# Real-Time Vessel Behavior Prediction

Dimitrios Zissis\*, Elias Xidias & Dimitrios Lekkas

Abstract—Vessel Traffic Management Systems (VTMS) and Vessel Traffic Monitoring Information Systems (VTMIS) have been available for a number of years now. These systems have significantly contributed to increasing the efficiency and safety of operations at sea. However, nowadays, risks at sea are once again on the rise, thus demanding an evolution in VTMS&VTMIS, such that they can support a human operator's better understanding of the complex reality at sea and enhance his or her decision-making in light of danger. A critical requirement of such systems is that they exhibit the ability to for-see unfolding cautious and potentially hazardous situations, so as to propose measures of danger avoidance. In this study, we employ machine learning, and specifically Artificial Neural Networks (ANN), as a tool to add predictive capacity to VTMIS and specifically MarineTraffic.com. The main concept of this study is to implement a publicly accessible, web-based system capable of real time learning and accurately predicting any vessels future behavior with low computational time. This work describes our approach, design choices, implementation and evaluation details, while we present a proof of concept prototype system. Our proposal can potentially be used as the predictive foundation for various intelligent systems, including vessel collision prevention, vessel route planning, operation efficiency estimation and even anomaly detection systems.

Keywords-- marine traffic prediction; machine learning; intelligent systems; artificial neural networks,

#### I. INTRODUCTION

In 2012, the public's awareness of maritime safety was heightened due to two high profile maritime incidents, the sinking of the Costa Concordia, off the shore of Italy, and the loss of the Rabaul Queen ferry, off Papua New Guinea, both leading to the loss of life at sea. In less than 50 years, cargo grew 14 times, fleet capacity grew 8 times, while oil tankers became twenty times bigger, average tankers 7 times bigger and dry bulk vessels ten to fifteen times bigger (ANAVE, 2013; Stopford, 2009). Ultra large cruise ships are now dwarfing the Titanic, as their capacity has risen to 6,000 passengers, posing various unique safety challenges. Today shipping is highly concentrated into modern sea-lanes, as vessels navigate between major ports to optimize efficiency (Allianz Global Corporate, 2012). Vessel growth in size, speed and volume in relation to shipping lane size, has made vessels difficult to maneuver, especially around major ports and inland waterways requiring a change in collision avoidance techniques (Westrenen & Praetorius, 2012). All of the above indicate a significant increase in traditional maritime risks, but also highlight the introduction of unique challenges in maritime shipping at various levels.

For many years, practitioners and researchers from the field of maritime safety have turned towards Information and Communication Technologies (ICT) in order to reduce risk at sea. As such, the Automatic Identification System (AIS) was developed, primarily as a tool for maritime safety and vessel collision avoidance and is an integral component of various Vessel Traffic Services (VTS), Vessel Traffic Management Systems (VTMS) and Vessel Traffic Monitoring Information Systems (VTMIS). A number of vessel tracking systems are open to the public through the Internet; as such, MarineTraffic (MT) is part of an open, community-based project that provides real-time information, regarding vessel movements and port traffic across the coastlines of many countries around the world. MT is at any time tracking more than 460,000 vessels and processing more than 50 million position reports per day, covering more than 10,000 ports and marinas across the globe. Information regarding these vessels is collected from over 1600 AIS receivers located at the coastline of more than 150 countries. A wide range of maritime stakeholder's access such information on a daily basis, in an attempt to increase the efficiency and the safety of their operations at sea. For example,

- Port authorities, coast guard, border controls, search and rescue teams combine vessel tracking information with other proprietary solutions accurately monitor and assess threats at sea;
- Pilots, tug operators, towage and salvage make use of vessel tracking information for assistance when navigating to a
  distress call to accurately track a vessel entering a port;
- Insurance companies use vessel historical data for incident investigation. The current condition, route and port calls of a vessel may also affect the insurance policy applied.
- Crewmembers, families of seafarers, recreational sailors and even passengers frequently access such information to learn about a specific vessels position, route, and estimated time of arrival.

Overall, stakeholders and a variety of end users make use of vessel tracking information to increase their own understanding and perception of reality at sea, to support their decision making and situation management. At a cognitive level, situation management is a goal directed process of (a) collecting information, (b) perceiving and recognizing situations, (c) analyzing past situations and predicting future situations and (d) realistic reasoning, planning and implementing actions so that desired goal

situation is reached with some pre-defined constrains (Rothblum, 2002). Intelligent systems have a great potential for addressing decision-making problems, because they can model the involved players and produce good results with low computational time(L. Gomes, Faria, Morais, Vale, & Ramos, 2014). Intelligent decision support systems use data and mathematical models, that possess the characteristics of flexibility, adaptability, comprehension, and the capacity to manage uncertain and constantly changing information (Krishnakumar, 2003), so as to support stakeholders decision making. They aim at automating steps (a) to (c) while providing human operators with proposals in support of their own decision making (d).

Maritime Domain Awareness (MDA) is the effective understanding of activities, events and threats in the maritime environment that could impact global safety, security, economic activity or the environment (Santos & Lunday, 2009). The major challenge faced today by MDA, is developing the ability to identify patterns emerging within huge amounts of data, fused from various sources (information fusion) and generated from monitoring thousands of vessels a day, so as to act proactively to minimize the impact of possible threats. Recent advancements in ICT have created opportunities for increasing MDA, through better monitoring, but most importantly understanding vessel movements. Statistical inference and machine learning algorithms can provide crucial help in this process. Achieving situational awareness, perceiving and comprehending elements and their contextual meaning in the environment within a given volume of time and space, while projecting their status into a future timeframe(Endsley, 1988), is a critical element of MDA(US Department of Homeland Security, 2005). Increasing Maritime Domain Awareness in light of safety and efficiency can be viewed as three-step process where,

- Accurately assessing the maritime environment: Assessment of objects and their relations, amongst themselves and their environment, to provide a better understanding of the current situation. Supporting an operators (a) and (b) processes.
- Impact Assessment: Projections of possible future situations and evaluations regarding evolving situations in an attempt to determine possible threats. Supporting human operators (c) analysis of past situations and predictions of future situations.
- Proactive hazard prevention and increased efficiency through process optimization. Supporting a human operator's decision making (d).

Deploying tools targeting at increasing MDA can potentially lea to improvements in various fields including,

- (1) Security and safety: The development of intelligent ship tracking systems opens possibilities of advancing maritime security far beyond simple collision prevention. Anomaly detection can be defined as a method that supports situation assessment process by indicating objects and situations that, in some sense, deviate from the expected, known or "normal" behavior and thus may be of interest for further investigation (Laxhammar, Falkman, & Sviestins, n.d.). Deviations from "normal" behavior may be a response to environmental conditions (weather condition, busy sea lanes/ports), in which case the system should support rerouting nearby vessels, but also threats related to terrorism, illegal trafficking, fishing, piracy and others.
- (2) Energy and economic efficiency of vessels: The growth in international maritime transport has resulted in increased air emissions. Exhaust gases, which include carbon dioxide (CO2), are a significant source of negative environmental impacts. Current research proves that a combination of better fleet management, logistics, voyage optimization and energy management can lead up to 50% decrease in emissions (International Maritime Organization, 2009). Prediction of vessel voyage speed and energy consumption can lead to voyage optimization. Forecasts of congestion in port can be used to invoke "Virtual Arrival" clauses, which reduce the speed of a ship en route to a port to coincide arrival with improved slot availability.
- (3) Autonomic vessel navigation: In the prospect of the deployment of autonomous vessels (MUNIN Collaboration, 2012) operating as multi-agent systems in dynamic, non-deterministic, continuous and inaccessible environments, such as the conditions at sea, it is critical that an agent can obtain a complete, accurate and up-to-date depiction of its operational environment. The ability of any agent to operate effectively within its environment is rooted in its ability to make sound decisions. Intelligent systems with predictive capabilities can provide the foundation for autonomous agent's decision-making.

In this manuscript, we describe our work on employing machine learning methods and specifically artificial neural networks, as a basis for accurately predicting a vessels future behavior with an emphasis on solution practicality. To this end, we focus on deploying a web-based infrastructure that can produce good results in low computational time. This work is meant to sit on the fence between theoretical computer science and software engineering that can provide practical solutions to everyday problems (applied soft computing). Predicting a vessels behavior with ANNs raises a number of unique design challenges. A balance needs to be sought between prediction accuracy and training times. Another challenge is processing data of such volume and velocity (AIS messages regarding vessels are received every 20-90 seconds). As the predictive capacity will be added to MarineTraffic, a system operating constantly and tracking thousands of vessels at any given time, special attention needs to be paid to data related design choices. Our overall objective is to design and develop a system, which exhibits the characteristics identified below,

- Is capable of learning vessels behavioral pattern from previous historical data available from MarineTraffic
- Is capable of real-time vessel behavior prediction on user request, in low computational time
- Is publicly accessible through the World Wide Web and can overlay vessel predictions on an interactive map for visualization purposes
- Has the capacity to operate as the foundation for vessel collision avoidance and anomaly detection systems

To overcome various obstacles and accurately address our stated requirements, a number of design choices were made that are documented throughout this report. In the following sections of our manuscript, we first describe our steps towards data

preparation and model building. In the successive section we present our training and evaluation results. Following this, we provide design and deployment details of our prototype system. We conclude our manuscript with an example case study investigating vessel collision detection.

#### II. PREDICTING A VESSELS BEHAVIOR

## A. Model Selection

Data Mining is an analytic process designed to explore large amounts of data in search of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data (Hill, T. & Lewicki, 2007; StatSoft, 2007). Machine learning provides the technical basis of data mining. In the maritime environment, machine learning techniques can be employed to "learn" vessel behaviors and project their status into a future time frame. Vessel behavior can be defined as the sum of all characteristics defining vessels movement, such as vessel position, course, heading and speed, observed over a given period of time. By definition, a pattern is composed of recurring events that repeat in a predictable manner. Vessel behavior monitored over a long period, gives insights into the navigational patterns followed by each vessel on specific routes.

An Artificial Neural Network (ANN) is a machine learning information-processing paradigm inspired by biological nervous systems. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons), which work in unison to solve specific problems (Bevilacqua, 2006). In general, a neural network is a parallel system, capable of resolving problems that linear-computing cannot (Verber, 2012). ANN have a broad applicability to various real world problems including classification and pattern recognition, data processing, control, robotics but also prediction. The unique characteristics of ANNs – adaptability, nonlinearity, arbitrary function mapping ability – make them quite suitable and useful for forecasting tasks. According to Karlaftis & Vlahogianni, ANNs have been mainly used as data analytic methods because of their ability to work with massive amounts of multi-dimensional data, their modeling flexibility, their learning and generalization ability, their adaptability and their good predictive ability (Karlaftis & Vlahogianni, 2011).

Forecasting has long been considered as the field of research in the domain of linear statistics. Traditional approaches, such as Box-Jenkins or ARIMA method (Box, G.E.P., Jenkins, 1976; Pankratz, 1983), assume that real world observations are generated from linear processes (Zhang, Eddy Patuwo, & Y. Hu, 1998). However, they may be totally inappropriate if the underlying mechanism is nonlinear. It is unreasonable to assume a priori that a particular realization of a given time series is generated by a linear process (Zhang et al., 1998). In fact, real world systems are most often nonlinear (Granger, 1993). When modeling complex datasets with possible nonlinearities or missing data, ANN's are often regarded as more flexible compared to statistical models. Neural networks are less sensitive to error term assumptions and can tolerate noise, chaotic components, and heavy tails better than most other methods (Kaastra & Boyd, 1996). Throughout academic literature, neural networks have been employed to predict various evolving patterns based on real world observations such as traffic flow prediction (Smith & Demetsky, 1994), bankruptcy prediction [25], 26], power electric demand (Connor, 1996), and many more.

In the maritime domain ANNs have been employed for tasks such as forecasting traffic flow at the Suez canal (Mostafa, 2004), predicting wave influence on the yaw motion of a ship (Nicolau, Aiordachioaie, & Popa, 2004) and vessel classification (Lagerweij, Vries, & Someren, 2009). Lagerweij, Vries, & Someren, analyze moving object trajectories from maritime vessels and classify vessels into three categories based on AIS data. In their work, they perform the tasks of clustering, classification and outlier detection from vessel trajectory data with the goal of identifying irregular vessel behaviors. In their work Perera, Oliveira and Soares, propose an ANN as the mechanism for detecting and tracking multiple vessels based on radar/laser tracking data (Perera, Oliveira, & Guedes Soares, 2012a).

Less work however has been conducted in the field of vessel movement prediction using an artificial neural network, mostly due to the lack of data. Ebada developed an Artificial Intelligent (A.I.) system, capable of predicting accurately the turning tracks of ships (A. Ebada, 2005). The physical and operational data of a ship are described and used as inputs into the system in order to predict the turning maneuvers. Closely related to this work is Simsir and Ertugrul (Simsir & Ertugrul, 2009) study, with the aim of predicting the future coordinates of a manually controlled vessel using a trained ANN in the Bosporus Straits. The ANN was trained by using position and speed data collected from vessels, which navigated manually in the Strait. They were able of accurately predicting vessel positions in a three-minute ahead window. Rhodes Bomberger, Seibert and Waxman developed a Fuzzy ARTMAP classification neural network architecture where normal vessel speeds for different regions in a port area are learned by clustering (Bomberger, Rhodes, Seibert, & Waxman, 2006). New data that is not recognized by the network during online operation is considered anomalous. The same research group has also proposed and implemented associative learning of motion patterns for anomaly detection, where associative neural networks learn to predict future vessel locations in a port given a current (Laxhammar, n.d.; Rhodes, Bomberger, & Zandipour, 2007).

## B. Exploration and Data Preparation

Neural networks are only as good as the data they are given and the questions that are asked of them (Azoff, 1994). One of the major constrains of applying machine learning techniques to vessel position prediction in the past, has been the lack of data necessary for training the ANN. The data used for this study is provided by MarineTraffic.com and is based on AIS. AIS transmissions can be defined as spatial time series (Rhodes et al., 2007), describing the movements of vessels across geographic regions. An AIS message contains the vessel's Maritime Mobile Service Identity (MMSI) – a unique nine digit identification number; Navigation status; Rate of turn; Speed over ground; Positional accuracy; Course over ground; True heading; True bearing at own position; UTC Seconds. Additionally messages may contain Radio call sign, Vessel Name, Vessel Type; Vessel Dimensions; Vessel Draught, Vessel Destination and vessels estimated time of arrival. Ships broadcast original position reports at time intervals that vary between 3 seconds and more than 30 seconds, depending on their speed and the type of their AIS transponder. These reports are used for the real-time map display, but MT only archives them every 2 minutes for each vessel, as this interval is enough for tracking purposes and the applications are not used for navigation purposes. For the neural network training, we decided that even a smaller frequency of position reports is enough for prediction accuracy. An interval of 15 minutes would be enough to detect an anomaly or course collision in vessel tracks while at the same time we ensure that the time series will be uninterrupted even for areas where the AIS signal reception is poor and the collected data is not as frequent as in fully covered areas. The following SQL query is used to achieve the described data down sampling:

```
select a.mmsi,a.speed,a.longtitude,a.latitude,a.course,a.timestamp from archive a inner join (select min(timestamp) as timestamp,mmsi from archive c where timestamp>='YYYY-MM-NN HH:MM' and timestamp<='YYYY-MM-NN HH:MM' and mmsi in (999999999) group by datepart(year,timestamp),datepart(month,timestamp),datepart(day,timestamp),datepart(ho ur,timestamp),datepart(minute,timestamp)/intervalXXXX, mmsi) b on a.mmsi=b.mmsi and a.timestamp=b.timestamp order by a.mmsi, a.timestamp
```

#### C. Model Building

Unfortunately, vessel data is so voluminous that it is impossible to train the ANN on raw past vessel data. To this end, we exploit the fact that specific vessel types follow repetitive patterns in short periods. We concentrate our study on passenger vessels that perform repetitive voyages within a given timeframe (such as around the Aegean Islands where voyages are performed within a few hours or days). Predicting the behavior of a vessel engaged in tramp trade or similar trade, would require a totally different approach, as on many occasions these vessels may not have performed a similar voyage in their short past data.

During the design of an ANN a balance needs to be sought between training times and the ANNs structure, as increases in the ANNS structure have a heavy impact on training times (Perera, Oliveira, & Guedes Soares, 2012b). The number of input nodes corresponds to the number of variables in the input vector used to forecast future values. We exploit the fact that a number of behavioral parameters can be implied and computed based on the predicted vessel positions. Thus, our forecasting neural network is required to predict a vessels latitude and longitude at a future point in time, while we programmatically calculate other values (such as bearing and speed as discussed in following sections).

For neural networks implemented to predict future values in a time series dataset, the number of input neurons corresponds to the number of lagged observations required to discover the underlying pattern (Zhang et al., 1998). Neural networks can be trained to predict on-step-ahead (based on the time interval data is collected in the dataset prediction will occur for the following value) or multi-step-ahead forecasting. The conventional approaches for multistep ahead forecasting, include iterative or direct methods of forecasting. In iterative approaches, forecast values are used as inputs for the preceding calculation, while direct methods require neural networks to have several output nodes to directly forecast each step into the future (Zhang et al., 1998). For our given dataset we experimented with a backward window size of four; thus requiring a structure that could handle 8 input neurons, for the short term prediction and ten lagged observations for long term prediction; thus requiring a structure that could handle 20 input neurons. For short term prediction (15 minutes) we trained the neural network to output a single prediction while for long term prediction [5 steps ahead (5\*15 minutes) and 10 steps ahead (10\*15 minutes)], we experimented with iterative and direct prediction approaches.

An important element of the neural networks structure is their net inputs by using a scalar-to-scalar function called "the activation function or threshold function or transfer function", output a result value called the "unit's activation" (Karlik & Olgac, 2010). In general, the activation function introduces a degree of nonlinearity that is valuable for most ANN applications. The predicted output of our ANN is in the range [-1, 1]; thus we selected the hyperbolic tangent function as an activation function for the hidden and output layers (G. S. da S. Gomes & Ludermir, 2013). Training is the means by which neural network weights are adjusted to give desirable outputs. The propagation training algorithm will go through a series of iterations that will most likely

improve the neural network's error rate by some degree (Heaton, 2011). The error rate is the percentage difference between the actual output from the neural network and the ideal output provided by the training data (Heaton, 2008). The Mean Square Error (MSE) is an Error Calculation method used in describing how well a machine learning method, typically a regression model, represents the data being modeled (Heaton, 2011). This process is repeated until the error for each training pattern drops under a certain accepted level. The training time is a major bottleneck for such systems as it can be exceptionally high for large datasets. It is challenging to strike a balance between model prediction accuracy and training time. Exploiting the fact that passenger vessels repeat their voyage within a given 24-48 hours timeframe, permitted restricting the training data to a number of days prior to prediction, without a major loss in vessel behavior information. This tradeoff allowed for much shorter training times (5-10 seconds) and a desired MSE of 0.01. This of course limits solution applicability to vessels performing repetitive voyages within a given period, but was necessary to reduce the size of the data and training times. The optimal tradeoff depends on the selected vessel.

## D. Model Evaluation

We implement our ANN in C# using the Encog3 machine learning library (Heaton, 2008). Encog is an open source advanced machine learning framework that supports a variety of algorithms, as well as support classes to normalize and process data. Most Encog training algorithms are multi-threaded and scale well to multicore infrastructure. To train and test our ANN we made a selection of passenger vessels and loaded data regarding these from the previous 48-72 hours (depending on prediction).

The conventional approach to evaluating a ANN accuracy usually involves randomly setting aside a portion of the dataset e.g. 70% for training and 30% for testing. The training data set is used exclusively for model development and then the test sample is used only to assess the trained network. After training was completed, we evaluated the trained ANN by feeding it data that was excluded from the original dataset. This data was collected in the following 24 hours and was previously unseen to the ANN during training. This data was pre and post-processed in the same method with the training data. In the following figures we report on evaluation results regarding predictions



Figure 1. Future (15 minute ahead steps) latitude prediction for a vessel sailing around the Aegean, Greece.

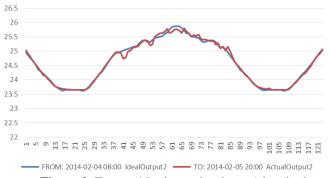


Figure 2. Future (15 minute ahead steps) longitude prediction for a vessel sailing around the Aegean, Greece.

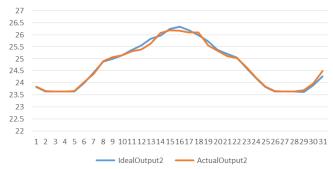
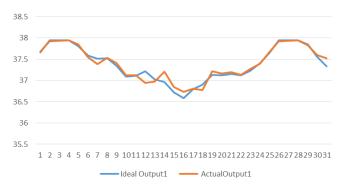


Figure 3. Future (4 hour's ahead steps) latitude prediction for a vessel sailing around the Aegean, Greece.

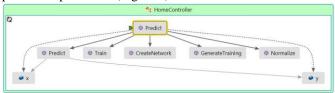


**Figure 4.** Future (4 hour's ahead steps) longitude prediction for a vessel sailing around the Aegean, Greece.

#### III. SOLUTION WEB ARCHITECTURE AND DEPLOYMENT

#### A. Deployment and Design Choices

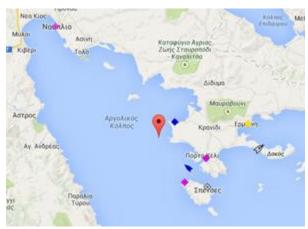
Consequently, to successfully testing and evaluating the ANN model we chose to design and deploy a prototype web application, which would meet the previously identified requirements. Due to the nature of our application, the web application was built using the ASP.NET MVC 5 (Model View Controller) framework. The ASP.NET MVC 5 is a framework for building scalable, standards-based web applications using well-established design patterns, that places an emphasis on a loosely coupled application architecture and highly maintainable code (Chadwick, Snyder, & Panda, 2012; Galloway, Haack, Wilson, & Allen, 2012). The Model-View-Controller pattern is an architectural pattern that encourages strict isolation between the individual parts of an application (loose coupling). The MVC pattern splits an application into three layers: the model, the view, and the controller. The model represents core business logic and data. Models encapsulate the properties and behavior of a domain entity and expose properties that describe the entity. The proposed solution contains models encapsulating vessel data. The view in MVC is responsible for transforming a model or models into a visual representation. In web applications, this most often means generating HTML to be rendered in the user's browser, although views can manifest in many forms such as AJAX interactive map mashups as in our solution. The controller, essentially C# code, controls the application logic and acts as the coordinator between the view and the model. Controllers receive input from users via the view, and then work with the model to perform specific actions, passing the results back to the view. A user can select any vessel from the interactive map initiating the prediction procedure (Figure 5)



**Figure 5.** Code map of MVC5 Home Controller and prediction function (generated by MS Visual Studio 2013 code analysis)

The "Prediction" function is capable of retrieving vessel data from the database, normalizing this, training the ANN and

returning the predicted geographical vessel position. Predictions from the neural network are transformed to the original data scale before positioned on the map at (http://mob0.marinetraffic.com/). The system architecture employs technologies that support the interoperability between loosely coupled components. In particular, the system design follows the principles of Service Oriented Architectures exposing SOAP or REST (web services) interfaces. Data from the vessel position database but also vessel predictions are returned as a JSON or XML web service.



**Figure 6.** Passenger vessel future position (latitude & longitude) prediction overlaid on an interactive map for user visualization (http://mob0.marinetraffic.com)

As opposed to going through a lengthy training process and only deploying the neural network after the process has been successful, we choose to train our network per vessel on user request in real time. When a user selects a vessel or specific geographic area, according to the required prediction, previous data is loaded. This data is fed into the neural network for training until its training rate is below 0.01. As this is a demanding CPU process, we have deployed a cloud infrastructure to support scaling on demand. Following this, the system will store the trained vessel ANN and recall it in the near future (24 hours) for following predictions. To guarantee the solutions quality of service and practical applicability a number of functional requirements had to be met. Several design choices were made to successfully address these as presented in the following table.

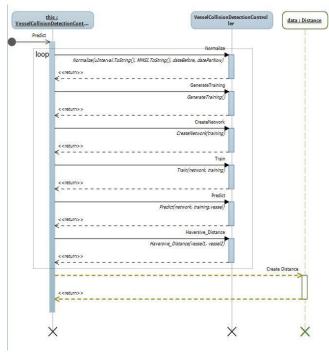
**Table 1.** Requirements mapped to design choices

	Table 1. Requirements i
REQUIREMENTS	DESIGN CHOICES
Solution Performance	Data used for training is limited to the previous 48-72 hours as a balance was necessary between training time and prediction accuracy. This was adequate for vessels performing repetitive patterns within a given timeframe.
	An ANN is trained on user request for a vessel of choice. Following this the system will store the trained vessel ANN and recall it in the near future (24 hours) for following predictions
Scalability of solution	ANN training is a CPU intensive and time-consuming process. We deploy our proposed system on a cloud infrastructure capable of scaling up and out on demand.
Availability & Reliability	The cloud infrastructure guarantee's high solution availability and resiliency. Additionally the solution is available as a Software as a Service (SaS) following the

	principles of Service Oriented architectures exposing SOAP or REST (web services) interfaces.
Complexity of ANN	Many behavioral characteristics are inferred so as to increase training speed, such as heading, course and speed. We implement the Haversive formula in c# and kinematic equations to calculate behavioral variables between positions

# B. Solution Case Study

Our proposal can potentially be used as the predictive foundation for various intelligent systems, including vessel collision prevention, vessel route planning, operation efficiency estimation and even anomaly detection. In this section, we present a case study investigating vessel collision detection. A user is permitted to select a geographical area containing a number of passenger vessels. In such circumstances, if no trained ANN is present for each vessel, the training process is initiated and a predicted vessel track is returned for each vessel. We compute a vessels bearing and speed between geographical points (Latitude and Londtitude). We call an implementation function of the Harversive formula in C#, that is capable of calculating the distance and bearing between two geographical points. If at any point, the calculated distance is smaller than a predetermined minimum an alert is generated by the system, informing the human operator. The same function is capable of predicting a vessels collision with ground when overloaded with the vessel prediction array and nautical information (an array of land coordinates). The following UML sequence diagram depicts the flow of messages between functions and other components in such a scenario.



**Figure 7.** Sequence Diagram depicting flow of messages in vessel collision detection scenario (generated by Visual Studio 2013 code analysis)

# IV. CONCLUSION

In this manuscript, we report on our ongoing work in adding predictive capacity to VTMIS and specifically MarineTraffic.com. In this, we describe our work on employing machine learning methods and specifically neural networks, as a basis for accurately predicting a vessels future behavior with an emphasis on solution practicality. To this end, we focus on deploying a web-based

infrastructure that can produce good results in low computational time. Further improvements are currently being performed in order to accommodate contextual information during the machine learning process.

## V. REFERENCES

- A. Ebada, M. A. M. (2005). Prediction of Ship Turning Manoeuvre Using Artificial Neural Networks. In 8th Numerical Towing Tank Symposium.
- Allianz Global Corporate. (2012). Safety and Shipping 1912-2012. From Titanic to Costa Concordia. An insurer's perpective from Allianz Global Corporate & Specialty.
- ANAVE. (2013). Merchant Marine and maritime transport 2012/2013.
- Azoff, E. M. (1994). Neural Network Time Series Forecasting of Financial Markets. Retrieved from http://dl.acm.org/citation.cfm?id=561452
- Bevilacqua, V. (2006). Hidden Markov Models for Recognition using Artificial Neural Networks. In *International Conference on Intelligent Computing: Intelligent computing* (p. 1331). Springer. Retrieved from http://books.google.com/books?id=7sH4RsXYu7cC&pgis=1
- Bomberger, N., Rhodes, B., Seibert, M., & Waxman, A. (2006). Associative Learning of Vessel Motion Patterns for Maritime Situation Awareness. In 2006 9th International Conference on Information Fusion (pp. 1–8). IEEE. doi:10.1109/ICIF.2006.301661
- Box, G.E.P., Jenkins, G. M. (1976). Time Series Analysis: Forecast- networks and robust time series prediction. *IEEE Transactioning*.
- Chadwick, J., Snyder, T., & Panda, H. (2012). Programming ASP.NET MVC 4: Developing Real-World Web Applications with ASP.NET MVC (p. 492). O'Reilly Media. Retrieved from http://www.amazon.co.uk/Programming-ASP-NET-MVC-Developing-Applications/dp/1449320317
- Connor, J. T. (1996). A robust neural network filter for electricity demand prediction. *Journal of Forecasting*, 15(6), 437–458. doi:10.1002/(SICI)1099-131X(199611)15:6<437::AID-FOR634>3.0.CO;2-H
- Cramer, R., Gennaro, R., & Schoenmakers, B. (1997). A Secure and Optimally Efficient Multi-Authority Election Scheme. In *EUROCRYPT'97*.
- Endsley, M. R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 32(2), 97–101. doi:10.1177/154193128803200221
- Galloway, J., Haack, P., Wilson, B., & Allen, K. S. (2012). *Professional ASP.NET MVC 4 (Wrox Professional Guides)* (p. 504). John Wiley & Sons. Retrieved from http://www.amazon.co.uk/Professional-ASP-NET-MVC-Wrox-Guides/dp/111834846X
- Gomes, G. S. da S., & Ludermir, T. B. (2013). Optimization of the weights and asymmetric activation function family of neural network for time series forecasting. *Expert Systems with Applications*, 40(16), 6438–6446. Retrieved from http://www.sciencedirect.com/science/article/pii/S0957417413003515
- Gomes, L., Faria, P., Morais, H., Vale, Z., & Ramos, C. (2014). Distributed, Agent-Based Intelligent System for Demand Response Program Simulation in Smart Grids. *IEEE Intelligent Systems*, 29(1), 56–65. doi:10.1109/MIS.2013.2
- Granger, C. W. J. (1993). Strategies for modelling nonlinear time-series relationships. The Economic Record, 69(206), 233–238.
- Heaton, J. (2008). *Introduction to Neural Networks for C#*, 2nd Edition (p. 428). Heaton Research, Incorporated. Retrieved from http://www.amazon.com/Introduction-Neural-Networks-2nd-Edition/dp/1604390093
- Heaton, J. (2011). *Programming Neural Networks with Encog3 in C#*, 2nd Edition (p. 240). Heaton Research, Incorporated. Retrieved from http://www.amazon.com/Programming-Neural-Networks-Encog3-Edition/dp/1604390263
- Hill, T. & Lewicki, P. (2007). (2007). What is Data Mining. In STATISTICS: Methods and Applications. StatSoft, Tulsa, OK.
- International Maritime Organization. (2009). Second IMO GHG Study 2009. 4 Albert Embankment, London SE1 7SR.
- Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10(3), 215–236. Retrieved from http://www.sciencedirect.com/science/article/pii/0925231295000399
- Karlaftis, M. G., & Vlahogianni, E. I. (2011). Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, 19(3), 387–399. Retrieved from http://www.sciencedirect.com/science/article/pii/S0968090X10001610

- Karlik, B., & Olgac, A. V. (2010). Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks. *International Journal of Artificial Intelligence and Expert Systems*, 1(4), 111.
- Krishnakumar, K. (2003). Intelligent systems for aerospace engineering An overview.
- Lagerweij, R., Vries, G. de, & Someren, M. van. (2009). Learning a Model of Ship Movements.
- Laxhammar, R. (n.d.). Anomaly detection for sea surveillance, 1–8. Retrieved from http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=4632192
- Laxhammar, R., Falkman, G., & Sviestins, E. (n.d.). Anomaly detection in sea traffic A comparison of the Gaussian Mixture Model and the Kernel Density Estimator, 756–763. Retrieved from http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=5203766
- Mostafa, M. M. (2004). Forecasting the Suez Canal traffic: a neural network analysis. *Maritime Policy & Management*, 31(2), 139–156. doi:10.1080/0308883032000174463
- MUNIN Collaboration. (2012). MUNIN Maritime Unmanned Navigation through Intelligence in Networks. Retrieved from http://www.unmanned-ship.org/munin/wp-content/uploads/2013/01/MUNIN-Brochure.pdf
- Nicolau, V., Aiordachioaie, D., & Popa, R. (2004). Neural network prediction of the wave influence on the yaw motion of a ship. In 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541) (Vol. 4, pp. 2801–2806). IEEE. doi:10.1109/IJCNN.2004.1381100
- Odom, M. D., & Sharda, R. (1990). A neural network model for bankruptcy prediction. In 1990 IJCNN International Joint Conference on Neural Networks (pp. 163–168 vol.2). IEEE. doi:10.1109/IJCNN.1990.137710
- Pankratz, A. (1983). Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. John Wiley, New York.
- Perera, L. P., Oliveira, P., & Guedes Soares, C. (2012a). Maritime Traffic Monitoring Based on Vessel Detection, Tracking, State Estimation, and Trajectory Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), 1188–1200. doi:10.1109/TITS.2012.2187282
- Perera, L. P., Oliveira, P., & Guedes Soares, C. (2012b). Maritime Traffic Monitoring Based on Vessel Detection, Tracking, State Estimation, and Trajectory Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), 1188–1200. doi:10.1109/TITS.2012.2187282
- Rhodes, B. J., Bomberger, N. A., & Zandipour, M. (2007). Probabilistic associative learning of vessel motion patterns at multiple spatial scales for maritime situation awareness. In 2007 10th International Conference on Information Fusion (pp. 1–8). IEEE. doi:10.1109/ICIF.2007.4408127
- Rothblum, A. M. (2002). Human Error and Marine Safety. In 2ND INTERNATIONAL WORKSHOP ON HUMAN FACTORS IN OFFSHORE OPERATIONS (HFW2002).
- Salchenberger, L. M., Cinar, E. M., & Lash, N. A. (1992). Neural Networks: A New Tool for Predicting Thrift Failures. *Decision Sciences*, 23(4), 899–916. doi:10.1111/j.1540-5915.1992.tb00425.x
- Santos, B., & Lunday, K. (2009). Maritime Domain Awareness. Coast Guard Journal of Safety & Security at Sea, Proceedings of the Marine Safety & Security Council, 66(2). Retrieved from http://trid.trb.org/view.aspx?id=900816
- Simsir, U., & Ertugrul, S. (2009). Prediction of manually controlled vessels' position and course navigating in narrow waterways using Artificial Neural Networks. *Applied Soft Computing*, *9*(4), 1217–1224. Retrieved from <a href="http://www.sciencedirect.com/science/article/pii/S1568494609000362">http://www.sciencedirect.com/science/article/pii/S1568494609000362</a>
- Smith, B. L., & Demetsky, M. J. (1994). SHORT-TERM TRAFFIC FLOW PREDICTION: NEURAL NETWORK APPROACH. *Transportation Research Record*, (1453). Retrieved from http://trid.trb.org/view.aspx?id=424677
- StatSoft, I. (2007). Electronic Statistics Textbook.
- Stopford, M. (2009). *Maritime Economics 3e* (p. 840). Routledge. Retrieved from http://www.amazon.com/Maritime-Economics-3e-Martin-Stopford/dp/041527558X
- US Department of Homeland Security. (2005). National Plan to Achieve Maritime Domain Awareness (MDA).
- Verber, D. (2012). Implementation of Massive Artificial Neural Networks with CUDA. In C. Volosencu (Ed.), *Cutting Edge Research in New Technologies*. InTech. doi:10.5772/2431

- Westrenen, F., & Praetorius, G. (2012). Maritime traffic management: a need for central coordination? *Cognition, Technology & Work*. doi:10.1007/s10111-012-0244-5
- Zhang, G., Eddy Patuwo, B., & Y. Hu, M. (1998). Forecasting with artificial neural networks: The State of the Art. *International Journal of Forecasting*, *14*(1), 35–62. Retrieved from http://www.sciencedirect.com/science/article/pii/S0169207097000447