論文の要旨

題 目　LONG-TERM SHIP POSITION PREDICTION USING AUTOMATIC IDENTIFICATION SYSTEM (AIS) AND DEEP LEARNING

（AIS データとディープラーニングを用いた船舶の長期的位置予測に関する研究）

氏 名　IBADURRAHMAN

Maritime transportation is recognized for its central role in the global supply chain considering it accounts for 90% of international trade by volume and 70% by value. The United Nations predicted that the total volume of seaborne trade worldwide would increase by 3.2% from 2019 to 2022. Therefore, establishing safety and security in maritime transportation is essential. Maritime situational awareness (MSA) is a critical aspect of maritime safety and security that can be achieved through tracking, surveillance, and position prediction of ships. Once a prediction of the ship position is obtained, decision making and action planning can be supported at different information levels. However, ship position prediction for MSA requires a longer time interval than other tasks such as collision avoidance and maritime traffic monitoring, which mainly use a short-term prediction from a high-precision real-time forecast spanning less than one hour. Accordingly, we considered a prediction with a time interval Δt longer than 12 h as a long-term prediction, and a prediction with a time interval between short- and long-term threshold as a medium-term prediction.

Studies on the long-term prediction of vessel position remain scarce despite its considerable potential for maritime applications, while almost all previous studies have focused on either near real-time predictions (short-term prediction) or predictions with a time interval lesser than 12 h (medium-term predictions). Long-term ship position prediction is required not only for MSA but also for efficient allocation of ships by shipping companies in accordance with global freight demand. It can be utilized to monitor and assist the fleet, specifically, when communication with a ship operator breaks down owing to poor weather conditions or when a ship is in distress. It can also be implemented by shipping insurance and maritime investigators for investigation purposes. The ability to predict the long-term position of a fleet could prove not merely necessary but vital for strategy formulation in the fast-changing dynamics of maritime industries.

Maximizing the potential of maritime big data is essential for predicting vessel positions. The automatic identification system (AIS) is a self-reporting message system on board a vessel that records its position and condition. Each record of the AIS message contains the static and voyage-related information of the vessel and its dynamic information such as longitude, latitude, speed over ground, course over ground, and heading. As of July 2008, all commercial vessels above 300 GT serving international routes were regulated to be outfitted with an AIS Class A device by the International Maritime Organization (IMO). Prior to the widespread use of maritime big data (i.e., AIS data), studies on vessel position prediction were conducted using data from radar or laser sensors, such as in the work of Perera et al. (2012). Since 2015, studies in the field have started using AIS as a historical data source for ship position information. Czapiewska and Sadowski (2015) conducted position prediction experiments using linear and nonlinear motion functions for location data compression of AIS records. The high volume of AIS data accumulated over the years is a potential asset that needs to be explored, especially in the current age of artificial intelligence. With this big data of vessel position records, it is possible to predict vessel position by applying advanced machine learning (ML) techniques such as deep learning.

This study aims to achieve a 24 h interval and more ship position prediction for MSA. Predicting a ship's long-term position can be game-changing in maritime industries. The application of the long-term position prediction may not be the end goal. The results can be fed to another machine learning system, along with many other signals. Then, a business objective of the machine learning pipeline can be established broadly, such as for ship allocation, shipping investment, or global economic projection. For instance, with the objective of shipping investment, this downstream system will determine whether it is worth investing in a certain area and future time or not by predicting direct components to the measurement of the Baltic Dry Index (BDI), which has a predictive ability for a range of stock markets. This work provides a valuable benchmark for future studies.

Previous studies focused mainly on short- and medium-term prediction, ranging from 30 min to 10 h intervals as the AIS data were dense (closely packed between short time intervals) but in a limited period or timespan. Naturally, a vessel position prediction with long time intervals (e.g., 24 h) requires data with a long timespan, given that state-of-the-art ML algorithms, such as deep learning, require large amounts of data to make accurate predictions for longer intervals. Accordingly, nine-year AIS data for capesize bulk carriers worldwide and deep learning (DL) were used to accomplish long-term prediction. End-to-end learning refers to training a learning system represented by a single deep network model that supersedes the preprocessing stages typically present in traditional pipeline designs. By using extensive fleet data and computing power, the model can learn robustness to noise and generalization to data variation; this result confirms that an accurate long-term prediction of vessel position can be generated with the straightforward method. The utilization of nine years of AIS data worldwide and the development of a generalized DL model from an uneven time-interval dataset for the long-term prediction of ship positions constitute the novelty of this study.

This research demonstrated that predictions with an average time interval of 24 h are possible, confirming that the straightforward motion-based method can generate an accurate long-term prediction of the vessel position. The DL models generated more accurate predictions than the geodesic calculation as the baseline model in all areas. Predictions in the open ocean areas yielded higher accuracy than in the chokepoint areas; however, compared to the geodesic calculation, the improvement scores were higher in the chokepoint areas than in the ocean areas since the geodesic calculation failed to predict vessel behavior near the ports and congested waters. The DL model can predict the complex movement of ships near ports and congested routes, whereas conventional calculations fail, with no information regarding vessel status, historical trajectory, or destination. The last experiment demonstrated that the DL model appears to have a sense of the dimension of the geographic coordinate system that can be or is often passed, wherein the chokepoint areas rely more on the input features of the current latitude and longitude of the vessels. Moreover, on the dataset with varying and uneven time-interval distribution and without a trajectory reconstruction, the proposed MLP model generated predictions as accurate as the LSTM with faster training times.

Results from this research show that the geodesic calculation as a non-ML baseline failed when the time interval average was magnified beyond 24 h, and the performance of the previous DL model as the baseline became worse in large coverage areas compared to the new larger networks. The DL model delivered a satisfactory performance compared to the baseline models and succeeded in responding to the task of an increasingly large time interval dimension and coverage area; thus, the larger networks generalized better. Meanwhile, the optimum coverage size for all time-interval models is 39 million square kilometers, corresponding to the dataset's high dispersion rate. By capturing the densest region on the observed location, which leads to a proportional ratio of the low-distance to the high-distance intervals, optimum coverage for long-term position prediction can be achieved. The deep learning models generated worse predictions in Europe and East Asia than in the other observed locations (i.e., South America, the Southern Region of Africa, and South and Southeast Asia). The locations where importers of bulk shipments constitute the largest portion appears to be harder to predict than those that constitute bulk export regions, which corresponds with the fact that Capesize dry bulk carriers' unloading time takes longer with a wider approximate range than loading time.