ABSTRACT
Modeling vessel movement in a maritime environment is an extremely challenging task given the complex nature of vessel behavior. Several existing multiagent maritime decision making frameworks require access to an accurate traffic simulator. We develop a system using electronic navigation charts to generate realistic and high fidelity vessel traffic data using Generative Adversarial Networks (GANs). Our proposed Ship-GAN uses a conditional Wasserstein GAN to model a vessel’s behavior. The generator can simulate the travel time of vessels across different maritime zones conditioned on vessels’ speeds and traffic intensity. Furthermore, it can be used as an accurate simulator for prior decision making approaches for maritime traffic coordination, which used less accurate model than our approach. Experiments performed on the historical data from heavily trafficked Singapore strait show that our Ship-GAN system generates data whose statistical distribution is close to the real data distribution, and better fit than prior methods.

KEYWORDS
Generative Adversarial Networks; Maritime Traffic Simulation

2 MARITIME TRAFFIC DOMAIN
Figure 1 shows the electronic navigation chart (ENC) of Singapore strait. The waterway is divided into multiple zones, which can be
of different types such as TSS. Anchorages among others, as color coded in Figure 1. A zone is a polygon, not necessarily convex, with its typical length around 5 km. The traffic separation scheme (TSS) zones are the maritime highways which allow traffic to enter Singapore port, or transit through. The anchorage zones are regions which lead vessels to berths at the port terminal from TSS and vice-versa. By far, the bulk of the traffic is in TSS zones, and this is where dangerous traffic conditions can arise, sometimes leading to collisions. Therefore, our goal is to learn movement pattern of vessel in TSS zones. These are also the zones which have been used to regulate traffic in previous works [10]. Our traffic simulator (in Figure 2) is based on such geographical features present in the ENC.

Data Description: We have obtained the vessel data from the company Marinetrace thus gaining access to the AIS (Automatic Identification System) data from vessel AIS calls. The data captures 2.5 million AIS calls from about 9000 vessels in the Singapore Straits for one month. A typical vessel broadcasts AIS information every 2 minutes when in transit. The data provides information about position of vessels, time stamp, speed, course, heading, and other features (e.g., destination, vessel size).

Problem Statement: We aim to simulate travel times of ships across zones. The travel time is stochastic and depends on many factors that are unknown. The goal is to learn this travel time distribution for each zone (marginalized over unknown factors) using deep generative models conditioned on the vessel’s speed and the traffic ahead when entering a zone. The processed data is of the form $\langle r, (\nu, n_{t}(z)) \rangle$ where $r$ is the time taken to travel across the zone, $\nu$ is the speed of vessel while entering the zone and $n_{t}(z)$ the number of ships in zone $z$ at the time of entry which denotes the traffic intensity.

3 GENERATIVE MODELING OF TRAVEL TIME

For each zone, we learn both unconditional and conditional travel time distribution using generative models. There are two popular techniques in deep generative models: generative adversarial networks (GANs) [5] and variational auto-encoders (VAEs) [8]; GANs performed best for us. We used Wasserstein GAN [2], an improved variant of the original GAN.

Evaluation Metric: GANs are notoriously challenging to evaluate quantitatively [4]. The The Kolmogorov-Smirnov (KS) distance provides a measure of difference between two univariate distributions [7]. Table 1 shows the average KS distance for the GAN, VAE and other baselines based on distributions such as Binomial and Gaussian (used in previous works [10, 11]). These results are also averaged over all the 19 zones we consider. Lower KS values indicate a better fit to the real data. These results clearly show that GANs are a better fit than other baselines.

In Table 2, we extend the KS measure for the conditional generation of the data. The GAN and VAE take as input a vessel’s speed as it enters a zone, and the current traffic conditions (count of vessels in the selected zone), and generates a travel time. Such capabilities are not considered in previous simulation models [10, 11] which can only simulate travel time without taking into account such conditions. These results also show that GANs provide a better fit for conditional modeling of the data than VAE.

Figure 3 provides a qualitative comparison of the distributions learned for the unconditional case for two zones for each of the four methods in the unconditional distribution fitting. The x-axis is the travel time, and y-axis shows how many vessels had the travel time within a particular interval on the x-axis. The results are shown using 1000 samples each for real and fake distribution. The results reveal visually that GAN matches the real distributions much better than fitting known distributions. Also, between GAN and VAE, GANs fits much better than VAE.

Conclusion: We demonstrated the superior performance of deep generative model based simulation of key aspects of ship traffic in a given area via quantitative KS metric and qualitative results. Our approach enables more realistic simulations for decision making models for maritime traffic management. Our demo video is at: https://youtu.be/xZVzr0WcefU

Acknowledgment: This research is supported by the Ministry of Education, Singapore under its MOE Academic Research Funding Tier 2 (Award MOE2018-T2-1-179). Any opinions, findings, and conclusions are those of the author(s) and do not reflect the views of the Ministry of Education, Singapore.

Figure 2: A screenshot from our simulator based on ENC.

Figure 3: Unconditional fitting qualitative results
REFERENCES


