WHITEPAPER

Digitization &Business Analytics for Ports

An Approach Note

Table of Contents

1. Ports: Evolution into Digitization & Business Analytics

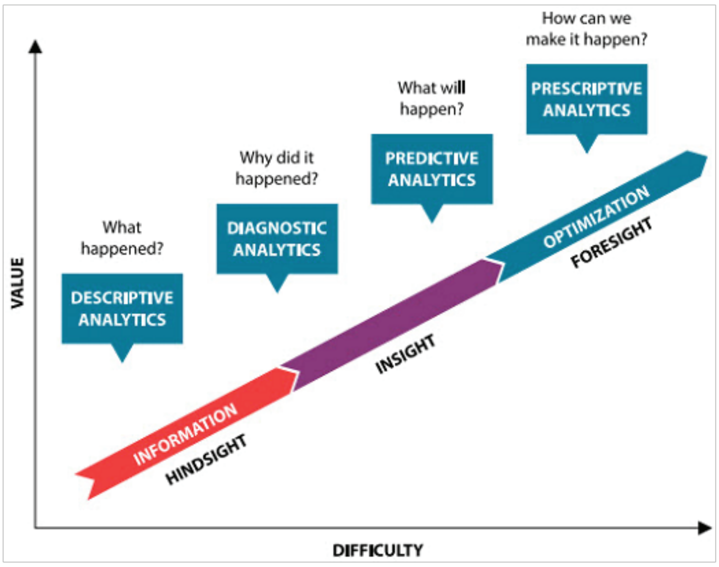
Since the beginning of containerization, the digital transformation of port operations has become indispensable for driving innovation and modernization in maritime ports. The ability to share information between involved actors and to track cargo is critical for reducing uncertainties, increasing reliability, and improving the coordination in integrated transport processes. Moreover, advanced information systems can provide a basis for addressing environmental sustainability in maritime ports.

An emerging impact of big data can be identified in the maritime industry, which mainly refers to different technologies and techniques to process and analyse large and complex sets of data that exceed the capacity or capability of conventional methods and systems. The Maritime Port Authority of Singapore (MPA), for instance, is collaborating to tap big data solutions for improving maritime and port operations, e.g., through a prediction of vessel arrival times and a better detection of movements, authorized activities (e.g., pilotage), and unauthorized activities (e.g., illegal bunkering). Although the term is often misinterpreted or used as a buzzword in the industry (e.g., as a substitute for data mining or business intelligence, BI), the growing interest in big data is also reflected in products and services of cargo-handling equipment providers and software vendors. Several instances include:

- **A cloud-based platform** is being built to display real-time productivity and operational data as well as maintenance information.
- **Terminal operating systems (TOS)** is being built to have a Terminal BI Portal to better understand the historical and real-time performance of terminal operations and further aims to use machine learning to gain additional insights from the TOS data.

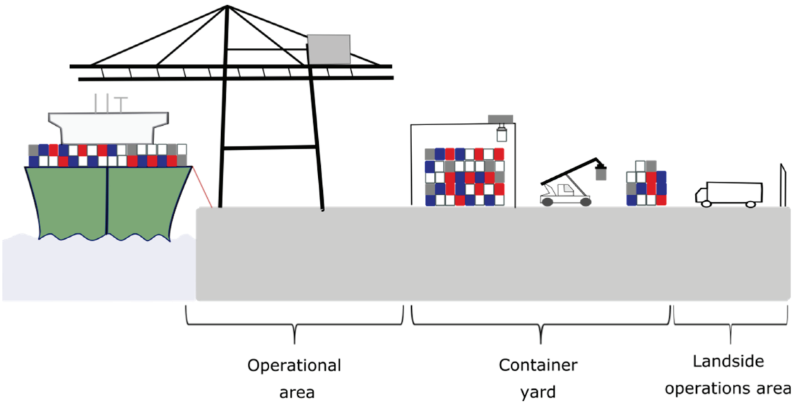
2. Business Analytics: A Brief Introduction

A common definition of business analytics is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models so that managers gain improved insights about their business operations and make better, fact-based decisions.



3. Data Mining Applications in Container Terminals

Given the fact that means of descriptive and predictive approaches have become increasingly important in the current phase of digitalization, this section provides an overview on potential applications with respect to the main operation areas in container terminals at the quayside, yard, and landside. This includes a brief overview of academic works applying data mining methods to produce more accurate forecasts as well as to better understand and address certain problems in container terminals



3.1 Quayside Operations Area

At the interface between seaside and landside operations, the main focus of the quayside operations area is on the discharging and loading of sea-going vessels using quay cranes (i.e., ship-to-shore gantry cranes). Moreover, it involves the horizontal transport of containers between quay wall and the yard operations area, e.g., using AGVs or straddle carriers. Besides providing modern equipment ensuring high productivity, it is important to efficiently allocate and schedule resources (e.g., berths, quay cranes, vehicles). In this regard, the quayside planning is dependent on many (external) factors, such as vessel arrival times, vessel call patterns, peak demands, and the handling capacities and capabilities of the quayside equipment. Different information technologies and systems are specifically used to collect and manage operational data at the seaside, including:

- **Automatic identification system (AIS):** A technology that supplements radar systems for tracking vessel positions with the primary objective of avoiding vessel collisions. After enabling the communication with satellites, referred to as S-AIS, the technology nowadays supports a real-time monitoring of vessels. AIS data messages include information about the vessel (e.g., Vessel type, length, width, draught) and voyage (e.g., port of destination, speed and course over ground, heading). Using this data, several vessel tracking web services can be established.

- **Vessel traffic service (VTS):** A VTS includes functionality to collect, analyse, and disseminate data, especially to navigate vessels in busy, confined waterways and port areas. The information system integrates various subsystems and technologies, including AIS, vessel movement reporting systems, radar systems, radio communication systems, traffic signals, and video surveillance.

3.2 Vessel Arrival Times

While being important for an efficient planning of subsequent terminal operations, reliable forecasts about the actual arrival times of vessels are still scarce in many seaports. This may lead to unused terminal capacities and longer vessel waiting and turnaround times. Any means to predict arrival times further allow to operate vessels more efficiently in terms of emissions. In this context, slow steaming and virtual arrival policies are currently being used, taking into account, for instance, the impact of tides. In the context of container terminals, the prediction of vessel arrival times is being done by using a feed-forward neural network for estimating ship arrival times in order to better determine capacity demands, for which an optimization model is used. The approach aims to reduce the number of additional workers in working shifts that need to be planned to cover uncertain demands, considering aspects like weather conditions, such as account geostrophic wind speeds, wave heights, peak wave periods, and wave directions. Using the Gini importance measure, measuring the relevance of input variables, the high impact of weather conditions on vessel arrival uncertainty is highlighted. The objective is to use a modified framework of case-based reasoning (CBR) for the early detection of vessel delays using real-time S-AIS vessel tracking data in addition to historical data (e.g., data from bill of ladings). The approach allows to detect delays in real-time and predict movement patterns of a vessel until its arrival.

3.3 Berth Operations

To predict the performance of vessel loading and discharging operations, a neural network is being used that takes into account operational data (e.g., berthing time, number of containers, number of gangs, vessel beam size) as well as wind conditions (e.g., average wind speed, wind direction) during berthing of respective vessels.

For addressing the berth allocation problem in bulk terminals, seeking to identify the berthing position and berthing time of bulk carriers, a machine learning approach is being used for selecting optimization algorithms dependent on the scenario at hand. A k-nearest neighbour’s algorithm is being used to classify each problem instance based on its features.

Taking into account the historic performance of algorithms in solving similar problem instances, a ranking of algorithms is generated for each problem instance.

3.4 Yard Operations Area

Yard operations mainly involve storage and stacking logistics and serve as a buffer between seaside and landside operations. Several complex planning and optimization problems result from yard operations, such as yard allocation problems, post-stacking problems (e.g., remarshalling, pre-marshalling, and relocation problems), crane scheduling, etc. The performance of yard operations is constrained by several factors, including the quay wall throughput (per year / in peaks), the size and shape of the yard area, characteristics of containers (e.g. type, size, weight, destination port), the handling performance of crane systems, and handling equipment. These factors can have an effect on important performance indicators, such as on container dwell times (i.e., the time a container spends at the terminal), handling performance and utilization of equipment, and operational costs.

Different information systems and technologies are in place to support yard operations, such as:

- **Terminal operating system (TOS):** Functionality for registering new containers and tracking their position within the container yard is provided by the TOS. In particular, automated transfer cranes (ATC) rely on the availability and accuracy of job and container data from the TOS to autonomously perform yard moves.

- **Automated transfer points for trucks:** Some container terminals have implemented automated transfer points at the yard to identify and serve incoming trucks. After following the instructions, the driver must leave the cabin and confirm the yard operations by showing a driver's card at the bay station. The latter increases safety and enables the identification of containers based on job data stored on the smart card.

Truck turnaround and transfers

3.5 Container Dwell Times

Prolonged container dwell times result in a high storage yard occupancy and may result in adverse effects on the terminal productivity and throughput capacity. While reducing dwell times increases the yard throughput capacity, storing containers in the yard over a longer time may also result in higher revenues earned from demurrage fees. Different methods to predict dwell times at terminal yard operations areas are: naïve Bayes (NB), decision tree and a hybrid Bayesian decision tree (NB tree). Using the model, the system further assess the impact of changes in determinants on the container dwell times, yard throughput capacity, and terminal demurrage revenues using three scenarios: changing the status of containers from empty to full, closing truck gates in low volume conditions, and changing the ocean carrier.

3.6 Container Stacking

Container stacking policies for containers have been widely discussed in the literature. Due to the ever-growing requirements to better use the space of container terminals and the impact of larger vessels, a higher yard utilization and a reduction in the number of reshuffles are desired. Besides advanced optimization and simulation approaches, there is a need to incorporate data mining methods; based upon fuzzy ANNs for the regulation of container yard operation including the system status evaluation as well operation rule and stack height regulation. A two-phase approach is being used: the first phase of the regulation process forecasts the quantity of incoming containers. The second inference phase decides on the operation rule and stack height, addressed as a fuzzy multi-objective programming problem with the objective of minimizing a ship's waiting time and the operation time. A comparison between results of the proposed model and current operation in 30 days shows that the total ship waiting time is reduced from 64 h to 46 h.

3.7 Landside Operations Area

Landside operations involve internal transports, truck operations, and railway operations. Related horizontal transport operations rely on an efficient handover of containers at the yard or in dedicated handling areas (e.g., rail or barge terminal) and might be subject to inspections. Improving those operations leads not only to a better hinterland accessibility and inland connectivity, crucial for the competitiveness of ports, but also facilitates efficient connections to auxiliary and value-added logistics areas within seaports. The increasing container volumes, peak demands, and a lack of coordination, however, lead to growing traffic and congestion at container terminals and within port areas, especially in areas located in urban environments with limited space for port expansion. As those operations highly contribute to congestion, traffic accidents, emissions, and noise, they have a great impact on the sustainable development of ports. In recent years, a large number of publications has been devoted to study and improve landside and hinterland operations, such as concerning gate/truck appointment, extended gate concepts / dry ports, and inter-terminal transportation.

- **Gate/truck appointment systems:** To better balance the workload and reduce waiting times at terminal gates, many container terminals require truck companies to pre-register containers and to book an available pickup or delivery time window. The planning of gate capacities and time windows requires a good understanding of truck arrival patterns and demand. Trucks that provide all documents in advance and arrive within the time window can therefore expect a guaranteed access to the terminal and a fast clearance process. Moreover, self-service stations have been introduced allowing the truck driver to complete missing data before arrival. Some ports penalize no-shows and late arrivals or charge a fee for day-shift or peak-hour appointments.

- **Port traffic management / intelligent transportation systems (ITS:** Some ports have implemented modern port road and traffic control systems to monitor and control traffic flows within the port area. For this purpose, different technologies, in particular sensors and actuators, are applied (e.g., laser vehicle detection systems, induction loops, etc.).

The collection and analysis of traffic-related data build not only the basis to analyse motion patterns, infrastructure bottlenecks, and areas with high accident risks, but also allow to react promptly to certain traffic conditions, e.g., by adapting electronic traffic signals and displaying relevant information on electronic display for traffic information and control, the performance of truck movements to explore movement bottlenecks, and to determine the frequency, costs, and environmental burden of recurring events. More accurate weather data and forecasts can be used to better control the traffic and warn vehicle drivers according to certain weather conditions. Moreover, the demand for an efficient parking space management is growing. In this regard, it becomes increasingly important to make reliable predictions about the availability of parking spaces in certain areas of the port. By identifying individual motion patterns and preferences of truck drivers, context-aware recommendations can be provided.

- **Mobile applications:** Mobile devices allow a direct interaction between actors involved in port operations and are equipped with powerful computing and sensing capabilities. Analysing contextual data may not only help to understand the situation of individuals but also predict forthcoming events in order to provide guidance and individual recommendations (e.g., recommended travel speed to reduce emissions and to benefit a series of green traffic lights). In many ports, new mobile apps have been introduced in recent years, especially for truck drivers.

- **Rail traffic management:** Besides the truck transport, a large part of cargo movements is handled via rail transport requiring information systems to efficiently manage rail operations. An example of a corresponding information system is transPORT, which is a new rail traffic management system of the Hamburg port railway. The system provides data on train locations, train movements, wagon sequences, track occupations, wagon destinations, un-loading/loading schedules etc. In the context of synchro modality, for example, analysing available sources of (real-time) data may be useful for predicting prices, available capacities, and the performance of alternative modalities.

3.8 Port-related Truck Traffic

The majority of works is focused on the prediction of truck-related cargo volumes in seaports. One of the first series of works looks at cargo flows and modal split in seaports of Florida (US) in order to support strategic planning regarding the prioritizing of public funds for roadway upgrades. More specifically, the works propose back-propagation ANN models to determine relevant factors and to predict inbound and outbound heavy-truck volumes in the Port of Miami (US) and, additionally, to determine the daily modal split between inbound and outbound rail and truck cargo volumes in the Port of Jacksonville (US). In general, the proposed models use seaborne import and export freight data of respective ports. By considering the dwell time of containers in the container terminals, representing the lead and lag times (in days) depending on the direction of cargo, the works were able to improve the accuracy of predictions.

3.9 Waiting Times and Turnaround Times

Besides the volume of cargo, the planning of landside and hinterland operations requires reliable indicators for waiting and turnaround times, for instance for a more efficient vehicle routing. While exceeding waiting and turnaround times may greatly affect the schedule of truck drayage and hinterland operators, there is a proposed concept for developing a decision support system based on truck arrival rates and predicted truck gate waiting times. While the focus is primarily on the system architecture and user interfaces, the systems apply an ANN model based on actual truck waiting times from an empty container depot in Northern Germany. In the experiments, with considering weekdays, daytimes, and public holidays in the set of input variables and eliminating night periods, the model accuracy was increased.

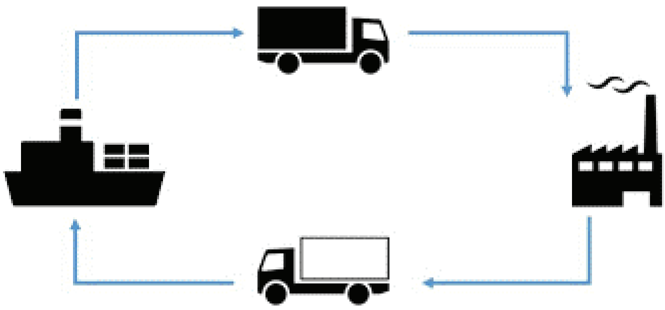
3.10 Truck Delays

For short-term and long-term planning, identifying the causes of inefficiencies at container terminals is at least as important as the prediction of future developments. Some works apply three decision tree models to identify causes of abnormally high truck turn times at the BCT (US), including a chi-squared automatic interaction detector (CHAID), classification and regression tree (CART), and a decision tree. As a data basis, the models use transactional data from gate operations (e.g., arrival at the gate queue, terminal entry time, use of chassis, etc.) and yard operations data concerning quayside operations, drawn from the TOS, over a period of eight months. Due to the higher priorities of quayside operations, the terminal operator wanted to know, for instance, whether vessel operations pose a conflict to drayage operations. The models are formulated as binary classification problems, where the indicator function is one (1) if the truck turn time (TT) is greater than one hour, and zero (0) otherwise. By analysing the resulting decision trees, main causes for high truck turn rates at BCT could be identified.

In another implementation, data of webcams was used to observe truck queuing patterns and to analyse the distribution of truck processing times, truck inter-arrival times, and truck queuing times at the entry gate of container terminals to better understand reasons of inefficient truck queuing. The implementation conducted goodness-of-fit tests to identify best-fit distributions using data of two container terminals. Several implications are drawn from the distributions, such as reasons for long queues in front of the gate. For example, long queues can be observed at the opening hour as truck drivers aim to perform as many moves per day as possible (usually they are paid per container) or in case of long turn times of other trucks within the gate.

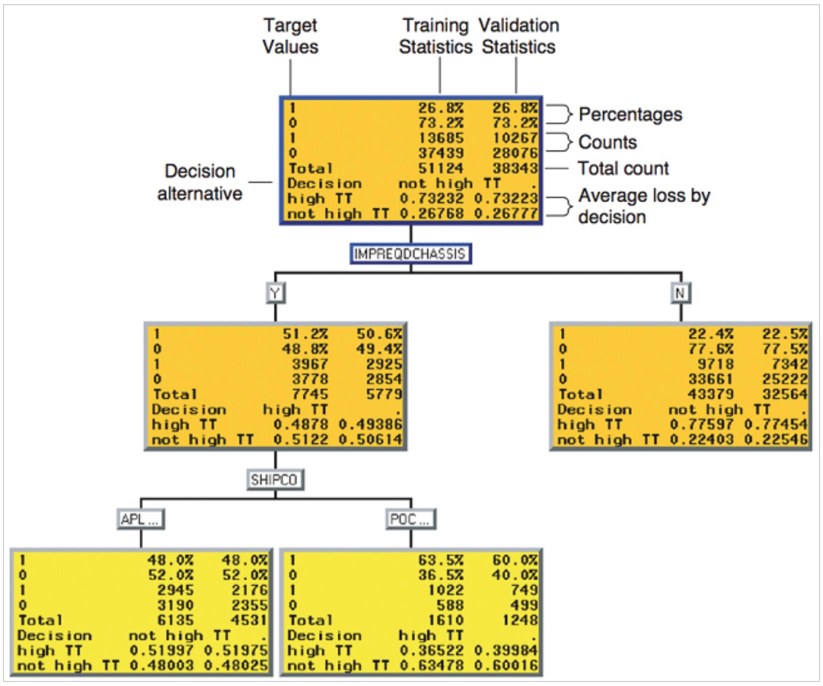
- First, at some terminals, the queues at the opening hour could be extensive because of the drayage drivers' desire to make their first move at the beginning of the day to allow for more time for subsequent moves later in the day; most drivers are paid by the move.

- Second, there is extensive queuing during the lunch hour at some terminals because of the policy to close for lunch. Moreover, analysing those distributions allows to identify peak hours (e.g., arrival of a new vessel) and daily/weakly variations. The models further demonstrate that truck queuing is higher during heavy rains, thus indicating an impact of weather on terminal operations



3.11 Conclusion and Outlook

In recent decades, maritime ports and container terminals have invested in automation and digitalization to improve the productivity and operational efficiency of related processes. Following the developments of the current generation of digital transformation, the amount of complex data is growing at a fast pace, while remaining mostly under-processed or under-analysed if not handled appropriately. Business analytics represents a concept to be able to use better information and knowledge in decision making processes, e.g., supported by means of optimization methods, it is essential to first process and analyse operational data. To put it concisely: a 'data-driven' perspective needs to enrich the traditional 'optimization' perspective.



Author

Amitabha Sinha
Vice President
Delivery and Analytics
India and SAARC

Amithaba has over 20 years of experience in technology and strategic consulting. Both on the technology as well as the digital strategy side, he has run transformation projects with Fortune 100 companies in geographies like US, Canada, UK, and Singapore. Amitabha started his career with IBM Watson Lab and has 5 patents with IBM. Thereafter, he worked for consulting majors like PwC (UK) and EY (US) and has run large BFSI ODCs based out of Singapore and Hong Kong.

Amitabha is a First Class First President's Gold Medal Winner from IIT Delhi, has an MS from Illinois Tech and a PGP from IIM Bangalore.