PREDICTION OF DELAYS IN SUPPLY LOGISTICS OF OFFSHORE PLATFORMS

Fun-sang Cepeda, Maricruz A. [1]; da Silva, Rafael Basilio[2]; Caprace, Jean-David[3]

[1]University Federal of Rio de Janeiro, Rio de Janeiro/Brazil.maricruzcepeda@oceanica.ufrj.br
[2]University Federal of Rio de Janeiro, Rio de Janeiro/Brazil. rafaelbasilio@poli.ufrj.br
[3]University Federal of Rio de Janeiro, Rio de Janeiro/Brazil. jdcaprace@oceanica.ufrj.br

ABSTRACT.

Nowadays, the oil offshore industry is making every effort to improve the logistic in oil production. Supply of offshore platforms require a high level of service using minimal logistic resources. Weather conditions and vessel off-hire are the main variables that induces delays in operations planning. These delays postpone or disturb the oil production in offshore platforms and FPSOs. Actual cargo logistic maritime fulfillment indicator shows that there is a huge potential for improvement. The objective of this paper is to forecast the delays that are occurring during the delivery of the supplies for offshore platforms operating in Brazil. In order to achieve this goal, a database of 2 851 offshore supply vessels travels have been processed with several data mining tools such as apriori, decision tree and multi-layer perceptron. The type of cargo (dry bulk, liquid bulk or container), the cargo priority, and type of operation (load, backload or transshipment) are some of the input parameters of the models developed in this study. The findings provide a new way to address efficiency and performance supply logistics of offshore platforms even if some future model improvements are required. Knowing in advance were the delays are more susceptible to occurs in the supply chain allows the planners to anticipate their strategies and delivery routes.

Keywords: Offshore, Supply Chain, Logistic, Offshore Supply Vessel, Offshore Platform

1. INTRODUCTION.

Logistics plays a fundamental role in the petroleum offshore industry as the distance between offshore units is constantly increasing. The offshore logistics uses huge infrastructure to maintain and develop operations of offshore units, composed by airports, ports, hubs, warehouses, specialized vessels, among other resources.

Supply logistics between a land depot and the offshore installations can be a complex logistics problem to solve for optimality (Nordbø, 2013).

One of the greatest challenges for offshore operations is getting a suitable level of operations planning to use a minimum amount of resources but with high service level. Weather conditions and vessel off-hire are the main variables, which affects the operations planning. Cargo transportation to and from offshore units is one of the main offshore logistics activities, which include personnel transportation, storage, inland transportation and cargo handling, towing and mooring rigs, among many others. The main types of cargo to be transported are food, industrial water, diesel, fluids, dry bulks, pipes, riser and general cargo. The type of vessel which is widely used to transport cargo is the Platform Supply Vessel (PSV). PSV is a vessel specifically designed to supply oil platforms in deep seas (Leite, 2012).

The present paper analyzed a database, which contains a list of cargo supplied to offshore units in the period from January 2014 to July 2014 in Campos Basin in Brazil. Association rules (A-priori), decision trees and multilayer perceptron neural network have been implemented to predict the delays of the logistic supply chain.

2. BUSINESS UNDERSTANDING.

A Maritime Cargo Fulfillment Indicator (MCFI) have been developed in this study. This indicator measures the percentage of cargo supplied before deadline in relation to total amount of cargo supplied to offshore units for a period of one month. For logistics operations, the major challenge is to keep the indicator above 76%. Various issues affect this indicator such as weather conditions, vessels off-hire (near 10%), deck area and number of berths limits, unexpected failures, multiple goods, and priority of some cargoes.

Environmental conditions affect widely operations performance, as PSVs propellers have not enough power to keep its position within the specified range and offshore cranes does not able to withstand strong gusts of wind. Another alternative is to use Anchor Handling Tug Supply

(AHTS) vessels to supply cargoes to offshore units. This type of vessel has a higher propulsion capacity that allows operating in weather with adverse conditions. However, AHTS vessels have a smaller cargo capacity and are more expensive than PSVs. Therefore, there is a great challenge to optimize the offshore logistics operations.

The present paper is focusing the analysis in the logistic of Campos Basin in Brazil. Distance to offshore units is one of the main factor that is affecting the logistic operational performance as Brazil oil production is concentrated more and more in deep seas, Figure 1

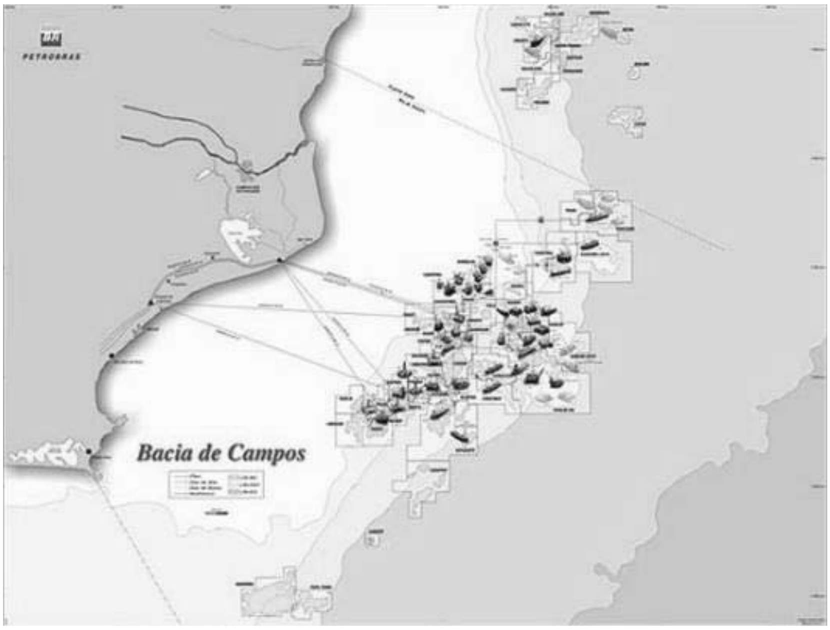

Figure 1 - Offshore Units in Campos Basin, (Ferreira Filho, 2014)

Vessel fleet dedicated to perform offshore operations is composed by different dedicated type of vessels (PSV, AHTS, etc.). This is not efficient, because the vessels are rarely fully loaded (either the deck is not used, or the tanks underneath the deck are not used) and represents a costly solution. In general, logistics system of offshore support is composed by a supply flow which begins with national and foreign suppliers, passes through warehouses and ports and ends with the supply to offshore units as shown in Figure 3, **(Ferreira Filho, 2014)**. The logistics process takes place according to the workflow shows in Figure 2. Materials for the units can be stored either in own warehouses or supplier warehouses. Both warehouses must provide materials via land transportation performed by trucks to the supplier bases. Each request is turned into transport requests.

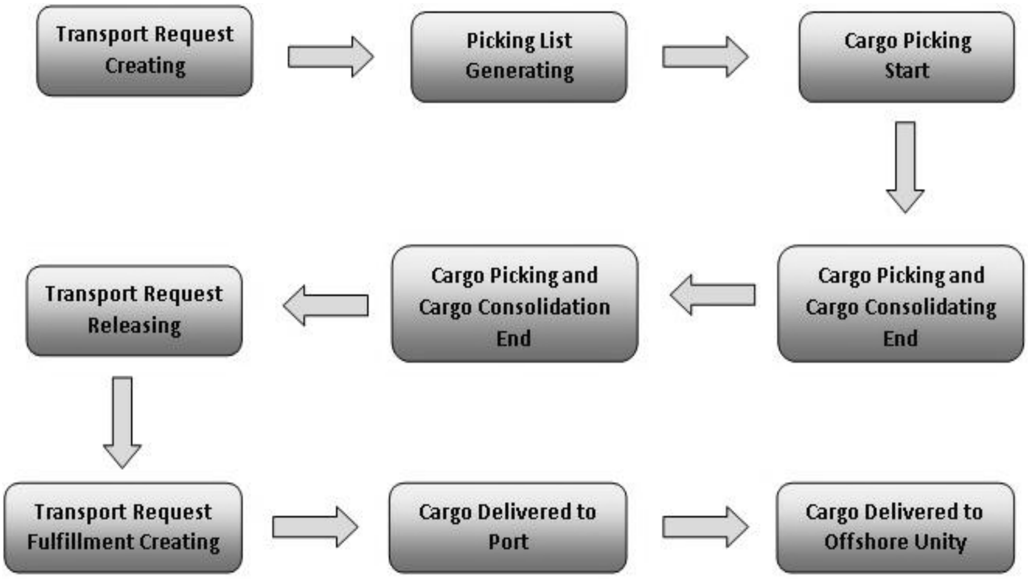

Figure 2 - Logistics Process Workflow

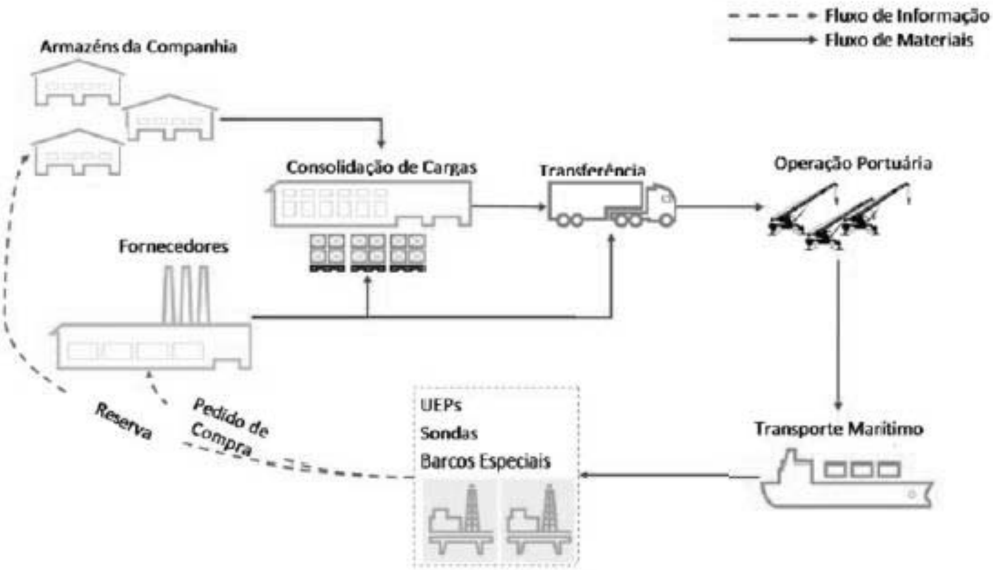

Figure 3 - Typical Offshore Logistics System, (Ferreira Filho, 2014)

After analyze and consolidation of the data, the requests orders are used to plan the available supplier vessels. Cargo weight and dimensions are important information in scheduling process as they limit the amount of cargo to be loaded to deck or underneath the deck into tanks. Furthermore, it is very important to taking into account the delivery deadline. The earliest date and latest date are associated to each request. The service quality will be affected whether the cargo is delivered respectively before or after the earliest date or latest date. **(Ferreira Filho, 2014)**.

The principal objective of this paper is to analyze the request order database and recovers hidden information that can be found through data mining. Three different algorithm have been applied. First, association rules has been detected among available transactions on the database such as cargo weight versus delay status, cargo weight versus type of operations, etc.; Then, decision trees and multilayer perceptron neural network models were used to carry out predictions on cargo delays.

3. METHODOLOGY.

Data Mining (DM) is the analysis of observational data sets to find unsuspected relationships and to summarize the data in ways understandable and useful to the data owner **(Hand, Mannila, & Smyth, 2001)**.

When carrying out a DM project, it recognizes that activities outside the main process are consuming more time and have great influence on the ultimate success of the project. Therefore, many publications on DM discuss the construction or application of algorithms to extract knowledge from these data. DM process generally emphasizes at the stage of data analysis and uses different algorithms to predict information about entire project. Finally, DM is a multi-disciplinary field that is at the intersection of statistics, machine learning, database management, and data visualization **(Feelders, Danielsa, & Holsheimer, 2000)**.

In this paper, we use Apriori association rules, Multilayer Perceptron Neural Network, and Decision Trees algorithms to analyze a database of offshore supply vessels to predict delays in supply logistics of offshore platforms.

4. DATA UNDERSTANDING, DATA EXPLORATION, AND DATA PREPARATION.

The analyzed database consists of a list of cargoes orders, which are supposed to be delivered to offshore units by PSVs. Each one of 239 921 rows of the database represent a transport request item and have key information about cargo which will be supplied to offshore units presented in Table 1.

ITEM	DESCRIPTION
Transport ID	Fulfillment ID number of a transport request performed by a vessel in a trip. Each one has many supplied cargoes, which are identified by a number of transport request
TR Type	Type of a transport request, R1 means the own company is the proprietary of the cargo. R2 means the cargo belongs to a supplier
TR Number	Is the quantity of a transport request
TR Items	Quantity of a transport request item. As an example, many pumps are loaded in a box inside a container, in this case, each box is an item and the transportation requisition (TR)
Priority	Classification of priority of cargo. Normal priority (N) means the cargo will be supplied to unit according a supplying schedule. Emergency priority (E) means that cargo will be supplied as soon as possible from the moment they are requested.
Source	Location where the vessel is departing. It can be a port or offshore unit (OU)
Destination	Location where the vessel is arriving. It can be a port or OU
Description	Vessel name which supplies cargo to OU
Creation date	Date when transport request is created
Later date	Date-line when cargo is supposed to be delivered to OU by the vessel
Earliest date	Earliest date from which cargo should be delivered to OU by the vessel. Represents the OU availability to receiving the vessel
Release date scheduled	Date when transport request is released to programming
Expected date	Date scheduled to initiate supplying by the vessel
Delivery date	Date when cargo is delivered to OU
Total weight	Total weight of loaded cargo in each transport request
Cargo Classification Type	Cargo class, it can be General Cargo(GC), Liquid bulk Cargo(LC), or Dry Bulk Cargo (DC)
Cargo Sub-Classification Type	Cargo subclass (tubes, cladding, risers, etc.)
Cargo Type	Type of cargo (toxic/infectious substance, corrosive, common, etc.)
Type of shipment	Type of operation (load, backload or transshipment). "Load" is a travel where destination is an OU and departure is a port. Backload is a travel where destination is a port and departure is an OU. Transshipment is a travel where destination and departure are OU.

Table 1 - Information about loads database

Figure 4 shows the offshore logistics operations performance measured by Maritime Cargo Fulfillment Indicator (MCFI). The figure cover the period from January 2014 to July 2014.

This indicator shows the fulfillment of the delivery time showing that the goal was not reached.

Days of delay are the difference between delivery date and deadline. Frequency of delays is plotted in Figure 5. The higher frequency is presented between 1 to 3 days. A classification of the delays has been defined as follow: 0 to 1 days is [Much Lower], 2 to 3 days is [Lower], 4 to 5 days is [About the same], 6 to 7 days is [Higher], and more than 7 days is [Much Higher]. Absolute frequencies distributions of the delays status were plotted in Figure 6.

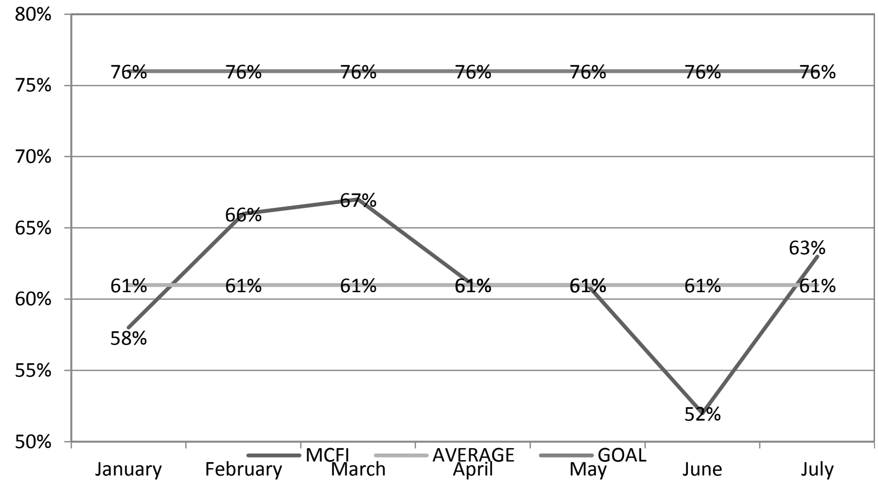

Figure 4 - Logistic Operations Performance

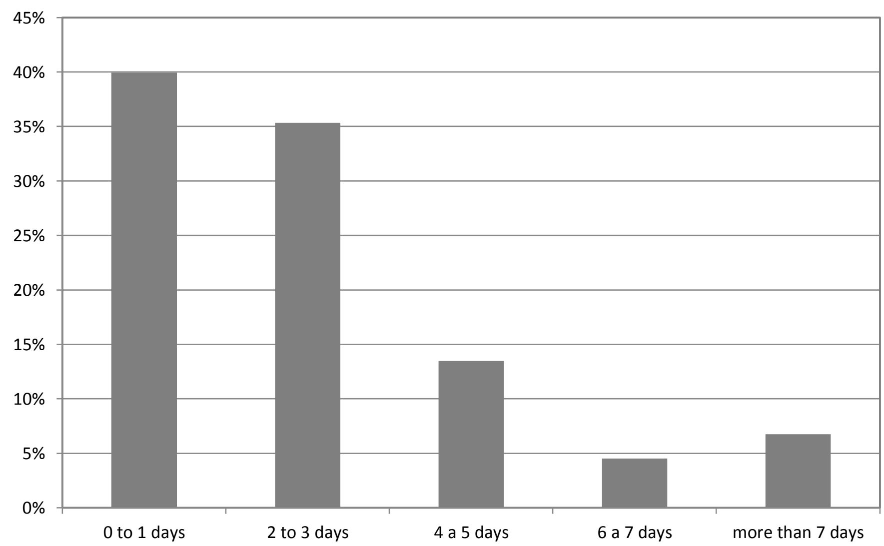

Figure 5 - Relative Frequency of Delays

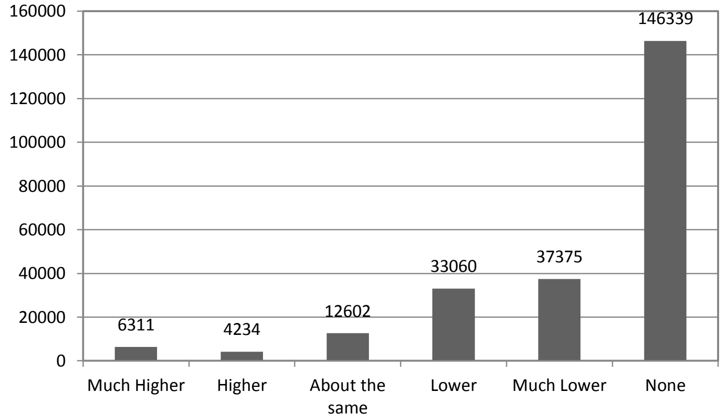

Figure 6 - Absolute Frequencies Distribution of the Delays Status

Extreme outliers' values were excluded using Equation 1 and 2. Outliers and missing values represent 24 116 records on 239 921 which means 10% of the database.

$$p < Q_1 - (3 \cdot IQR) \quad \textbf{(1)}$$

$$p < Q_3 - (3 \cdot IQR) \quad \textbf{(2)}$$

Where p data points,

Q_1 the lower quartile,

Q_3 the upper quartile,

IQR the distance between Q_1 and Q_3.

In addition to the fact that logistics service has not performed above the desired goal (MCFI > 76%), the percentage of cargoes delayed plus one day is around 40% of the total amount of delays, which can be considered very high. In Figure 7 can be seen the proportion of normal and emergency priorities over the analyzed period. Figure 8 shows the proportion of each type of operations (Load, Backload and Transshipment) over the analyzed period.

As shown, cargoes movements are considered one of the most important operations, because it could affect oil production. Indeed, some cargoes have a huge impact on the development of the production and the habitability (foods and water) of the offshore units. We can also conclude that the amount of transport requests items delivered after deadline is higher for load than backload and transshipment operations. The proportion of delayed items is respectively 48%, 18% and 23% for load, backload and transshipment operations.

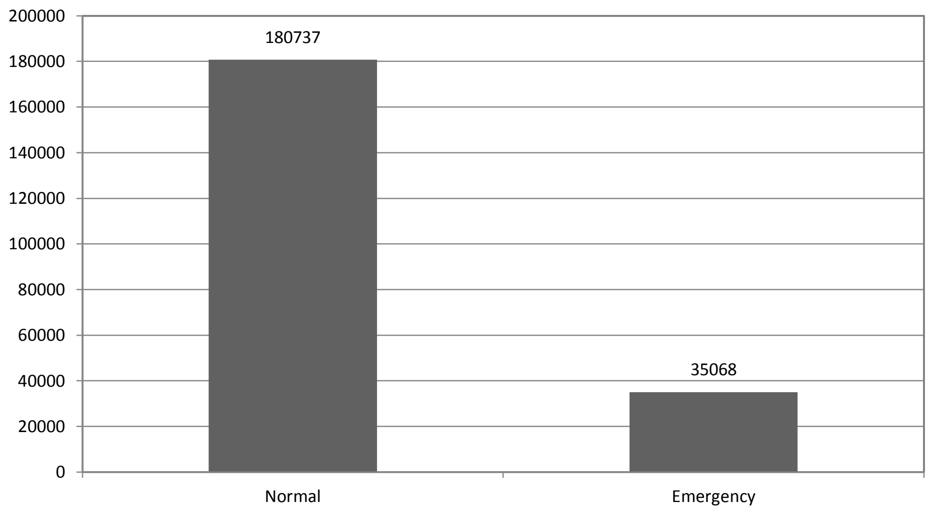

Figure 7 - Proportion of Normal and Emergency Priorities

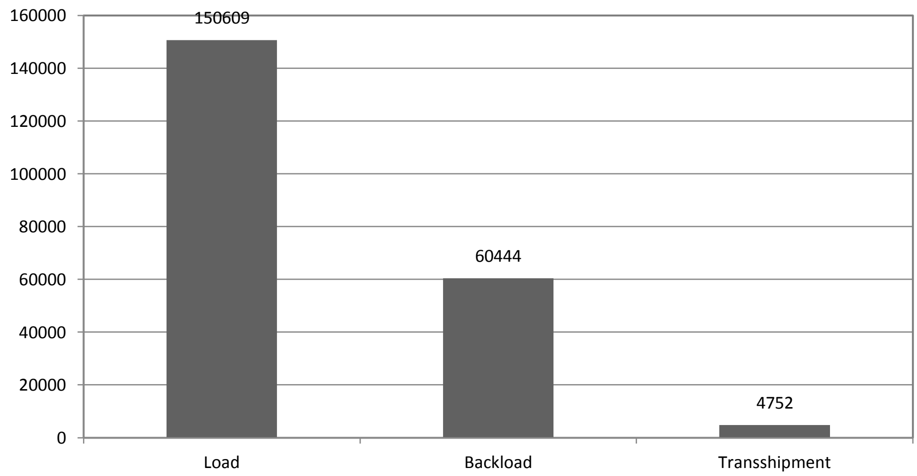

Figure 8 - Proportion of Load, Backload and Transshipment Operations

Figure 9 shows the load frequency per weekday over the covered period. It can be seen that the distribution of delays is almost the same over the week. In Figure 10, we can see a histogram of the status of each trip performed over the covered period. This indicates whether cargoes were delivered before the deadline (within of time) or not (out of time).

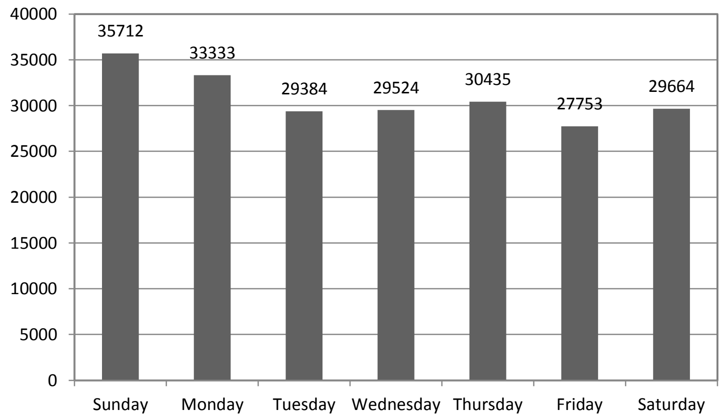

Figure 9 - Proportion of Weekdays for Cargo Delivery

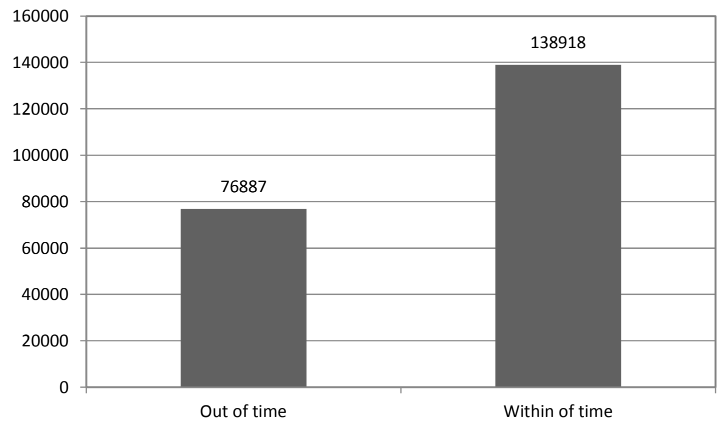

Figure 10 - Distribution of Term Status

5. RESULTS AND KNOWLEDGE DISCOVERY.

In this section, the models and results are described and discussed. Modeling step has been divided in three steps explained in this section. First, association rules inference algorithm (apriori) has been applied on the database to find possible relations between transactions such as cargo weight versus delay status, cargo weight versus type of operations, etc. Then, multilayer perceptron neural network models were used to carry out predictions on cargo delays starting from voyages features such as weekdays, origin, destination, cargo class and type of operations. Finally, decision tree model has been used to undertake predictions about delay status and type of operations (load, backload and transshipment).

In order to feeds the different models, the load order request database (239 921 rows) presented in Table 1 have been gathered in 2 851 voyages as shown in Table 2.

ITEM	DESCRIPTION FOR EACH TRIP
Quantity of ID	It counts the number of transport request items
Summary of Cargo weights	It sums the cargo weights of transport request items
Average Delay days	It averages days delays of transport request items
Normal priority	it sums the amount of transport request items with normal priority
Emergency priority	it sums the amount of transport request items with emergency priority
General Cargo	it sums the amount of general cargo transport
Liquid bulk Cargo	it sums the amount of liquid bulk cargo transport
Dry Bulk Cargo	it sums the amount of dry bulk cargo transport
Total of Transshipments Operations	it sums the amount of transport request items supplied by transshipment operations
Total of Load Operations	it sums the amount of transport request items supplied by load operations
Total of Backload Operations	it sums the amount of transport request items supplied by backload operations
Out of time	it sums the amount of transport request items supplied after deadline(later date)
Within of time	it sums the amount of transport request items supplied before deadline(later date)
Name of Ship	is the name of the vessel that perform operations of supplying to offshore units

Table 2 - Description of voyages data (2 851 records)

Selected input and outputs of the different models are presented in Table 3.

Item number	Inputs		Outputs
	Inputs name	Description	
1	Quantity of cargo items	It counts the number of transport request items.	1. Apriori: Association rules 2. Multilayer Perceptron: Average of delay days 3. Decision Trees: Average of delay day
2	Total weight of Cargo items	It sums the cargo weights of transport request items.	
3	Average Delay days	It averages day's delays of transport request items.	
4	Priority	Classification of priority of cargo. Normal priority (N) means the cargo will be supplied to unit according a schedule supplying schedule. Emergency priority (E) means that cargo will be supplied as soon as possible from the moment they are request on.	
5	Cargo Classification Type	Cargo class, it can be General Cargo (GC), Liquid bulk Cargo (LC), or Dry Bulk Cargo (DC).	
6	Type of shipment	Type of operation which vessel performs. Load corresponds when the source is a port and destination is an OU. Backload corresponds when the source is an OU and destination is a port. Transshipment corresponds when the source and destination are OU.	
7	Status of delay	If cargo is within of time or out of time of term.	
8	Ship ID	Ship identification use only in Multilayer Perceptron and Decision Trees.	

Table 3 – Input and output values of the models

Apriori Algorithm.

Apriori proceeds by identifying the frequent individual items in a transaction database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database to determine association rules, which permit to determine trends in the database.

Support and Confidence are major indicators in this technology. Support is defined as fraction of transactions that an item set, i.e, is the frequency of occurrence of an item set. An item set is one which support is equal or greater than a defined threshold. Thus, confidence measures how often items in Y appear in transactions that contain X (Agrawal R., 1994). For this work we used an apriori algorithm. Confidence and support values have respectively set to 90% and 80%.

Apriori model have provided relations, which was not known previously. Table 4 shows the fifteen association rules considered satisfactory from an operational perspective. For instance backload operations with dry bulk cargo and transshipment operations with liquid bulk cargo has a probability to suffer higher delays.

Consequent	**Precedent**
Delays average	Backload, Dry bulk
Delays average	Emergency
Delays average	Transshipment, Liquid Bulk
Transshipment	Expired
Backload	Delays average, Emergency
Transshipment	Expired, Delays average
Delays average	Emergency, Transshipment
Expired	Emergency
Backload	Expired, Emergency
Backload	Delays average
Delays average	Transshipment
Delays average	Dry Bulk
Dry Bulk	Backload
Backload	Delays average, Expired
Backload	Emergency, Delays average

Table 4 - Apriori Results: Association Rules

Multilayer Perceptron.

A multilayer perceptron (MLP) neural network is a feed-forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called back-propagation for training the network (Rosenblatt, 1961), (Rumelhart, 1986). MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable. (Cybenko, 1989).

The learning set of the MLP has been set to 70% (1 995 records) of the data while test set used 30% (856 records) of the data available. MLP has been used to predict Average of delay days for each trip. The prediction carried out for delays average on logarithm base shows good correlation as we can see in Figure 11. R^2 is 0.991, mean absolute error is 0.007, mean squared error is 0.001, root squared error is 0.026, and mean signed difference is 0.001. This give a good confidence in the results of the model.

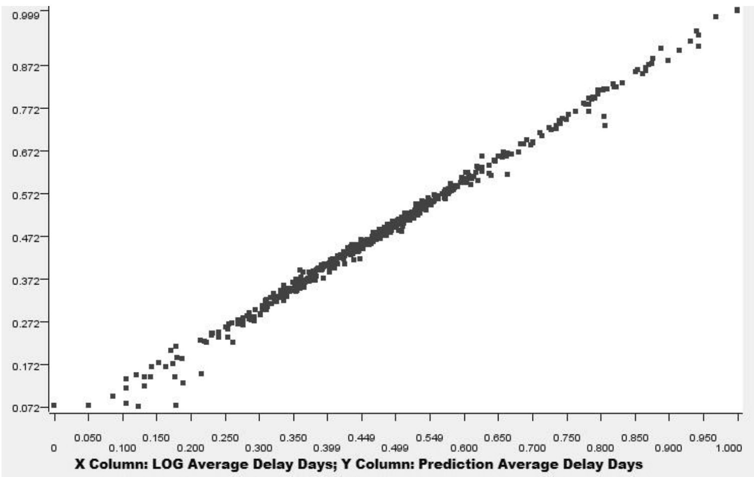

Figure 11 – Multilayer Perceptron Neural Network error of the test set (856 records)

Decision Trees.

A decision tree can be used as a model for sequential decision problems under uncertainty. This describes graphically the decisions to be made, events that may occur, and outcomes associated with combinations of decisions and events. Probabilities are assigned to events, and values are determined for each outcome. A major goal of the analysis is to determine the best decisions (Middleton, 2015). Decision trees are commonly used in operational research, specifically in decision analysis, to help identify a strategy most likely to reach a goal. Another use of decision trees is as a descriptive means for calculating conditional probabilities. (Quinlan, 1987), (Y. Yuan, 1995).

The learning set of the decision tree has been set to 70% (1 995 records) of the data while test set used 30% (856 records) of the data available. Decision tree have been used to predict average of delay days for each trip. Figure 12 shows results for this model, which provided a good accuracy (99.87%). Indeed, error is 0.121% in relation to prediction delays average, and wrong classified 1 on 824.

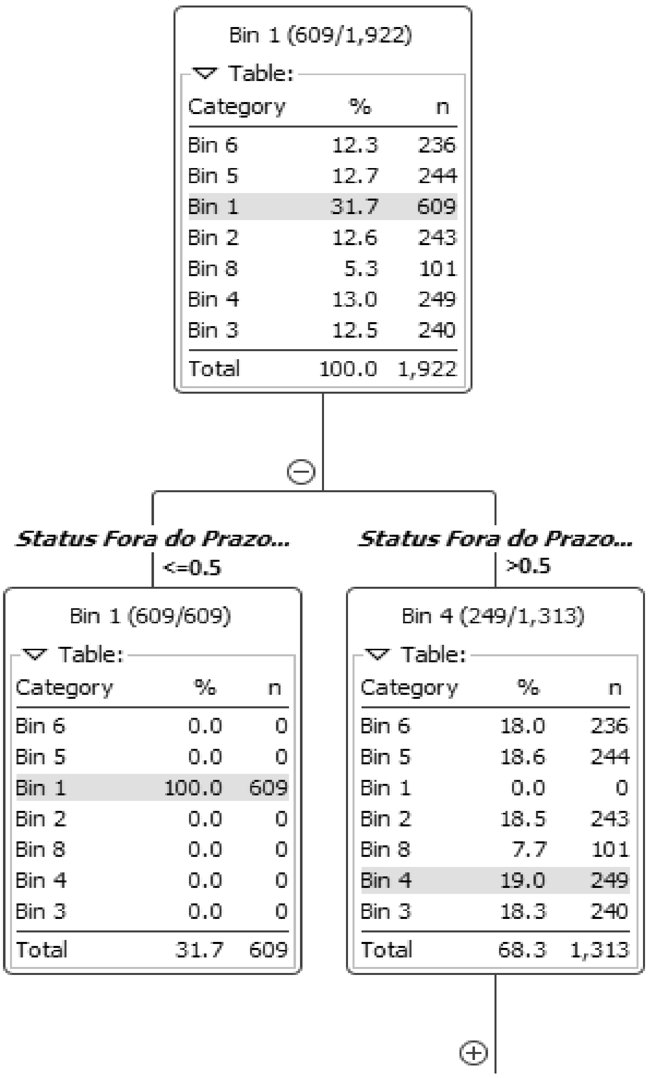

Figure 12 – Decision Trees Results: View and Accuracy

6. CONCLUSIONS AND FUTURE WORK.

This paper presented an application of machine learning algorithms on the analysis and prediction of offshore supply delays. Three models have been used, i.e. apriori association rules, multilayer perceptron neural network and decision tree. All presented a great reliability (accuracy better than 95%) during testing (test set have been set up to 30% of available data, ie. 856 records).

Knowing in advance were the delays are more susceptible to occurs in the supply chain, allows the planners to anticipate their strategies and delivery routes. These findings provide a new way to address efficiency and performance supply logistics of offshore platforms even if some future model improvements are required.

It is interesting to highlight that the most of working time have been spent for data collection, data preparation and data cleaning. Before to deploy the results in the industry we recommend to investigate more carefully the results provided here. For instance a time attribute could be included as well as more input data, e.g. one year span instead of six months.

7. REFERENCES.

Agrawal R., S. R. (1994). Fast Algorithms for Mining Association Rules. Proceedings of the 20th VLDB Conference. Santiago.

Banker, R., Charnes, A., & Cooper, W. (1984, September). Some models for the estimation of technical and scale inefficiencies in Data Envelopment Analysis. Management Science, 30(9), 1078–1092.

Caprace, J. -D., & Coronel, J. R. (2013). Measuring the efficiency of a ship fleet. Libro de Ponencias y Conferencias del XXIII Congreso Panamericano de Ingenieria Naval, Costa Afuera e Ingenieria Portuaria COPINAVAL 2013 (pp. 665-679). Margarita: IPEN.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 429 - 444.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function . Mathematics of Control, Signals, and Systems, 2(4), 303–314.

Feelders, A., Danielsa, H., & Holsheimer, M. (2000, August). Methodological and practical aspects of data mining. Information & Management, 37(5), 271–281. doi:10.1016/S0378-7206(99)00051-8

Ferreira Filho, V. (2014). Gestão de Operações e Logística na Produção de Petróleo. Rio de Janeiro: COPPE/UFRJ.

Hand, D., Mannila, H., & Smyth, P. (2001). Principles of Data Mining. Cambridge, Massachusetts: MIT Press.

Leite, R. (2012). Maritime transport of deck Cargo to Petrobras fields in Campos Basin: an empirical analysis, identification and qualification of improvement points. Rio de Janeiro: PUC-RIO.

Middleton, M. R. (2015). Introduction to Decision Trees. In M. R. Middleton, TreePlan Tutorial (pp. 157-168).

Nordbø, H. (2013). Optimal configuration of supply logistics for remote oil and gas fields. Trondheim: Norwegian University of Science and Technology.

Quinlan, J. R. (1987). Simplifying decision trees. International Journal of Man-Machine Studies 27 (3): 221. doi:10.1016/S0020-7373(87)80053-6.

Ragsdale, C. (2012). Spreadsheet Modeling and Decision Analysis: A Practical Introduction to Business Analytics. Virginia: Joe Sabatino.

Rosenblatt, F. x. (1961). Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Washington DC: Spartan Books.

Rumelhart, D. E. (1986). Learning Internal Representations by Error Propagation (Vol. Volume 1: Foundations.). MIT Press.

Wikipedia, t. f. (2014, September 14). Decision Tree. Retrieved from http://en.wikipedia.org/wiki/Decision_tree

Xiao, F., Ligteringen, H., Gulijk, C., & Ale, B. (2015, February). Comparison study on AIS data of ship traffic behavior. Ocean Engineering, 95, 84–93.

Y. Yuan, S. M. (1995). Induction of fuzzy decision trees. Fuzzy Sets and Systems 69, 125–139.

Zubcoff, J., Pardillo, J., & Trujillo, J. (2009, June). A UML profile for the conceptual modelling of data-mining with time-series in data warehouses. Information and Software Technology, 51(6), 977–992.