Optimizing Maritime Vessel Service Time with Adaptive Quay Crane Deployment Through Level 4 Hard-Soft Information Fusion

Ashwin Panchapakesan
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa
Email: apanc006@uottawa.ca

Rami Abielmona
Research and Engineering Division
Larus Technologies
Ottawa, Canada
Email: rami.abielmona@larus.com

Emil Petriu
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa
Email: petriu@uottawa.ca

***Abstract*—Commercial maritime ports must maintain high service throughput in order to remain profitable. One of the most critical operations in a commercial maritime port is the loading and unloading of shipping containers on a vessel (i.e. servicing a vessel) and storing them in the storage yard. A delay in this process would cause cascading delays in servicing further vessels, causing delays in moving cargo across land, rail, and sea. Furthermore, the port itself may incur fines for allowing such delays in their operational procedure. This work highlights a Fuzzy System optimized by a Genetic Algorithm to adaptively control the deployment of quay cranes (and their operators and all other supporting equipment and personnel) to optimize the time required to service a vessel, while simultaneously reducing the operational costs of doing so.**

I. Introduction

Seaborne trade accounts for approximately 90% of global trade [1]. Therefore, seaports must operate at optimal efficiency in order to not hinder trade, but also to remain profitable. This specifically includes the turnaround time of a berthed vessel, during which period, shipping containers are removed from the vessel (and subsequently stored in the storage yard) and new containers from the storage yard are loaded onto the vessel. The total time between when a vessel berths (at which point, the container unloading process presumably begins) and when it leaves the berth (after it has been loaded with new shipping containers) is the turnaround time or vessel service time. During its service time, the vessel occupies berth space, blocking any other vessel from berthing there. As a result, slower service results in an increase in the backlog of vessels that require servicing. Such delays in vessel service slows down global trade and incur fines for the port as well. This problem can be solved by increasing the resources deployed to service the vessel in order to reduce the service time. However, this comes at an added cost to operate the required equipment and wages to pay the required personnel.

While a fixed, one-time optimization will provide a contextually appropriate resource deployment scheme, such an optimization would become less optimal over time, due to changes in vessel arrival schedules and other transient contextual properties. Therefore a constantly adaptive system is called for. Systems with such capabilities fall under Level 4 of the JDL/DFIG data fusion model, and such a system is developed in this work. While Level 3 deals with generating courses of action based on projected assessed impact (in this case, vessel service delay), Level 4 of the JDL/DFIG data fusion model (Process Refinement) considers the historic performance of the course-of-action generator and optimizes the algorithm that generates courses of action to be more effective in the future. For example, a Level 3 algorithm could be a hand-tuned fuzzy system to optimally deploy resources to service incoming vessels. Yet, as the schedule intensifies (or undergoes other relevant changes), the optimized fuzzy systems may perform less optimally. Simultaneously, a Level 4 algorithm might be a meta-heuristic algorithm that optimizes the parameters of the Level 3 fuzzy system using historic data, to optimize the fuzzy system so that the overall system's (the Level 3 fuzzy system and the Level 4 fuzzy-system optimizer) performance may continue to be optimal over time.

In this work, we explore a fuzzy system with a fixed rule base to optimize the resource deployment to service incoming vessels. The variable parameters of this fuzzy system (namely, the parameters of the fuzzy membership functions) are optimized with a genetic algorithm (GA). Using a dataset comprised of real-world data and real-world-inspired synthetic data (which include both hard and soft data), the optimization has been shown to outperform current industry practice. While other factors such as weather and sea state would also influence the arrival and departure times (and therefore delays) of a given vessel, those factors are out of the scope of this work and are good candidates for future study.

The remainder of this paper is organized as follows: a survey of current methods is presented in Sec. II and the generation of the dataset is described in Sec. III-A. The optimization is explained in Sec. III-B, the results of which are discussed in Sec. IV before concluding with some directions of future research in Sec. V.

II. Previous Work

While ports do have outbound land and/or rail traffic, global shipping container traffic is primarily through maritime vessel traffic. When a vessel does enter a port, it must berth at one of many berthing locations at the port. The optimization of scheduling berths to incoming (and outgoing) ships based on their schedules is the well-known Berth Allocation Problem (BAP) [2], of which there are three major variants. Ultimately, BAP is related to resource allocation problems involving the various personnel and equipment at port, in the optimization of processing each vessel, thereby optimizing port throughput.

Berthing Space

The berthing locations of a vessel entering port may be discrete or continuous. Each has its own advantages and disadvantages in port operation optimization.

Discrete Berthing: A discrete berth is a section of the port at which a vessel may berth. It is also surrounded by an area in which the vessel may not berth [2]. Thus, the berthing locations are discrete along the port's quay [2].

Continuous Berthing: A continuous berth is a section of the port at which a vessel may berth. It is typically larger than any one vessel will require for berthing space, and can accommodate multiple ships at once. Additionally, a dynamic berth allows for a vessel to berth anywhere within it, which allows for multiple ships of various sizes to berth in an ad hoc manner, without having to allocate a separate berth for each [2].

Vessel Arrivals

Typically, approximate vessel arrival times are known by the destination port, and the captain of each vessel continually updates the destination port on updated and more accurate arrival times. As such, the literature discusses two vessel arrival schedules, namely static and dynamic.

Static Vessel Arrivals: Static vessel arrivals refer to vessel arrival schedules that are known a priori, a generalization over having all processable vessels within the port's waters [3]. Despite industry-wide behavior of vessels periodically announcing their updated arrival times as they approach a port, the eventual actual arrival time is rarely known a prori with both high accuracy and high confidence. This is because vessel arrival schedules are subject to change, given weather patterns and other operational delays, which necessitate the sending of an updated arrival schedule. This in turn causes operational delays on the port's side.

Dynamic Vessel Arrivals: Dynamic vessel arrivals refer to vessel arrival schedules that are not known a priori, a generalization over having only a fraction of the processable vessels within the port's waters [3]. This also refers to uncertainties in vessel arrival schedules and the induced necessity to modify vessel handling processes on the fly, as vessels arrive.

Vessel Handling Times

Vessel handling time refers to the amount of time required to discharge, maintain, and reload a vessel, i.e. the amount of time from when the vessel berths, to when the vessel leaves the berth (and subsequently, the port).

Static Vessel Handling Time: Static vessel handling time refers to cases in which the handling time for each vessel is known before the arrival of the vessel [3]. These are therefore considered inputs to any optimizer.

Dynamic Vessel Handling Time: Dynamic vessel handling time refers to cases in which the handling time for each vessel is not known beforehand and must therefore be computed based on the vessel attributes and environment variables [3]. These are therefore not considered inputs to an optimizer.

A. Vessel Loading and Unloading and Container Storage

Vessel loading and unloading are at the core of port operations, and require much operational time (this is vessel handling time). Therefore, optimizing vessel handling times promises to yield improvements in port operations. Indeed, the vessel handling operations are associated with their own optimization parameters, which rely on the optimal storage conditions of the shipping containers, the presence of hazardous materials, the duration of time for which the containers are expected to stay within the port before being loaded onto another ship, or rail or a truck for delivery. For instance, storing containers by outbound date causes intra-port traffic bottlenecks, leading to suboptimal performance [4], [5], [6]. Similarly, vessel loading and unloading are well studied problems, that account for vessel balance and cargo priority [7], [8]. Note that while the literature typically discusses containers in twenty-foot equivalent units (TEUs), the specific operational handling times of shipping containers are agnostic to the actual shipping container size, as long as the collection of shipping containers to be processed is homogeneously sized. Homogeneous sizing allows for equipment to remain at a single setting pertaining to container size throughout the vessel service duration, as opposed to necessarily altering crane bracket sizing, etc in response to changes in the sizes of processed containers.

While others have attempted to solve this problem with Mixed Integer Programming (MIP) solutions, these solutions suffer from a "planning horizon" [9] and can therefore not be deployed in a continuous manner. Further, they force a discretization on equipment operating time, making the model less flexible. Other solutions formulate this as a processor scheduling problem [10]. Yet, both approaches are *offline* and require prior knowledge of vessel schedules. In contrast, the methodology proposed in this work is continuous and can function in an *online* manner, and can therefore tolerate vessel delays and other operational faults.

III. Methodology

In order to optimize vessel service, a dataset containing information on the vessels to be serviced is first required. Data in this dataset must include the movement of serviced vessels into and out of ports, the number of shipping containers that were moved onto and off the vessel during vessel service, and the total service time. These attributes are required for each vessel in the optimization. As well, these attributes are required to be known as ground truth in order to establish the effectiveness of the optimization. The creation of this dataset is discussed in Sec. III-A.

Once this dataset is created, a GA-optimized fuzzy system is used to optimize resource deployment on this dataset. The implementation of the optimization is discussed in Sec. III-B2. Since the fuzzy system itself is optimized by a GA by analyzing historic data, the GA optimization can be run periodically to ensure the consistent optimality of the deployed fuzzy system.

A. Dataset Generation

Maritime vessels use Automated Information System (AIS) messages to broadcast critical operational information about themselves. Such information includes time-stamped geospatial coordinates, their speed, heading, destination, and their estimated arrive time (ETA) at the destination [11]. Such AIS information is available from sources such as Lloyd's register [12], MarineTraffic [13], etc. From this information, the trajectory of each vessel (i.e. its track) was correlated and vessels that stopped at known ports (i.e. had a reported speed of under 5 knots [14], within sufficient proximity of the cranes at the port). The total duration for which vessels were stopped is noted as their service time (an example vessel track is shown in Fig. 1 and the mined vessel service times are illustrated in Fig. 8).

The ports also publish their incoming vessel schedule (an example is seen in [15]), which was then used to filter out unscheduled vessels from the unfiltered AIS data. These provide the hard (structured) data, which is then used in the optimization.

The AIS messages also contain vessel draught information, along with its physical dimensions (length and width). This is used to calculate the volume of water displaced (as seen in Eq. 1 and Eq. 2), which when multiplied by the density of water gives the mass of displaced water (as seen in Eq. 3). This is the total mass of the vessel including fuel, crew, and the mass of the shipping containers on board. Assuming that the remaining weights are negligible in the absence of more specific information (or otherwise allowing for such values within a margin of error), the mass of the displaced water is the combined mass of the shipping containers on board the vessel. The number of shipping containers on board (given the draught and physical dimensions of a vessel) is therefore computed by dividing the mass of displaced water by the average mass of a shipping container (as seen in Eq. 4). Yet, since some vessels may be loaded with heavier containers than others on average, the average mass of a shipping container is drawn from a normal distribution with μ =75 tons and σ =10 tons [11], capturing the range of weights of a loaded shipping container.

$$totalDraught(v) = d_o(v) + d_i(v) \tag{1}$$

$$volume_{water} = totalDraught(v) \times L(v) \times W(v) \tag{2}$$

$$mass_{water} = V_{water} \times D_{water} \tag{3}$$

$$numContainers = \frac{mass_{water}}{\mathcal{N}(\mu = 75, \sigma = 10)} \tag{4}$$

Where

$d_i(v)$ is the draught (in meters) of vessel v as it enters the port

$d_o(v)$ is the draught (in meters) of vessel v as it exits the port

$L(v)$ is the length (in meters) of vessel v as reported in AIS messages

$W(v)$ is the width (in meters) of vessel v as reported in AIS messages

D_{water} is the density of water (1000Kg/m^3)

Finally, the real-world resource deployment is mined in order to benchmark the performance of this optimization against established practices in the real world. Given the known start and end times of each vessel service, the number of containers processed during that service, and the fact that a single crane can load or unload a single container from a vessel requires two minutes [9], the total resource deployment is computed as the number of individual cranes required to process the computed number of shipping containers over all vessels being serviced. Note that this must be computed per eight-hour shift, in order to capture the notion that personnel may be scheduled to work in an integer number of eight hour shifts.

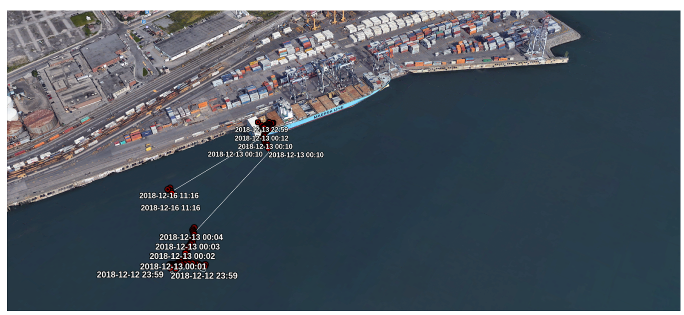


Fig. 1. An Example Vessel Service Track

Ultimately, the dataset includes the hard and soft data listed in Tab. I.

TABLE I
DATASET FEATURES

Feature	Data Type	Data Source
Vessel arrival schedule	Hard data	Port website
Number of containers to be processed on each vessel	Hard data	Computed from AIS (hard) data
Port situation report	Soft data	Synthesized based on real-world example (soft) data

1) Port Situation Reports: Ports write daily situation reports (an example of this is seen in [16]). Since such reports are extremely sensitive to the port and are not publicly available, similar reports were manually synthesized for use in this study (examples of the relevant sections of such synthetic reports are listed in Tab. II, and an example situation report is seen in Tab. III). From the example report [16], it is evident that reports of any foreseeable delays in port operations are only present when such a delay is expected. When no delay is expected, these reports are left empty. The specifics of the usage of these synthetic reports are discussed in Sec. III-B2.

The current resource deployment is known at any given time. The difference in time between the the given moment and the start of the current shift is trivially computed, and computing the number of two minute intervals in this period gives the number of containers processed thus far in this shift. While this calculation yields the number of remaining containers to be processed, the incoming vessel schedule provides another source of shipping containers to be processed. Adding the data from these two streams together yields the size of the backlog (i.e. the number of remaining containers to be processed) at any given time. An increase in the size of this backlog at the end of a given day indicates impending vessel service delay, a the situation report is synthesized to include a comment indicative thereof. On the other hand, if the backlog decreases, the comment section is left empty.

It is therefore clear that situation reports must be generated adaptively, in response to the current backlog, which is dictated by the current resource deployment; which in turn is dictated by the optimizer (explained in Sec. III-B). Therefore, the generation of these situation reports occurs as part of the fitness evaluation function of the multi-objective evolutionary algorithm (described in Sec. III-B2) in response to the size of the current backlog at the end of each simulated day). Further, as discussed in Sec. III-B2, these situation reports are ingested as soft data in order to optimize resource deployment.

TABLE II
EXAMPLES OF SYNTHETIC SITUATION REPORTS

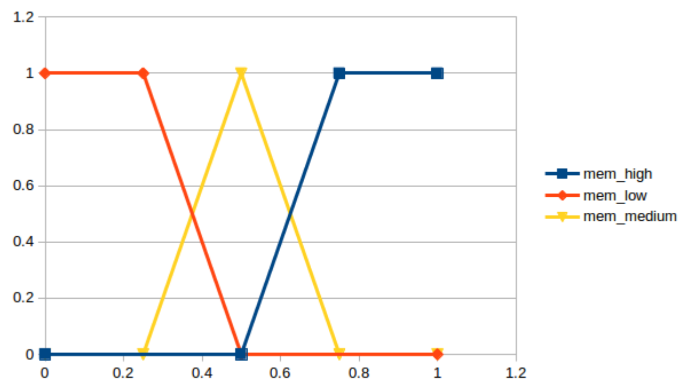
Example Situation Report
extreme delay due to weather
delay due to unavailable port-side resources
congestion-induced delay
inbound delay due to weather
service delays caused outbound delay

B. Optimization

It is intuitive that as the inter-vessel proximity worsens, the backlog of vessels to be served would increase. As a result of this, the deployment of equipment and personnel is needed to also increase throughput and throttle the backlog formation. Describing the optimization as such makes it well suited to be solved by a fuzzy system. However, it is unclear as to what the membership functions of such a fuzzy system should be, in order to function optimally. Therefore, these membership functions will be evolved by a multi-objective evolutionary algorithm (MOEA) as described in this section.

1) Fuzzy System: A Mamdani fuzzy system is used to solve the adaptive resource utilization problem, and its application is discussed in this section.

a) Fuzzifier: This fuzzy system accepts as input, the crisp value of the projected vessel service delay metric. The fuzzifier accepts this crisp input and outputs the fuzzy membership values for "low", "medium", and "high" delay. These membership values are computed through the membership functions evolved by the MOEA, which are either trapezoidal or Gaussian membership functions (examples are illustrated in Fig. 2)



(a) Trapezoidal Membership Function

Fig. 2. Fuzzy Membership Functions to "low", "medium", and "high" Delay

b) Inference Engine: The fuzzy system has the following fuzzy rulebase:

1) If `delay` is low and `delta_delay` is low, decrease deployment
2) If `delay` is medium and `delta_delay` is low, decrease deployment
3) If `delay` is high and `delta_delay` is low, maintain deployment
4) If `delay` is low and `delta_delay` is unchanging, decrease deployment
5) If `delay` is medium and `delta_delay` is unchanging, maintain deployment
6) If `delay` is high and `delta_delay` is unchanging, increase deployment
7) If `delay` is low and `delta_delay` is high, increase deployment
8) If `delay` is medium and `delta_delay` is high, increase deployment
9) If `delay` is high and `delta_delay` is high, increase deployment

c) Defuzzifier: Given the fuzzy membership functions and the inference engine, the fuzzy system uses a center-of-gravity calculation to determine the change in equipment and personnel deployment at a given time, to respond to the foreseen change in vessel service backlog

TABLE III
EXAMPLES OF REAL-WORLD SYNTHETIC SITUATION REPORTS

VESSEL	CODE / VOY	ROTATION	LAST PORT	CURRENT POSITION	ETA NEXT PORT	FORWARD SCHEDULE	COMMENTS
Imua II	IMA 057	AKL-LTK-SUV-APW-PPG-RAR-AIT-VAV-TBU-AKL Auckland,Lautoka,Suva,Apia,Pago,Rarotonga,Aitutaki,Vavau,Nukualofa,Auckland	APW ARR 07:18, 08 Aug SLD 13:42, 08 Aug	PPG ARR 10:12, 08 Aug Lt ETD 14:00, 08 Aug Lt	RAR ETA 06:00, 11 Aug Lt	AIT 14 Aug Lt	
Liloa II	LL2 023	AKL-LTK-SUV-APW-PPG-RAR-AIT-NIU-TBU-AKL Auckland,Lautoka,Suva,Apia,Pago Pago,Rarotonga, Aitutaki, Niue, Nuku'alofa, Auckland	TBU ARR 02:24, 06 Aug SLD 08:24, 06 Aug	AT SEA	AKL ETA 12:00, 10 Aug	LTK 15 Aug SUV 16 Aug	Delay Nukualofa to Auckland with bad weather.

2) Multi-objective Evolutionary Algorithm: In order to optimize the fuzzy system, a MOEA is used to evolve the fuzzy membership functions. The structure and functioning of this MOEA is described in this section.

a) Individual Structure: Each individual is comprised of three chromosomes which encode the values of the vertices of a trapezoidal fuzzy membership, respectively to `low delay`, `medium delay`, and `high delay`. An example of this tri-chromosomal structure is shown in Fig. 3.

	A	B	C	D
$\mu_{\texttt{low\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{high\_delay}}$	0.33	0.5	0.66	1

Fig. 3. Chromosomal Structure for Trapezoidal Individual

$$\text{mem\_low}(x) = \begin{cases} 1 & x < 0.33 \\ \frac{100-100x}{17} & 0.33 \leq x < 0.5 \\ 0 & 0.5 \leq x \end{cases} \tag{5}$$

$$\text{mem\_med}(x) = \begin{cases} 0 & x < 0.33 \\ \frac{6}{|x-\frac{1}{2}|} & 0.33 \leq x < 0.66 \\ 0 & 0.66 \leq x \end{cases} \tag{6}$$

$$\text{mem\_high}(x) = \begin{cases} 1 & x < 0.5 \\ \frac{100x-100}{17} & 0.5 \leq x < 0.66 \\ 0 & 0.66 \leq x \end{cases} \tag{7}$$

On the other hand, individuals encoding Gaussing membership functions use the tri-chromosomal structure shown in Fig. 4, which encodes for the three membership functions in Eq. 8, Eq. 9, Eq. 10.

A population of 100 individuals was used in this study, with 0.95 crossover probability and 0.05 mutation probability.

	μ	σ
$\mu_{\texttt{low\_delay}}$	0	0.1
$\mu_{\texttt{medium\_delay}}$	$\frac{1}{2}$	0.1
$\mu_{\texttt{high\_delay}}$	1	0.1

Fig. 4. Chromosomal Structure for Gaussian Individual

$$\text{mem\_low}(x) = \frac{1}{0.1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{0.1}\right)^2} \tag{8}$$

$$\text{mem\_med}(x) = \frac{1}{0.1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\frac{1}{2}}{0.1}\right)^2} \tag{9}$$

$$\text{mem\_high}(x) = \frac{1}{0.1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{0.1}\right)^2} \tag{10}$$

b) Fitness: Given an individual of the MOEA which encodes a set of fuzzy membership functions, the projected delays are computed as follows. First, the known vessel arrival times (based on ground truth, previously mined from AIS data) are simulated. Incoming vessels are then serviced with the currently deployed resources. Additionally, hard data in the form of the incoming vessel schedule is ingested and the impending shipping container processing load is computed. At the end of the current day (three eight-hour shifts), the backlog of containers to be processed is computed and compared with the backlog from the end of the previous day. The daily situation report is then synthesized using this information (as previously described in Sec. III-A1).

This situation report (soft data) is then ingested to compute the projected delay on a $[0, 1]$ scale as follows. A keyword search is performed on the situation report, over a lexicon including the synonyms of the words "`congestion`" and "`delay`". If "`delay`" is found in the situation report, then "`congestion`" is not searched for, since mentions of delays describe impending vessel service delays more accurately than mentions of congestion which are simply a reason thereof. If no keyword from the lexicon is found in the situation report, a projected delay of 0 is reported. If a keyword is found, however, a projected delay of 0.5 is noted, pending further analysis. Once the keyword is found, any adjectives modifying the keyword are searched for by simply searching the words appearing prior to the found keywords. If not adjectives are found, then the previously noted delay of 0.5 is reported. However, if an adjective is found, then it is tested for membership within two disjoint lexicons, namely "diminutive" and "exaggerative", mined from real-world situation reports. The "diminutive" lexicon includes adjectives that decrease the intensity of the word they modify such as "`small`", "`minor`", "`some`", etc. On the other hand, the "exaggerative" lexicon includes adjectives that increase the intensity of the word they modify such as "`significant`", "`large`", etc. Given these lexicons, if the found adjective is a member of the "diminutive" lexicon, then the reported delay value must be reduced from the previously noted value of 0.5. In order to achieve this effect, it is multiplied by a uniformly distributed random number in $[0.7, 1]$. Conversely, if the found adjective is a member of the "exaggerative" lexicon, then the reported delay value must be increased from the previously noted value of 0.5. In order to achieve this effect, it is multiplied by a uniformly distributed random number, drawn from $[1, 1.3]$. Finally, the modified delay value is reported as the projected delay on a $[0, 1]$ scale, which is then used as the crisp input to the fuzzy system described by the given individual.

The fuzzy system then returns the resource deployment which is then used to compute the number of shipping containers processed in the following shifts (as previously described). The fitness of the given individual is then reported as a point on a two dimensional non-dominated front spanning the sum of the service times for all vessels, and the total resource deployment (summed over all shifts).

c) Selection: Since crossover requires two individuals, a selection mechanism is required to select these two individuals from the population. First, the fitness of each individual is computed as previously described, and the individuals are sorted into fitness fronts. Next, two parents are independently, randomly selected such that a random individual from front i is selected with probability $\frac{1}{i+1}$ (the $i+1$ in the denominator allows for the selection of an individual in the Pareto front to be selected with probability 0.5 instead of 1) [17].

d) Crossover: Crossover is performed by means of a fitness proportional variant of the Weighted Arithmetic Crossover [18], which we will call "FitWAM". However, since a fitness proportional crossover requires a scalar (i.e. uni-dimensional) fitness measure, a relative fitness measure (F_r) is defined. This is computed as the average distance from the mean of a given individual, along each dimension of the two-dimensional fitness plane, from the population's average along each dimension. This is more formally defined in Eq. 11 and Eq. 12.

$$\overline{F_i} = \frac{\sum_{x=1}^{|P|} F_i P_x}{|P|} \tag{11}$$

$$F_r(P_x) = \frac{1}{2} \cdot \Big[\sum_{i=1}^{2} F_i(P_x) - \overline{F_i} \Big] \tag{12}$$

where

F_r is the relative fitness measure
$F_i(P_x)$ is the fitness of individual P_x along objective i
P is the population of all individuals

This gives us a relative scalar measure with which to compare individuals P_1 and P_2. Next, the child individual is computed by assigning the value of each gene to be the weighted average of the corresponding genes of each parent, weighted by the relative fitness measure. An example of this is seen in Fig. 5.

e) Mutation: Mutation is defined as changing the value of one gene in one chromosome to another random, feasible value. For example, mutating the value for B in a trapezoidal individual results in a different value of B, while still being constrained to $A \leq B \leq C$, as shown in Fig. 6.

	A	**B**	**C**	**D**
$\mu_{\texttt{low\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{high\_delay}}$	0.33	0.5	0.66	1

(a) Parent Individual P_1 ($F_r = 1$)

	A	**B**	**C**	**D**
$\mu_{\texttt{low\_delay}}$	0	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0	0.5	0.66	1
$\mu_{\texttt{high\_delay}}$	0	0.5	0.66	1

(b) Parent Individual $P_2 (F_r = 3)$

	A	**B**	**C**	**D**
$\mu_{\texttt{low\_delay}}$	0.0825	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0.0825	0.5	0.66	1
$\mu_{\texttt{high\_delay}}$	0.0825	0.5	0.66	1

(c) Child Individual

Fig. 5. Example Crossover

	A	**B**	**C**	**D**
$\mu_{\texttt{low\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0.33	**0.5**	0.66	1
$\mu_{\texttt{high\_delay}}$	0.33	0.5	0.66	1

(a) Original Individual

	A	**B**	**C**	**D**
$\mu_{\texttt{low\_delay}}$	0.33	0.5	0.66	1
$\mu_{\texttt{medium\_delay}}$	0.33	**0.45**	0.66	1
$\mu_{\texttt{high\_delay}}$	0.33	0.5	0.66	1

(b) Mutant (mutated `B` value in $\mu_{\texttt{medium\_delay}}$)

Fig. 6. Example Mutation (notice the changed `B` value in $\mu_{\texttt{medium\_delay}}$)

f) Termination: Since the optimum value is unknown, it is impossible to know when a new Pareto front will be generated. Therefore, the MBGM stopping methodology [19] is used here. Specifically, the frequency at which new Pareto fronts are generated is tracked. This is described by the number of generations between the creation of each subsequent new Pareto front, the largest of which is the maximum number of generations between any two successive new Pareto front creations. If at any given time, more than twice as many generations have passed without the creation of a new Pareto front, it is unlikely that any new Pareto front will be generated. Thus, the MOEA is stopped, and the current Pareto front is returned as the result of evolution. However, since the search space is substantially large, fitness improvements were not immediately observed, this termination criterion was applied only after 100 generations.

IV. Results

In order to compare the performance of a MOEA-optimized fuzzy system to control resource deployment, they must be compared with the assessment of current industry practice. As mentioned in Sec. III-A, the real-world resource deployment (ground truth) is mined from the available AIS data. These are compared with the performance of the optimized system in Tab. IV.

The MOEA described in Sec. III-B2 is run 30 independent times (for statistical validity) and the Pareto fronts of each of those runs (along with the best performant fuzzy system over the entire evolutionary process) is reported. The mean performance (along with their 95% confidence intervals) of the individuals from these 30 Pareto fronts (as well as the 30 best performant individuals) is shown in Tab. IV. Further, the mean characteristics of the trapezoidal membership functions along with their 95% confidence intervals are shown in Tab. V and illustrated in Fig. 7.

TABLE IV
PERFORMANCE OF EVOLVED FUZZY SYSTEMS

	Mined, Real-World Performance	Optimized Performance
Number of Crews Used	63	23.654 ± 0.05
Total Service Time	77 days, 12 hours, 51 min, and 55 sec.	4 days, 4 hours, 27 min, 53 sec

TABLE V
MEAN CHARACTERISTICS OF EVOLVED FUZZY SYSTEMS

`delay`	$\mu_{\texttt{low}}$	$\mu_{\texttt{medium}}$	$\mu_{\texttt{high}}$
A	0	0	0.377 ± 0.008
B	0	0.289 ± 0.008	0.682 ± 0.009
C	0.366 ± 0.009	0.529 ± 0.01	1
D	0.690 ± 0.008	0.771 ± 0.007	1

Δ`delay`	$\mu_{\texttt{low}}$	$\mu_{\texttt{medium}}$	$\mu_{\texttt{high}}$
A	-1	-1	-0.170 ± 0.019
B	-1	-0.311 ± 0.017	0.446 ± 0.017
C	-0.188 ± 0.018	0.135 ± 0.017	1
D	0.373 ± 0.016	0.570 ± 0.015	1

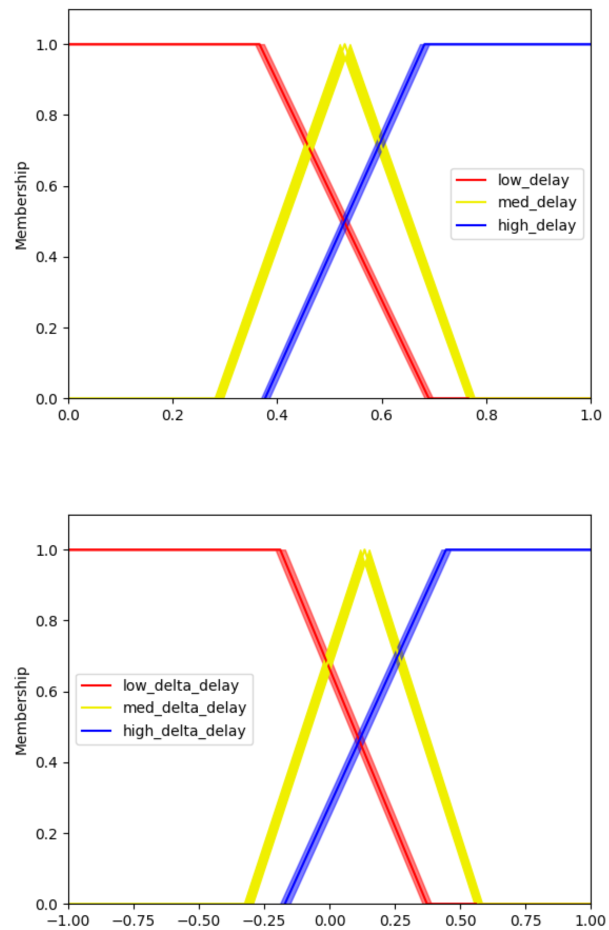


Fig. 7. Mean Fuzzy Membership Functions

The resource deployment mined from real-world AIS data shows that a total of 63 crews were used for overall vessel service. Each crew is comprised of one quay crane operator and all supporting personnel and equipment. Each quay crane requires 3-5 internal shunt trucks (to transport the shipping container within the port), each of which requires one operator. Each shunt truck, in turn, requires one forklift operator (to load a container between a storage pile and shunt truck) [20]. Therefore, each quay crane operator requires an additional 6-10 personnel. Therefore a single crew comprises of 7-11 personnel, deployed for an integer number of shifts.

The MOEA-optimized fuzzy system shows that 23-24 crews are required to perform the same vessel service, accounting for a $62.45\% \pm 0.0008\%$ increase in per-crew productivity. Thus, this MOEA-optimized fuzzy system for crew deployment would save a port 62% in just personnel wages.

Such improvements in resource deployment classically come at the cost of worsened execution time (in this case, vessel service time). However, this does not appear to be the case in this study. A reduction of the overall service time from 1860 hours to 100 hours shows that not only does this system perform the same work with fewer resources, but it does so with approximately 94.6% improvement in service time (the service schedule mined from real-world data and the service schedule induced by the optimized fuzzy system are shown in Fig. 8). This is most likely due to the fact that the optimized fuzzy system deployed multiple crews for certain shifts whereas the mined real-world data shows a maximum of only one crew on any given shift.

V. CONCLUSIONS AND FUTURE WORK

A MOEA-optimized fuzzy system, fusing hard and soft data from multiple sources has been shown to outperform current industry practice. As a direction of further study, it may be augmented to use different types of fuzzy membership functions and an optimized rule set. Additionally, using more advanced Natural Language Processing techniques to parse of situation reports could yield a delay metric with higher veracity, improving upon this optimization.

Further, the inclusion of satellite imagery could be used to corroborate the number of containers on board a vessel at a

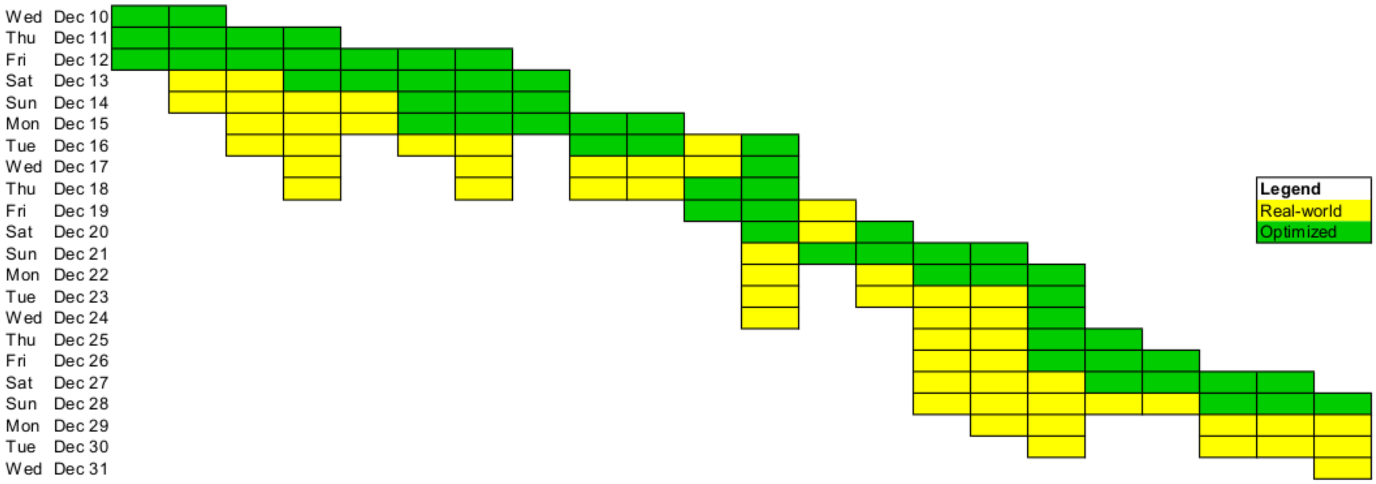


Fig. 8. Vessel Service Schedules

given time. This could be used to determine the effects of other contributing factors to vessel draught, such as ballast, etc, and otherwise contribute to a more accurate estimate of the number of containers on board a vessel at a given time.

As previously mentioned, vessel service delays are caused by a multitude of factors including weather patterns that are non-conducive to the safe operation of heavy equipment. These are considered out of the scope of the current study and are good candidates for further study on the topic.

ACKNOWLEGMENT

We would like to thank SciNet and SOSCIP on whose infrastructure our experiments were run. We would also like to thank Montreal Gateway Terminals Partnership and their on-staff domain experts, and professional highway truck driver Alyssa Fred Wai-Yi Wong, for sharing their expertise used in data creation and validation.

REFERENCES

[1] F. Cheraghchi, I. Abualhaol, R. Falcon, R. Abielmona, B. Raahemi, and E. Petriu, "Big-data-enabled modelling and optimization of granular speed-based vessel schedule recovery problem," in *Big Data (Big Data), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1786–1794.

[2] I. B. G. Rodrigues, R. d. A. Rosa, T. C. Gomes, and G. M. Ribeiro, "Mathematical model for the Berth Allocation Problem in ports with cargo operation limitations along the pier," *Gestão & Produção*, vol. 23, no. AHEAD, p. 0, 2016.

[3] M. M. Golias, G. K. Saharidis, M. Boile, S. Theofanis, and M. G. Ierapetritou, "The berth allocation problem: Optimizing vessel arrival time," *Maritime Economics & Logistics*, vol. 11, no. 4, pp. 358–377, 2009. [Online]. Available: http://link.springer.com/10.1057/mel.2009.12

[4] Y. Hirashima, K. Takeda, S. Harada, D. Mingcong, and A. Inoue, "A Q-learning for group-based plan of container transfer scheduling," *JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing*, vol. 49, no. 2, pp. 473–479, 2006.

[5] Y. A. Saanen and R. Dekker, "Intelligent stacking as way out of congested yards? part 1," *Port Technology International*, vol. 31, pp. 87–92, 2007.

[6] ——, "Intelligent stacking as way out of congested yards? part 2," *Port Technology International*, vol. 31, pp. 87–92, 2007.

[7] D. Ambrosino, A. Sciomachen, and E. Tanfani, "Stowing a containership: the master bay plan problem," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 2, pp. 81–99, 2004.

[8] R. Dekker, P. Voogd, and E. van Asperen, "Advanced methods for container stacking," *OR spectrum*, vol. 28, no. 4, pp. 563–586, 2006.

[9] C. Zhang, Y.-W. Wan, J. Liu, and R. J. Linn, "Dynamic crane deployment in container storage yards," *Transportation Research Part B: Methodological*, vol. 36, no. 6, pp. 537–555, 2002. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.463.6610{&}rep=rep1{&}type=pdf

[10] R. I. Peterkofsky and C. F. Daganzo, "A branch and bound solution method for the crane scheduling problem," *Transportation Research Part B: Methodological*, vol. 24, no. 3, pp. 159–172, 1990.

[11] A. Panchapakesan, R. Abielmona, R. Falcon, and E. Petriu, "Prediction of Container Damage Insurance Claims for Optimized Maritime Port Operations," in *Advances in Artificial Intelligence: 31st Canadian Conference on Artificial Intelligence, Canadian AI 2018, Toronto, ON, Canada, May 8–11, 2018, Proceedings 31*. Springer, 2018, pp. 265–271.

[12] "Lloyd's Register," 2016. [Online]. Available: http://www.lr.org/en/

[13] "MarineTraffic: Global Ship Tracking Intelligence | AIS Marine Traffic." [Online]. Available: https://www.marinetraffic.com/en/ais/home/centerx:-12.0/centery:25.0/zoom:4

[14] I. Abualhaol, R. Falcon, R. Abielmona, and E. Petriu, "Data-Driven Vessel Service Time Forecasting using Long Short-Term Memory Recurrent Neural Networks," in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 2580–2590.

[15] "Vessel schedule - Future." [Online]. Available: https://www.mtrtml.com/vs/vessel-visit-future-sommaire.php

[16] RFairbrother, "MATSON Daily Sitrep," Tech. Rep., 2018. [Online]. Available: www.matson.co.nz

[17] S. Wesolkowski, N. Francetic, and S. C. Grant, "TraDE: Training device selection via multi-objective optimization," *Evolutionary Computation (CEC), 2014 IEEE Congress on*, pp. 2617–2624, 2014.

[18] S. N. Sivanandam and S. N. Deepa, *Introduction to genetic algorithms*. Springer Science & Business Media, 2007.

[19] L. Marti, J. Garcia, A. Berlanga, and J. M. Molina, "A cumulative evidential stopping criterion for multiobjective optimization evolutionary algorithms," in *Proceedings of the 9th annual conference companion on Genetic and evolutionary computation*. ACM, 2007, pp. 2835–2842.

[20] A. Karam, A. B. Eltawif, and N. A. Harraz, "An improved approach for the quay crane assignment problem with limited availability of internal trucks in container terminal," in *2014 IEEE International Conference on Industrial Engineering and Engineering Management*. IEEE, dec 2014, pp. 112–116. [Online]. Available: http://ieeexplore.ieee.org/document/7058611/