

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Computing and Information Systems

School of Computing and Information Systems

1-2023

Learning large neighborhood search for vehicle routing in airport ground handling

Jianan ZHOU

Yaoxin WU

Zhiguang CAO

Singapore Management University, zgcao@smu.edu.sg

Wen SONG

Jie ZHANG

See next page for additional authors

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research



Part of the [Databases and Information Systems Commons](#)

Citation

ZHOU, Jianan; WU, Yaoxin; CAO, Zhiguang; SONG, Wen; ZHANG, Jie; and CHEN, Zhenghua. Learning large neighborhood search for vehicle routing in airport ground handling. (2023). *IEEE Transactions on Knowledge and Data Engineering*. 35, (9), 9769-9782.

Available at: https://ink.library.smu.edu.sg/sis_research/8192

This Journal Article is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Author

Jianan ZHOU, Yaoxin WU, Zhiguang CAO, Wen SONG, Jie ZHANG, and Zhenghua CHEN

Learning Large Neighborhood Search for Vehicle Routing in Airport Ground Handling

Jianan Zhou, Yaoxin Wu, Zhiguang Cao, Wen Song, Jie Zhang, and Zhenghua Chen

Abstract—Dispatching vehicle fleets to serve flights is a key task in airport ground handling (AGH). Due to the notable growth of flights, it is challenging to simultaneously schedule multiple types of operations (services) for a large number of flights, where each type of operation is performed by one specific vehicle fleet. To tackle this issue, we first represent the operation scheduling as a complex vehicle routing problem and formulate it as a mixed integer linear programming (MILP) model. Then given the graph representation of the MILP model, we propose a learning assisted large neighborhood search (LNS) method using data generated based on real scenarios, where we integrate imitation learning and graph convolutional network (GCN) to learn a destroy operator to automatically select variables, and employ an off-the-shelf solver as the repair operator to reoptimize the selected variables. Experimental results based on a real airport show that the proposed method allows for handling up to 200 flights with 10 types of operations simultaneously, and outperforms state-of-the-art methods. Moreover, the learned method performs consistently accompanying different solvers, and generalizes well on larger instances, verifying the versatility and scalability of our method.

Index Terms—Data-driven Optimization, Learning to Optimize, Airport Ground Handling, Large Neighborhood Search, Graph Neural Network, Deep Learning.

1 INTRODUCTION

DUETO the growing demands for air travelling, the airports worldwide are becoming much busier and more crowded, causing severe flight delays and substantial economic loss. To counteract the undesirable impacts, one attainable measure is ameliorating the efficiency of airport ground handling (AGH) based on the historical data, which hinges on how to efficiently dispatch the vehicle fleets for serving the flights by performing the needed operations when they land in the stand for turnarounds. Dispatching the vehicle fleets could be naturally formulated as a vehicle routing problem (VRP), however, the one in AGH is hard to be solved. On the one hand, it always needs to serve a large number of flights at a time and each needs multiple (types of) operations, such as disembarking, refueling, cleaning and so on. On the other hand, each (type of) operation may need a fleet of vehicles to perform, and complex precedence relations between different operations are always required, e.g., the one in Fig. 1. Therefore, solving the VRP in AGH is nontrivial yet challenging for the aviation industry. Although some commercialized exact solvers based on the branch and bound framework (or its variants) have been

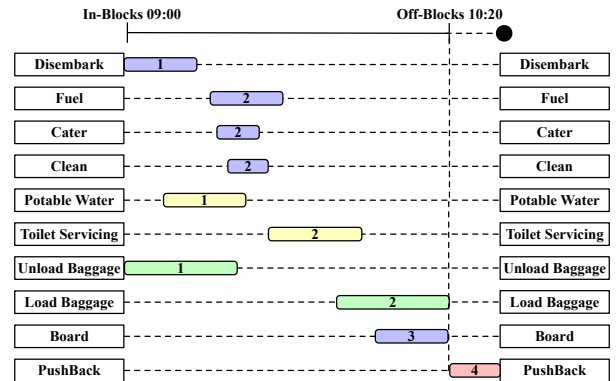


Fig. 1. An example of AGH operations and their precedence. Operations with the same color have precedence relation and the one with a lower value needs to perform earlier, e.g., fueling, catering and cleaning with the same priority precede boarding.

developed to optimally solve combinatorial optimization problems including VRP, e.g., CPLEX [1] and Gurobi [2], they may cost prohibitively long time when directly solving the routing problems in AGH given their intractability.

A practical alternative to tackle such hard problems is the use of heuristic methods. They have the potential to deliver sub-optimal solutions in reasonable computation time. In those methods, certain hand-crafted rules are always designed to cope with the hard constraints so that the computation could be expedited when searching a solution of satisfactory quality. Following this principle, a number of heuristic methods have been attempted to handle the simple variants of VRP in AGH, which only consider a single operation, i.e., de-icing [3], towing [4], catering [5], fueling [6], trailering [7], and baggage loading and unloading [8], respectively. Although some success has been achieved, they fail to consider those operations simultaneously and their

- Jianan Zhou and Jie Zhang are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. (emails: jianan004@e.ntu.edu.sg, zhangj@ntu.edu.sg)
- Yaoxin Wu is with the Department of Information Systems, Faculty of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Netherlands. (emails: wyxacc@hotmail.com)
- Zhiguang Cao is with the School of Computing and Information Systems, Singapore Management University, Singapore. (email: zhiguang-cao@outlook.com)
- Wen Song is with the Institute of Marine Science and Technology, Shandong University, China. (email: wensong@email.sdu.edu.cn)
- Zhenghua Chen is with the Institute for Infocomm Research (I2R), Agency for Science Technology and Research (A*STAR), Singapore. (email: chen0832@e.ntu.edu.sg)
- Corresponding authors: Yaoxin Wu, Wen Song.

precedence relation, which may render them less effective for solving the real world routing problems in AGH. Among this line of research, a notable work considers a relatively hard variant of VRP in AGH with multiple operations [9]. To solve this problem, it first leverages constraint programming (CP) to decompose the whole problem into VRP with time window (VRPTW) for each operation, and then exploits the large neighborhood search (LNS) to solve the respective sub-problems. Recently, another work exploits genetic algorithm (GA) to tackle the decomposed routing problems in AGH, which encodes the sequence of vehicle allocation as chromosome [10]. However, these methods may need massive domain knowledge for AGH and heavy hand-engineering for the key designs, e.g., how to properly decompose the original problems; how to effectively design operators in LNS or genetic algorithms.

On the other hand, the industrial counterpart might be keen on the off-the-shelf solvers for a fast deployment rather than designing complex heuristics for every new problem, although directly using them often suffers from poor performance on large or complex problems. Hence, it would be more desirable and practical to learn improving the performance of those solvers from their externals, e.g., the mixed integer linear programming (MILP) solvers like CPLEX [1] or OR-Tools [11], for solving VRP in AGH. Ideally, the proposed method will, 1) learn to automatically decompose the problems to better take the advantage of the solvers, without much domain knowledge for AGH; 2) also circumvent the design of labor intensive rules for the resulting subproblems, given the direct use of solvers with default settings.

To this end, in this paper, we present a learning based LNS method to directly solve the MILP model of the VRP in AGH, which can tackle 200 flights and 10 operations (e.g. the one in Fig. 1). Specifically, we first model the problem as a multiple-fleet VRP with various constraints pertaining to capacity, time windows, precedence, etc., and formulate it as a MILP model. Then, in our LNS method, we leverage an imitation learning and a bipartite graph convolutional network (GCN) to learn the *destroy* operator so that it would automatically select the decision variables in the MILP model at each step, which would be subsequently reoptimized by adopting an off-the-shelf solver, i.e., CPLEX or OR-Tools, as the *repair* operator. In doing so, the original hard problem is automatically decomposed into a series of sub-MILP problems with less complexity, which is supposed to considerably save the computation. Our contributions are threefold as follows.

- 1) We propose a LNS method to solve the routing problems in AGH, where we exploit destroy operator to select decision variables in the MILP model and repair operator to reoptimize them. Thus, the original problem is decomposed into sub-problems.
- 2) We leverage the imitation learning and bipartite GCN to learn the destroy operator given the graph representation of the MILP model, which circumvents the handed-crafted rules in the conventional heuristic methods, and also enables desirable versatility for the repair operator.
- 3) We conduct extensive experiments under realistic

settings of the well-known CHANGI airport. Results show that our method allows for handling up to 200 flights and 10 types of operations with superior performance to the conventional ones.

The remainder of this paper is organized as follows. Section 2 reviews existing works related to ours. Section 3 introduces the vehicle routing problem in AGH with multiple operations, and formulates it as a MILP model. Section 4 elaborates our learning based LNS method. Section 5 presents empirical results and analysis. The conclusion and future work are finally stated in Section 6.

2 RELATED WORKS

In this section, we review the existing works pertaining to airport ground handling (AGH), and typical LNS methods for solving VRP and its variants, respectively. Recent studies that combine LNS with learning methods are also discussed.

2.1 Airport Ground Handling

AGH, as one of the main airport activities, plays an important role in raising the operational efficiency of airports and the service quality of air transportation. Most of early attempts on the AGH problem focused on scheduling vehicles for performing a single (type of) operation. Norin et al. modeled the scheduling problem for the de-icing operation as a VRPTW, with the objective to reduce the delay of flights and traveling distance of de-icing vehicles [3], [12]. They proposed a greedy randomized adaptive search procedure (GRASP), which was regarded as a simplified variant of LNS, and delivered superior results in comparison with other types of greedy heuristics. Du et al. coped with a scheduled problem for the towing operation by leveraging the column generation heuristic to solve a MILP problem [4]. Ho and Leung evaluated the performance of both tabu search and simulated annealing for solving the scheduling problem with respect to catering operation [5]. Du et al. modeled the scheduling problem for fueling operation as a multi-objective VRPTW and solved it by an ant colony optimization (ACO) heuristic [6]. Zhou et al. tackled the trailer scheduling problem by developing a multi-target scheduling model with rolling windows [7]. Guo et al. performed the scheduling of baggage transport vehicles by leveraging both population diversity and fitness in a genetic algorithm [8]. Han et al. described the scheduling of ferry vehicles as a VRPTW, and designed three heuristic algorithms to solve the problem [13]. Nevertheless, the above methods ignore the underlying interplay between multiple operations and thus may deliver inferior solutions to AGH. On the other hand, the global optimization regarding AGH with multiple operations has received relatively less attention. Among the existing works, Padrón et al. proposed to simultaneously schedule multiple AGH operations at an airport [9], where constraint programming (CP) was adopted to assign time windows to vehicle fleets and then the resulting VRPTW for each fleet was solved by LNS, with the objective to reduce the waiting time of flights and the duration of turnarounds. This method is further ameliorated in terms of computational efficiency in their subsequent work [14]. Liu et al. considered scheduling each type of vehicle (fleet)

as a separate sub-problem, which was solved by the genetic algorithm (GA) [10]. Different from ours, it optimized both the number of vehicles and time cost, and attained superior performance to the multiobjective evolutionary algorithm based on decomposition (MOEAD) and the particle swarm optimization (PSO) algorithm. Recently, Zhang et al. used GA to solve the scheduling problem of supporting vehicles, which is similar to AGH but ignores the travel time of vehicles between aircraft. The method was evaluated with less than 10 aircraft, which is far from the ones in practical situations [15].

Besides the AGH, there exists literature on tackling other operational issues in airports. For example, Yıldız et al. developed an object detection system to monitor and analyze the ground service actions from video frame sequences [16]. Alomar and Tolujevs developed a simulation model to generate event streams and movements of the mobile objects, aiming to test control techniques in the airport [17]. Another line of research analyzes safety issues in airports, which are captured by the interplay between air and ground vehicles or the use of civil drones [18], [19]. Other less related works for manpower scheduling or airside operation research can be found in [20], [21], [22], [23].

2.2 Large Neighborhood Search

LNS is a classic metaheuristic based on local search, which has demonstrated competitive performance for solving various VRPs [24], such as the pickup and delivery problem [25], the dial-a-ride problem [26], and the capacitated VRP (CVRP) [27], [28], [29]. With destroy and repair procedures, LNS decomposes the original intractable problem into a series of subproblems, which are iteratively solved to fast improve the solution. However, in LNS, a number of rules have to be manually designed for the respective problems, e.g., heuristics in operators, the choice of operators, destroy degree, etc. Regarding the routing problems in AGH with multiple operations, it could be represented as a multiple-fleet VRP with complex constraints, which is more computationally intractable than the common VRPs as mentioned above. Designing rules for such problem usually needs substantial domain knowledge. Existing works [9], [14] handle a similar variant of this problem by decomposing it into VRPTW for each respective fleet using constraint programming, and then sequentially solve them by LNS. In contrast, we directly leverage LNS to decompose the MILP model of the problem, by exploiting imitation learning to learn desirable behaviors of the destroy operator. We also apply existing solvers as repair operators to save manual efforts in designing meticulous rules otherwise. In doing so, our method greatly reduces the need for much domain knowledge and tuning work in comparison to classic LNS methods, which not only accelerates the algorithmic development but also promotes the eventual performance.

On the other hand, some recent attempts have been performed to learn destroy or repair operators in LNS. Hottung et al. utilized an attention model to learn repair operators for standard CVRP [30]. Gao et al. adopted recurrent neural network (RNN) and graph attention network (GAT) to learn a reinsertion operator for CVRP with time windows [31]. However, neural networks involved in these methods are

only applicable to a specified type of VRP. Instead of learning problem-dependent operators, Song et al. proposed to destroy a solution by selecting variables at each step in LNS, and reoptimize them by an existing solver [32]. This method is relatively general and relies on less domain knowledge. However, it fixes the number of the variables that need to be reoptimized to a constant value, which is determined by prior trials for each problem. Also, they trained fully-connected neural networks with incidence matrices of same-sized instances as features, which limits its scalability. In contrast, our method allows reoptimizing a flexible number of variables at each step, and the GCN based policy network is less sensitive to problem scale. Besides, we would like to note that, our method is also different from the automatic neighborhood design, which usually groups entities (e.g., customers, commodities) from the MILP model and then optimizes the resulting small portions [33], [34]. Most of this line of works pursue a good feasible solution with one-step clustering by manually designing metrics to compute distances between the entities. In contrast, we aim to improve solutions sequentially with imitation learning and enable the training of GCN to automatically select variables for optimization at each step.

3 PROBLEM STATEMENT AND MILP MODEL

In this section, we describe the vehicle routing problem in AGH for performing multiple operations and also present its MILP formulation.

Pertaining to our problem, each type of operation is conducted by a vehicle fleet. Therefore, we describe it as a multiple-fleet VRP with constraints of *capacity*, *time windows* and *precedence*. Formally, the problem is defined on an undirected graph $G = (\Omega, E)$ with the node set $\Omega = \{0, 1, \dots, n, \hat{n}\}$ and edge set $E = \{(i, j) | i, j \in \Omega; i \neq j\}$. In the node set, 0 and \hat{n} represent the initial and final depot node (refer to the same one), and $\Omega^* = \Omega \setminus \{0, \hat{n}\}$ represent flights with demand $q_i, i \in \Omega^*$, where we set $q_0 = q_{\hat{n}} = 0$. Each edge $(i, j) \in E$ bears an associated cost c_{ij}^k for operation k , where we set $c_{0\hat{n}}^k = 0, \forall k \in \mathcal{K}$. The heterogeneous vehicle fleets $\mathcal{K} = \{1, \dots, K\}$ are responsible for different operations, and $k_1 \prec k_2 (k_1, k_2 \in \mathcal{K})$ means the operation by k_1 precedes that of k_2 . Each fleet $k \in \mathcal{K}$ has V^k homogeneous vehicles $\mathcal{V}^k = \{1, \dots, V^k\}$ with the same capacity Q^k . Regarding a vehicle in the fleet k , s_i^k denotes its service duration for flight i , and $s_0^k = s_{\hat{n}}^k = 0; t_{ij}^k$ denotes the travel time from flight i to j , and $t_{0\hat{n}}^k = 0; [a_i^k, b_i^k]$ represents the time window for servicing flight i . Naturally, the problem entails two classes of decision variables, 1) the binary variable x_{ijv}^k , which is equal to 1 if a vehicle v in fleet k travelled from flight i to j , and 0 otherwise; 2) the real variable T_{iv}^k , which represents the start time for servicing flight i by vehicle v in fleet k . Accordingly, we formulate the AGH problem as follows,

$$\min. \sum_{k \in \mathcal{K}} \sum_{v \in \mathcal{V}^k} \sum_{(i,j) \in E} c_{ij}^k x_{ijv}^k \quad (1)$$

$$\text{s.t.} \sum_{i \in \Omega} \sum_{v \in \mathcal{V}^k} x_{ijv}^k = 1, \forall j \in \Omega^*, k \in \mathcal{K} \quad (2)$$

$$\sum_{i \in \Omega \setminus \{\hat{n}\}} x_{iuv}^k = \sum_{j \in \Omega \setminus \{0\}} x_{ujv}^k, \forall u \in \Omega^*, v \in \mathcal{V}^k, k \in \mathcal{K} \quad (3)$$

$$\sum_{j \in \Omega^*} \sum_{v \in \mathcal{V}^k} x_{0jv}^k \leq V^k, \forall k \in \mathcal{K} \quad (4)$$

$$\sum_{j \in \Omega^*} \sum_{v \in \mathcal{V}^k} x_{0jv}^k = \sum_{i \in \Omega^*} \sum_{v \in \mathcal{V}^k} x_{inv}^k, \forall k \in \mathcal{K} \quad (5)$$

$$\sum_{i \in \Omega \setminus \{0\}} \sum_{v \in \mathcal{V}^k} x_{i0v}^k = \sum_{j \in \Omega \setminus \{n\}} \sum_{v \in \mathcal{V}^k} x_{njv}^k = 0, \forall k \in \mathcal{K} \quad (6)$$

$$\sum_{i \in \Omega^*} q_i \sum_{j \in \Omega} x_{ijv}^k \leq Q^k, \forall v \in \mathcal{V}^k, k \in \mathcal{K} \quad (7)$$

$$x_{ijv}^k (T_{iv}^k + s_i^k + t_{ij}^k - T_{jv}^k) \leq 0, \forall v \in \mathcal{V}^k, k \in \mathcal{K} \quad (8)$$

$$a_i^k \leq T_{iv}^k \leq b_i^k, \forall v \in \mathcal{V}^k, k \in \mathcal{K} \quad (9)$$

$$T_{iv}^{k_1} + s_i^{k_1} \leq T_{iv}^{k_2}, \forall k_1, k_2 \in \mathcal{K}, k_1 \prec k_2 \quad (10)$$

$$x_{ijv}^k \in \{0, 1\}, \forall (i, j) \in E, v \in \mathcal{V}^k, k \in \mathcal{K} \quad (11)$$

$$T_{iv}^k \geq 0, \forall i \in \Omega^*, v \in \mathcal{V}^k, k \in \mathcal{K}. \quad (12)$$

The objective function in Eq. (1) aims to minimize the total cost (distance). Eq. (2) ensures that each flight is only served by one vehicle in each fleet. Eq. (3) ensures the continuity of the route. Eq. (4) ensures that the number of used vehicles will not exceed the available vehicles in each fleet. Eq. (5) ensures the equivalence between the number of vehicles leaving and returning to the depot. Eq. (6) ensures that all routes are from the initial depot node to final depot node. Eq. (7) defines capacity constraints in each fleet. Eq. (8) and (9) ensure that the start time of a service is within the respective time windows. Eq. (10) imposes that the operation by a fleet must precede that of another if they have precedence relation, as illustrated in Fig. 1. Since Eq. (8) is nonlinear, we transform it as follows,

$$T_{iv}^k + s_i^k + t_{ij}^k - T_{jv}^k \leq C(1 - x_{ijv}^k), \forall (i, j) \in E, k \in \mathcal{K} \quad (13)$$

where C is a large constant, and we set it to 10^6 . In doing so, we model the problem as the form of a mixed integer linear programming (MILP).

We would like to mention that the presented formulation is more complicated than most of conventional VRPs such as CVRP, VRPTW, etc. [35], [36]. We not only tackle practical constraints in AGH, i.e., time windows, capacity or precedence constraints, but also deal with heterogeneous operations for a large volume of aircraft. We also note that the problem definition is not a mere extension of conventional VRPs but well reflects the complex context in practical AGH. All constraints in our formulation have their own physical meaning. Meanwhile, the AGH has its own exclusive attributes, e.g., it often requires decisions on both routes and start times of operations, which is much more practical and challenging than conventional VRPs. Solving AGH with such a complex context is fairly challenging. Thus, designing hand-crafted heuristics to solve this problem might be intractable, which is usually accompanied with substantial domain knowledge and trial-and-error efforts. In contrast, we turn to automatically learn LNS policy to solve the MILP model, which is expected to save large amounts of labour and be more friendly to non-expert users.

4 METHODOLOGY

In this section, we present the learning based LNS method to solve the routing problem in AGH. We first introduce the general LNS framework which is characterized by a destroy

operator and repair operator. Then we design a simple yet efficient heuristic destroy operator which would be used to yield demos for the subsequent imitation learning. Finally, given the graph representation of our MILP model, we exploit the imitation learning and graph neural network to guide the destroy operator to automatically select variables, which will be reoptimized by the repair operator. In doing so, the original problem is decomposed into a series of sub-problems, and the heavy computation would be considerably mitigated. The proposed LNS framework with the learned destroy operator is overviewed in Fig. 2.

4.1 The LNS Framework

As a local search based metaheuristic, LNS defines its solution neighborhood by destroy and repair operators. Formally, given an instance p of a combinatorial optimization problem with a minimization objective function, the set of feasible solutions is denoted as $X(p)$. We aim to find a solution $x^* \in X(p)$ such that $c(x^*) \leq c(x)$, $\forall x \in X(p)$, where $c(\cdot)$ is the objective function. Generally, LNS solves the problem in an iterative manner. In each iteration, the destroy operator \mathcal{D} destructs part of the current solution \tilde{x} , and the repair operator \mathcal{R} rebuilds the destroyed solution to obtain a new solution \tilde{x}' . Since stochasticity exists in both destroy and repair operators, the neighborhood $N(\tilde{x}) \subseteq X(p)$ of the solution \tilde{x} is defined as the set of solutions that can be derived by first applying \mathcal{D} and then \mathcal{R} , so that $\tilde{x}' \in N(\tilde{x})$. The new solution \tilde{x}' will replace the current solution according to the acceptance criterion, such as a simple scheme of accepting only improved solutions or the one in simulated annealing [25], [37]. The above process is repeated until the termination condition is satisfied. The initial solution \tilde{x}_0 is obtained by either simple heuristics or exact solver with short runtime. The incumbent solution x_b is updated along the solving process of LNS.

In our method, we leverage LNS to directly solve the MILP model in section 3. Specifically, assuming that \mathcal{X} denotes the set of decision variables x_{ijv}^k , we employ the destroy operator in each iteration to select a subset of variables \mathcal{X}_r for reoptimization, while fixing the other variables $\mathcal{X} \setminus \mathcal{X}_r$ to the values in the current solution \tilde{x} . A natural way for reoptimization is to use an off-the-shelf solver as the repair operator, e.g., CPLEX, especially given that the resulting problem has been scaled down. Obviously, our proposed LNS framework brings two major benefits, 1) we target on variable-level destroy operator in the context of MILP, and it eschews massive problem-specific manual rules; 2) the original problem is decomposed into a series of sub-problems of MILP, and it allows tractable computation for state-of-the-art exact solvers. Accordingly, the pseudo-code of our LNS framework is summarized in Algorithm 1.

4.2 Heuristic Destroy Operator

To provide demonstrations (demos) for the learning based destroy operator in section 4.3, we first design a heuristic destroy operator in consideration of *destroy degree* and *stochasticity*. The destroy degree is a key factor for a destroy operator, which is defined as $\mathcal{D}_d = |\mathcal{X}_r|/|\mathcal{X}|$ with $|\cdot|$ denoting the cardinality of a set. It determines the percentage of decision variables to select for reoptimization,

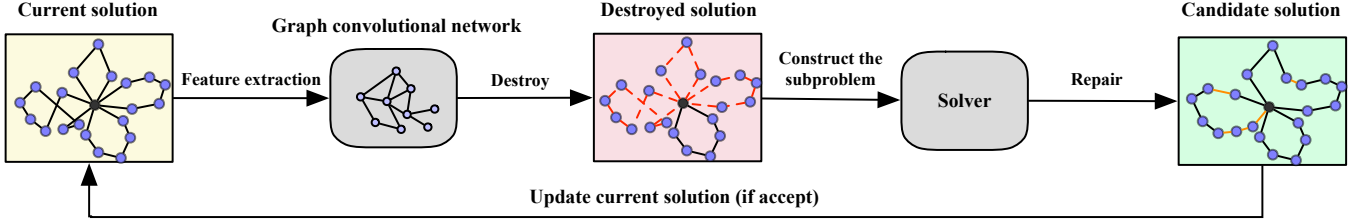


Fig. 2. **An overview of the learning based LNS framework.** Given an instance and the current solution, the trained policy selects groups of integer decision variables associated to certain vehicles, and then the repair operator (e.g. CPLEX) solves the sub-MILP, where the selected variables are reoptimized and others are fixed to the values in the current solution. Once a new solution is computed, the current solution is updated according to an acceptance criterion. The process is repeated until a termination condition is met.

Algorithm 1: Formulation based LNS Framework

- 1: Input: an instance of MILP problem p with initial solution \tilde{x}_0 ; number of iterations n ; destroy and repair operators \mathcal{D}, \mathcal{R} ; percentage of decision variables to destroy \mathcal{D}_d .
 - 2: $x_b = \tilde{x} = \tilde{x}_0$
 - 3: $best_obj = c(\tilde{x}_0)$
 - 4: **for** $t = 1, \dots, n$ **do**
 - 5: $\mathcal{X}_r = \mathcal{D}(p, \tilde{x}, \mathcal{D}_d)$
 - 6: $\tilde{x}' = \mathcal{R}(p, \tilde{x}, \mathcal{X}_r)$
 - 7: **if** $accept(\tilde{x}', \tilde{x})$ **then**
 - 8: $\tilde{x} = \tilde{x}'$
 - 9: **if** $c(\tilde{x}') < best_obj$ **then**
 - 10: $x_b = \tilde{x}'$
 - 11: $best_obj = c(\tilde{x}')$
 - 12: **return** $x_b, \tilde{x}', \mathcal{X}_r$
-

and controls the size of the resulting sub-MILP at each step of LNS. If the subset of the selected variables is too small, the repair operator may fail to sufficiently explore the solution space. If the subset is too large, the sub-problem would be almost the same to the original problem, which may lead to prohibitively heavy computation. Besides, the stochasticity is typically required for a destroy operator to avoid the stagnation, i.e., the situation where almost the same subsets of variables are optimized in consecutive steps with the solution improved negligibly. Keeping both the destroy degree and stochasticity in mind, we propose a *Vehicle-random* destroy operator to solve the MILP model in section 3, which selects the integer decision variables related to the randomly chosen vehicle(s) given the destroy degree, i.e., $\mathcal{X}_r = \{x_{ij}^k | k \in \mathcal{K}_s \subset \mathcal{K}, v \in \mathcal{V}_s^k \subseteq \mathcal{V}^k, \forall i, j \in \Omega\}$, where \mathcal{K}_s is the indices of selected fleets and \mathcal{V}_s^k is the indices of selected vehicles in the fleet k . All the continuous decision variables T_i^k will be reoptimized together with the selected integer variables \mathcal{X}_r by the repair operator, since continuous ones require much less computation. Although *Vehicle-random* destroy operator is simple and intuitive, we could apply it multiple times to an instance and retrieve the best one, which would be adequately good to serve as demonstrations for the imitation learning.

Note that most of existing LNS methods rely on the relatedness or cost on each node (i.e., flights in our AGH problem) to design operators for specific VRPs [25]. Normally, they need substantial domain knowledge or manual efforts, and may not be directly applied to select variables from the

mathematical programming formulation. In contrast, both our heuristic and learning based destroy operators have favorable potential to alleviate those issues.

4.3 Learning Destroy Operator

We present the learning based destroy operator for LNS to automatically select the variables. To this end, we first formulate the LNS as a Markov Decision Process (MDP), and then parameterize the policy by a bipartite GCN since the MILP model could be represented as a graph (an illustration of our network architecture is depicted in Fig. 3), which is trained by the imitation learning algorithm.

4.3.1 MDP Formulation

For an instance of MILP model p with integer decision variables \mathcal{X} , the solving process by LNS can be represented as a MDP. Specifically, we consider the destroy operator as the *agent* and the remainder of LNS as the *environment*. At the t^{th} iteration, the *state* s_t consists of both the formulation information and the current solution. The *action* a_t is to select a subset of variables among all integer decision variables \mathcal{X} according to the *policy* $\pi(a_t | s_t)$. After the action is taken, the environment returns a new solution by reoptimizing the selected variables via the repair operator, and also updates state to s_{t+1} . The reward is represented as the difference between the objective values of the new and current solution, i.e., $r(s_t, a_t) = c(x_{t+1}) - c(x_t)$, where x_t and x_{t+1} are reflected in states s_t and s_{t+1} . This interaction between the agent and environment continues until the pre-defined termination condition is met.

4.3.2 State Representation

We represent the state as a bipartite graph $\mathcal{G} = (\mathcal{C}, \mathcal{V}, \mathcal{E})$, with constraints $c_i \in \mathcal{C}$, variables $v_j \in \mathcal{V}$ and edges $e_{ij} = (c_i, v_j) \in \mathcal{E}$, which is depicted in Fig. 3(a). The red nodes \mathcal{C} refer to constraints, i.e., rows in the adjacency matrix of the MILP formulation. They are represented by a feature matrix $\mathcal{M}_c \in \mathbb{R}^{|\mathcal{C}| \times n_c}$, where n_c is the dimension of constraint features. The blue nodes \mathcal{V} refer to variables, i.e., columns in the adjacency matrix of the MILP formulation. They are represented by a feature matrix $\mathcal{M}_v \in \mathbb{R}^{|\mathcal{V}| \times n_v}$, where n_v is the dimension of variable features. The edges \mathcal{E} refer to connections between variables and constraints (i.e., an edge e_{ij} exists if the coefficient of variable v_j is nonzero in constraint c_i). They are represented by a feature matrix $\mathcal{M}_e \in \mathbb{R}^{|\mathcal{C}| \times |\mathcal{V}| \times n_e}$, where n_e is the dimension of edge features. We represent the features for each constrain

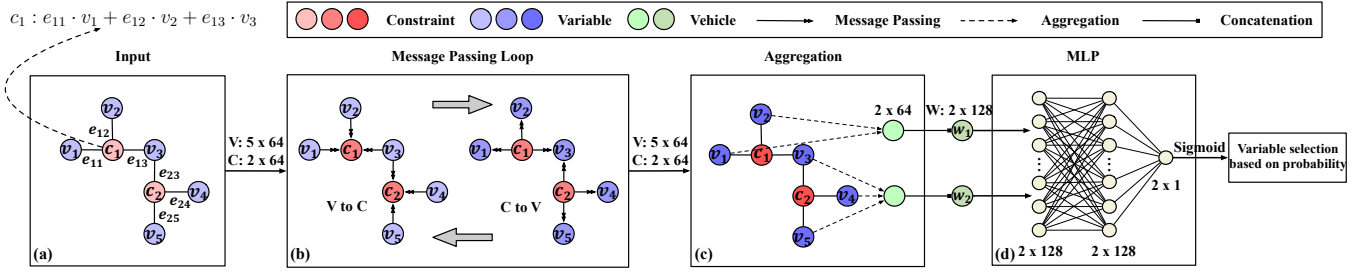


Fig. 3. An illustration of policy network π_θ with 5 variables and 2 constraints. (a) The state of constraints and variables is represented as a bipartite graph; (b) The message passing from variables \mathcal{V} (\mathcal{C}) to constraints \mathcal{C} (\mathcal{V}) updates the constraint (variable) embeddings; (c) The variable embeddings are aggregated into vehicle nodes \mathcal{W} by mean pooling over variables related to the same vehicle; (d) The forward passing of the aggregated embeddings is further processed through the MLP. The embeddings for each kind of nodes are being updated as the color deepens.

TABLE 1

Features for constraint, edge, variable and vehicle (Dim represents the dimension of each feature).

	Features	Description	Dim
\mathbf{c}_i^0	obj_cos_sim	Cosine similarity with objective.	1
	bias	Normalized bias value.	1
	type	The constraint form ($=, \leq$).	2
\mathbf{e}_{ij}^0	coef	Normalized constraint coefficient.	1
	type	Type (Integer, continuous).	2
\mathbf{v}_j^0	coef	Normalized objective coefficient.	1
	sol_val	Solution value.	1
	inc_val	Value in incumbent.	1
\mathbf{w}_h^0	is_used	Whether used or not in a solution.	2
	sol_cost	Normalized cost (distance) in a solution.	1

c_i , variable v_j , edge e_{ij} as \mathbf{c}_i^0 , \mathbf{v}_j^0 , \mathbf{e}_{ij}^0 , respectively, and describe them in TABLE 1, where \mathbf{w}_h^0 refers to the feature of each vehicle w_h (denoted by the green nodes in Fig. 3(c)).

4.3.3 Policy Network

We parameterize the policy by a graph convolutional network (GCN) π_θ , given its desirable capability in learning representation over graph-structured inputs and generalizing to different sizes. Specifically, given the state representation in Fig. 3(a), the network first executes the graph convolution by applying two sequential message passing, i.e., from variable to constraint nodes (\mathcal{V} to \mathcal{C}), and from constraint to variable nodes (\mathcal{C} to \mathcal{V}), respectively. Such convolution process is expressed as follows,

$$\mathbf{c}_i \leftarrow f_1([\mathbf{c}_i; g_1(\sum_{v_j \in \mathcal{E}} \mathbf{e}_{ij}^0 \mathbf{v}_j)]), \forall c_i \in \mathcal{C} \quad (14)$$

$$\mathbf{v}_j \leftarrow f_2([\mathbf{v}_j; g_2(\sum_{c_i \in \mathcal{E}} \mathbf{e}_{ij}^0 \mathbf{c}_i)]), \forall v_j \in \mathcal{V} \quad (15)$$

where \mathbf{c}_i and \mathbf{v}_j represent the embeddings of constraint c_i and variable v_j , which are initialized by linearly projecting their raw features \mathbf{c}_i^0 and \mathbf{v}_j^0 into 64-dimensional feature spaces, respectively; $[\cdot]$ means the concatenation operation; f_1, f_2, g_1 , and g_2 are multilayer perceptrons (MLPs) with a hidden layer and a ReLU activation. We keep all MLPs having hidden dimensions 128 and output dimensions 64, so that all embeddings are 64-dimensional. Accordingly, the variable and constraint embeddings are updated by the above convolutions, as shown in Fig. 3(b).

Since we aim to select partial variables for reoptimization, we aggregate the variable embeddings by *mean pooling* to nodes \mathcal{W} , each of which corresponds to a group of variables with the same index of fleet and vehicle. Then, the aggregated embeddings over each group of variables $w_h \in \mathcal{W}$ are concatenated with the 64-dimensional linear projection of the feature matrix $\mathcal{M}_w \in \mathbb{R}^{|\mathcal{W}| \times n_w}$ of nodes \mathcal{W} . The raw features of each vehicle in fleets (\mathbf{w}_h^0) are contained in each row of \mathcal{M}_w and described in TABLE 1. The above aggregation process (in Fig. 3(c)) is expressed as follows,

$$\mathbf{w}_h \leftarrow \left[\frac{1}{|\eta(w_h)|} \sum_{v_j \in \eta(w_h)} \mathbf{v}_j; \mathbf{w}_h \right], \forall w_h \in \mathcal{W} \quad (16)$$

where \mathbf{w}_h denotes the embedding of w_h and is initialized by linear projection of its raw features \mathbf{w}_h^0 ; $\eta(\cdot)$ is a function that gathers the set of variables with respect to a node w_h . Thus, the embedding \mathbf{w}_h represents a group of variables related to the same vehicle in a problem instance. As depicted in Fig. 3(d), we finally process these embeddings by another MLP (f_3) that is similar to the previous ones but with a single output, followed by sigmoid activation, such that,

$$P(w_h) = \text{Sigmoid}(f_3(\mathbf{w}_h)), \forall w_h \in \mathcal{W}. \quad (17)$$

Consequently, we attain the probability for each group of variables $w_h \in \mathcal{W}$, with which we could further sample variables related to certain vehicles for reoptimization.

4.3.4 Imitation Learning

We train our policy by imitation learning to mimic desirable behaviors of a heuristic destroy operator, which avoids learning from scratch in the huge combinatorial search space. Here, we apply our *Vehicle-random* multiple times and retrieve the best one to serve as demos for imitation learning. The well-known behavior *cloning* is a typical imitation learning algorithm [38], but it suffers from quadratic-growth compounding errors along the time horizon. In our method, we exploit a *forward training* algorithm, which is able to correct the cascading errors [39] by adaptively collecting demos based on the learned policy. The training details in one epoch are displayed in Algorithm 2. Particularly, at the t^{th} LNS iteration, we collect demos D_t by running *Vehicle-random* destroy operator n_s times on each training instance, and retrieving the ones with the best objectives (Lines 3-16). The demos with non-improved solutions x_j on some instances will not be collected (Line 13). Accordingly, $D_t = \{(p_u, x_u, a_u)\}_{u=1}^{\mathcal{N}_t}$ with \mathcal{N}_t denoting the number of

Algorithm 2: Forward Training for LNS

```

1: Input: a set of MILP instances  $\{p_i\}_{i=1}^m$  with initial
   solutions  $\{x_i\}_{i=1}^m$ ; number of iterations  $n_i$ ; number of
   sampling for an instance  $n_s$ ; Vehicle-random destroy
   operator  $\mathcal{D}$ ; repair operator  $\mathcal{R}$ ; percentage of decision
   variables to destroy  $\mathcal{D}_d$ .
2: for  $t = 1, \dots, n_i$  do
3:    $D_t = \square$ 
4:   for  $i = 1, \dots, m$  do
5:      $best\_obj, best\_D = c(x_i), \text{None}$ 
6:     for  $j = 1, \dots, n_s$  do
7:        $_, x_j, \mathcal{X}_r = \text{Algorithm 1}(p_i, x_i, 1, \mathcal{D}, \mathcal{R}, \mathcal{D}_d)$ 
8:       if  $c(x_j) < best\_obj$  then
9:          $best\_D = \mathcal{X}_r$ 
10:         $best\_obj = c(x_j)$ 
11:       if  $best\_D$  is not None then
12:          $D_t.append((p_i, x_i, best\_D))$ 
13:        $\pi_t = \text{SUPERVISE\_TRAIN}(\pi_{t-1}, D_t)$ 
14:       for  $i = 1, \dots, m$  do
15:          $_, x_i, _ = \text{Algorithm 1}(p_i, x_i, 1, \pi_t, \mathcal{R}, \mathcal{D}_d)$ 
16: return  $\pi_1, \pi_2, \dots, \pi_{n_i}$ 

```

collected demos in t^{th} iteration. Each element in D_t is a 3-tuple that contains the problem instance, current solution and action (i.e., the selected variables). We then represent them by the respective state as described in section 4.3.2, and thus obtain $D_t = \{(s_u, a_u)\}_{u=1}^{N_t}$. Thereafter, the policy network is trained with the collected data in a supervised manner (Line 17), to minimize the binary cross-entropy loss as follows,

$$\mathcal{L}(\theta) = -\frac{1}{\mathcal{N}_t |\mathcal{W}|} \sum_{(s,a) \in D_t} \sum_{w_h \in \mathcal{W}} [y(w_h) \cdot \log P(w_h) + (1 - y(w_h)) \cdot \log(1 - P(w_h))], \quad (18)$$

where $y(w_h) = 1$ if the group of variables w_h is selected otherwise $y(w_h) = 0$. In other words, we regard the action of selecting a subset of variables as determining whether a group of variables related to the same vehicle is selected. Given a state, it comes down to a binary classification problem for each group of variables. After the policy network is updated with the demos in iteration t , the new solution to each instance is attained by running the LNS with the trained policy π_t (Lines 18-20), and we continue collecting new data in the next iteration to repeat the training.

4.3.5 Deployment

Given the trained policy, we could apply it in the LNS framework as depicted in Fig. 2. Regarding how to select the groups of variables according to the output probabilities, we could deploy it in three different ways, i.e., *fix-sized sampling*, *adaptive sampling* and *disjoint sampling*, which accordingly yields three LNS variants as follows.

IL-sample: Given the destroy degree, we sample a fix-sized subset of variables in each LNS iteration, based on a softmax over probabilities for each group of variables w_h , such that, $\sigma_{w_h} = e^{P(w_h)} / \sum_{\kappa=1}^{|\mathcal{W}|} e^{P(w_\kappa)}$, $\forall w_h \in \mathcal{W}$. This deployment selects variables with a high potential to improve the solution.

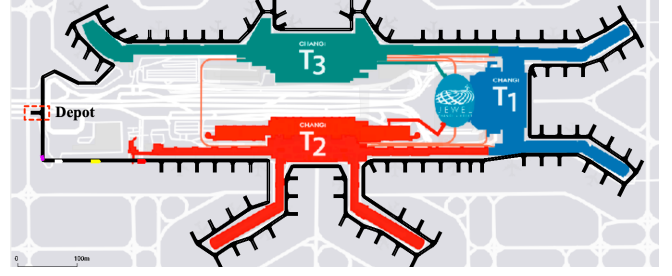


Fig. 4. The map of CHANGI airport with 3 terminals and 90 gates (where the most left point is the depot).

IL-sampleA: Each group of variables is sampled independently based on its output probability $P(w_h)$, $\forall w_h \in \mathcal{W}$, such that the number of variables to be optimized in each iteration is stochastic.

IL-sampleD: It follows the way in IL-sample, but imposes that the sampled subsets between consecutive LNS iterations are disjoint.

5 EXPERIMENTS

In this section, we evaluate our method for solving the vehicle routing problem in AGH with more realistic scenarios, where instances of 20, 50, 100 flights and 10 types of operations are first considered. We then compare our learning based LNS with the state-of-the-art commercial solver CPLEX (i.e., also the repair operator in our LNS) and conventional LNS heuristics to verify its superiority. We also show that our learning based LNS is able to generalize well to large instances even if the policy is trained on small instances. Subsequently, we verify the versatility of our method by adopting a different solver, i.e., OR-Tools, as the repair operator. Finally, we further evaluate the scalability of our method on larger instances of up to 200 flights. The code is publicly available at: <https://github.com/RoyalSkye/AGH>.

5.1 Settings

Instance Generation. Our experiments are conducted on a well-known international airport in Singapore, i.e., CHANGI airport, whose topology is shown in Fig. 4. It contains 3 terminals with 90 gates where the operations are executed. We load the map¹ into SUMO [40], by which the distances between gates and the depot are calculated automatically. On top of the map, we first generate AGH instances of 20, 50 and 100 flights which are referred as AGH20, AGH50 and AGH100, respectively. Particularly, the instances are generated based on statistics of the real scenario in CHANGI airport. For the flight density, we follow the scheduling table in one day², and randomly generate the number of hourly aircraft arrivals from [5, 25], where 5 and 25 correspond to the flight quantities in the idlest and busiest periods, respectively. The duration of a turnaround for each aircraft is sampled from [30, 60] minutes. We consider three types of aircraft with respective precedence relations according to their manufacturers [41], [42], the first

1. <https://www.changiairport.com/en/maps.html>
2. <https://www.changiairport.com/en/flights.html>

TABLE 2
Comparison with Baselines.

Methods	AGH20 (6k/15k)				AGH50 (30k/140k)				AGH100 (120k/700k)									
	Mipgap	Limit	\mathcal{D}_d	Obj.	Gap	Time	Mipgap	Limit	\mathcal{D}_d	Obj.	Gap	Time	Mipgap	Limit	\mathcal{D}_d	Obj.	Gap	Time
CPLEX	1e-04	1m	-	86384	12.61%	1m	1e-04	5m	-	365353	152.80%	5m	1e-04	30m	-	696762	178.58%	30m
CPLEX ($\times 5$)	1e-04	5m	-	77630	1.65%	5m	1e-04	25m	-	341397	134.65%	25m	1e-04	2.5h	-	684509	173.82%	2.5h
Decomposition [14]	1e-04	45m	-	107269	41.44%	45m	1e-04	47m	-	215620	47.70%	47m	1e-04	57m	-	369966	48.01%	57m
GA [10]	-	1h	-	115706	51.71%	1h	-	1h	-	261137	77.60%	1h	-	1h	-	438542	76.74%	1h
Random	0.1	10s	0.4	151562	99.37%	1m	0.1	1m	0.4	332537	128.76%	5m	0.2	10m	0.4	491285	96.42%	30m
Vehicle-random	0.1	10s	0.4	86301	13.04%	1m	0.1	1m	0.4	166735	14.66%	5m	0.2	10m	0.4	311035	24.21%	30m
Vehicle-worst	0.1	10s	0.4	98383	28.31%	1m	0.1	1m	0.4	223397	53.43%	5m	0.2	10m	0.4	663911	164.62%	30m
Vehicle-worst-random	0.1	10s	0.4	90477	18.32%	1m	0.1	1m	0.4	170875	17.58%	5m	0.2	10m	0.4	316498	26.31%	30m
IL-sample (Ours)	0.1	10s	0.4	85942	12.53%	1m	0.1	1m	0.4	166080	14.24%	5m	0.2	10m	0.4	302512	20.86%	30m

one of which has been shown in Fig. 1. The others are similar to the first type yet much harder, i.e., the priority of catering could be 1 or 2 for the second type; the cleaning could be prioritized as 1 or 2 and catering can be executed anytime before pushback for the third type. The three aircraft types are uniformly across all the flights, which are also uniformly assigned to the gates. For the 10 operations to be scheduled, their service durations, time windows and vehicle speeds are set individually following [43], [44]. The number of vehicles in each fleet is randomly sampled from [10, 20]. We randomly generate the flight demands for each operation from [5, 15], and set the ratio of demand to capacity to a random value in [0.7, 0.9] following [45].

Baselines. We compare our learning based LNS with the state-of-the-art exact solver for MILP, and various conventional LNS heuristics. In specific, the baselines include, 1) the exact solver CPLEX [1] (we also use OR-Tools [11] instead to verify the versatility of our method), which is also the repair operator in the LNS methods; 2) *Random*, which destroys the current solution by randomly picking a subset of integer decision variables to reoptimize; 3) *Vehicle-random*, which is used to collect demos for imitation learning as described in section 4.2; 4) *Vehicle-worst*, which sorts all used vehicles according to their travelling distance in the current solution, and selects the variables related to those vehicles with the longest distances; 5) *Vehicle-worst-random*, which selects $\mathcal{D}_d/2$ of variables with *Vehicle-random* and $\mathcal{D}_d/2$ of variables with *Vehicle-worst*, given a destroy degree \mathcal{D}_d ; 6) *Decomposition*, which tackled a similar problem to ours and reported favorable performance by first decomposing it into sub-problems for each fleet using constraint programming (CP) and then applying LNS heuristic to solve them [14]. We adapted it to solve the same problem as ours. Note that the work in [14] is the recent one for solving AGH problems whose settings are most similar to ours [23]; 7) *GA*, which solved the AGH problem by a genetic algorithm [10]. While its original work only considers one type of aircraft, we adapt the GA method to our setting with heterogeneous aircraft. Other specialized approaches for VRPs with individual operations are not considered as baselines, since they are less relevant to the studied problem in this paper.

For LNS heuristics and our method, we generate initial solutions with a simple heuristic based on nearest neighbor

insertion [46], which is commonly used to initialize the solutions for routing problems [47], [48], [49]. Specifically, each route, starting from the depot, is iteratively extended with the nearest unserved flight (which also satisfies the basic constraints), until no such flights exist and then a new route starts. Meanwhile, the routes are constructed for each fleet following the precedence relation among each other. We note that the above insertion heuristic basically picks the nearest aircraft which satisfies all constraints and it scarcely needs domain knowledge for the problem. Alternatively, we can also simply run a solver (e.g. CPLEX) with a time limit to attain initial solutions.

Hyperparameters in LNS Heuristics. We use CPLEX 12.10 as the solver in baseline 1, and as the repair operators in baselines 2-5 and our method. We set two parameters in CPLEX (i.e., mipgap³ and timelimit) as stopping conditions in each LNS iteration. The current LNS iteration stops when either of the two conditions is met. In particular, we set mipgap to 0.1 for AGH20 and AGH50, 0.2 for AGH100. The timelimit is set to 10s, 1m and 10m respectively, though they could be longer than the real runtime in iterations. The acceptance criterion is that in each iteration, we only accept the solution whose objective value is smaller or at most 1% larger than that of the incumbent. Typically, we find that all LNS heuristics considerably improve solutions between 5 and 10 iterations, so we empirically set the total runtime to 1m, 5m and 30m for AGH20, AGH50 and AGH100, respectively. Since the destroy degree cannot be too large or small (as analyzed in Section 4.2), we have tried 0.3, 0.4, 0.5 and 0.6 in our experiment. We observe that when it equals to 0.4, all LNS heuristics (including *Vehicle-random* LNS that generates demos for the policy network) perform relatively well on the validation set. Therefore, we set it to 0.4 for all LNS heuristics on AGH20, AGH50 and AGH100.

Hyperparameters in Learning based LNS. We train the model only on instances of AGH20 to learn the policy (which consumes about five days), and test it on instances of AGH20, AGH50 and AGH100 to evaluate its performance. Regarding AGH20, we generate 70 instances for training, 30 instances for validation and 50 instances for testing. For each of other sizes, we also generate 50 instances for testing. During the training, we set the epoch number to 10 according

3. The relative gap between the lower and upper objective bound.

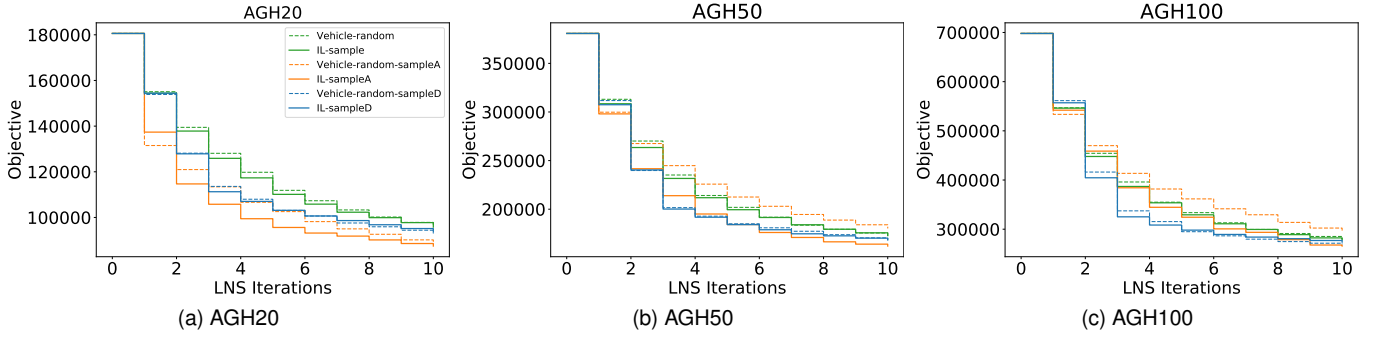


Fig. 5. The objective values of incumbents against iterations for the learned LNS and heuristic counterparts.

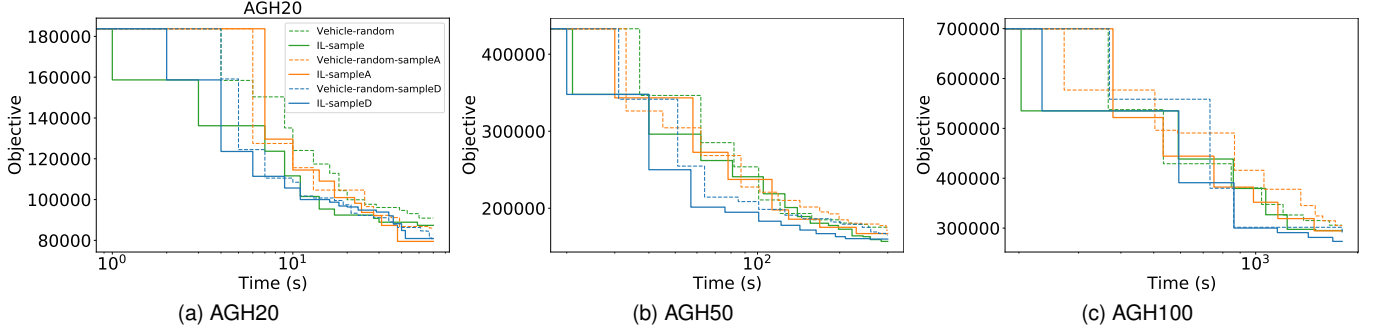


Fig. 6. The objective values of incumbents against runtime for the learned LNS and heuristic counterparts.

to the convergence of the training curve. In each epoch, we set the iteration number of LNS $n_i = 5$, since we observe its performance cannot be significantly improved afterwards. In each iteration of the forward training algorithm (i.e., each LNS iteration), we run Vehicle-random LNS $n_s = 10$ times for each instance in the training set, and keep the best one as a demo. Accordingly, the stopping condition, acceptance criterion and destroy degree are set following the hyperparameters of Vehicle-random LNS, as mentioned above. After the demos are collected, we train the neural network by the stochastic gradient descent (SGD) optimizer [50], with the learning rate equals to 10^{-4} and the batch size equals to 16 due to GPU memory limit. For validation, we set the iteration number $n_i = 10$ and the destroy degree $D_d = 0.4$. For testing, we keep the destroy degree same as the training, and run the learned LNS with the same time as other baselines. All experiments are conducted on a machine with 32 AMD EPYC 7601 32-Core Processors.

5.2 Comparison Study

We compare the learning based LNS with other baselines including CPLEX solver, conventional LNS heuristics (with CPLEX as the repair operator), the decomposition method in [14] and the GA method adapted from [10]. We set the total runtime of the above methods to 1m, 5m, 30m for AGH20, AGH50 and AGH100, respectively, since beyond which we found that solutions in LNS methods only have marginal improvement. Regarding CPLEX, we not only apply it to solve the instances of the MILP model using the same time with other methods, but also allow it to run as 5 times long as others. For the method in [14], we reproduce the algorithm using its original parameters, and allow it to run much longer time than ours. We also allow the GA

method to run a long time, i.e., 1h for all problem scales. For a fair comparison, we only deploy the learned destroy operator with the fix-sized sampling strategy as stated in section 4.3.5, i.e., IL-sample, and keep $D_d = 0.4$. Then we compare the average objective value of the best found solutions over the testing set for all methods, and the primal gap⁴ [51] since we are unable to acquire the optimal solution to calculate the optimality gap. All results are gathered in Table 2, where we also display the average number of constraints and variables (in brackets) over the MILP formulations for different problem sizes, along with some key parameters for each method. We can easily observe that, 1) all LNS methods including our learning based one, with the same runtime, can significantly outperform the CPLEX solver on AGH50 and AGH100 in terms of objective value and gap; 2) though CPLEX has an obvious advantage over other LNS methods on AGH20, the learned LNS by our method still outstrips it with the same runtime; 3) when running with much longer time, CPLEX can achieve the best result on AGH20 since the instances are relatively small and tractable for the exact solver, while it is inferior to LNS methods on AGH50 and AGH100, with a growing gap as the problem scales up; 4) for LNS methods, random destroy operator has clearly larger gap than the three vehicle based destroy operators, where Vehicle-random and Vehicle-worst-random demonstrate similar performance and outperform Vehicle-worst; 5) the learned LNS surpasses all other LNS heuristics on AGH20, AGH50 and AGH100, indicating that our method is able to learn a more effective destroy operator to guide the LNS algorithm; 6) despite the extra runtime, the

4. We first compute $|c^T \bar{x} - c^T \bar{x}^*| / \max\{|c^T \bar{x}|, |c^T \bar{x}^*|\} \cdot 100\%$ for each instance, where \bar{x} is the solution by a method and \bar{x}^* is the best one out of all methods, and then average the gaps over the testing set.

TABLE 3
Results with OR-Tools.

Methods	AGH20		AGH50		AGH100	
	Obj.	Gap	Obj.	Gap	Obj.	Gap
OR-Tools	80073	4.07%	150835	3.12%	323762	30.47%
Random	135183	75.86%	261293	78.77%	466692	88.29%
Vehicle-random	88205	14.34%	159680	9.15%	313815	26.63%
Vehicle-worst	88413	14.85%	184750	26.33%	502614	102.82%
Vehicle-worst-random	88208	14.53%	162148	10.81%	352973	42.20%
IL-sample(Ours)	86479	12.27%	151488	3.59%	300160	20.97%

decomposition and GA method are generally comparable to some LNS heuristics but far inferior to our learnt one.

The method in [14] adopts constraint programming (CP) to decompose the original problem into a VRPTW for each fleet, then applies LNS to solve each resulting VRPTW. In the LNS, the random pivot and small routing operators [52], [53] are particularly designed to remove flights or edges in the route based on various criteria (e.g., geographic and temporal proximity), and CPLEX solver is used to rebuild the destroyed solution. In contrast, we exploit learned LNS to decompose the original problem by directly selecting variables from the MILP formulation, which circumvents much domain knowledge. It is clearly demonstrated that our method is much more effective. A possible reason is that the one in [14] cannot handle well the impact on subsequent operations when scheduling for the early or current ones. While ours is able to simultaneously optimize any combination of variables with respect to different operations, vehicles and so on. On the other hand, the GA method needs to maintain a population of solutions, with higher demands in computational time and space (memory usage), especially for large-scale instances. Meanwhile, the design of crossover and mutation operators requires much domain knowledge to satisfy the complex constraints and ensure the feasibility of offspring. We circumvent such issues by directly using the off-the-shelf solver and learning to decompose the computation-heavy problem into a series of small sub-problems.

5.3 Learned LNS vs Heuristic Counterparts

In our method, rather than manually designing complex operators, we attempt to automatically learn destroy operators with imitation learning, the demos of which are selected among multiple solutions by a heuristic, i.e., Vehicle-random LNS. We wonder that whether the learned destroy operator, based on the GCN representation of problem instances and solving states, could achieve better performance than the corresponding heuristic one. To verify this point, we compare the learning based LNS with Vehicle-random LNS, which is used as the provider of demos in imitation learning and also has shown better performance than other LNS heuristics on AGH20, AGH50 and AGH100 in Table 2. As discussed in section 4.3.5, we can deploy the learned destroy operator in two more ways besides the IL-sample used in section 5.2, i.e., IL-sampleA and IL-sampleD. For fair comparisons, we also apply Vehicle-random destroy operator in similar manners as follows, 1)

TABLE 4
Results on AGH200 (450k/4000k).

Method	Mipgap	Timelimit	Destroy degree	Obj.	Gap	Time
CPLEX	1e-04	2h	-	1322757	152.72%	2h
Decomposition	1e-04	2.2h	-	663254	26.73%	2.2h
Random	0.2	30m	0.2	1321166	152.41%	2h
Vehicle-random	0.2	30m	0.2	666801	27.33%	2h
Vehicle-worst	0.2	30m	0.2	1080257	106.66%	2h
Vehicle-worst-random	0.2	30m	0.2	741957	41.27%	2h
IL-sample (Ours)	0.2	30m	0.2	652405	24.59%	2h
OR-Tools	-	2h	-	1288353	146.28%	2h
Random	-	30m	0.2	1318479	151.97%	2h
Vehicle-random	-	30m	0.2	1084422	107.24%	2h
Vehicle-worst	-	30m	0.2	1120329	114.09%	2h
Vehicle-worst-random	-	30m	0.2	1087186	107.76%	2h
IL-sample (Ours)	-	30m	0.2	1015181	94.05%	2h

Vehicle-random-sampleA, which applies Vehicle-random by using flexible destroy degree in each iteration, i.e., selecting a subset of variables with the percentage sampled from a range⁵; 2) Vehicle-random-sampleD, which is applied on top of Vehicle-random destroy operator, and imposes that the selected subsets between consecutive LNS iterations are disjoint. For IL-sampleA and IL-sampleD, we directly use the trained model of IL-sample with their respective sampling strategies in the deployment. We evaluate the three different deployments for both the learned LNS and Vehicle-random LNS with 10 iterations on testing instances of all problem sizes. Fig. 5 shows the average objective values by each method against iterations.

In general, we observe from Fig. 5 that IL-sampleA consistently achieves better solutions than the three different heuristic LNS on all problem sizes. It means that we can learn a more effective destroy operator from the simple Vehicle-random LNS, which guides LNS to search better solutions more efficiently. Also, the three deployments of the learned policy are better than or on par with their heuristic counterparts, which are summarized as below,

IL-sample vs Vehicle-random. With the same destroy degree (i.e., $\mathcal{D}_d = 0.4$), the results indicate that the learned policy can select more suitable variables in each iteration, which generally leads to more efficient improvement of the solutions for problem instances of different sizes.

IL-sampleA vs Vehicle-random-sampleA. Both methods dynamically select a subset of variables to be optimized in each iteration. However, IL-sampleA significantly outperforms Vehicle-random-sampleA, which suggests that the learned policy, based on GCN representation of the instance and solving state, is able to construct desirable sub-MILPs that lead to better solutions along the solving process.

IL-sampleD vs Vehicle-random-sampleD. When the selected variables between consecutive iterations are exclusive, IL-sampleD is on par with its counterpart for all problem sizes. It is fairly acceptable, given the fact that we do not

5. We set it to [0.2, 0.7], [0.2, 0.5], [0.2, 0.4] for AGH20, AGH50, AGH100 respectively, to render sub-MILPs tractable for each problem size and also prevent infeasible solutions.

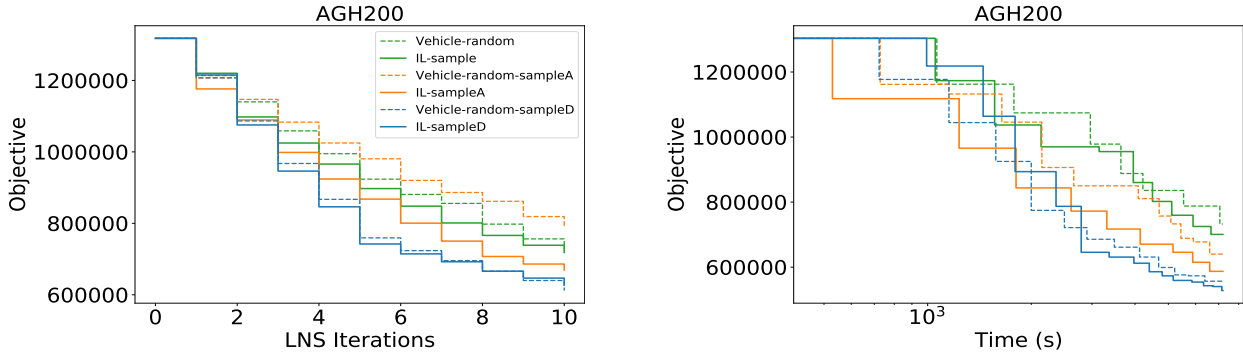


Fig. 7. Comparison between the learned LNS and heuristic counterparts (left: the objective values of incumbents against iterations; right: the objective values of incumbents against runtime).

intentionally involve any disjoint sample during training, since it directly uses the trained model of IL-sample.

In addition, we further evaluate the trends of solution improvement along the runtime for each deployment. Specifically, we run all these methods with the same time as in Table 2, i.e., 1m, 5m, 30m for AGH20, AGH50, AGH100, respectively. Fig. 6 shows the incumbent solutions against the runtime. We observe that the deployments of the learned destroy operator can generally retain the incumbents along the solving process for all instances, which means they are more efficient than heuristic destroy operators to achieve better solutions. It is also revealed that there is no single deployment dominating all the rest. On AGH20, IL-sample and IL-sampleD alternately achieve incumbents at the early stage, while finally IL-sampleA attains the lowest objective value. On AGH50, IL-sampleD retains the incumbent solution along almost the whole solving process. On AGH100, IL-sample, IL-sampleA and IL-sampleD yield the incumbent in early, middle and late stages, respectively.

Considering that the policy network is only trained with demos of Vehicle-random on AGH20, the above results indicate that our method can deliver a more effective destroy operator than the heuristic counterpart, and the deployments with different sampling strategies may also further boost the performance, i.e., adaptive or disjoint sampling.

5.4 Versatility Analysis

One desirable potential of the proposed method is that we can integrate it with any off-the-shelf solver, which saves considerable efforts in manually designing the repair operator. To verify this merit of versatility, we replace CPLEX with OR-Tools (v8.1), i.e., a popular open source solver, as the repair operator in the LNS methods. In specific, we train the policy network on instances of AGH20, and test the learned LNS and LNS heuristics on instances of AGH20, AGH50 and AGH100 as similarly done for CPLEX in section 5.2. We keep the destroy degree of each method the same as the ones in Table 2, and ignore the mipgap since it is not explicitly accessible in OR-Tools. For all LNS methods, we set the time limit in each iteration to 25s, 120s and 500s on AGH20, AGH50 and AGH100, and the total runtime is set to 4m, 30m and 1.5h, respectively. We record the average objective values and the primal gap over the testing set, and display them in Table 3. Note that we only consider the deployment of IL-sample as the learning method for demonstration purpose.

We observe that IL-sample still works fairly well as a learning scheme to guide OR-Tools. Compared to other LNS baselines which are generally inferior to OR-Tools (except Vehicle-random on AGH100), we can see that IL-sample performs consistently better on all problem sizes. On AGH20, IL-sample significantly outperforms other LNS heuristics, and on AGH50 it becomes comparable to the standalone OR-Tools. On AGH100, IL-sample achieves considerably smaller objective value and gap than all other methods, revealing the beneficial property of our method on the large-scale problem. Therefore, the versatility of our framework is well verified, which can yield competitive performance even if we adopt different solvers as the repair operator.

5.5 Scalability Analysis

Our method has already demonstrated desirable scalability property, however, we would like to further apply it to solve much larger instances, i.e., AGH200. Although the most recent literature on applying deep (reinforcement) learning to solve CVRP or CVRPTW [30], [31], [54], [55] also deal with large instances, the problem settings in those works are quite standard or simple, which are far easier than the ones in this paper. Therefore, we are keen on evaluating the performance of our method on AGH200. To this end, we generate 20 instances with 200 flights, and solve them by solvers (i.e. CPLEX and OR-Tools), the decomposition method [14] following previous setting, and LNS methods with CPLEX and OR-Tools as the repair operators, respectively⁶. For all these methods, we set the time limit and total runtime to 30m and 2h, respectively. In specific, we set mipgap to 0.2 for CPLEX in each iteration, and set the same destroy degree $\mathcal{D}_d = 0.2$ for all LNS methods to keep sub-problems tractable. For the learned LNS, we still use the one trained on AGH20 and only consider fix-sized sampling strategy for inference, i.e., IL-sample. Then, we compute the average objective and primal gap over the tested instances for each method. All results are displayed in Table 4, where the upper and lower half show the results for methods associated with CPLEX and OR-Tools, respectively. It is clear that no matter which solver is used as a repair operator, our method significantly outperforms other baselines, with the smallest average objective and gap. Most of the LNS methods demonstrated salient advantages over CPLEX and

6. We did not compare with GA here since it consumes too much memory to keep the population for AGH200, and fails in our device.

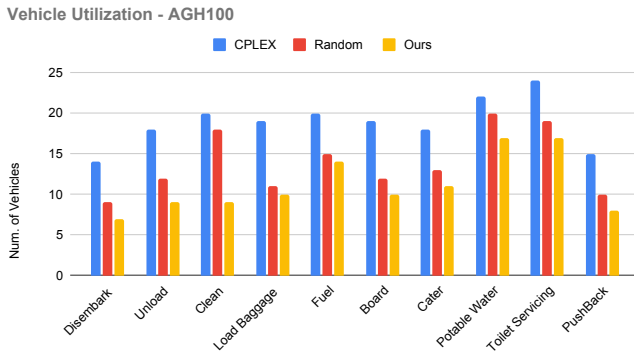


Fig. 8. The average number of occupied vehicles in different fleets.

OR-Tools, since they can iteratively search promising solutions by solving sequential sub-problems. The decomposition method with slightly longer runtime surpasses LNS heuristics, but it is still far inferior to IL-sample (Ours). The superiority of our method comes from the fact that, in comparison with LNS heuristics, it can effectively guide LNS by determining more appropriate sub-MILPs that lead to better solutions. It is also suggested that the policy trained for small size of instances can generalize well to large ones, which may result from the desirable scaling property of GCN for learning the underlying representation of MILP.

To further verify that the learned policy is superior to the heuristic on AGH200, we compare their different deployments as did in section 5.3. Specifically, we evaluate the three deployments on test instances with 10 iterations, and the curves of average objective for each deployment are shown in the left panel of Fig. 7. As shown, IL-sample is able to improve the incumbent more efficiently than Vehicle-random, which provides the demos for training. We also find that though we do not involve other schemes (i.e. adaptive and disjoint sampling) during training, the results achieved by IL-sampleA and IL-sampleD are still better than or on par with their counterparts, i.e., Vehicle-random-sampleA and Vehicle-random-sampleD, respectively. Furthermore, we also present curves to show the objective against 2h runtime. From the right panel of Fig. 7, we observe that the deployments of the learned LNS generally outperform their heuristic counterparts, which again indicates that the LNS with the learned destroy operator is more effective in boosting the solution quality.

5.6 Effect on Vehicle Utilization

Despite that we merely set the objective to explicitly minimize the total tour length of vehicles, we empirically observe that our method is able to concurrently achieve good performance in terms of the vehicle utilization. To manifest this point, we solve AGH100 instances in the testing set by CPLEX, Random LNS and our method as did in section 5.2, and record the average number of occupied vehicles in every fleet over all the instances. The results are displayed in Fig. 8. It is clear that our method attains a better solution which employs significantly less vehicles in each fleet, in comparison with CPLEX and Random LNS. Our method implicitly enables a reasonable vehicle utilization in each

fleet, serving all aircraft on time with a small number of vehicles. Considering that an airport generally possesses a limited number of vehicles for AGH, our method is more suited to make the best usage of this critical resource.

We would like to note that this paper mainly contributes to a learning based framework to solve AGH with multiple heterogeneous operations, which has been verified superior to various baselines. Nevertheless, the adaptation of our method to AGH with different objective functions is straightforward, by modelling them with an off-the-shelf solver and substituting the objective in this paper. In the future, we will apply our method to solve AGH with other potential objectives to further verify its effectiveness, such as explicitly minimizing vehicle utilization [56] and maximizing safety to prevent accidents [18], [19].

6 CONCLUSIONS AND FUTURE WORKS

To our knowledge, we are the first one to leverage learning based LNS framework to solve the complex vehicle routing problem in AGH with multiple operations and hundreds of flights. In this framework, we integrate imitation learning and GCN to guide the destroy operator to automatically select variables in each iteration, and adopt the off-the-shelf solver as the repair operator to reoptimize the selected variables, which circumvents substantial human efforts for designing the rules in conventional LNS heuristics. Results demonstrate that our method has a strong capability in delivering high-quality solutions compared to the exact solvers and conventional heuristics, with desirable properties in versatility and scalability. Hence, the proposed method has a favorable potential that we could fast deploy an off-the-shelf solver and automatically raise its performance without much manual work or domain knowledge.

As mentioned, this work is an early attempt on applying learning based method to tackle the complex vehicle routing problem in airports. In future, we will, 1) investigate deep reinforcement learning to learn the destroy operators, and test our method on other datasets; 2) study how to enable the destroy operator to automatically decide the number of variables (i.e., the size of subproblems) to be optimized; 3) take into account the safety factor to avoid potential accidents in the organization of traffic; 4) extend the policy network to optimize multiple objectives in the MILP model.

ACKNOWLEDGEMENTS

This research was conducted in collaboration with Singapore Telecommunications Limited and supported by the Singapore Government through the Industry Alignment Fund - Industry Collaboration Projects Grant. It was also supported by the National Natural Science Foundation of China (Grant No. 62102228), and the Natural Science Foundation of Shandong Province (Grant No. ZR2021QF063).

REFERENCES

- [1] IBM ILOG, "Cplex," 2022. [Online]. Available: <https://www.ibm.com/analytics/cplex-optimizer>
- [2] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2022. [Online]. Available: <https://www.gurobi.com>

- [3] A. Norin, D. Yuan, T. A. Granberg, and P. Värbrand, "Scheduling de-icing vehicles within airport logistics: a heuristic algorithm and performance evaluation," *Journal of the Operational Research Society*, vol. 63, no. 8, pp. 1116–1125, 2012.
- [4] J. Y. Du, J. O. Brunner, and R. Kolisch, "Planning towing processes at airports more efficiently," *Transportation Research Part E: Logistics and Transportation Review*, vol. 70, pp. 293–304, 2014.
- [5] S. C. Ho and J. M. Leung, "Solving a manpower scheduling problem for airline catering using metaheuristics," *European Journal of Operational Research*, vol. 202, no. 3, pp. 903–921, 2010.
- [6] Y. Du, Q. Zhang, and Q. Chen, "Aco-ih: An improved ant colony optimization algorithm for airport ground service scheduling," in *IEEE International Conference on Industrial Technology*, 2008, pp. 1–6.
- [7] Z. Zhou, S. Liu, and K. Huang, "Research on airport trailer emergency scheduling model based on genetic simulation annealing algorithm," in *IOP Conference Series: Materials Science and Engineering*, vol. 383, no. 1. IOP Publishing, 2018, p. 012044.
- [8] W. Guo, P. Xu, Z. Zhao, L. Wang, L. Zhu, and Q. Wu, "Scheduling for airport baggage transport vehicles based on diversity enhancement genetic algorithm," *Natural Computing*, vol. 19, no. 4, pp. 663–672, 2020.
- [9] S. Padrón, D. Guimarans, J. J. Ramos, and S. Fitouri-Trabelsi, "A bi-objective approach for scheduling ground-handling vehicles in airports," *Computers & Operations Research*, vol. 71, pp. 34–53, 2016.
- [10] Y. Liu, J. Wu, J. Tang, W. Wang, and X. Wang, "Scheduling optimisation of multi-type special vehicles in an airport," *Transportmetrica B: Transport Dynamics*, pp. 1–17, 2021.
- [11] L. Perron and V. Furnon, "Or-tools," Google. [Online]. Available: <https://developers.google.com/optimization/>
- [12] A. Norin, T. A. Granberg, P. Värbrand, and D. Yuan, "Integrating optimization and simulation to gain more efficient airport logistics," in *Eighth USA/Europe Air Traffic Management Research and Development Seminar*, 2009, p. 10.
- [13] X. Han, P. Zhao, and D. Kong, "A bi-objective optimization of airport ferry vehicle scheduling based on heuristic algorithm: A real data case study," *Advances in Production Engineering & Management*, vol. 17, no. 2, 2022.
- [14] S. Padrón and D. Guimarans, "An improved method for scheduling aircraft ground handling operations from a global perspective," *Asia-Pacific Journal of Operational Research*, vol. 36, no. 04, p. 1950020, 2019.
- [15] J. Zhang, X. Chong, Y. Wei, Z. Bi, and Q. Yu, "Optimization of apron support vehicle operation scheduling based on multi-layer coding genetic algorithm," *Applied Sciences*, vol. 12, no. 10, p. 5279, 2022.
- [16] S. Yıldız, O. Aydemir, A. Memiş, and S. Varlı, "A turnaround control system to automatically detect and monitor the time stamps of ground service actions in airports: a deep learning and computer vision based approach," *Engineering Applications of Artificial Intelligence*, vol. 114, p. 105032, 2022.
- [17] I. Alomar and J. Tolujevs, "Optimization of ground vehicles movement on the aerodrome," *Transportation Research Procedia*, vol. 24, pp. 58–64, 2017.
- [18] A. V. Shvetsov, "Analysis of accidents resulting from the interaction of air and ground vehicles at airports," *Transportation research procedia*, vol. 59, pp. 21–28, 2021.
- [19] S. V. Shvetsova and A. V. Shvetsov, "Ensuring safety and security in employing drones at airports," *Journal of Transportation Security*, vol. 14, no. 1, pp. 41–53, 2021.
- [20] Y. S. Gök, D. Guimarans, P. J. Stuckey, M. Tomasella, and C. Ozturk, "Robust resource planning for aircraft ground operations," in *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Springer, 2020, pp. 222–238.
- [21] Y. S. Gök, M. Tomasella, D. Guimarans, and C. Ozturk, "A simheuristic approach for robust scheduling of airport turnaround teams," in *2020 Winter Simulation Conference (WSC)*. IEEE, 2020, pp. 1336–1347.
- [22] G. Andreatta, L. Capanna, L. De Giovanni, M. Monaci, and L. Righi, "Efficiency and robustness in a support platform for intelligent airport ground handling," *Journal of Intelligent Transportation Systems*, vol. 18, no. 1, pp. 121–130, 2014.
- [23] K. Ng, C. K. Lee, F. T. Chan, and Y. Lv, "Review on meta-heuristics approaches for airside operation research," *Applied Soft Computing*, vol. 66, pp. 104–133, 2018.
- [24] P. Shaw, "Using constraint programming and local search methods to solve vehicle routing problems," in *International conference on principles and practice of constraint programming*. Springer, 1998, pp. 417–431.
- [25] S. Ropke and D. Pisinger, "An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows," *Transportation science*, vol. 40, no. 4, pp. 455–472, 2006.
- [26] S. Jain and P. Van Hentenryck, "Large neighborhood search for dial-a-ride problems," in *International Conference on Principles and Practice of Constraint Programming*. Springer, 2011, pp. 400–413.
- [27] G. M. Ribeiro and G. Laporte, "An adaptive large neighborhood search heuristic for the cumulative capacitated vehicle routing problem," *Computers & operations research*, vol. 39, no. 3, pp. 728–735, 2012.
- [28] D. Sacramento, D. Pisinger, and S. Ropke, "An adaptive large neighborhood search metaheuristic for the vehicle routing problem with drones," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 289–315, 2019.
- [29] J. Xiao, T. Zhang, J. Du, and X. Zhang, "An evolutionary multiobjective route grouping-based heuristic algorithm for large-scale capacitated vehicle routing problems," *IEEE transactions on cybernetics*, 2019.
- [30] A. Hottung and K. Tierney, "Neural large neighborhood search for the capacitated vehicle routing problem," in *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI)*, 2020.
- [31] L. Gao, M. Chen, Q. Chen, G. Luo, N. Zhu, and Z. Liu, "Learn to design the heuristics for vehicle routing problem," *arXiv preprint arXiv:2002.08539*, 2020.
- [32] J. Song, R. Lanka, Y. Yue, and B. Dilkina, "A general large neighborhood search framework for solving integer linear programs," *Advances in Neural Information Processing Systems*, vol. 33, 2020.
- [33] G. Ghiani, G. Laporte, and E. Manni, "Model-based automatic neighborhood design by unsupervised learning," *Computers & Operations Research*, vol. 54, pp. 108–116, 2015.
- [34] T. Adamo, G. Ghiani, E. Guerriero, and E. Manni, "A learn-and-construct framework for general mixed-integer programming problems," *International Transactions in Operational Research*, vol. 27, no. 1, pp. 9–25, 2020.
- [35] L. Feng, Y. Huang, L. Zhou, J. Zhong, A. Gupta, K. Tang, and K. C. Tan, "Explicit evolutionary multitasking for combinatorial optimization: A case study on capacitated vehicle routing problem," *IEEE transactions on cybernetics*, vol. 51, no. 6, pp. 3143–3156, 2020.
- [36] J. Duan, Z. He, and G. G. Yen, "Robust multiobjective optimization for vehicle routing problem with time windows," *IEEE Transactions on Cybernetics (Early Access)*, 2021.
- [37] G. Schrimpf, J. Schneider, H. Stamm-Wilbrandt, and G. Dueck, "Record breaking optimization results using the ruin and recreate principle," *Journal of Computational Physics*, vol. 159, no. 2, pp. 139–171, 2000.
- [38] D. Pomerleau, "An autonomous land vehicle in a neural network," *Advances in neural information processing systems*, vol. 1, 1998.
- [39] S. Ross and D. Bagnell, "Efficient reductions for imitation learning," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 661–668.
- [40] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic traffic simulation using sumo," in *The 21st IEEE International Conference on Intelligent Transportation Systems*, 2018. [Online]. Available: <https://elib.dlr.de/124092/>
- [41] S. Airbus, "A320 airplane characteristics for airport planning," *Issue: Sep*, vol. 30, p. 85, 1995.
- [42] B. C. Airplanes, *767 airplane characteristics for airport planning*. Boeing Commercial Airplanes Seattle, WA, USA, 2005.
- [43] 3rd, "Airside driving theory handbook," Ph.D. dissertation, 2010.
- [44] S. S. De Vicente, "Ground handling simulation with cast," Ph.D. dissertation, 2010.
- [45] I.-M. Chao, B. Golden, and E. Wasil, "A computational study of a new heuristic for the site-dependent vehicle routing problem," *INFOR: Information Systems and Operational Research*, vol. 37, no. 3, pp. 319–336, 1999.
- [46] M. M. Solomon, "Algorithms for the vehicle routing and scheduling problems with time window constraints," *Operations research*, vol. 35, no. 2, pp. 254–265, 1987.
- [47] M. Battarra, M. Monaci, and D. Vigo, "An adaptive guidance approach for the heuristic solution of a minimum multiple trip vehicle routing problem," *Computers & Operations Research*, vol. 36, no. 11, pp. 3041–3050, 2009.
- [48] C. Yin, L. Bu, and H. Gong, "Mathematical model and algorithm of split load vehicle routing problem with simultaneous delivery and

pickup," *International journal of innovative computing, information and control*, vol. 9, no. 11, pp. 4497–4508, 2013.

- [49] O. Bräysy, W. Dullaert, and M. Gendreau, "Evolutionary algorithms for the vehicle routing problem with time windows," *Journal of Heuristics*, vol. 10, no. 6, pp. 587–611, 2004.
- [50] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.
- [51] E. B. Khalil, B. Dilkina, G. L. Nemhauser, S. Ahmed, and Y. Shao, "Learning to run heuristics in tree search," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 2017, pp. 659–666.
- [52] D. Guimarans, "Hybrid algorithms for solving routing problems," Ph.D. dissertation, PhD thesis, Autonomus University of Barcelona, 2012.
- [53] L.-M. Rousseau, M. Gendreau, and G. Pesant, "Using constraint-based operators to solve the vehicle routing problem with time windows," *Journal of heuristics*, vol. 8, no. 1, pp. 43–58, 2002.
- [54] Y. Xu, M. Fang, L. Chen, G. Xu, Y. Du, and C. Zhang, "Reinforcement learning with multiple relational attention for solving vehicle routing problems," *IEEE Transactions on Cybernetics (Early Access)*, 2021.
- [55] L. Xin, W. Song, Z. Cao, and J. Zhang, "Multi-decoder attention model with embedding glimpse for solving vehicle routing problems," in *Proceedings of 35th AAAI Conference on Artificial Intelligence*, 2021, pp. 12 042–12 049.
- [56] P. Zhao, W. Gao, X. Han, and W. Luo, "Bi-objective collaborative scheduling optimization of airport ferry vehicle and tractor," *International Journal of Simulation Modelling*, vol. 18, no. 2, pp. 355–365, 2019.



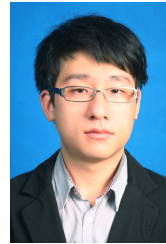
Jianan Zhou received the B.Eng. degree in software engineering from Northeastern University, Shenyang, China, in 2019, and the M.Sc. degree in artificial intelligence from Nanyang Technological University, Singapore, in 2021. He is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interest includes machine learning with combinatorial optimization problems.



Yaoxin Wu received the B.Eng degree in traffic engineering from Wuyi University, Jiangmen, China, in 2015, the M.Eng degree in control engineering from Guangdong University of Technology, Guangzhou, China, in 2018, and the Ph.D. degree in computer science from Nanyang Technological University, Singapore, in 2023. He was a Research Associate with the Singtel Cognitive and Artificial Intelligence Lab for Enterprises (SCALE@NTU). He joins the Department of Information Systems, Faculty of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, as an Assistant Professor. His research interests include combinatorial optimization, integer programming and deep learning.



Zhiguang Cao received the Ph.D. degree from Interdisciplinary Graduate School, Nanyang Technological University. He received the B.Eng. degree in Automation from Guangdong University of Technology, Guangzhou, China, and the M.Sc. in Signal Processing from Nanyang Technological University, Singapore, respectively. He was a Research Fellow with the Energy Research Institute @ NTU (ERI@N), a Research Assistant Professor with the Department of Industrial Systems Engineering and Management, National University of Singapore, and a Scientist with the Agency for Science Technology and Research (A*STAR), Singapore. He joins the School of Computing and Information Systems, Singapore Management University, as an Assistant Professor. His research interests focus on learning to optimize (L2Opt).



Wen Song received the B.S. degree in automation and the M.S. degree in control science and engineering from Shandong University, Jinan, China, in 2011 and 2014, respectively, and the Ph.D. degree in computer science from Nanyang Technological University, Singapore, in 2018. He was a Research Fellow with the Singtel Cognitive and Artificial Intelligence Lab for Enterprises (SCALE@NTU). He is currently an Associate Research Fellow with the Institute of Marine Science and Technology, Shandong University. His current research interests include artificial intelligence, planning and scheduling, multi-agent systems, and operations research.



Jie Zhang received the Ph.D. degree from the Cheriton School of Computer Science, University of Waterloo, Canada, in 2009. He is currently a Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. He is also a Professor at the Singapore Institute of Manufacturing Technology. During his Ph.D. study, he held the prestigious NSERC Alexander Graham Bell Canada Graduate Scholarship rewarded for top Ph.D. students across Canada. He was also a recipient of the Alumni Gold Medal at the 2009 Convocation Ceremony. The Gold Medal is awarded once a year to honour the top Ph.D. graduate from the University of Waterloo. His papers have been published by top journals and conferences and received several best paper awards.



Zhenghua Chen (Member, IEEE) received the B.Eng. degree in mechatronics engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2011, and the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University (NTU), Singapore, in 2017. He is working with NTU as a Research Fellow. He is currently a Scientist with Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore. His research interests include sensory data analytics, machine learning, deep learning, and transfer learning and related applications.