Decision Making for Improving Maritime Traffic Safety Using Constraint Programming

Saumya Bhatnagar , **Akshat Kumar** and **Hoong Chuin Lau**
School of Information Systems, Singapore Management University
{saumyab, akshatkumar, hclau}@smu.edu.sg

Abstract

Maritime navigational safety is of utmost importance to prevent vessel collisions in heavily trafficked ports, and avoid environmental costs. In case of a likely near miss among vessels, port traffic controllers provide assistance for safely navigating the waters, often at very short lead times. A better strategy is to avoid such situations from even happening. To achieve this, we a) formalize the decision model for traffic hotspot mitigation including realistic maritime navigational features and constraints through consultations with domain experts; and b) develop a constraint programming based scheduling approach to mitigate hotspots. We model the problem as a variant of the resource constrained project scheduling problem to adjust vessel movement schedules such that the average delay is minimized and navigational safety constraints are also satisfied. We conduct a thorough evaluation on key performance indicators using real world data, and demonstrate the effectiveness of our approach in mitigating high-risk situations.

1 Introduction

With the increase in maritime vessel traffic in recent years, navigational safety in port waters is of paramount importance. Busy waterways such as Tokyo Bay, Straits of Malacca and Singapore observe heavy movement of very large container and cargo ships (VLCC) on a daily basis. They are also susceptible to frequent traffic bottlenecks and congestion, thereby making vessel movement difficult and increasing the importance of maritime safety and collision avoidance [Evers and Gerke, 2006]. Maritime collision incidents endanger human lives, and frequently cause chemical and oil spillage, adversely affecting the marine ecosystem. Controlling oil spill is often among the top most contingency plans of a port authority [MPA Singapore, 2018]. While a number of studies have been conducted in the maritime domain that focus on future trajectory prediction [Hexeberg *et al.*, 2017], risk calculation and assessment [Balmat *et al.*, 2009], relatively less work has been conducted towards a proactive management of maritime traffic to prevent hotspots from forming in the first place. Our work addresses this gap, and also fits under the goal 14 of the United Nations Sustainable Development Goals [United Nations, 2016].

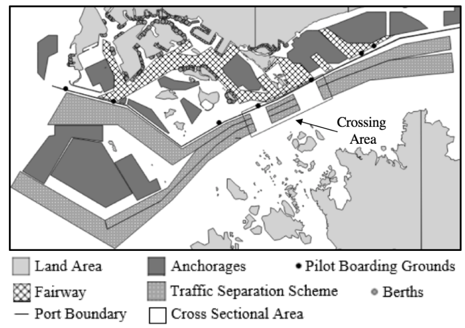


Figure 1: Waterway region showing demarcated Electronic Navigational Chart (ENC) features.

Domain setting: Figure 1 shows as an example the e-navigation chart for Singapore Straits. To better regulate vessel movement, the waterways are divided into different regions. Traffic Separation Schemes (TSS) are uni-directional lanes where all vessels in a given lane follow the same direction. Vessels traversing through fairways do not have restrictions on the directionality of movement. Other areas include anchorages, pilot boarding grounds, berths and crossing area. The crossing area is often susceptible to high traffic movements and hotspots since this area observes vessels traversing through the Straits as well as the ones that are making a transition from the Straits to the port waters and vice versa.

Contributions: We first formalize the decision making model for maritime safety after consultations with maritime domain experts from one of the busiest cargo ports of the world. Using insights from them, we aim to reschedule maritime traffic in crossing areas where vessels going in different directions pass through (as shown in Figure 1). The goal is to develop a maritime traffic scheduling approach to avoid formation of regions of high traffic intensity (hotspots) and near miss situations (when two or more vessels are within a very close and high risk proximity situation) while at the same time minimizing the delay incurred due to adjustments in the vessel schedule. We take into account maritime notions of safe navigation (as used by ship captains, such as cable lengths), and vessel movement constraints (speed limits, acceleration etc.). Next, we incorporate all such domain requirements into a resource constrained scheduling framework (RCPSP) [Hartmann and Briskorn, 2010]. The standard RCPSP model cannot be directly applied to our setting as it uses fixed activity durations, whereas activities (which denote vessel movements) have variable durations in our case. Multi-mode RCPSP allows for variable activity durations [Coelho and Vanhoucke, 2011], wherein each activity i can run in one of M_i modes each having a different duration. Typically, there are a few dozen different modes considered in the literature [Coelho and Vanhoucke, 2011]. For mode assignment, their scheme enumerates all activity mode variables for each activity. However, such an approach is computationally intractable in our case. An activity in the maritime case denotes the act of crossing a zone by a vessel. There will be a mode for every possible duration of an activity between T_{min} (minimum time to cross a zone) and T_{max} (maximum time to cross a zone). Such a strategy is infeasible as the difference between T_{min} and T_{max} is quite large due to time granularity being 2 seconds in our use case.

Use case: Our approach can be used as a decision support system by traffic controllers and port operators to help diffuse high-risk situations up to 30 minutes in advance. Cargo ships of 500 Gross Tonnage and above are required to provide a Pre-Arrival Notification (PAN) at least 24 hours before the ship's arrival in Singapore [MPA, 2019]. Using such a passage plan and historical movement data, a near term traffic forecast can be generated [Xiao *et al.*, 2017]. Such predicted traffic can be used to identify impending hotspots, and our scheduling approach can be triggered to recommend adjusted vessel movements to mitigate such hotspots. This strategy reduces the cognitive burden of port watch operators who can use the output schedule of our solver, and notify vessel captains as necessary.

Related work: Various studies have been conducted in the literature to develop safe maritime traffic management approaches. [Kelareva *et al.*, 2012] propose constraint programming and mixed-integer programming (MIP) models with the objective to maximize cargo throughput at a port. Their work takes into account constraints related to availability of tug boats and time-varying ship drafts that are environment-dependent. [Zhang *et al.*, 2017] propose a mathematical model for ship scheduling to minimize weighted average of mean and maximum waiting time for restricted waterways. However, their method assumes that all ships are navigating with the same speed, which may not always be true in all scenarios. Given deterministic shipping routes and sequence of ports to be visited, a shortest path approach to minimize fuel consumption was presented by [Fagerholt *et al.*, 2010]. All such previous approaches either restrict the problem setting (e.g., a single unidirectional waterway channel is assumed in Kelareva *et al.*), or optimize other parameters (such as fuel efficiency) rather than safety. In contrast, our approach models maritime traffic safety in one of largest cargo ports with special focus on crossing areas which are most dangerous.

In a related direction, [Agussurja *et al.*, 2018] propose MIP based models for congestion avoidance in the maritime domain for *macro-level* planning (24 hours in advance) for the entire waterway in Singapore Straits, and [Singh *et al.*, 2019] extend such macro-level planning to incorporate navigation under uncertainty. Our framework and problem setting have several differences from this previous work. *First*, their objective is to optimize overall traffic over the entire Straits. Mitigating short-term, localized hotspots, which is a primary concern of port watch operators, is not their goal. In contrast, our approach is tailored for such near term hotspots, and as a result is of immediate use to traffic controllers. *Second*, our work provides much more detailed modeling using several actual domain constraints. Such modeling aspects are not taken into account by Agussurja *et al.* Extending their Bender's approach with domain constraints we use is highly non-trivial. *Third*, their approach is impractical for our setting which requires near real-time solutions. Their MIP model requires a long time ($\approx$1hr) to get good solutions. Our proposed CP approach provides high quality solutions significantly faster (in less than 5 minutes for the majority of instances). *Lastly*, resource capacity is modeled as a hard constraint in their approach, whereas, we incorporate it as a soft penalty in our objective to minimize the peak resource usage. This is required as enforcing capacity as a hard constraint often makes the problem infeasible in our short-term scheduling case. Therefore, we minimize the peak usage of resources.

2 Solution Approach

The RCPSP is a well studied problem that aims to minimize the project duration (makespan), and consists of a set of tasks (or activities) and resources with a fixed capacity [Hartmann and Briskorn, 2010]. Each task has a duration and requires a given amount of resources at any given time. The resultant schedule must satisfy a set of precedence constraints that determine the order in which activities should be carried out. The cumulative consumption of any resource by different activities must not exceed the resource capacity at any time point. Unlike the RCPSP model where activity durations are known in advance, the duration of activities themselves are variables in our model. An activity in the maritime domain involves navigating from one sea zone to the neighboring one. Each activity may have certain lower and upper bounds (corresponding to minimum and maximum vessel speeds). Our goal is to optimize activity durations of each vessel subject to maritime domain constraints described in the next sections.

2.1 Zone Structure and Vessel Dataset

We first divide the planning area considered for scheduling into regular hexagons using the open source Quantum Geographical Information System (QGIS) application. The planning region considered in this work is the crossing area as shown in Figure 1. Our proposed approach is a generic one and can be applied to other planning areas as well. The width and height of the zones are kept as parameters. As per our discussion with domain experts, these values are kept as 555 metres which translates to 3 cable lengths [Wikipedia, 2018], a commonly followed precautionary distance among two vessels by operators in the maritime domain. This way, 216 hexagonal zones are generated for the planning area.

Vessel Dataset The data used in this work is a 4-month vessel dataset of Singapore Straits and Port Waters. Each file contains information of about 2700-3500 unique vessels and

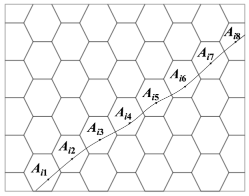


Figure 2: Sequence of activities A_{i1} to A_{i8} as the vessel i moves from one zone to next successive zone. Each zone is represented by a regular hexagon. Arrows indicate the direction of vessel movement.

an average of around 16 million records. Since the dynamic information, i.e., latitude, longitude, speed (SOG) and course of the vessels (COG) is not available at uniform time intervals in the raw data, an interpolation step is required before vessels trajectories can be constructed. In this work, we linearly interpolate the data on a 2-second interval basis, so that we have data points containing dynamic information at every subsequent time step. After data has been interpolated, we then construct trajectories of the individual vessels, where each trajectory is of the form:

$$\langle datapoint_1\rangle\ \langle datapoint_2\rangle........\langle datapoint_n\rangle$$

where each $datapoint_k$ at time step k is a 5-tuple:

$$\langle timestamp_k, longitude_k, latitude_k, SOG_k\ , COG_k\rangle$$

where SOG is vessel's speed-over-ground, and COG is the course-over-ground (or the direction of movement).

2.2 Problem Formulation

Our scheduling approach is modeled using constraint programming. Traversable sea space is considered as a resource, which is limited in capacity. As noted earlier, we divide the planning area into a grid of regular hexagon-shaped zones. Each zone is considered as a resource with capacity as 1. The length of the largest cargo vessels is around 300 meters, therefore, in our zones sized 555 meters in length, a single vessel fits safely. The resource limits for pilot boarding grounds (where a pilot boards the vessel to navigate it to a berth), and anchorages can be set accordingly.

Activities: Let R be the set of resources and V be the set of vessels. The set of vessels of the type tankers and cargos is represented by $V_{t,c} \subset V$. We primarily consider tanker and cargos for rescheduling as these are large vessels which are hard to manoeuvre at short notice. We still take the trajectories of other vessels (e.g., tugs, barges etc.) into consideration while rescheduling tanker and cargos.

During the time duration a vessel is present in a given zone, it is assumed to consume one unit of that resource. Whenever a vessel traverses through a zone, it is said to perform an activity. Thus, as the vessel moves through the planning area and crosses successive zones, it performs a set of activities A. Figure 2 shows an example of a vessel i performing activities A_{i1} to A_{i8} as it moves from one to the next consecutive zone. As mentioned earlier, we assume that the set of activities (and the time to finish such activities) performed by all vessels in the planning area is predicted ahead of time (e.g., 30 mins in advance). In an operational scenario, such a prediction can be done using previous approaches as [Xiao *et al.*, 2017]. For the testing of our model, we extract each vessel's activity from the historical data, and coordination is performed based on such activities. Notice that our approach only requires prediction accuracy at the level of hexagon-shaped zones, exact lat-long coordinates are not required. Thus, it is robust to minor inaccuracies in trajectory predictions.

CP variables: Let i be the index for vessels and j be the index for activities. The release time of a vessel i is denoted by $\mathrm{ReleaseTime}_i$ and is calculated as the start time of the vessel's first activity, i.e., the time at which the vessel first enters a zone in the planning area. Since we use IBM ILOG CP Optimizer to implement our model, we briefly describe the modeling concepts using terminology from [Laborie *et al.*, 2018]. An interval variable a is a decision variable used to represent activities to be scheduled over time whose domain dom(a) is a subset of $\{\perp\} \cup \{[s,e)|s,e \in Z, s \leq e]\}$. An interval variable is said to be fixed if its domain is reduced to a singleton, i.e., $a = \perp$ when interval is absent or $a = [s,e)$ if interval is present. In this case, s and e denote the start and end of interval and its length $l = e - s$. We use interval variables to represent each activity of a vessel, denoted as $\mathrm{act}_{i,j} \forall i \in V, \forall j \in A_i$, where A_i is the set of all activities of the vessel i. Our model does not use optional activity interval variables, therefore, $\{\perp\} \notin$ dom(a) in our case.

The *start time* of activity interval variables are the *decision variables* in our problem instance. Table 1 shows the declaration of CP interval variables in our model. We declare these activity interval variables with a range for the size in $[\mathrm{T}_{min}, \mathrm{T}_{max}]$, as shown in equation (1). $\mathrm{Speed}^{\mathrm{avg}}_{\mathrm{act}_{i,j}}$ is the average speed during activity $\mathrm{act}_{i,j}$. This parameter either can be predicted, or computed using historical data as in our case in (2). Let SOG^{ij}_{d} denote the historical speed obtained from data point d, and n denote the number of data points associated with the activity $\mathrm{act}_{i,j}$. Tankers and cargos having activity speed exceeding a certain value of $\mathrm{MinSpeed}$ are considered for scheduling. If a tanker/cargo's speed is below the $\mathrm{MinSpeed}$, based on our observations from the data, it implies that the vessel is slowing down to stop (e.g., in an anchorage or at a pilot boarding ground). Thus, during slowing down stage, we do not recommend any schedule adjustment.

For vessels to be rescheduled, the parameter T_{min} is computed from the distance travelled by the vessel during that activity, as obtained from data, denoted by $\mathrm{Distance}_{\mathrm{act}_{i,j}}$ and $\mathrm{MaxSpeed}$, which is a parameter in our model. The $\mathrm{MaxSpeed}$ refers to the maximum allowed safe speed in Straits. The upper bound T_{max} is similarly computed using the parameter $\mathrm{MinSpeed}$, which is the minimum cruising speed a vessel must maintain if it is not stopping (determined from data). If a vessel moves slower than $\mathrm{MinSpeed}$, it risks overturning due to water currents. Equations (3) and (4) denote this logic.

For activities of tankers and cargos, where the activity speed is less than or equal to $\mathrm{MinSpeed}$, we compute the historical time taken to complete these activities denoted by $\mathrm{T}_{hist_{i,j}}$. As shown in (5), these activity interval variables have T_{min} and T_{max} exactly equal to $\mathrm{T}_{hist_{i,j}}$. The duration of activity interval variables of other vessel types like tug boats and pilot boats is defined in a similar manner, as shown in (6). This is because these vessels are not considered for rescheduling but are still part of the problem instance and

$$\mathrm{act}_{i,j} = intervalVar(\mathrm{T}_{min}, \mathrm{T}_{max}) \quad (1)$$

$$\mathrm{Speed}^{\mathrm{avg}}_{\mathrm{act}_{i,j}} = \sum_{d=1}^{n} (\mathrm{SOG}^{ij}_{d})/n \quad (2)$$

$\forall i \in V_{t,c}, \forall j \in A_i \mid \mathrm{Speed}^{\mathrm{avg}}_{\mathrm{act}_{i,j}} > \mathrm{MinSpeed}:$

$$\mathrm{T}_{min} = (\mathrm{Distance}_{\mathrm{act}_{i,j}}/\mathrm{MaxSpeed}) \cdot 3600 \quad (3)$$

$$\mathrm{T}_{max} = (\mathrm{Distance}_{\mathrm{act}_{i,j}}/\mathrm{MinSpeed}) \cdot 3600 \quad (4)$$

$\forall i \in V_{t,c}, \forall j \in A_i \mid \mathrm{Speed}^{\mathrm{avg}}_{\mathrm{act}_{i,j}} \leq \mathrm{MinSpeed}:$

$$\mathrm{T}_{min} = \mathrm{T}_{hist_{i,j}}, \mathrm{T}_{max} = \mathrm{T}_{hist_{i,j}} \quad (5)$$

$\forall i \in V \setminus V_{t,c}, \forall j \in A_i:$

$$\mathrm{T}_{min} = \mathrm{T}_{hist_{i,j}}, \mathrm{T}_{max} = \mathrm{T}_{hist_{i,j}} \quad (6)$$

Table 1: CP Variable Declaration

hence, their activities require resources. E.g., a tug boat and a cargo vessel cannot be present in the same zone at the same time.

CP Model Validation: The goal of our CP model is that vessels reach their intended position with minimum delay (incurred due to recommended schedule adjustments) while at the same time congestion is reduced. As metrics for congestion, we measure resource violations and Distance at Closest Point of Approach (DCPA) in CP solver's output schedule, and compare them with respective values before coordination. Resource violations happen when the cumulative consumption of a resource at a timepoint exceeds its capacity. DCPA is defined as the minimum distance between two vessels at any timepoint throughout their trajectories. As discussed with domain experts, DCPA is often considered as a contributing factor for risk, and higher DCPA is better. We also take into account that navigational constraints that are significant from an operational point of view are satisfied in the output schedule. These include minimum and maximum vessel speeds, maximum speed change, vessels should not stop mid-TSS and limiting the number of vessel having schedule adjustments. The last constraint is highly desired to reduce the cognitive burden on operators, and ease of real world implementation.

CP Model: Table 2 shows the CP objectives and constraints. Different notations are explained next. δ_k denotes weight parameter penalising resource violations of zone k. y_i is an array of boolean variables where element value is 1 if the change in the overall schedule of vessel i is greater than or equal to γ, and is 0 otherwise. γ denotes the value in seconds at and after which a vessel is considered delayed and is parametrized in our model. The cumulative consumption of the resource k by all activities at any given time is given by $\mathrm{cumResource}_k$, and $\mathrm{Capacity}_k$ denotes the capacity of zone/resource k. To limit the number of vessels that have a change in their schedule, we use the limit parameter θ. $\mathrm{changeExpr}_i$ denotes the change due to rescheduling of an individual vessel i. The parameter for minimum speed and maximum speed of tankers and cargos is given by $\mathrm{MinSpeed}$ and $\mathrm{MaxSpeed}$; $\mathrm{time}_{\mathrm{act}_{i,j}}$ is a variable that denotes the time taken by the vessel i to complete its activity j.

We now describe built-in constructs provided by CP Optimizer that we use in our model. a represents an interval variable and s and e denote its start and end respectively.

$$\mathrm{Minimize}(\sum_{i}^{|V|} \mathrm{le}_i + \sum_{k}^{|R|} (\delta_k \cdot \mathrm{cost}_k)) \quad (7)$$

$$\forall i \in V: \mathrm{le}_i = \mathtt{endOf}(\mathrm{act}_{i,|A[i]|-1}) \quad (8)$$

$$\forall i \in V: \mathrm{changeExpr}_i = \mid \mathrm{le}_i - \mathrm{HistCT}_i \mid \quad (9)$$

$$\forall i \in V: \mathrm{ReleaseTime}_i = \mathtt{startOf}(\mathrm{act}_{i,0}) \quad (10)$$

$$\forall i \in V, \forall j \in |A_i|-1: \mathtt{endAtStart}(\mathrm{act}_{i,j}, \mathrm{act}_{i,j+1}) \quad (11)$$

$\forall k \in R: \mathrm{cumResource}_k =$

$$\sum_{i}^{|V|} \sum_{j}^{|A_i|} \mathtt{pulse}(\mathrm{act}_{i,j}, \mathrm{Consumption}^{k}_{\mathrm{act}_{i,j}}) \quad (12)$$

$$\forall k \in R: \mathrm{cost}_k \geq 0 \quad (13)$$

$$\forall k \in R: \mathrm{cumResource}_k \leq \mathrm{Capacity}_k + \mathrm{cost}_k \quad (14)$$

$\forall i \in V: \mathtt{ifThenElse}$

$$(\mathrm{changeExpr}_i \geq \gamma, \mathrm{y}_i = 1, \mathrm{y}_i = 0) \quad (15)$$

$$\sum_{i}^{|V|} \mathrm{y}_i \leq \theta \quad (16)$$

$$\forall i \in V_{t,c}, \forall j \in A_i: \mathrm{time}_{act_{i,j}} = \mathtt{lengthOf}(\mathrm{act}_{i,j})/3600 \quad (17)$$

$\forall i \in V_{t,c}, \forall j \in |A_i|-1:$

$$|\mathrm{speed}_{\mathrm{act}_{i,j}} - \mathrm{speed}_{\mathrm{act}_{i,j+1}}| \leq \mathrm{MaxChange} \quad (18)$$

$$\forall i \in V_{t,c}, \forall j \in A_i: \mathrm{speed}_{\mathrm{act}_{i,j}} = \mathrm{Distance}_{\mathrm{act}_{i,j}}/\mathrm{time}_{\mathrm{act}_{i,j}} \quad (19)$$

Table 2: CP for the Maritime Traffic Scheduling problem

- $\mathtt{startOf}(a)$ and $\mathtt{endOf}(a)$ are functions provided by CP optimizer that return integer expressions representing the start and end of interval variable a respectively whenever the interval variable is present and 0 otherwise.
- $\mathtt{endAtStart}(a, b)$ restricts the relative temporal position of interval variables a and b such that b starts exactly at the end of a.
- $\mathtt{lengthOf}(a)$ returns the length of interval variable a, i.e., $e - s$ if $a = [s, e)$.
- $\mathtt{pulse}(a, h)$ is a CP Optimizer function F such that $F(t) = h$ if $t \in [s, e)$ and $F(t) = 0$ otherwise. Here, s and e denote the start and end of interval variable a.
- $\mathtt{ifThenElse}(c1, c2, c3)$ is a method that enforces the constraint that if the first constraint $c1$ is satisfied, it implies the second constraint $c2$ and if $c1$ is violated, it implies the third constraint $c3$. More formally, $(c1 \Rightarrow c2)$ and $(\neg c1 \Rightarrow c3)$.

The first objective term is to minimize the sum of last activity end times of all vessels in the problem instance. This ensures that vessels are not delayed unnecessarily. Let le_i denote the end time of last activity of vessel i and it is a variable expression in our model. This is computed as shown in equation (8). For a given vessel i, we define historical completion time HistCT_i as the sum of its release time and the historical processing time of all its activities $\mathrm{T}_{hist_{i,j}}$ as derived from data. The change due to rescheduling of this vessel, $\mathrm{changeExpr}_i$, is then calculated as the absolute difference between the last activity end time and the historical completion time, as shown in equation (9). Based on our testing with historical data, we observed that almost all vessels are delayed (i.e. asked to slow down from their historical speeds) in the cross sectional area. As a result, checking last activity end time is sufficient to capture if a vessel's schedule has changed.

We experimented with computing the change expression (9) for each activity/vessel also. It made solver run slower ($\gg$5 mins), and did not impact the final solution. The second objective term is to minimize the resource violations, i.e., total number of time periods when the aggregate resource consumption by all vessels exceeds the capacity of that resource. We put this risk-based term in the objective since making this as a hard constraint can make the problem infeasible (in any feasible solution that respects hard constraints such as speed limits, some resource violations may always occur). Being a bi-objective optimization problem, we formulate it using weighted sum method [Marler and Arora, 2010] with 1 and δ_k being the scalar weights for first and second terms in the objective function respectively. The value of δ_k was determined by testing with different values on several historical near miss cases provided by them. The final objective is as shown in the equation (7). The constraints formulated in the model include the following:

- **Constraint for Release Time**: Constraint (10) specifies that the first activity of a vessel must begin exactly at its release time.
- **Precedence Constraint**: For any two consecutive activities of a vessel, the next activity of the vessel can begin only and exactly at the end of its first activity. Constraint (11) is used to denote this. This is to enforce that there is no time lag between consecutive activities of the same vessel.
- **Capacity Constraint**: This constraint is used to specify that the cumulative consumption of a resource by all vessels at any given time point should not exceed its resource capacity. This is implemented as a soft constraint by adding it as a part of objective function with an appropriate penalty term. Introducing it as a hard constraint results in infeasible model (no solution) as this conflicts with other hard constraints such as release time and precedence constraints. The resource consumption and capacity constraint are modelled using CP Optimizer's pulse function. Since in our formulation, we assume each vessel is occupying 1 unit of resource, the value of $\mathrm{Consumption}^k_{\mathrm{act}_{i,j}}$ is equal to 1, if vessel i is in zone k during its activity j and 0 otherwise. For every resource k, we have $\mathrm{cumResource}_k$ which is used to store the sum of individual pulse contributions of activity interval variables.

  By adding cost_k in the objective function, we minimize the peak usage of the resource k. It does not take into account how many times the usage of zone k exceeds its capacity. Nevertheless, our computational results in Section 3.1, Figure 3 clearly show reduction in resource violations before and after the scheduling. The constraints (12), (13) and (14) are used to model capacity constraint.
- **Limit Constraint**: Constraints (15) and (16) are maritime domain specific constraints used to specify an upper bound on the total number of vessels that may have a change in schedule. This is significant as the maritime port operators may wish to recommend speed changes to only a few number of vessels and not all the vessels.
- **Speed Constraint**: The minimum and maximum speed constraints are domain specific and are applied only to vessels of the type tankers and cargos. Equations (3) and (4) in Table 1 are used to enforce limit values on maximum and minimum duration of activities of tankers and cargos.
- **Speed Change Constraint**: Constraint (18) is also domain specific and is applied to only vessels of the type tankers and cargos. It is used to specify limit values on successive speed changes of tankers and cargos. We model this as absolute difference in vessel speed for successive activities. This is significant since the port operators recommend gradual changes in vessel speed rather than sudden speed variations. Equation (17) denotes the time taken to complete the activity, denoted by $\mathrm{time}_{\mathrm{act}_{i,j}}$. Equation (19) denotes the calculation for activity speed $\mathrm{speed}_{\mathrm{act}_{i,j}}$. This term is formulated as a variable expression in our model.

3 Experiments

For each day, we divide 24 hours into 48 slots of 30 minutes each. We then find out the busiest half-hour slot (maximum number of vessels) for each day for the planning area. We generated a total of 64 instances which represented hotspot situations. Minimum and maximum number of vessels in the problem instances tested is 27 and 53 respectively. Average number of activities per vessel is 15. Number of vessels scheduled is 13 (averaged over all 64 instances). Total number of zones in the planning area considered is 216.

For our experiments, the planning area considered is the cross sectional area as shown in Figure 1. This region is along the Traffic Separation Schemes (TSS) and port boundary. This area is considered to be a static hotspot that observes high traffic movements throughout the day, and therefore is susceptible to congestion and near miss situations.

Configurable Hyperparameters: We set different configurable parameters of our model (used in Table 2) after analyzing the data. More details about them are in the supplemental material[1]. We set CP optimizer's maximum allowed runtime to 10 minutes. In practice, CP solver provided a good solution within 2-3 minutes for most instances, and spent rest of the time in proving the optimality of the solution.

3.1 Experimental Results

The experimental results and criteria used to evaluate our proposed approach are described next.

Vessel Delay: We record the maximum and average delay incurred after scheduling using our approach. Delay of a vessel is computed as the difference between end time of last activity of the vessel (variable in the formulation) and historical completion time. This is the additional time a vessel takes after getting rescheduled. We then take maximum and average over delays of all vessels in that problem instance. Figure 4(a) shows the maximum and average vessel delays for the top 30 busiest days (instances having the maximum number of vessels). Results show that for 30-minute simulations, the maximum delay for these days comes out as less than 12 minutes. The average delay is about 5-6 minutes, which shows that only minor schedule adjustments are needed. In Figure 4(b), we show the maximum vessel delay on the X axis and

[1] https://www.dropbox.com/s/08midi341njh3tj/appendix.pdf?dl=0

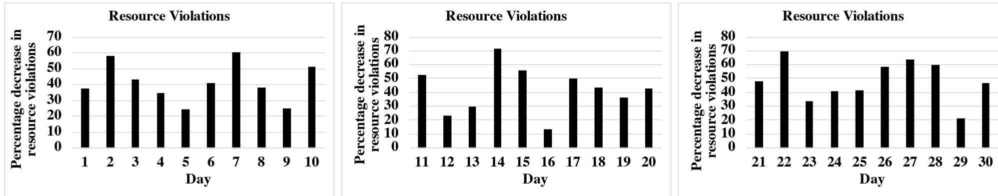
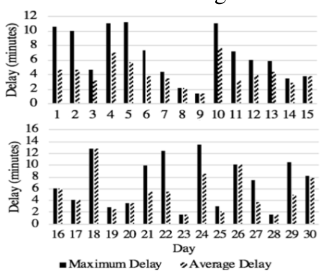
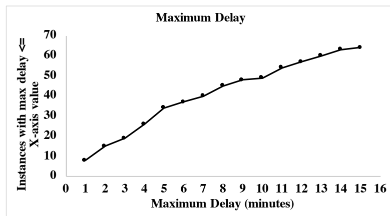


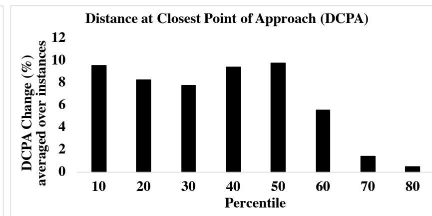
Figure 3: Plots for percentage decrease in resource violations for 30 busiest days.



(a) Delay for 30 busiest days.



(b) Plot for maximum vessel delay.



(c) Distance at Closest Point of Approach

Figure 4: Vessel Delay and DCPA Plots

the number of instances that have a maximum delay of less than or equal to the corresponding X axis value. As noted from the figure, for around 50 instances, the maximum delay comes out to be less than or equal to 10 minutes.

Maximum Resource Violations: This is regarded as a key performance indicator for congestion/hotspot. Resource violations happen when cumulative consumption of a resource at a timepoint exceeds its capacity. We compute the resource violations from the historical data (say v_{hist}) for each instance, and the resource violation for the new schedule computed by the CP optimizer (say v_{cp}). We then plot the percentage improvement in the resource violation by our method or $(v_{hist} - v_{cp})/v_{hist}$ in Figure 3 for 30 busiest days. The results clearly show that resource violations are decreased significantly by our method consistently across all test instances. The detailed results for delay and resource violations for the remaining days are included in the supplemental material.[1]

Distance at Closest Point of Approach (DCPA): This is defined as the minimum distance between two vessels at any timepoint throughout their trajectories. The details of computation are presented in the supplemental material.[1] As per domain experts, DCPA is often considered as a major contributing factor for navigation risk. Although we do not directly optimize DCPA, minimizing resource violations is observed to increase DCPA. For every problem instance, we calculate pairwise DCPA of every vessel pair for before (d_{hist}) and after coordination (d_{cp}). We then compute the different percentiles values from this distribution, such as 10^{th}, 20^{th} and 80^{th} percentile. Next, we compute the percentage *increase* in DCPA $(d_{cp} - d_{hist})/d_{hist}$ at every percentile value (higher DCPA is better). Figure 4(c) shows this plot. X axis denotes the different percentile values and Y axis values denotes the corresponding DCPA percentage change after doing an average over top 30 busiest days/instances. Positive DCPA change values indicate risk reduction after scheduling using our approach. Improvement in DCPA is higher for lower percentiles, which is a positive sign as we want to increase the minimum distance among vessels after coordination.

Runtime: Time taken by CP solver to generate a solution is of critical importance from the perspective of end users. For 44 out of 64 test instances, CP optimizer provided optimal solution within 5 minutes with proof of optimality. For rest of the instances, CP solver provided a good feasible solution within 5 minutes, and spent the rest of the time to prove its optimality. For these 20 instances, the maximum average delay is ≈6 minutes using the solution within 5 minutes. The average decrease in resource violations is $\approx 50\%$. These results show that our approach is fast enough to be used in near real-time by the end users.

Overall, our results show that there is a significant potential to increase maritime safety of navigation by adjusting schedule of vessels while also satisfying several operational constraints. The fact that DCPA was consistently higher after coordination even in a small crossing region was considered highly positive. Such safer navigation combined with low delay for vessels in resulting schedules validate the key promises of our model.

4 Conclusion

We developed a CP based scheduling approach to reduce traffic hotspots in maritime domain. We modeled the problem as a variant of resource constrained scheduling with the objective to minimize resulting vessel delay due to changes in schedule, and resource violations, which is a measure for congestion. We also incorporated realistic maritime navigational constraints. Empirical tests on a real dataset containing positional information of vessel movements showed significant potential of our approach to make maritime navigation safer, thereby protecting marine ecosystem.

Acknowledgments

This research is supported by the Agency for Science, Technology and Research (A*STAR), Fujitsu Limited and the National Research Foundation Singapore as part of the A*STAR-Fujitsu- SMU Urban Computing and Engineering Centre of Excellence.

References

[Agussurja *et al.*, 2018] Lucas Agussurja, Akshat Kumar, and Hoong Chuin Lau. Resource-constrained scheduling for maritime traffic management. In *AAAI Conference on Artificial Intelligence*, 2018.

[Balmat *et al.*, 2009] Jean-Francois Balmat, Frederic Lafont, Robert Maifret, and Nathalie Pessel. Maritime risk assessment (marisa), a fuzzy approach to define an individual ship risk factor. *Ocean engineering*, 36(15-16):1278–1286, 2009.

[Coelho and Vanhoucke, 2011] Jose Coelho and Mario Vanhoucke. Multi-mode resource-constrained project scheduling using RCPSP and SAT solvers. *European Journal of Operational Research*, 213(1):73 – 82, 2011.

[Evers and Gerke, 2006] Hans-Dieter Evers and Solvay Gerke. The strategic importance of the Straits of Malacca. 2006.

[Fagerholt *et al.*, 2010] Kjetil Fagerholt, Gilbert Laporte, and Inge Norstad. Reducing fuel emissions by optimizing speed on shipping routes. *Journal of the Operational Research Society*, 61(3):523–529, 2010.

[Hartmann and Briskorn, 2010] Sönke Hartmann and Dirk Briskorn. A survey of variants and extensions of the resource-constrained project scheduling problem. *European Journal of operational research*, 207(1):1–14, 2010.

[Hexeberg *et al.*, 2017] Simen Hexeberg, Andreas L Flåten, Edmund F Brekke, et al. AIS-based vessel trajectory prediction. In *Information Fusion (Fusion), 2017 20th International Conference on*, pages 1–8. IEEE, 2017.

[Kelareva *et al.*, 2012] Elena Kelareva, Sebastian Brand, Philip Kilby, Sylvie Thiébaux, Mark Wallace, et al. CP and MIP methods for ship scheduling with time-varying draft. In *ICAPS*, 2012.

[Laborie *et al.*, 2018] Philippe Laborie, Jérôme Rogerie, Paul Shaw, and Petr Vilím. IBM ILOG CP Optimizer for scheduling. *Constraints*, 23(2):210–250, 2018.

[Marler and Arora, 2010] R Timothy Marler and Jasbir S Arora. The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*, 41(6):853–862, 2010.

[MPA Singapore, 2018] MPA Singapore. Over 250 participate in Joint Oil Spill Exercise to Test Responsiveness to Oil Spills at Sea. https://www.mpa.gov.sg/web/portal/home/media-centre/news-releases/mpa-news-releases/detail/091cd124-ca60-4f34-bdb6-a0967f82defd, 2018.

[MPA, 2019] MPA. Pre-Arrival Notification (PAN) . https://www.mpa.gov.sg/assets/app/ePANS/epan.html/, 2019.

[Singh *et al.*, 2019] Arambam James Singh, Duc Thien Nguyen, Akshat Kumar, and Hoong Chuin Lau. Multi-agent decision making for maritime traffic management. In *AAAI Conference on Artificial Intelligence*, 2019.

[United Nations, 2016] United Nations. Conserve and sustainably use the oceans, seas and marine resources. https://www.un.org/sustainabledevelopment/oceans/, 2016.

[Wikipedia, 2018] Wikipedia. Cable Length. https://en.wikipedia.org/wiki/Cable_length, 2018.

[Xiao *et al.*, 2017] Zhe Xiao, Loganathan Ponnambalam, Xiuju Fu, and Wanbing Zhang. Maritime traffic probabilistic forecasting based on vessels' waterway patterns and motion behaviors. *IEEE Transactions on Intelligent Transportation Systems*, 18(11):3122–3134, 2017.

[Zhang *et al.*, 2017] Jinfen Zhang, Tiago A Santos, C Guedes Soares, and Xinping Yan. Sequential ship traffic scheduling model for restricted two-way waterway transportation. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 231(1):86–97, 2017.