Analytics of Big Data in Maritime Industry

Rahul. D[1]*
[1]Indian Maritime University, Mumbai Port Campus, Marine Engineering and Research Institute, Mumbai.
*rahuldurai1998@gmail.com

Abstract
Shipping industry handles a vast amount of data. Most of it is not stored or secured and get lost over a period of time .But the need for those data is crucial in the time of any incident. The ships in future will be in need of Big Data Analytics for efficient condition monitoring, auto-piloting, freight tracking, ship building etc. The advancement of Big Data will enable the ship to communicate with another through intelligent condition monitoring systems and engines. Thus Big Data Analytics enhances the safety as well as the efficiency of the maritime industry. The classification society and ship-builders data must be stored periodically for future references and advancements where Big Data Analytics plays a major role.

Keywords: Condition-monitoring, Data Analytics, Performance-monitoring, Cyber-threat, IT

1. Introduction

Big data has the potential to transform our industry. Through application and insights, big data is creating new opportunities to drive innovation and deliver tangible operational efficiencies across the shipping world. But information alone is not enough. It is the analysis of this data and the actionable insights it provides that will move our industry forward and determine our future. Shipping is doing brilliantly at taking advantage of new technology. Maritime, not so much. Retailers are already disrupting shipping by taking greater control over logistics: they are moving from last mile delivery to first mile – and the first mile is supposed to be the remit of maritime.

This objective research brings together all the latest news and thinking to provide a comprehensive overview of big data in the marine sector. It includes a look at how and where these technologies are being implemented and the key application areas that will deliver future operational efficiencies for ports and terminals. The study also explores some of the key barriers to change, such as investment and skills, and how to overcome these challenges. Overall what does emerge from this benchmarking study is that while many shipping and logistics players recognize the importance of transformative technology, the maritime industry lags behind other sectors of the global economy when it comes to adopting it. In the context of shipping, the economic outlook remains tough, so finding new ways to optimize performance and remain competitive is a priority. Big data needs to be part of the solution.

2. What is Big Data?

The term Big Data is used to describe large and complex data sets that are difficult to process and analyze using traditional data processing techniques and applications. It refers to the collection and subsequent analysis of any significant large collection of unstructured data that may contain hidden insights .Big data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or does not fit the structures of existing database architectures. To gain value from this

data, there must be an alternative way to process it". The growth of big data increases daily and there is no limit in sight. The annual data generation is predicted to increase by 4300% by 2020. Big data analytics is not just about gathering data but also covers the process of analysis to discover hidden insights, trends and correlations. That information has the potential to bring a competitive advantage to any industry.

2.1. Big Data Characteristics

Big data covers information from various sources i.e. sensors. There are many difficulties associated with capturing, sorting, analyzing and managing data. Fig.1 illustrates the features of big data according to IBM. Big data has four main characteristics and referred to as the 4V's (Volume, Variety, Velocity and Veracity) .

- Volume refers to the massive quantity of the data. Nowadays, sensors produce a massive amount of data in terabytes, petabytes and beyond.
- Variety refers to the form of the data. In big data, the datasets are stored across in multiple formats. Data variation differentiates big data from traditional data.
- Velocity defines the speed of the data creation and movement. The data is created at different rates and must be stored for processing. Generally, a huge amount of data is created in real-time and data flow rates are themselves increasing rapidly.
- Veracity refers to the data accuracy and trustworthiness. Datasets from different sources may use different scales to measure the same variable and this raises issues of how to maintain data quality. Veracity needs to be addressed and maintained throughout the data lifecycle.

3. Challenges in Big Data Analytics

3.1. Cyber threats

Marine IT and telecommunication infrastructures are at high risk of penetration from cyber criminals, terrorists or other malevolent interests. According to CyberKeel, a Danish cyber security firm, more than 90% of the largest container lines are vulnerable to hackers. ESC Global Security's head of cyber security division, Joseph Carson, also mentioned that big data will increase the vulnerability to cyber-crime in the maritime industry. There is the potential for a major cyber-attack on the maritime industry, which may lead to disruption in food and energy supplies, as shipping transports 90% of the world's total trade. ClassNK has been taking a leading role in the implementation of cyber security standards for maritime data analysis. It has collaborated with the United States Maritime Resource Center to improve its understanding of existing and emerging cyber security challenges due to big data across the marine transportation system.

3.2. Misreporting of data

The following are highlights of a report by Windward, published in 2014: 1% of all ships broadcast fake IDs, i.e. a ship claiming to be a completely different vessel Less than half of all vessels report their next port of call accurately 55% of ships misreport their actual port of call throughout their journey Misreporting of data can cause concerns, as it may lead to incorrect analysis and inappropriate decision-making. "Although the marine sector recognizes the advantages of big data in terms of value to business, human element, environmental protection, offshore activities and so on, it hasn't yet found ways to overcome the challenges of data ownership and sharing." Maria Kouboura Senior Technical Advisor, IMarEST.

3.3. Slowdown in investment in big data analytics due to existing challenges

The shipping industry has been facing numerous disturbances and challenges such as market fluctuations, over supply, margin pressures and labor shortages. These challenges are expected to impact the industry's profitability. According to a report published by IHS in November 2015, in the coming 5–10 years, the industry will experience slow growth. In such an unstable environment, players are uncertain about implementing big data as it is a relatively new technology in the industry. As a result, investments by these companies in big data analytics technologies is decreasing.

3.4. Lack of cross-enterprise technology implementation

Currently, ship builders, ship owners and ports are solely focusing on running reasonably efficient operations and not on running a highly flexible, responsive trading business of 'container-as-a-community'. This means that there is a lack of cross-enterprise processes. Companies are concentrating on automating processes within functional silos instead of taking a holistic view of the enterprise. This prevents the true potential of big data from being realized.

3.5. Lack of big data-skilled workforce / Skills shortage

Ensuring enough quantity and quality of human resources is essential for developing the use of big data solutions for maritime. There is a shortage of highly trained data scientists. This shortage is expected to further increase in the future. According to McKinsey, by 2018, there will be more than 490,000 data science jobs in the US alone. However, there will only be 200,000 available data scientists to fill those positions. Demand for data scientists is expected to exceed supply by more than 50% by 2018.

3.6. Other Challenges

• Data Reliability depends on collection methods and definition of measurements. Data reliability depends on how the raw data was collected. The strength of analysis is related to data quality and reliability.

• Data Availability relates to on-time accessibility. Data needs to be available at any time for analysis. If sensors provide data for a measurement at a particular sample rate then the raw data would be accessible for a specific period of time. Archive or historical data should be available for collection and analysis.

• Linking Data refers to how to connect the data together. Usually, the data is collected for a specific task or purpose and might require different types of data from various datasets. If a single user accesses multiple datasets then a primary key is needed to connect the data together.

• Data Presentation is related to how the data may be viewed to facilitate meaningful interpretation. Poor data presentation could obstruct data analysis and reduce the value of the original data. Data presentation for different datasets has become important for efficient operations.

• Dataset Scalability defines the capability of a system to support any type of dataset. All datasets have their own structure, type, semantics and accessibility. The analytical

algorithms applied to big data must have the capability to support increasingly expanding and complex datasets.

• Data Compression refers to the filtering of data to reduce the volume. The data generated from sensor networks generally contains a high level of redundancy. That data must be filtered and compressed at orders of magnitude to reduce redundancy.

• Data Life Cycle Management relates to how frequently and which data will be stored or discarded. Generally, the value of big data depends on data freshness. Sensor network systems can generate data at unprecedented rates and scales, so processing and storage become key issues for such massive datasets.

• Data Confidentiality refers to safety and protection of data. Data may be analyzed or shared with third parties increasing the potential safety risk. The data needs to be protected from any risks.

4. Common Questions and Solutions:

Is This Just An Expensive IT Solution?

No – it's not about IT at all; remember the term “Garbage In = Garbage Out?" That's too-much data, not Big Data. A company needs to determine what they want to do...reduce costs? Increase revenue? Optimize transit times? These are all serious strategic issues for both a carrier and a shipper, whose implementation – or not – affects profitability on every voyage. The carrier needs to sit with a company like ours in order to figure out the questions in shipping analytics that need to be asked in order to affect positive outcomes, and determine what is the traditional and non-traditional data that goes into profitability.

What Is Traditional Data Vs. Non-Traditional Data?

Traditional data is look-back data from ordinary business systems; fuel costs, transit times, wages, insurance, revenue per TEU, TEU rates. It's accounting data that's used to determine the profitability of a voyage. But non-traditional data is time-sensitive data – weather and traffic delays, port strikes, unexpected repairs. It’s also extremely large volumes of data being generated from sensors, GPS devices, RFID tags and traffic management systems. If Big Data can assist in helping forecast or avoid problems, the money saved goes straight to profitability.

Would This Sort Of Information Be Equally Important For The Shipper?

Absolutely; whether you’re shipping the goods as a vendor, or receiving them as the buyer; getting them to the final destination on time, undamaged, and in a cost-effective manner is important. Whether the data comes from the carrier, or is transmitted via RFID; it's data the shipper can use to quantify his shipping metrics and work on improvement.

Big Data and the Ocean Freight Market

When properly used, Big Data will enable management to take a gut-feel decisions and quantify it, which lets them take that good idea and make it work. For ocean freight shipping, it's more than just 'look-back information', of say how many TEUs you shipped in May from Shanghai to Rotterdam, but rather pulling the data necessary to benchmark and ship them at a more competitive price, faster, or to achieve whatever metric you need addressed.

There’s always more to be said about Big Data and among what should be emphasized is the fact that it includes both internal and external data. Combining both of these types of data, analyzing and deriving solutions based on your company’s strategy is the goal of Big Data. What data to include is also a consideration. One school of thought is to gather all data, internal and external and then see what the computer spits out as a result. Another thought is to identify beforehand what needs to be analyzed such as buyer and seller habits, customer satisfaction or lack of, market trends across all markets and industries, fuel costs, rates, material costs and so on. Like everything else in the supply chain market, no one size fits all so it will depend on your company’s specific needs and goals.

The software allows users to measure their own freight rates and their suppliers' versus market movements for various port to port routes, allowing them to see averages over time, gains and losses, and percentage movements.

Despite what may seem to be an improving ocean freight market, the need is great to monitor it almost on a daily basis. Congestion at ports, bankruptcies, acquisitions, changes within alliances and potentially new ones can all impact your ocean freight invoice.

How about insight into the market six months from now or just next month?

Almost impossible without Big Data. While it may not give the complete forecast, it can certainly be used to mitigate potential risks.

Is There A Place For Big Data In Supply Chain Management?

Absolutely; in fact it's what makes advanced SCM possible. It's using Big Data to start replacement inventory moving before inventory is exhausted; it's taking look-back vendor reliability data and instead of just disqualifying those who missed shipments, rewarding those with good reliability as well as others who stepped up to cover for those vendors who failed. It's bringing information together from multiple sources to make predictions and real-time decisions to keep the logistics pipeline flowing. Remember, Big Data is used to quantify what you, the customer, has determined is important to your business.

Final Thoughts on the Use of Big Data Shipping Analytics For Shippers And Carriers

For the past few years most industries have been leveraging the power of data or "big data" to increase their pace of innovation and efficiency. The shipping and logistics industry, though usually a bit slower to pick up on digitization, is now realizing the importance data and technology has on making it relevant, competitive and susceptible to change. While the sheer volume of data available to a company today is daunting, the first step in leveraging 'big data' is deciding on the problem you want solved, or the issues you need addressed, and then finding the proper data necessary.

5. Application in the Maritime Industry

A future data-driven maritime might include:

• **Remote Sensing:** Ships will be monitored continuously from remote locations. Data will be collected autonomously by using remote sensor networks. A robust wireless network with high transmission capabilities would be required for the shipping industry. The real-time sensor data will come to the database and be distributed to the interested parties giving them up-to-date information on what is happening onboard.

• **Voyage Planning:** Ship operators or charters will be able to implement voyage planning after analyzing the route, vessel performance and meteorological data. The voyage planning will be based on the vessel performance on the same and different routes. A reliable forecast of wind and ocean current data will be required for voyage planning. Data analytics will help to identify the most efficient route for the journey, accurately estimated arrival time and alternative routes can be planned to avoid any delay or disturbance.

• **Intelligent Traffic Management:** Port authorities will have access to the ship data for safety. Intelligent traffic management systems will be introduced as data-oriented applications in the shipping industry. Ship current position, cargo and personnel information will be transferred to the port so that the port authorities would be able to monitor the congestion and improve cargo handling performance.

• **Operational Predictability:** Vessel operational performance can be monitored in real-time by analyzing ship data. Ship operators will gain the capability to predict the vessel performance based on current operational conditions. This predictability will assist in making decisions on maintenance.

• **Energy Management:** Shipping is moving towards flexible and alternative energy systems. Energy production, storage and re-use will become part of the energy management system; in fact, battery operated vessels have already been introduced in the industry. The ship energy management system will be run on the basis of real-time data of

load requirements and power availability from all sources. The system will distribute and balance ship and shore power.

• **Environmental legislation monitoring:** Most emissions are related to fuel type. Ship operators need to comply with environmental legislations including the requirement to switch fuel in emission controlled areas i.e. the Sulphur content of the fuel should not be more than 0.1% on and after 1 January 2015. New data-oriented regulations are coming in to force in the EU to monitor CO2 emissions and the system will give an indication of fuel switching as well as monitoring the current emissions. The ship operators will also be able to set the KPI for the vessel after analyzing the emissions data.

• **Performance Monitoring and Optimization:** Automation expands the capability of the control of machinery and vessel optimization. A range of data measurements are required for monitoring vessel performance and optimizing performance. Access to historical data is essential for optimizing and forecasting.

• **Vessel Safety and Security:** The use of wireless sensors and extensive satellite communication systems will increase vessel safety. Sensor data analytics will provide information on vessel maneuvering to avoid collision. The inbuilt sensors on machinery will provide real-time information about their current condition; information that will be useful for crew safety. Vessel safety and security will be increased by the adoption of innovative technology.

• **Condition Monitoring:** This will be improved by analyzing asset data. This will be applicable to machinery such as engines, pumps, boilers and compressors. Sensors will help to monitor the machinery and give early warning of the need for maintenance. It will determine the condition of machinery and record data during operation.

• **Predictive Maintenance System:** This will detect the need for maintenance to avoid potential failure. The potential failures would be detectable and measurable. The system will record all data and indicate risk of failure by synchronizing related data such as engine data, fuel consumption, running hours etc. It will reduce the cost of asset failures and minimize unscheduled downtime. The system relies on machine health monitoring to dictate when and what maintenance would be required before failure occurs so that the crew would not be required to spend additional time on maintenance and scheduling activities.

• **Automatic Mode Detection System:** This system detects the vessel's operational mode automatically based on sensor data (flow meter, GPS) . Using the auto-mode detection system, the crew would not be required to update the mode every time the ship changed operational state. The vessel operational profile will be created by analysis of real-time data of fuel consumption, distance travelled and speed. The auto-mode detection system will function without human intervention and will provide a summary report of fuel consumption for the individual engines, ship running hours and emissions in different modes. The onboard and onshore members of staff would be able to use this information to measure the vessel operational performance and KPI. This system will assist ship operators in meeting the EU MRV regulation by monitoring the fuel consumption and emissions for different vessel modes.

6. Conclusions

This paper illustrates how big data analytics have the potential to create a significant impact in the shipping industry. It also covers the associated applications and obstacles to the implementation of big data in the industry. The shipping industry will be faced with ever-increasing requirements for safety, environmental and efficiency performance beyond 2020 . New emerging technologies will be applied for efficient and safety operation. More data-oriented regulations are expected to come into force in the near future. So it will be an opportunity rather than a threat as this regulation will help the shipping industry to move into the big data era.

7. Acknowledgement

The author would like to acknowledge the support of Marine Engineering and Research Institute, Mumbai.

8. References

[1]WMU Journal of Maritime Affairs.

[2]Journal of Marine Science and Technology.

[3] European Commission, Integrating maritime transport emissions in the EU's greenhouse gas reduction policies, Brussels, 2013.

[4] International Maritime Organization, Third IMO GHG study 2014, London, 2015.

[5] DNV-GL, Maritime EU-MRV Regulation, Hamburg, 2015.

[6] Oxford Economics, The economic value of the EU shipping industry (2014) 6.

[7] A.E. Hassanien, A.T. Azar, V. Snasel, J. Kacprzyk, J. H. Abawajy, Big Data in Complex Systems-Challenges and Opportunities, Springer, London, 9 (2015).

[8] M. Bulger, G. Taylor, R. Schroeder, Data-Driven Business Models: Challenges and Opportunities of Big Data, University of Oxford (2014)19- 50.

[9] S. Mao, Y. Zhang,V. M. Leung, Big Data Related Technologies-Challenges and Future Prospects, Springer, London, 2014.

[10] IBM, The Four V's of Big Data, 2016.