BAE Lab
Bigdata Analytics & Engineering

부산대학교
PUSAN NATIONAL UNIVERSITY

An Evolutionary and Predictive Discrete Event Simulation for Port Operation

Kikun Park
Department of Industrial Engineering, Pusan National University
doublekpark@pusan.ac.kr

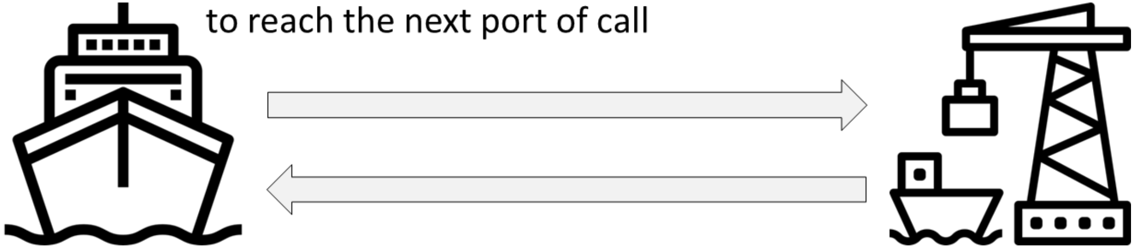
Contents

Introduction

- Maritime transportation is widely acknowledged as the crucial element of global trade, responsible for managing over 80% of the overall trade volume

- The congestion in container terminals (CTs) leads to a reduction in production, hence impacting the capacity to achieve the **Required Time of Departure (RTD)**

RTD: The minimum amount of time needed to reach the next port of call

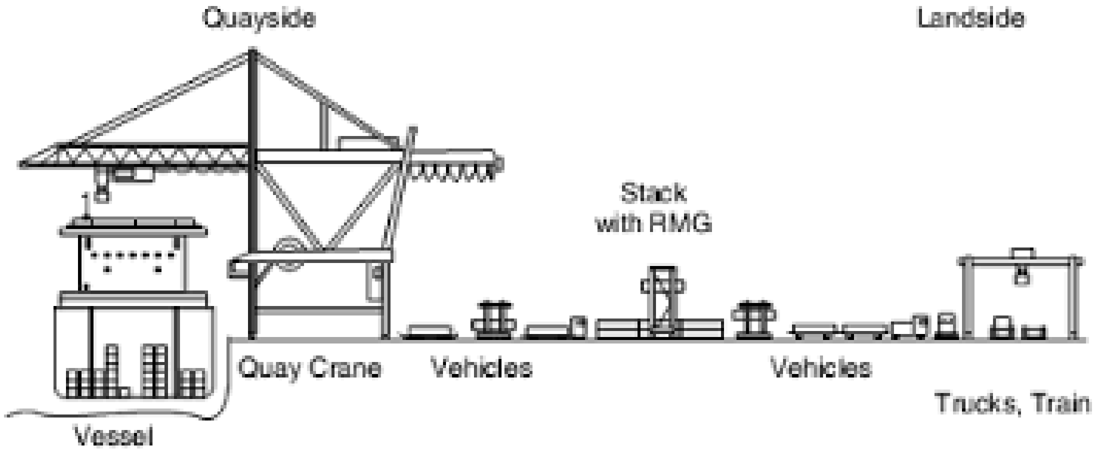


A strategic strategy for the vessel that guarantees adherence to the RTD : **Operational Plan**

Introduction

- The operational plan involves the allocation and scheduling of **Quay Crane (QC)**, the assignment and deployment strategy of **Internal Truck (IT)**, and the assignment of **Yard Crane (YC)**

- In order to meet the RTD, it is crucial that both QCs and YCs carry out their operations quickly for each container handling job



Introduction

- Delays in the operations of ITs, which are responsible for moving containers between QCs and YCs, can create obstructions and impede **adherence to RTD** in CT operations.
- In order to maintain efficient operations of the vessel, operators must strategically manage the allocation of ITs, taking into account the delicate **balance between reducing operating delays and managing expenses effectively**.

Large number of ITs:
- Low operating delays
- High operational cost

Limited number of ITs:
- High operating delays
- Low operational cost

Introduction

- This paper proposes an approach to address this issue by determining the optimal number of ITs.
- The **GAIML** method, which combines **Genetic Algorithm (GA)** with **Machine Learning (ML)**, in order to determine the optimal number of IT

ML

To reduce the search space of GA

GA

To explore the optimal number of IT

Determining the optimal number of IT

(Minimizing the IT deployments while maintaining compliance with the RTD)

Related works

1. Determining the number of Container Handling Equipment
2. Predictive Discrete Event Simulation
3. Genetic Algorithm
4. Non-dominated sorting algorithm
5. Computational Time Reduction Approach for Genetic Algorithm

Related works

1. Determining the Container Handling Equipment

- Studies on the **IT deployment plans** for improving the operational efficiency of container terminals have identified a constraint related to assuming a fixed number of equipment inputs throughout the planning phase. Therefore, the issue of determining the most suitable number of equipment deployments has not been resolved.

- Various research have presented techniques for determining the optimal number of IT deployments using **Discrete Event Simulation (DES)**.

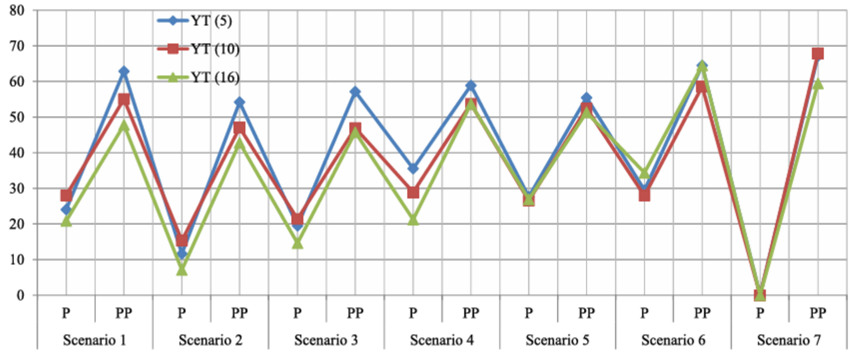


Fig. 4 QCs working time (%)

ORIGINAL PAPER Open Access

Determining the optimal number of yard trucks in smaller container terminals

Stojaković and Twrdy European Transport Research Review

Related works

1. Determining the Container Handling Equipment

- Studies on the **IT deployment plans** for improving the operational efficiency of container terminals have identified a constraint related to assuming a fixed number of equipment inputs throughout the planning phase. Therefore, the issue of determining the most suitable number of equipment deployments has not been resolved.

- Various research have presented techniques for determining the optimal number of IT deployments using **Discrete Event Simulation (DES)**.

Table 4. AUR of terminal equipment (the ratio of QC to RTG is 1 to 2.4)

Number of QC	Ratio of QC to ITT								
	1:4			1:4.5			1:5		
	AUR of QC (%)	AUR of ITT (%)	AUR of RTG (%)	AUR of QC (%)	AUR of ITT (%)	AUR of RTG (%)	AUR of QC (%)	AUR of ITT (%)	AU (%)
14	> 90	> 90	75–80	85–90	65–70	65–70	> 90	60–65	70–75
15	85–90	60–65	70–75	85–90	65–70	60–65	85–90	55–65	55–65
16	80–85	60–65	50–55	80–85	60–65	50–55	80–85	50–55	50–55
17	75–80	60–65	50–55	65–70	50–55	45–50	70–75	45–50	45–50
18	60–65	50–55	40–45	65–70	50–55	45–50	65–75	40–50	45–50
19	60–65	50–55	35–40	60–65	45–50	35–40	60–65	40–45	35–40

AUR, average equipment utilization rate; QC, quay cranes; RTG, rubber-tired gantry; ITT, intra-terminal trucks.

ARTICLE

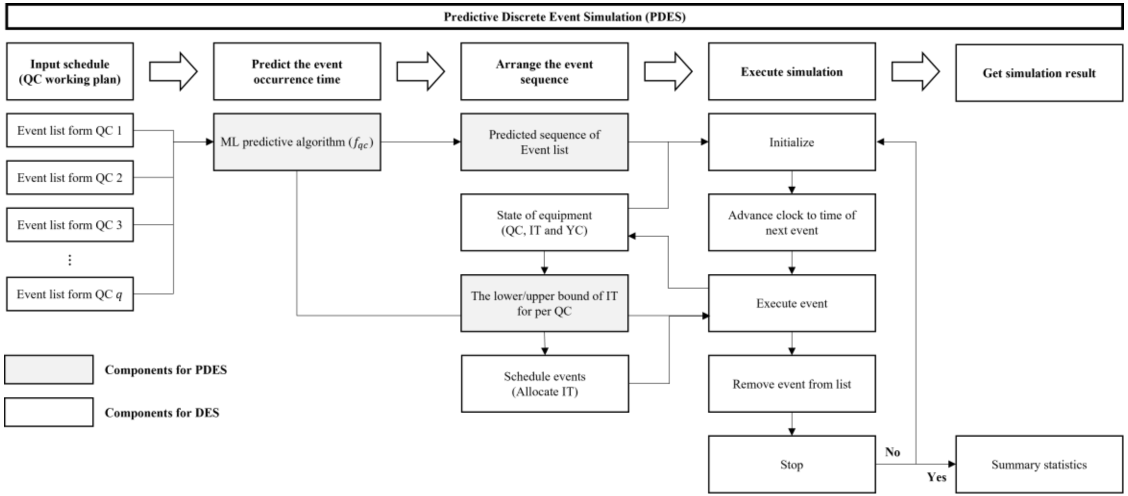
Simulation Model to Determine Ratios between Quay, Yard and Intra-Terminal Transfer Equipment in an Integrated Container Handling System

Journal of International Logistics and Trade

Related works

2. Predictive Discrete Event Simulation

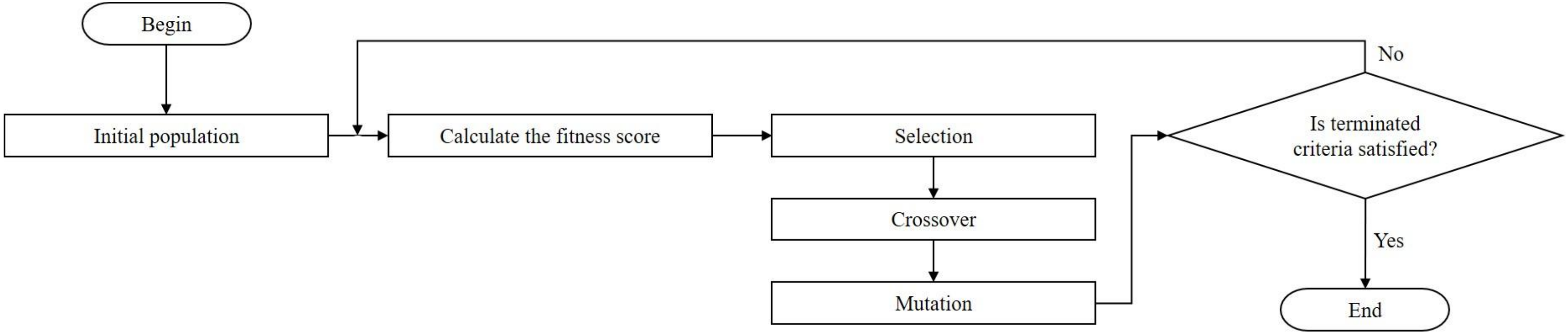
- **Predictive Discrete Event Simulation (PDES)** is utilized for determining the optimal number of IT
- **PDES** is a simulation model for predicting the port operational times that shows better performance than DES



Related works

3. Genetic Algorithm

- The **Genetic Algorithm (GA)** is a well-known stochastic algorithm that operates on a population-based approach. The main operators used are selection, crossover, and mutation



Related works

4. Non dominated sorting algorithm

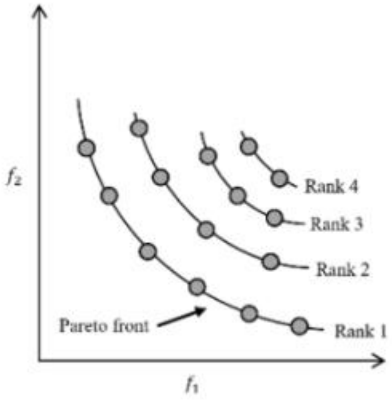
- **Non-dominated sorting** is an essential stage in **multi-objective evolutionary algorithms** for identifying efficient solutions.

Index	Number of IT	Makespan	Group	n_p	S_p	Pareto front level
A	**40**	**5500**	**Non-dominated**	**0**	**C, D**	**1**
B	**35**	**5700**	**Non-dominated**	**0**	**C, D**	**1**
C	48	6000	Dominated	2	D	2
D	52	6500	Dominated	3	-	3

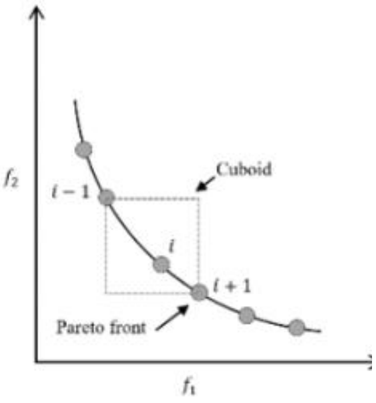
Related works

4. Non dominated sorting algorithm

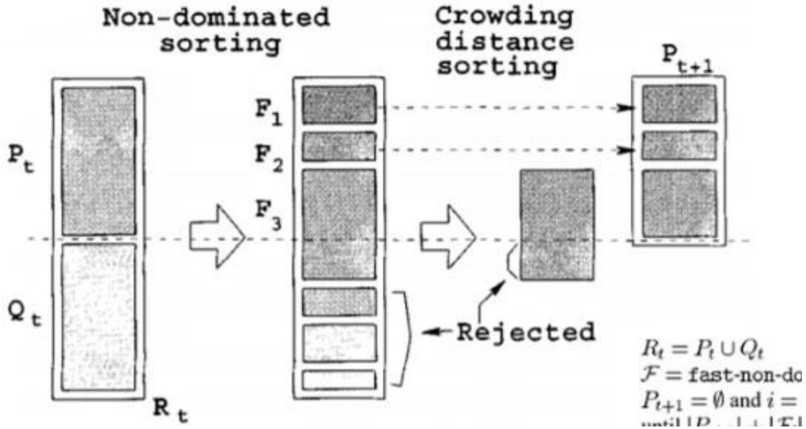
- **NSGA (Non-dominated Sorting Genetic Algorithm)** is well known multi objective evolutionary algorithm



a) Non-dominated sorting



b) Crowding distance calculation



Related works

5. Computational Time Reduction Approach for Genetic Algorithm

- An effective approach to decrease the amount of time required for GA is to skip the simulation stage and instead **utilize machine learning or deep learning methods to immediately calculate the fitness score**.

- This approach is defined as **Surrogate Model**

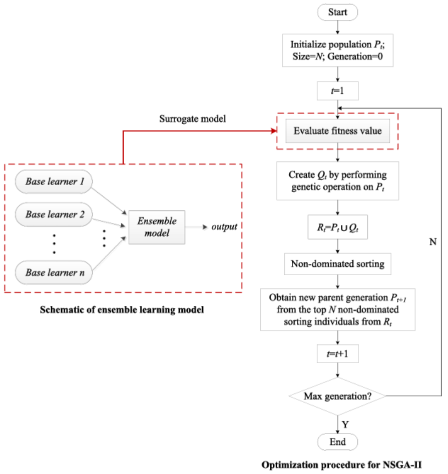


Fig. 3. Integration of surrogate model and NSGA-II.

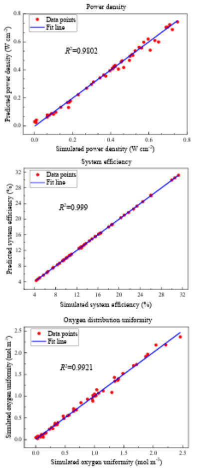


Fig. 5. Correlations between simulated and predicted data of the three indexes on the test set.

ELSEVIER
Contents lists available at ScienceDirect
Energy Conversion and Management
journal homepage: www.elsevier.com/locate/enconman

Multi-objective optimization of PEM fuel cell by coupled significant variables recognition, surrogate models and a multi-objective genetic algorithm

Related works

5. Computational Time Reduction Approach for Genetic Algorithm

- For container terminal operation, **surrogate model** is utilized for scheduling

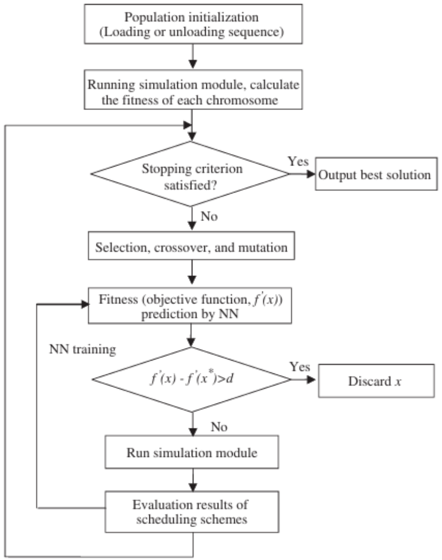


Fig. 3. Simulation optimization method of NN-based surrogate model.

- According to **proposed surrogate model**, until a sufficient of data required for training is collected, it remains identical to a standard GA
- Once the NN model is trained, the **GA operators are only activated when the anticipated values of freshly produced chromosomes surpass the current optimal values**. This helps to decrease the amount of computing time required.
- The use of a **surrogate model strategy is considered efficient**, provided that the predictive model reaches a high level of accuracy.

Integrating simulation and optimization to schedule loading operations in container terminals

Problem definition and data description

- **Problem 1**: Reduce the number of simulation for calculating fitness sore
- **Problem 2**: Reduce the range of initial population space
- **Problem 3**: Improve the exploitation strategies for finding optimal solution

Problem 1

- In order to address the problem 1, **ML predictive algorithms** are trained using the collected data from PDES

Problem 2

- In order to address the problem 2, **ML predictive algorithms** are trained using the actual container handling equipment deployment plan

Problem 3

- In order to address the problem 3, **complementary crossover** strategy is utilized

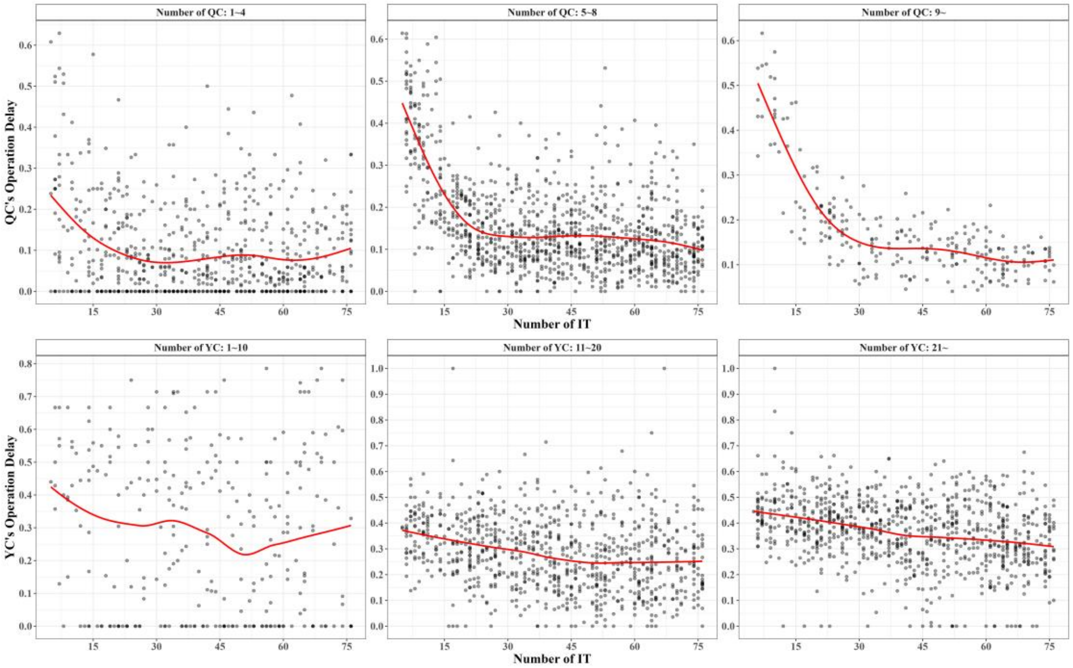
Problem definition and data description

- The data utilized for this study is described as below:

Notation	Feature	Notation	Feature
t	Time	y_t^q	**QC operation delay proportion at time t**
v	Vessel ID	y_t^c	**YC operation delay proportion at time t**
q	QC ID	y^s	The start time of jobs
c	YC ID	y^f	The finish time of jobs
i	IT ID		
x_t^v	Number of vessel		
x_t^q	Number of QC		
x_t^c	Number of YC		
x_t^i	Number of IT		
x_t^{job}	Amount of berth job		
x_t^{gate}	Amount of gate job		

Problem definition and data description

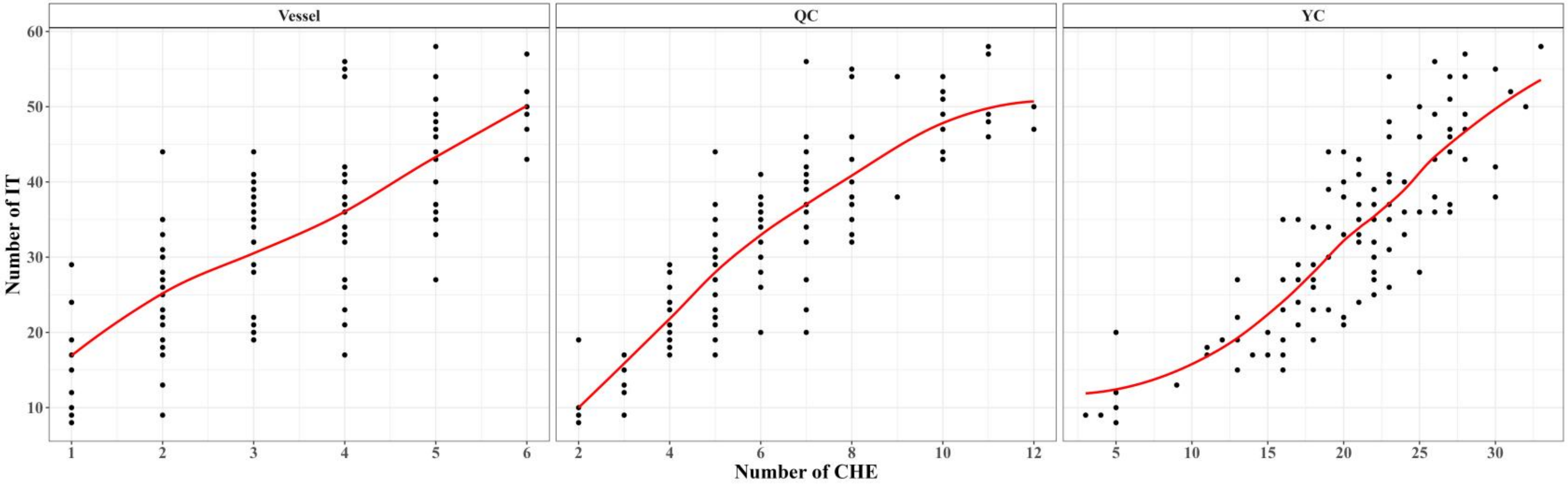
- To train the ML predictive algorithm for fitness score calculation, the simulated data from PDES is used



- Generally, **the more number of IT the less makespan**.
- Therefore, the **boundary is required for preventing over deployment**

Problem definition and data description

- To train the ML predictive algorithm for reducing the range of initial populations, the actual deployment plan is used



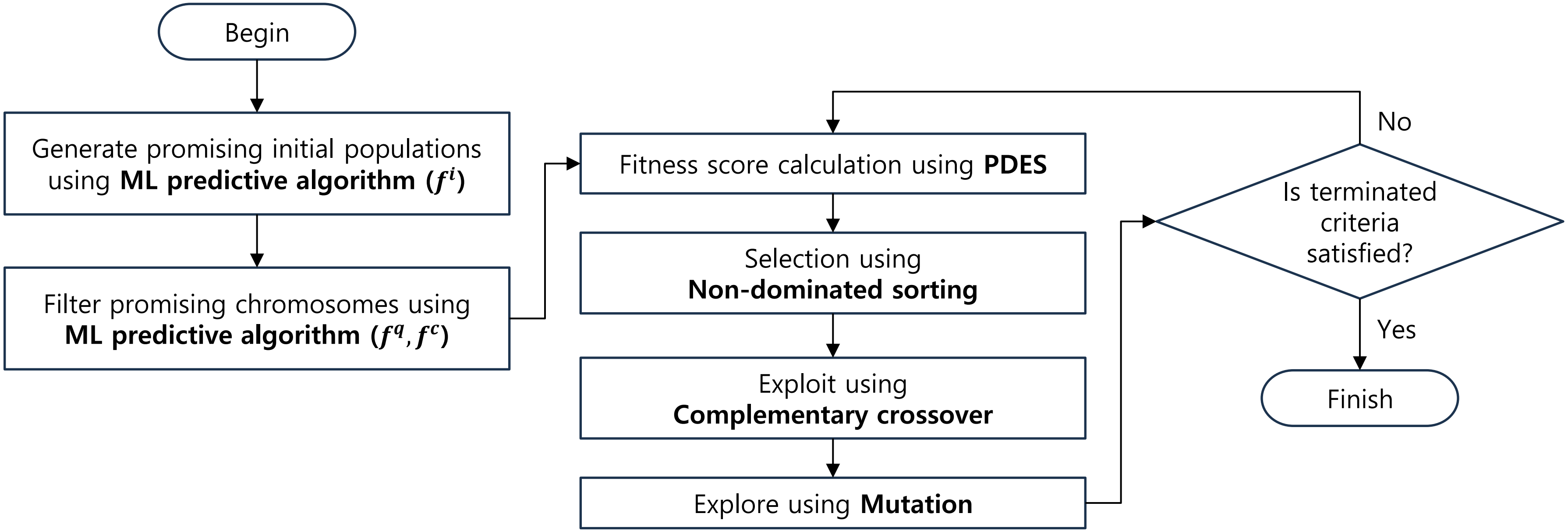
Proposed approach

- The proposed approach consists of three main steps in total

Proposed approach	
Generate the promising initial population	• **Step 1:** Train ML algorithms to predict fitness score
	Find the optimal predictive model for QC and YC delay proportion
	• **Step 2:** Train ML algorithms to predict the expert's experience
	Find the optimal predictive model for the expertise of an expert when determining the number of IT
Enhance the efficiency of exploration	• **Step 3:** Implement crossover strategies to enhance the efficiency of exploration
	Design complementary crossover strategies

Proposed approach

- The flowchart of proposed approach is depicted as:



Proposed approach

- **Chromosome encoding**

$$\boxed{x_0^i}\quad \boxed{x_1^i}\quad \cdots \quad \boxed{x_{n(T)}^i}$$

x_t^i: The deployed number of IT at time t

- **ML predictive model for fitness score calculation**

$$\hat{y}_t^q = \begin{cases} f^q\left(t, x_t^v, x_t^q, x_t^i, x_t^c, x_t^{job}, x_t^{gate}\right), & if\ \hat{y}_t^q \geq 0 \\ 0, \quad otherwise \end{cases}$$

$$\hat{y}_t^c = \begin{cases} f^c\left(t, x_t^v, x_t^q, x_t^i, x_t^c, x_t^{job}, x_t^{gate}\right), & if\ \hat{y}_t^c \geq 0 \\ 0, \quad oterwise \end{cases}$$

- **ML predictive model for reducing the range of chromosomes**

$$\hat{x}_t^i = \begin{cases} x_l^i, \quad if\ \ \hat{x}_t^i \leq x_l^i \\ f^i\left(t, x_t^v, x_t^q, x_t^c, x_t^{job}, x_t^{gate}\right), & if\ x_l^i < \hat{x}_t^i < x_u^i \\ x_u^i, \quad otherwise \end{cases}$$

x_l^i: Lower bound

x_u^i: Upper bound

Proposed approach

- **Random number for chromosome generation**

$$p(r_t) = \frac{1}{\sigma^{max} - \sigma^{min} + 1}, \qquad where\ \sigma^{min} \le r_t \le \sigma^{max}$$

- Because the deterministic models calculate the one numerical value, in order to ensure diversity, randomness is necessary

- Repeat Algorithm 1 until reaches the population size

Algorithm 1: Generate a chromosome for initial population

$t = 0;\ X^i = \{\};$

while $t \le n(T)$ **do**

$$\hat{x}_t^i = \begin{cases} x_l^i, & if\ \ \hat{x}_t^i \le x_l^i \\ f^i(t, x_t^v, x_t^q, x_t^c, x_t^{job}, x_t^{gate}), & if\ x_l^i < \hat{x}_t^i < x_u^i \\ x_u^i, & otherwise \end{cases}$$

$\hat{x}_t^i \leftarrow \hat{x}_t^i + r_t$

$$\hat{x}_t^i = \begin{cases} x_l^i, & if\ \ \hat{x}_t^i \le x_l^i \\ \hat{x}_t^i, & if\ x_l^i < \hat{x}_t^i < x_u^i \\ x_u^i, & otherwise \end{cases}$$

$X^i[t] = \hat{x}_t^i$

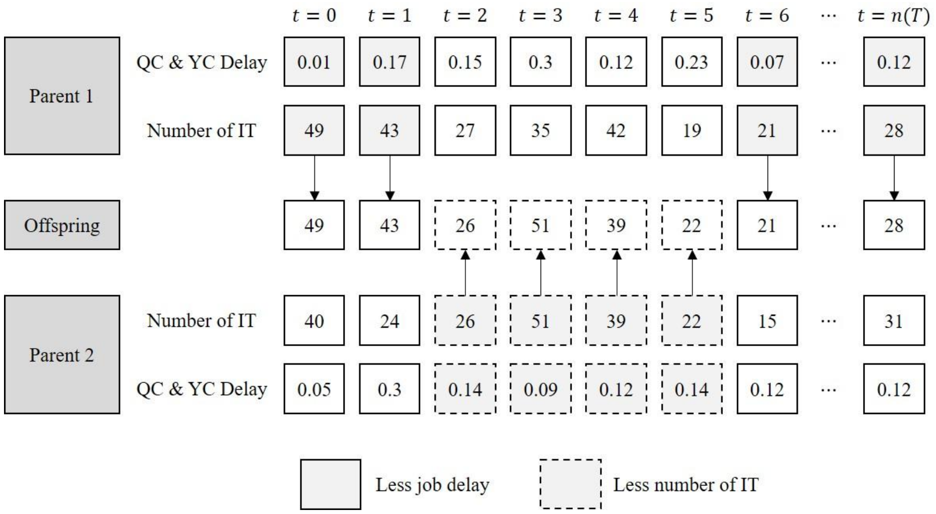
$t \leftarrow t + 1$

return X^i

Proposed approach

- **Complementary crossover**

Complementary crossover is for finding the optimal solution considering bi-objective



Experiment

- **Data used for experiments**

Train / Test	Dates	Number of Scenarios	Random simulation
Train	2020-02-01 ~ 2020-02-05	10	300 (for each scenario)
Test	2020-02-06 ~ 2020-02-15	20	N/A

- **Scenario information (Problem size)**

No	Vessel	QC	YC	Job	Gate	No	Vessel	QC	YC	Job	Gate
1	8	2~12	4~33	1218	3984	11	7	3~7	9~22	865	3758
2	10	5~11	19~27	1195	968	12	8	3~9	9~21	910	820
3	7	4~9	13~27	1115	4046	13	8	4~11	19~31	1452	3437
4	8	4~8	13~30	987	932	14	8	2~10	5~31	870	872
5	6	2~7	11~23	834	1551	15	6	1~6	3~22	639	3668
6	8	2~10	12~29	1232	278	16	3	2~8	5~23	795	1036
7	8	4~10	20~31	1347	544	17	7	2~9	8~33	1205	3617
8	8	1~9	2~32	1088	416	18	9	3~9	15~29	1040	882
9	7	2~10	8~25	1027	3689	19	6	3~8	5~24	790	1330
10	8	4~11	12~32	1035	985	20	8	1~6	7~26	524	150

- **Objective functions**

Objective 1: Minimize the total makespan (MS)

Objective 2: Minimize the average number of IT (AT)

Experiment

- **Experiment description**
  - **Experiment 1 (Exp 1):** Select optimal ML predictive algorithm
  - **Experiment 2 (Exp 2):** Determine the optimal number of IT
- **Experiment setting**

Setting	Description	Value
Generation	The maximum generation	10
Population size	The size of initial population	10
Candidate size	The size of promising chromosomes	150
Crossover probability	The probability of crossover operator	0.9
Mutation probability	The probability of mutation operator	0.1
Mutation change points	The number of genes which are changed when mutating	2
Random number (σ^{max}, σ^{min})	The range of random number	15, -10
Bound (x_l^i, x_u^i)	The lower, upper bound of number of IT	5, 58

Experiment

- **Exp 1:** Select the optimal ML predictive model

The optimal ML predictive algorithm for GA is described with **bold**

Model	QC's operation delay ($\hat{y}_t^q$)		YC's operation delay ($\hat{y}_t^c$)		The number of IT ($\hat{x}_t^i$)	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
LR	0.097	0.076	0.162	0.131	4.983	3.852
GAM (f^i)	0.101	0.078	0.156	0.126	**4.817**	**3.855**
DT	0.101	0.078	0.18	0.135	7.898	6.406
RF (f^q)	**0.092**	**0.073**	0.162	0.129	5.151	4.043
SVR (f^c)	0.104	0.081	**0.155**	**0.124**	5.382	4.083
XGB	0.109	0.084	0.194	0.148	6.392	5.258
BART	0.106	0.081	0.164	0.132	5.117	3.948
ANN	0.104	0.083	0.173	0.142	5.016	4.044

Experiment

- **Exp 2:** Determine the number of optimal IT

The models for comparative experiments are defined as below:

Models	Initial Population	Crossover	Mutation
QC (5)	N/A	N/A	N/A
QC (10)	N/A	N/A	N/A
NSGA	Random	Random	Random
Surrogate Model	Random	Conditional random (When the predictive values are better)	Conditional random (When the predictive values are better)
GAIML 1 (Using f^q, f^c)	ML algorithm	Random	Random
GAIML 2 (Using f^i)	ML algorithm	Random	Random
GAIML 3 (Using f^q, f^c, f^i)	ML algorithm	Random	Random
GAIML 4 (Using f^q, f^c, f^i)	ML algorithm	Complementary crossover	Random

Experiment

- **Exp 2:** Determine the number of optimal IT

The results of initial population follow as:

- QC (5) shows the least average number of IT
- GAIML3 shows the least total makespan
- The main objective of Container terminal is to minimize the makespan, in order to minimize the makespan, the more number of IT is required

No	QC (5)		QC (10)		NSGA		GAIML 1		GAIML 2		GAIML 3	
	MS	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*
1	67.3	39	64.1	54	72.4	25	63.4	38	64.8	36	62.9	41
2	46.8	32	46.9	51	51.2	22	46.5	39	46.7	40	46.8	39
3	40.5	26	37.6	49	43.3	30	38.1	40	38.8	32	38	36
4	50.2	30	50	52	51.2	26	49.9	44	49.8	39	49.8	36
5	33.6	24	33.5	46	33.5	34	33.6	34	33.9	28	33.5	33
6	39.4	32	38.9	49	40.8	26	39.7	36	39.3	38	39.3	39
7	48.5	35	47.5	56	51.8	25	51.3	35	47.4	38	48.2	40
8	33.6	28	33.5	46	36.3	29	34	33	33.8	30	33.6	33
9	32.2	22	30.1	38	31.2	28	30.1	36	30.5	25	30	29
10	37.4	32	37	51	38.5	29	36.6	35	37	35	37.4	34
11	31.9	23	32.2	45	34.2	26	32.1	34	31.9	28	32.1	28
12	44	28	43.8	50	43.6	41	43.7	32	43.8	30	43.6	33
13	59.8	32	58.8	49	63.3	28	58.7	39	58.8	36	58	41
14	34.1	27	33.8	44	34.1	39	33.7	44	33.9	33	33.7	34
15	25.1	19	24.7	37	25.7	24	25.5	36	25.1	24	24.5	27
16	27	22	26.8	39	26.9	26	26.8	40	27	25	26.9	29
17	56.3	31	52.4	51	52.9	31	49.8	41	50.6	34	48.5	39
18	45.4	27	45	48	50.2	28	44.7	40	44.7	30	44.8	36
19	35	23	34.8	44	34.8	29	34.6	35	34.7	26	34.7	27
20	20.8	14	20.7	28	20.8	29	20.6	36	20.8	22	20.6	27
Average	40.4	**27.3**	39.6	46.4	41.8	28.8	39.7	37.4	39.7	31.4	**39.3**	34.0
Times	27.3		26.8		244.1		245.8		243.8		245.03	

Experiment

- **Exp 2:** Determine the number of optimal IT

The results of optimal solution follow as:

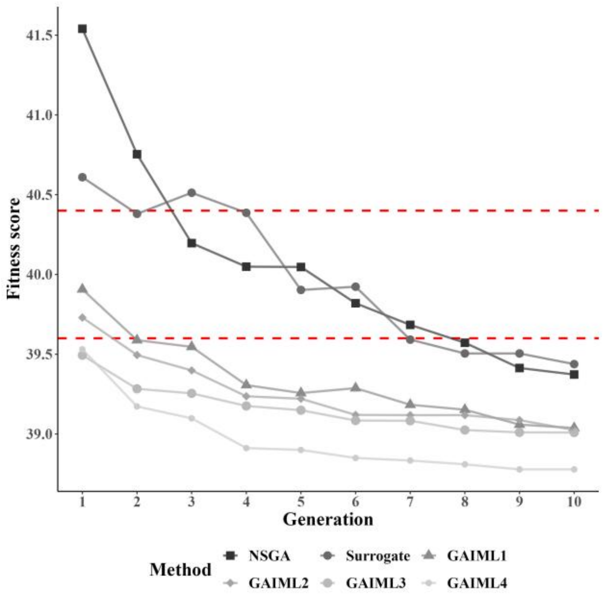
- Surrogate model shows the least average number of IT
- GAIML4 shows the least total makespan

No	NSGA		Surrogate		GAIML1		GAIML2		GAIML3		GAIML4	
	MS	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*	*MS*	*AT*
1	63.1	38	63.2	36	62.4	37	61.9	40	62.2	43	61.9	39
2	46.5	37	46.5	31	46.5	39	46.5	32	46.5	37	46.4	40
3	37.3	34	37	37	37	37	38.3	32	36.8	32	36.2	33
4	49.6	35	49.7	37	49.9	36	49.8	39	49.6	39	49.5	41
5	33.5	28	33.7	27	33.3	33	33.5	32	33.5	33	33.4	28
6	39.4	33	39.4	34	39.1	40	39.1	39	39.2	39	39	39
7	47.9	34	50.3	24	46.9	40	47	39	47	43	46.5	42
8	33.5	36	33.6	31	33.5	38	33.4	31	33.5	36	33.3	34
9	30.5	31	29.5	37	29.6	41	29.5	28	29.4	30	29.5	27
10	36.4	27	36.7	33	36.4	35	36.2	39	36.6	37	36.1	34
11	31.8	29	31.9	27	31.8	32	31.8	29	31.8	29	31.7	28
12	43.6	41	43.5	35	43.6	38	43.6	30	43.5	29	43.6	30
13	59.2	44	58.6	34	57.2	33	57.1	36	57.6	40	56	38
14	34.1	39	33.9	31	33.7	37	33.7	32	33.6	31	33.6	33
15	24.7	27	24.6	29	24.5	39	24.6	25	24.5	28	24.5	25
16	26.8	28	26.8	30	26.8	36	26.9	26	26.9	29	26.7	26
17	49.5	39	50.1	35	48.8	36	47.8	38	47.8	38	48	42
18	44.8	34	44.7	26	44.6	29	44.6	32	44.8	36	44.5	31
19	34.6	31	34.6	29	34.5	34	34.7	26	34.6	30	34.4	27
20	20.6	37	20.6	34	20.6	32	20.6	24	20.6	28	20.5	26
Average	39.4	34.1	39.4	**31.9**	39.0	36.1	39.0	32.5	39.0	34.4	**38.8**	33.2
Times	2,695		2,226		2,701		2,691		2,693		2,646	

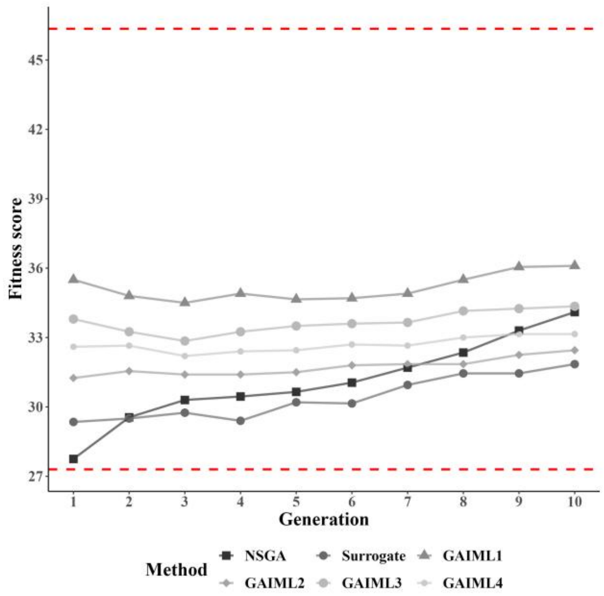
Experiment

The average fitness score per generation are depicted below:

Total Makespan

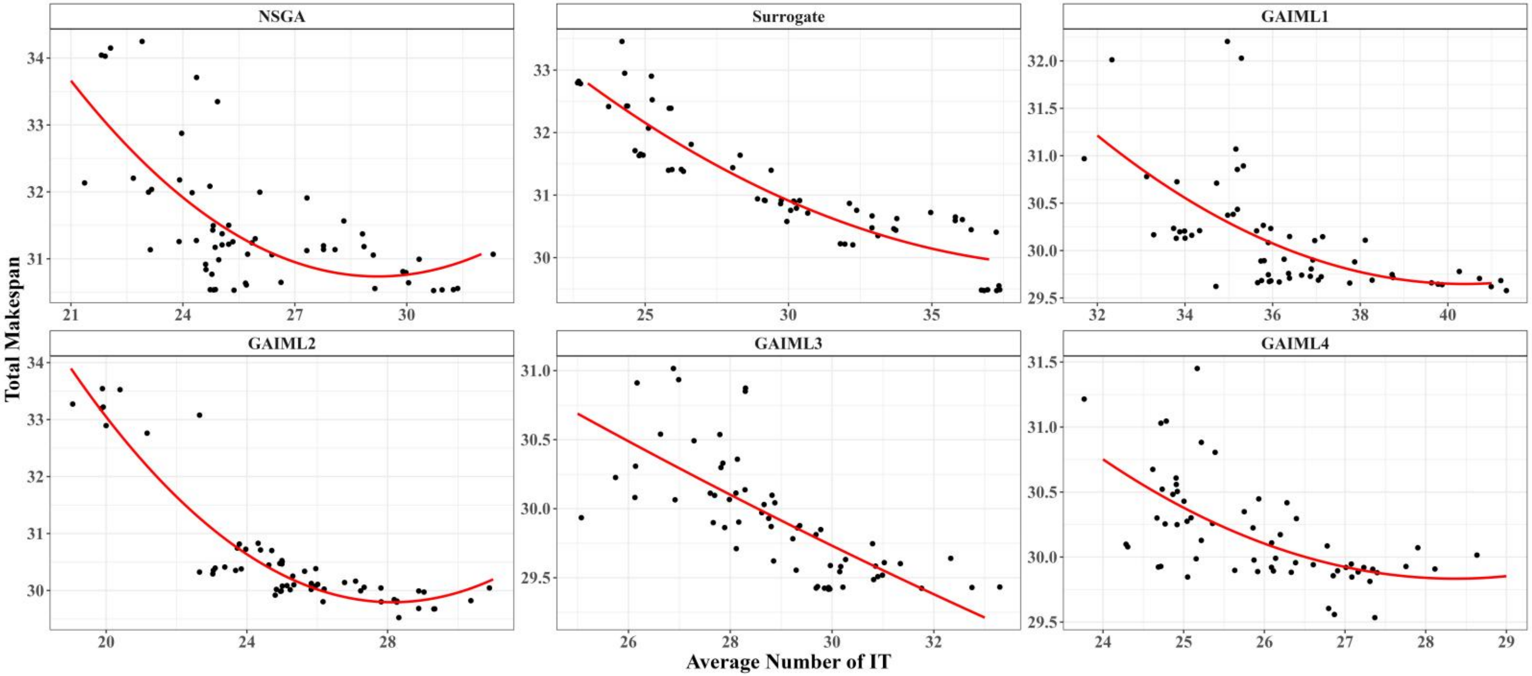


Average number of IT



Experiment

The pareto front is depicted below:



Thank you