With many new opportunities emerging from the current wave of digitalization throughout global logistics chains, terminal planning and management need to be revisited with a data-driven perspective.

The amount of operational data, such as from a terminal operating system (TOS), together with data from a variety of new data sources, such as sensors and mobile technologies, is growing fast but for the most part remains to a considerable extent too under-analyzed to be of real value.

Meanwhile, many current projects and initiatives in the port industry indicate a growing interest in data analytics solutions. One example of applying data analytics is the SAFER project of the Maritime and Port Authority of Singapore (MPA). Under the project, MPA has piloted three IBM analytics-based modules to improve the management of Singapore’s growing vessel traffic.

Another example is the Navis ATOM Labs, which is investigating the use of Machine Learning (ML) for the optimization and automation of terminal operations. In this technical paper, we provide a brief overview of potential applications of ML in container terminals and discuss some relevant challenges.

To establish a data-driven perspective on terminal planning and management, we analyzed the current state-of-the-art in academia regarding applications of ML in the context of container terminals. After giving a brief introduction to ML, we briefly summarize the scope of those works within the operational areas of a container terminal.

MACHINE LEARNING

One central aspect of human intelligence is the ability to learn, i.e., to infer stable relationships from observations which are often reflected in data. Given the growing volume and complexity of data, discovering patterns, regularities, or even irregularities has become increasingly important, not only in container terminals.

ML is a branch of artificial intelligence (AI) devoted to the development of algorithms that can automatically detect patterns in data and then use the uncovered patterns for supporting decision making, e.g. by predicting future data. Generally, two main types of ML are differentiated.

• Predictive/supervised learning: The goal is to infer a mapping from inputs $x$ to outputs $y$ from a labeled training set, which consists of set of past observations, known as training examples. The training set also includes the observed output values. An algorithm is then used to infer a function or model for predicting output values for new input values. The output can be categorial, known as classification or pattern recognition task, or nominal, known as a regression task. Thus, there are different ways of modeling dependent on the real-world case. A vessel delay prediction, for example, can be modeled as a regression task to predict the vessel delay either as real number in hours or minutes or as a class of lateness, for instance, “late”, “very late”, etc.

• Descriptive/unsupervised learning: The goal is to find interesting patterns from unlabeled data, meaning that the desired output for each input is not defined in
Discovering clusters in data is a common task of unsupervised learning. Besides the above, we find other terms that are often discussed in this context, such as deep learning, which extends commonly-known approaches of neural networks. Another inflationary term used in the port industry is big data, which mostly refers to the computing architectures and technologies required for storing and processing data that exceeds the processing capacities of conventional systems. Nowadays, affordable and highly scalable data storage and processing services, including those offered by public cloud providers as well as software frameworks like Hadoop, are available to handle big data. Still, performing big data analytics requires ML approaches to analyze data.

**QUAYSIDE MACHINE LEARNING**

The performance of 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. Disruptions and uncertainties may result from a lack of reliable information and forecasting. This comprises delays and over-punctual vessel arrivals, weather and tidal conditions, traffic congestion and equipment breakdowns. A strong research focus is on the analysis of satellite automatic identification system (S-AIS) data for identifying patterns and anomalies of vessel operations, for example, to avoid vessel accidents and to identify authorized activities like pilotage or unauthorized activities like illegal bunkering). Applications of ML in the quayside include:

- **Prediction of vessel arrival times**: To reduce the uncertainty of vessel arrivals, research has been conducted to evaluate different algorithms for predicting vessel arrivals, also by taking into account weather conditions, such as wind speeds and peak wave periods. Models can also be further used to identify causes of vessel delays based on an analysis of input variables.

- **Berth planning**: Existing approaches predict the performance of vessel loading and discharging operations taking into account operational data, such as berthing time, number of containers, vessel beam size and wind conditions. ML can also be used to improve the selection of optimization methods used for berth allocation planning.

**YARD MACHINE LEARNING**

Several complex planning and optimization problems result from yard operations, such as yard allocation problems, post-stacking problems, crane scheduling, etc. To efficiently plan operations in the yard, it is important to reduce uncertainties and predict future scenarios, e.g., regarding demand and dwell times. Applications of ML in the yard include:

- **Prediction container dwell times**: Different algorithms have been developed and evaluated in research to predict and determine the
determinants of container dwell times in the yard. Models can be also used to assess the impact of changing determinants on the container dwell times, yard capacity and terminal demurrage revenues. A classification of container data using different dwell time classes, for example in days, has been used to determine stacking policies.

- **Container stacking:** In combination with optimization methods, algorithms have been developed to predict the quantity of incoming containers and weight groups of containers in order to optimize the container stacking policy.

**LANDSIDE MACHINE LEARNING**

Improving landside operations can lead to better hinterland accessibility and inland connectivity, crucial for the competitiveness of container terminals.

The increasing container volumes and peak demands, however, lead to growing traffic and congestion at container terminals and within port areas. Modern technologies, such as sensors, actuators, and mobile technologies provide new sources of (real-time) contextual data, which can be used by ML approaches to better understand and coordinate traffic flows, for example within a gate appointment system. Discussed applications include:

- **Prediction of truck traffic:** Existing approaches determine relevant factors and predict inbound and outbound heavy-truck volumes, for instance, using geospatial sensor-based data from trucks.
- **Prediction of truck waiting and turnaround times:** Approaches to analyze truck arrival rates and predict gate waiting times have been proposed, taking into account temporal effects.
- **Prediction of truck delays:** Different algorithms have been proposed to identify causes of abnormally high truck turn times in container terminals.

**CHALLENGES**

The gap between the data, produced in and around terminal operations, and its use for terminal planning and management is growing. ML provides a set of methods to use better information and knowledge in decision making processes. As such, it provides means to establish a data-driven perspective, complementing and supporting the traditional optimization and automation perspective. Although there is currently a lack of academic studies, some promising ML applications can be found in the context of container terminals. However, there are some challenges before data-driven decision making can be realized including:

- **No free lunch:** There is usually no “one size fits all” regarding ML models. The development of appropriate models involves several recurring phases, where the importance and workload of the preliminary steps of selecting, preprocessing and transforming data is often underestimated. In these phases, the background knowledge of experts from the individual terminal is essential for setting the goals, defining the modeling approach and preparing the data.
- **Integration of systems and data sources:** The quality of predictive models can be influenced by information provided by the input data and its quality. However, many useful insights are hidden in internal subsystems or external systems. Examples include weather stations and port traffic systems.
- **Expertise:** People having the skills necessary to extract relevant insights from complex amounts of data, for example, by using ML, statistics and software tools as well as by communicating with domain experts, are in strong demand in all industries, but are nonetheless still rare. As seen in current ML projects, the collaboration with leading tech providers is necessary to obtain the required expertise, but collaborating with research institutes could be an alternative.

**REFERENCES**


**ABOUT THE AUTHORS**

Dr Leonard Heilig is a researcher and lecturer at the Institute of Information Systems at the University of Hamburg. His research interests are in cloud computing, machine learning, optimization methods, and related applications in maritime ports. He serves as guest editor in highly-ranked journals in the area of cloud computing, information systems, and maritime logistics. He spent some time at the University of St Andrews in Scotland and, most recently, at the University of Melbourne in Australia as a visiting scholar. His professional background includes positions at Airbus Group Innovations and Adobe Systems. He is a consultant for companies in different sectors and international projects.

Dr Eduardo Lalla-Ruiz is a researcher and lecturer at the Institute of Information Systems at the University of Hamburg and an Alexander von Humboldt fellow. His research interests lie in the field of operations research, artificial intelligence, game theory and maritime logistics. He also serves as editor in special issues in operations research and logistics journals and participates in several national and international projects including collaboration with companies.

**ABOUT THE ORGANIZATION**

The Institute of Information Systems of the University of Hamburg in Germany specializes in interdisciplinary research for supporting decision-making processes within various application areas. A strong research focus is on quantitative methods, data mining, and cloud computing for supporting the planning and management in port logistics. Numerous publications in journals emphasize the quality of the institute’s research. Several projects in the port industry have been successfully carried out in recent years.

**ENQUIRIES**

University of Hamburg
Institute of Information Systems
Von-Melle-Park 5
20146 Hamburg, Germany

+49-40-42838-5519
Email: leonard.heilig@uni-hamburg.de, eduardo.lalla-ruiz@uni-hamburg.de
Website: www.bwl.uni-hamburg.de/iwi