction technology of working condition," Transportation Research Part D: Transport and Environment, vol. 46, pp. 81-93, 2016.