Kong, Fanhui, Bin Jiang, Jian Wang, Huihui Wang, and Houbing Song. "Collaborative Delivery Optimization With Multiple Drones via Constrained Hybrid Pointer Network." IEEE Internet of Things Journal, 2023, 1–1. https://doi.org/10.1109/JIOT.2023.3318524.

https://doi.org/10.1109/JIOT.2023.3318524

Collaborative Delivery Optimization With Multiple Drones via Constrained Hybrid Pointer Network

Fanhui Kong, Bin Jiang, *Member, IEEE*, Jian Wang, *Member, IEEE*, Huihui Wang, *Senior Member, IEEE* and Houbing Song, *Fellow, IEEE*

***Abstract*—Drone participation in truck delivery is a potential booster for the last-mile logistics system, which has been an emerging hot research field. Among that, how to arrange a fleet of drones from the truck and optimize the vehicle routing problem with drones (VRPD) is a key issue. However, most existing studies fail to derive the feasible solutions due to unordered customer distributions and multi-variant drone feature constraints. In this paper, we propose a novel self-driven reinforcement learning structure, named constraint-based hybrid pointer network(CH-Ptr-Net) model, which is a hybrid pointer network approach composed of graph neural network(GNN) embedding and attention decoder. We go into developing the simpler embedding version for multiple drones-assisted truck delivery. The CH-Ptr-Net model tends to generate a set of optimal delivery sequence, after constructing the mixed integer linear program(MILP) formulation. Extensive numerical testing indicates that the proposed method performs better than recent exact and heuristic approaches for collaborative delivery routing optimization with the truck carrying multiple drones.**

***Index Terms*—Collaborative routing optimization, multiple drones delivery, deep reinforcement learning, constraint-based hybrid pointer network, graph neural network embedding.**

I. Introduction

UNMANNED aerial vehicle(UAV) has attracted great attention as assistant for truck delivery, which is known as a potential booster for improving the logistics system efficiency. Amazon was the earliest to raise the drone application in the field of logistics delivery [1]. With the development of drone technology, deploying a fleet of drones from the truck has gradually been promoted to more wider range of commercial activities, which enables the delivery routing sparingly [2]. Taking a certain area in Tampa, FL, USA as an example, Fig. 1 shows a typical overview comparison for the vehicle routing problem (VRP) with different number of drones. Compared with traditional single truck delivery, the advantages of the vehicle routing problem with drones (VRPD) are obvious. First of all, compared with ground transportation, the drone achieves transportation conservation and economical infrastructure construction by utilizing idle low altitude resources. Afterwards, the freight operating cost is relatively low, and it helps the intensive development for all factors in the operation management. In addiction, collaborative delivery type of truck and drone brings vehicle own advantages into full play, improving delivery efficiency and logistics service capabilities. In the end, drone delivery is more flexible in emergency situations, such as earthquakes, forest fires and flood disaster, getting rid of the natural environment limitations. However, deploying multiple drones from the truck delivery faces more complex network system representation learning pressure. Therefore, collaborative delivery routing optimization analysis is worth further studying.

Significant effort has gone into analyzing and optimizing the VRPD. Many researchers have shown deep interest in using mathematical exact and heuristic algorithms to optimize delivery path. The first research on truck carrying drone delivery was presented by Murray and Chu [3], and authors proposed the heuristic approaches to solve the flying sidekick traveling salesman problem. In recent published journal papers, Murray [4] continued to delve into the vehicle routing visualization problem for multiple drone delivery, specifically, Gantt charts, dynamic 3D videos and static maps were used for the visual drone scheduling. Buck *et al.* [5], [6] put forward the drone delivery mode for asset-intensive organizations, and heuristic solution frameworks for solving large-sized instances of drone delivery lost the effectiveness. Lee *et al.* [7] researched the multi-drone delivery routing framework, and the developed exact routing algorithms achieved a maximum of six times higher performance for faster delivery runtime. It can be seen that how to arrange a fleet of drones from the truck and optimize the total delivery routing is still an issue worth studying.

Deep reinforcement learning has been a compelling choice to solve complex objective optimization tasks. The essence of combinatoric optimization is a sequential decision problem, and pointer network(Ptr-Net) is an effective artificial intelligence method for solving sequential decision problems. It is a new network architecture generated based on sequence to sequence networks, which solves the mapping problem from one sequence to another [8]. The content of output sequence is completely consistent with the input sequence, and the sequence order has changed, which can effectively perform intelligent sorting on the output sequence. Therefore, Ptr-Net establishes the relationship between neural network and combinatorial objective optimization, and it optimizes the element order of variable length

This work was supported in part by Humanities and Social Sciences Research Project of the Ministry of Education of China under Grant 21YJC630051, National Natural Science Foundation of China under Grant 62102264, Youth Foundation of Shandong Natural Science Foundation under Grant ZR2023QG034, Youth Innovation University Team Project in Shandong under Grant 2022KJ062, Independent Innovation Fund of China University of Petroleum (East China) under Grant 22CX06056A and Talent Project of Qingdao University of Technology under Grant 901020220046. *(Corresponding author: Bin Jiang)*

F. Kong is with the School of Management Engineering, Qingdao University of Technology, Qingdao, 266520, China (email: kongfh@qut.edu.cn).

B. Jiang is with the College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao, 266580, China (email: jiangbin@upc.edu.cn).

J. Wang is with Department of Computer Science, College of Engineering and Natural Sciences, The University of Tennessee at Martin, Martin, TN 38238 USA (email:jwang186@utm.edu).

H. Wang is with Cybersecurity Program, St. Bonaventure University, St. Bonaventure, NY 14778 USA (e-mail: h.wang@ieee.org).

H. Song is with the Department of Information Systems, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (email:h.song@ieee.org).

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3318524

sequences or sets.

However, most existing studies on truck-drone delivery problems are defective by the randomness of network data. These problems can be summarized as follows: 1) Most routing optimization algorithms are only suitable for simple projects, but unsatisfactory for complex cooperative scheduling problems; 2) Most existing works focus on a single objective optimization problem, but truck-drone delivery may face multiple objective functions that may be inconsistent and conflicting with each other; 3) It is difficult to solve the dynamic VRPD because the various customer needs are constantly addressed with the transportation dynamically; 4) As vehicle routing optimization is NP-hard, few large instances are tested. Multiple depots and customer nodes take more time to search information, and model structures are uncontrollable. In order to overcome these shortcomings, some efforts must be made.

In this paper, the proposed deep reinforcement learning coupled with constraint-based hybrid pointer network is adopted to optimize a collaborative truck-drone routing problem. We take the delivery operation characteristics, especially energy consumption, vehicle routing availability, drone endurance, customer time windows and service sequence into account, and then decision variables and objective functions are set. Based on the analysis of single objective optimization, solution approach is extended to solve multi-objective optimization problem (MOP). Considering the dynamic node update capability, data representation learning depends on GNN embedding and attention decoder, which generate a string of routing. Contributions of this work can be concluded as follows.

1) This paper extends the VRP to multi-trip VRPD (MT-VRPD) that the single truck operates in coordination with a fleet of drones. The MT-VRPD is a challenging variant, which belongs to the class of NP-hard problem. In this paper, we fully consider the delivery characteristic factors of deploying drones from the truck, so as to build the mixed integer linear program (MILP) formulation. Then we develop the truck-drone collaborative routing problem from an objective optimization to multi-objective optimization problem for the minimum delivery duration and delivery cost.

2) We propose a novel self-driven reinforcement learning procedure, named CH-Ptr-Net model, which is a hybrid pointer network solution approach based on a sequence of delivery constraints. Solution approach can not only optimize the sequence combination order for an objective function, but also extensively provide multiple arrangement points, by constantly varying the weights for the MOP. Unlike mathematical exact and heuristic methods, the proposed reinforcement learning framework derives the rewards with updating node information dynamically.

3) GNN embedding is added to replace recurrent neural network (RNN) encoder, which simplifies the version of pointer network. RNN encoder is only helpful for input tasks related to order. With all inputs of the current step, GNN uses low dimensional vectors to represent graph nodes and topological structure, which improves the generalization ability and extract superiority.

This paper is organized as follows. Sec. II presents the related research work. Then the MILP formulation is built in Sec. III. In Sec. IV, the proposed constraint-based hybrid pointer network is presented. After that, the experimental results and analysis are given in Sec. V. Sec. VI concludes this paper.

II. Related Work

In order to further analyze the related studies on delivery routing optimization with truck in coordination with drone, we will focus on two necessary content in this section. Single drone delivery problem and joint delivery with multiple drones-assisted truck will be summarized. Especially, we will investigate the factors and mechanisms of drone energy consumption. Application of deep reinforcement learning in VRPD will be discussed.

A. Truck-Drone Delivery Routing Optimization

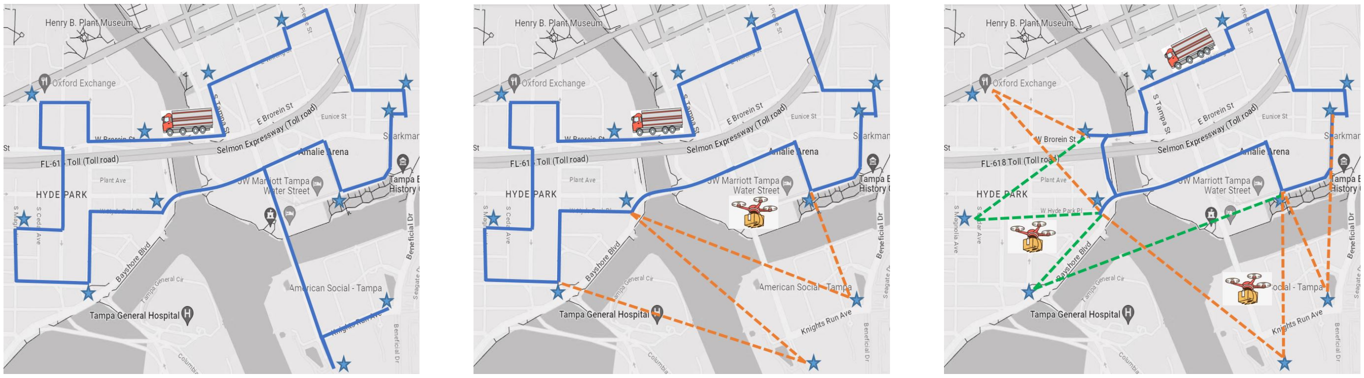
The drone is a new transportation tool for the logistics delivery process. Unlike traditional truck or aircraft delivery, drones face more complex delivery system conditions [9]. Preliminary researches have been studied on the classification based on the drone participation degree. Truck-drone delivery routing optimization problem contains the truck equipped with single drone delivery problem and joint delivery with multiple drones-assisted truck. The number of drones has a direct impact on constructing the decision variables and constraints.

1) Truck equipped with single drone delivery problem: Over the past years, several researchers have shifted their attention to single drone delivery problem. Archetti *et al.* [10] proposed the multimodal transportation plan for the VRPD. Specially, the operations research method became a new model for correctly distribution of complex unmanned aerial vehicle transportation system. Ensafian *et al.* [11], [12] researched on the drone reliable routing optimization, applied in the stochastic time-dependent public transportation network, which was an innovative exploration of single drone logistics delivery. An intelligent optimization algorithm based on one unmanned aerial vehicle delivery problem was proposed by Chen *et al.* [13]. They incorporated the Internet of things perspective into the network model during the last-mile drone delivery.

For single drone delivery constraints considered, Aurambout *et al.* [14] regarded the VRPD as the complex system optimization problems. Authors developed the single drone delivery problem for local shop, and the hop mode effectively overcame the limitation of drone energy consumption on flight distance. Bastone *et al.* [15], [16] focused on the impact of human resources on drone delivery performance. The level of drone energy consumption and carrying capacity affected the drones delivery capacity non linearly. For the spatial-temporal interaction constraints, Savkin *et al.* [17] decomposed the VRPD considering the large delivery scale and low operating cost factors. Authors constructed the complex integer programming model for single drone path optimization.

2) Joint delivery with multiple drones-assisted truck: In the case of joint delivery problem with truck and multi-drones, vehicle collaborative routing faced more complex optimization systems. Chase *et al.* [18], [19] proposed the collaborative truck and drone delivery model for emergency situation. Specially, non-contact package delivery during the epidemic was constructed, and the joint mode of drone and truck was an emergency exploration. For regular logistics delivery, Arafat *et al.* [20] researched the combined routing and charging strategy for multiple drones and truck, which considered both the differences and linkage between drone and truck transportation vehicles. Raissa *et al.* [21], [22] took use of parallel data and algorithm to design the parallel drone scheduling problem. Specially, authors generated the dynamic update model with multiple drones and vehicles at

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3318524



(a) Without drone (b) One drone (c) Two drones

Fig. 1. A typical overview comparison for the vehicle routing in Tampa region. (a)Only one truck without drone for delivery. (b) and (c) exist one truck with one and two drones respectively.

the same time. Through joint delivery, the operational efficiency of the delivery center was improved and service time was saved. Meng *et al.* [23] put forward the environmental and economic impacts as the important constraints, and the carbon market price became a performance indicator for measuring joint delivery.

In addition, hybrid multi-objective optimization problem with drone-assisted truck was formally defined by Luo *et al.* [24]. Authors proposed the pareto local search to shorten the delivery distance between trucks and drones. Considering flexible time windows factors, Elsokkary *et al.* [25], [26] took advantage of a fleet of heterogeneous drones for joint delivery, and they proposed the genetic algorithm approach for dynamic route planning. Yang*et al.* [27] addressed the planning robust truck-drone delivery routes prediction based on uncertain road traffic conditions. Das *et al.* [28] researched the synchronized delivery problem with multiple drones and truck. Trucks was mainly in charge of drones, without the delivery auxiliary function. Xue *et al.* [29], [30] put forward the new method with two-stage heuristic solution for joint delivery with multiple drones and trucks. Targeting the multi-depot collaborative pickup, it brought drone and truck transportation advantages into full play to reduce carbon emission costs.

Energy consumption is an essential constraint for drone delivery to provide services and reduce operational costs. Accurate estimation of drone energy consumption ensures the feasible and efficient delivery system. Various factors may affect the drone energy consumption [31], [32]. In published studies, drone design, environment, drone dynamics and drone operations are the four main categories. Drone design factor includes the drone itself and battery weight, the number of rotors, battery energy capacity, power transfer efficiency, lift-to-drag ratio and maximum speed and payload [33]–[35]. Environmental factor includes the air density, wind conditions, weather and gravity ambient temperature [36]. Drone dynamic factor includes the drone airspeed, drone motion, flight altitude, flight angle and angle of attack [37]. Drone operation factor includes the size of payload and payload weight, empty return, delivery mode and service area [38], [39]. Some of these factors are certain, such as battery energy capacity, size of payload and payload weight, etc., while others may be interdependent and dynamic during the drone delivery process.

Many researchers have established the energy consumption model for truck-drone delivery problem. Dorling *et al.* [40] used the linear approximation for the drone status including in hover, flight, takeoff and landing during logistics distribution. Figliozzi [41] calculated the power transfer efficiency involve energy loss from charging the battery and the power transmission efficiency. Stolaroff *et al.* [42] proposed the component energy consumption model by the forces of the weight, the parasite and induced drags. Total drone weight, parasite drag force and area perpendicular were considered fully. Therefore, how to determine drone energy consumption is an important component of the VRPD.

B. Application of Deep Reinforcement Learning in the VRPD

Optimization approaches for the VRPD can be divided into two types: traditional mathematical methods and intelligent algorithms. Among them, the exact algorithms and heuristic algorithms are commonly used mathematical methods, which have made some significant research progress. Yin *et al.* [43] conducted the branch-and-price-and-cut algorithm to extract the truck-based drone delivery routing problem. On the basis of considering the service time windows, authors focused on researching the improvement space of exact algorithms for path optimization. Hafza *et al.* [44] put forward another exact algorithm approach for solving MT-VRPD, which considered cost interactions among environmental multiple agents. For the heuristic algorithms, how to obtain the optimal solution has been a key issue for delivery routing optimization. Najy *et al.* [45], [46] explored the routing nonresistance sensitivity of heuristic algorithm structures. Optimization algorithm in different application instances was unstable, therefore, unreliable calculation results received challenge.

Intelligent algorithms, especially deep reinforcement learning algorithms, have made unprecedented progress in VRPD [47], [48]. As a new attempt of cross research, deep reinforcement learning has shifted the research field from image recognition and natural language processing to combinatoric optimization, helping to better solve combinatoric optimization problems. Choi *et al.* [49], [50] researched on reinforcement learning to build an Internet of Vehicles systems. Fu *et al.* [51], [52] also applied deep reinforcement learning algorithm to the Internet of Vehicles. Specially, authors optimized the vehicle scheduling problem in delay aware content delivery.

Essence of combinatoric optimization problem is to make sequential decisions, as generative neural network, pointer network suitable for solving combinatoric optimization problems.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3318524

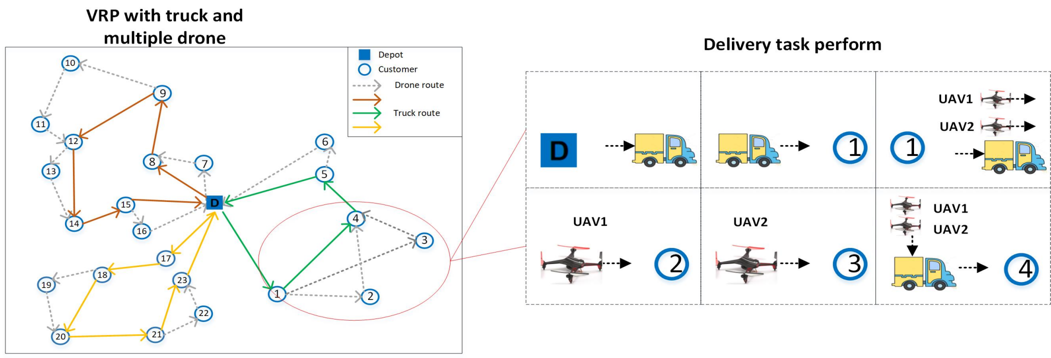


Fig. 2. The truck delivery schedule in coordination with a fleet of drones.

Regarding to the variants, Li *et al.* [53], [54] constructed the heterogeneous attentions to optimize vehicle delivery distance, and the attention mechanism enhanced the sequence convergence. So pointer network has emerged promising value for complex system sequence decision making.

III. MILP Formulation for the Collaborative Delivery Optimization

This section provides the MILP formulation for the collaborative delivery problem. While MILP contains decision variables and multiple constraints, especially energy consumption, optimal solution is used to find the minimum delivery duration and cost. We extend the truck-drone collaborative routing problem from a single objective optimization to multi-objective optimization problem. Customer locations are randomly distributed, and the parcels to be delivered are also unique. Generally, it is impossible to complete all customer services through one delivery route. The logistics network system consists of depot, customers, truck routes and drone routes. A set of customers are randomly distributed and has a service time window, which needs to be delivered by the truck or drone. To illustrate the cooperative delivery potential, Fig. 2 shows an example of how to arrange the feasible routes.

A. Problem Description and Assumptions

The MT-VRPD is formally defined as a single truck equipped with a fleet of drones for completing delivery tasks with a multi-trip. The truck is in charge of package delivery and serving multiple drones. The truck may return to the depot several times to load parcels. Besides traditional delivery activities, the truck also need to carry enough parcels to provide multiple drones. The truck trip is set as multiple visits until all customer nodes are serviced.

Multiple drones are launched from the single truck to deliver the customer parcels, and then the drones return to the truck for a new launch location or load a new parcel. Every drone has the limited battery capacity and payload weight, and it can only serve bits of customers. Multiple drone trips are defined as the flight solutions, containing the launch nodes, a sequence of customers serviced by drones, and the retrieval nodes.

Logistics network exists one depot and many customer locations distributed randomly, and parcel demand of each customer node is stochastic and can only be served by either the truck or a drone exactly once. Available service time of each customer is also specifically constricted. Objective function of MT-VRPD is to arrange a set of truck-drone routing to serve the customers to minimize total delivery time and cost.

In order to construct feasible MILP formulation, some specific assumptions should be set as follows:

1) Only one depot covers an area. Depot provides parcels and sufficient energy for the truck and multiple drones. The truck can return to the depot to load parcels and supply energy.

2) The single truck starts from and returns to the depot many times until the end of task performing. In addition to routine parcel delivery activities, the truck also need to launch and retrieve multiple drones. The truck has enough endurance to complete all the delivery activity.

3) The drone has limited battery capacity, so it can only maintain a short distance flight. Moreover, the drone payload is also relatively small, and it carries parcels with small weight. Despite all this, the single load capacity for every drone is greater than the parcel weight delivered by any customer. Drone can visit one or several nodes per sortie. If the drone energy is not enough to serve the next customer, it needs to fly back to the truck for recovery.

4) Customer locations obey the random distribution in a region, and their delivery demand quantities and available service time windows are different. Through collaborative delivery activities, every customer parcel can be serviced by the truck or drone.

5) If the truck and a fleet of drones return to the depot, it indicates that no customer has not been visited in the delivery network. When all the delivery tasks are completed, each customer has be served by the truck or the drone.

B. Decision Variables and Constraints

Before the MILP formulation is built, some mathematical parameter notations are shown in Table I. $C = \{1, 2, ..., c\}$ and $D = \{d_1, d_2, ..., d_n\}$ respectively represent the set of the customers and drones, and together with the depot constitute all network nodes in delivery system. All node settings can be represented as $N = \{0, 1, ..., c+1\}$, accordingly, setting of the departed nodes and arrived nodes are $\overrightarrow{N}$ and $\overleftarrow{N}$. The launch nodes, delivery nodes and rejoin nodes make up the delivery route $\langle i, j, h \rangle$. $x_{ijh} = 1$ represents that the drone is launched from node $i \in \overrightarrow{N}$, delivered to node $j \in C$, retrieved to node $h \in \overleftarrow{N}$. Let y_{ij} is a binary decision variable, and it represents that the truck travels from node $i \in \overrightarrow{N}$ to node $j \in C$. The travel time and service time of the truck and drones are defined, and the launch time and recovery time are ignored. Let e_{ijh} represent the endurance time of the drone.

To clarify the service sequence of collaborative delivery, the formal definitions are shown in Table II. Let $\mu^l_{d,i}$ represent whether the drone is launched from node i before another drone is launched from i. $\mu^l_{0,d,i}$ and $\mu^l_{d,0,i}$ represent whether the truck has completed service for node i before the drone is launched. Thus, let $\mu^r_{d,j}$ represent whether the drone is retrieved before another drone is retrieved from i. $\mu^r_{0,d,j}$ and $\mu^r_{d,0,j}$ represent whether the truck has completed service before the drone is retrieved from j. $\mu^1_{d,i}$ and $\mu^2_{d,i}$ indicate the launch and landing sequence between multiple drones. Above mathematical parameter notations can avoid the repeated delivery for customers.

1) Drone energy consumption: For the drone hovering state, Dorling *et al.* [40] put forward an energy consumption model for the multiple rotor drone. When the drone hovers, the airspeed is zero and the thrust balances the weight force. The thrust Γ can be expressed as Eq.1.

$$\Gamma = g\sum_{k=1}\vartheta_k \tag{1}$$

where ϑ_k is the weight to be lifted and g represents the gravity. According to the helicopter theory [55], energy consumption P of the drone to hover has the following expression.

$$P = \frac{\Gamma^{3/2}}{\sqrt{2n\rho\varrho}} = \frac{\left(g\sum_{k=1}\vartheta_k\right)^{3/2}}{\sqrt{2n\rho\varrho}} \tag{2}$$

where ρ represents the air density, n is the number of drone rotors and ϱ is the disc area of the spinning blade for each rotor. Based on hover model, we also need to consider additional energy consumption of flight, including the weight, the parasite and wind induced drags.

$$\Gamma = g\sum_{k=1}\vartheta_k + \frac{1}{2}\rho\sum_{k=1}\varepsilon_k\varsigma_k\eta^2_\alpha \tag{3}$$

Total drone weight brings out the first term, and the second term reflects the parasite and wind induced drags force. The induced drags force is subject to the drag coefficient ε_k. ς_k is vertical projected area for the drone components. When the drone is in hover, the air speed is zero. Therefore, Eq.3 is consistent with Eq.1, namely, $\Gamma = g\sum_{k=1}\vartheta_k$.

For the forward flight and wind effects, energy consumption formula is as Eq.4.

$$P = \frac{\Gamma\left(\eta_\alpha \sin\alpha + \eta_i\right)}{\ell} \tag{4}$$

where α represents the angle of the drone rotor with the airspeed, ℓ represents the power transfer efficiency and η_i is the induced speed that needs to be solved. The angle of attack α and the induced speed η_i can be expressed as Eq.5 and Eq.6.

$$\alpha = \tan^{-1}\left[\frac{1/2\rho\left(\sum_{k=1}\varepsilon_k\varsigma_k\right)\eta^2_\alpha}{g\sum_{k=1}\vartheta_k}\right] \tag{5}$$

$$\eta_i = \frac{g\sum_{k=1}\vartheta_k}{2n\rho\varrho\sqrt{\left(\eta_\alpha\cos\alpha\right)^2 + \left(\eta_\alpha\sin\alpha + \eta_\alpha\right)^2}} \tag{6}$$

So energy consumption P can be further sent, depending on hover and forward flight. By setting some parameter values, including the power transfer efficiency, drag coefficient, air density, weight to be lifted, and so on, the corresponding energy consumption can be determined and evaluated.

TABLE I
MATHEMATICAL PARAMETER NOTATIONS

Variable	Definition
D	Set of the drones, $D = \{d_1, d_2, ..., d_n\}$
C	Set of all customers in delivery system, $C = \{1, 2, ..., c\}$
C_d	Set of the customers serviced by drones, $C_d \subset C$
N	Set of all the network nodes is $N = \{0, 1, ..., c+1\}$, where 0 and $c+1$ represent the depot
$\overrightarrow{N}$	Set of the departed nodes of truck or drone, $\overrightarrow{N} = \{0, 1, ..., c\}$
$\overleftarrow{N}$	Set of the nodes that truck or drone may arrive during the delivery activities, $\overleftarrow{N} = \{1, 2, ..., c+1\}$
i	The launch node, the start of delivery process, $i \in \overrightarrow{N}$
j	The delivery node, accessible customer nodes for drones, $j \in \overleftarrow{N}$ and $j \neq i$
h	The rejoin node, located in customer nodes or depot, $h \in \overleftarrow{N}$, $h \neq i$ and $h \neq j$
x_{ijh}	$x_{ijh} \in \{0,1\}$, $x_{ijh} = 1$ represents that the drone is launched from node $i \in \overrightarrow{N}$, delivered to node $j \in C$, retrieved to node $h \in \overleftarrow{N}$
y_{ij}	Binary decision variable, $y_{ij} \in \{0,1\}$, $y_{ij} = 1$ represents that the truck travels from node $i \in \overrightarrow{N}$ to node $j \in \overleftarrow{N}$
δ_{ij}	Distance from node i to j traveled by the truck, $i \in \overrightarrow{N}$, $j \in \overleftarrow{N}$
δ'_{ij}	Distance from node i to j traveled by the drone, $i \in \overrightarrow{N}$, $j \in \overleftarrow{N}$
c_t	Transportation cost per unit of distance of the truck
c_d	Transportation cost per unit of distance of the drone
c_e	The cost in financial unit of a kJ of energy
p_i	Energy consumption from the drone after leaving node i, $i \in \overrightarrow{N}$
t_i	The time for the truck to arrive at node i, $t_i \geq 0$
t^d_i	The time for the drone to arrive at node i, $t^d_i \geq 0$
$\bar{t_i}$	The completion time for the truck at node i, $i \in \overrightarrow{N}$
$\bar{t^d_i}$	The completion time for the drone at node i, $i \in \overrightarrow{N}$
γ_{ij}	The travel time for the truck from node $i \in \overrightarrow{N}$ to node $j \in \overleftarrow{N}$
γ^d_{ij}	The travel time for the drone from node $i \in \overrightarrow{N}$ to node $j \in \overleftarrow{N}$
s_j	The service time for the truck at node $j \in \overleftarrow{N}$
s^d_j	The service time for the drone at node $j \in \overleftarrow{N}$
e_{ijh}	Endurance time of the drone during the route $\langle i, j, h\rangle$

2) Delivery routing availability: To ensure every delivery route is valid through, and the truck and multiple drones realize the collaborative delivery activities. The network graph constraints are expressed as Eq.7-9.

$$\sum_{i\in\overrightarrow{N}}\sum_{h\in\overleftarrow{N}} x_{ijh} + \sum_{i\in\overrightarrow{N}} y_{ij} = 1, j \in \{C, j \neq i\} \tag{7}$$

In Eq.7, the launch node, delivery nodes and rejoin nodes are the separate areas. Eq.7 guarantees that every customer in the delivery network is served exactly once, which prevents the occurrence of duplicate delivery.

$$\sum_{i\neq h}\sum_{j\in C} x_{ijh} \leq 1, i \in \overrightarrow{N}; \sum_{i\neq j}\sum_{h\in\overleftarrow{N}} x_{ijh} \leq 1, j \in C \tag{8}$$

For the smooth movement of drones, the above two inequalities state that every drone is launched and retrieved exactly once at a specific node. They effectively avoid the repeated launch and recovery at a customer node or the depot. Exact flight can economize energy consumption.

$$\sum_{i\in\overleftarrow{N}}\sum_{j\in\overleftarrow{N}} y_{ij} \leq 1, i \neq j; x_{dojh} \leq \sum_{i\in\overrightarrow{N}} y_{ih}, i \neq h \tag{9}$$

Eq.9 indicates that the truck may return the depot for loading the parcels until all the customer nodes are visited. This vehicle routing problem is mainly caused by the multiple return of the truck. Eq.9 requires that the truck must be assigned to h, when a drone lands at node h. So every drone can be retrieved by the truck at the appropriate node.

3) Drone endurance: Due to the small battery capacity γ, the flight time of each drone is also limited. In the process of delivery, launch, parcel delivery and recovery will deplete the drone endurance time. If the drone travels from i to j to h, the following constraints shall be met:

$$\gamma_{ij}^{d} + s_j + \gamma_{jh}^{d} \leq e_{ijh}, i \in \overrightarrow{N}, j, h \in \overleftarrow{N} \tag{10}$$

where γ_{ij}^{d} represents the travel time for the drone from node $i \in \overrightarrow{N}$ to node $j \in \overleftarrow{N}$. s_j is the service time at node $j \in \overleftarrow{N}$, and γ_{jh}^{d} is the travel time for the drone from node $j \in \overleftarrow{N}$ to node $h \in \overleftarrow{N}$. The above three decision variables are added together should be less than the endurance time of the drone. Inequality consisting of drone travel time and service time ensures the smooth delivery progress.

$$t_h^d - \bar{t_j^d} \leq e_{ijh} + M\left(1 - x_{ijh}\right) \tag{11}$$

In Eq.11, M represents a sufficiently large number, to get the upper bound of customer service time windows. Therefore, the right side of the inequality provides an upper bound on the maximum service time. The arrival time for the drone at node $h \in \overleftarrow{N}$ takes away the departure time at node j, and the result should be less than the drone endurance.

4) Timing constraints: Timing constraints of truck and drone are essential components. First of all, the truck timing constraints are related to arrival time, travel time, service time and completion time. Eq.12-14 are a reasonable specification of the truck delivery time, associated with the time windows and the drone locations.

$$\bar{t_i} + \gamma_{ij} - M\left(1 - y_{ij}\right) \leq t_j, i, j \in \overrightarrow{N} \tag{12}$$

Eq.12 prevents the truck visit the node j from the node i until the service is completed at node i. That is, the time when the truck arrives at node j is always later than that of node i.

$$t_h^d - M\left(1 - \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh}\right) \leq \bar{t_h}, h \in \overleftarrow{N} \tag{13}$$

$$\bar{t_h^d} - M\left(1 - \sum_{k \in C} \sum_{l \in \overleftarrow{N}} x_{hkl}\right) \leq \bar{t_h}, h \in \overrightarrow{N} \tag{14}$$

Eq.13 and Eq.14 state that if a drone is launched or retrieved at node h, the truck can not leave the node h. t_h^d and $\bar{t_h^d}$ are used to indicate the drone landing and takeoff time at node.

In the next place, the drone timing constraints also need to be considered, and it takes more factors into account than truck. A sequence of the drone timing constraints ensure the delivery activities running properly.

$$t_h^d - M\left(2 - \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh} - \sum_{k \in C} \sum_{l \in \overleftarrow{N}} x_{jkl}\right) \leq \bar{t_j^d}, j \in C \tag{15}$$

TABLE II
NOTATIONS FOR SERVICE SEQUENCE

Variable	Definition
$\mu_{d_1,d_2,i}^{l}$	$\mu_{d_1,d_2,i}^{l} \in \{0,1\}$, $\mu_{d_1,d_2,i}^{l} = 1$ represents that $d_1 \in D$ is launched from node i before $d_2 \in D$. Otherwise, $\mu_{d_1,d_2,i}^{l} = 0$
$\mu_{0,d,i}^{l}$	$\mu_{0,d,i}^{l} \in \{0,1\}$, $\mu_{0,d,i}^{l} = 1$ represents that the truck has completed service for node i before the drone is launched. Otherwise, $\mu_{0,d,i}^{l} = 0$
$\mu_{d,0,i}^{l}$	$\mu_{d,0,i}^{l} \in \{0,1\}$, $\mu_{d,0,i}^{l} = 1$ represents that the drone has been launched before the truck completes service for node i. Otherwise, $\mu_{d,0,i}^{l} = 0$
$\mu_{d_1,d_2,j}^{r}$	$\mu_{d_1,d_2,j}^{r} \in \{0,1\}$, $\mu_{d_1,d_2,j}^{r} = 1$ represents that $d_1 \in D$ is retrieved from node j before $d_2 \in D$. Otherwise, $\mu_{d_1,d_2,j}^{r} = 0$
$\mu_{0,d,j}^{r}$	$\mu_{0,d,j}^{r} \in \{0,1\}$, $\mu_{0,d,j}^{r} = 1$ represents that the truck has completed service for node j before the drone is retrieved. Otherwise, $\mu_{0,d,j}^{r} = 0$
$\mu_{d,0,j}^{r}$	$\mu_{d,0,j}^{r} \in \{0,1\}$, $\mu_{d,0,j}^{r} = 1$ represents that the drone has been retrieved before the truck completes service for node i. Otherwise, $\mu_{d,0,j}^{r} = 0$

To avoid multiple drones being launched and retrieved at the same node, Eq.15 makes a reasonable setting. If a drone is retrieved at node h from node j, it can not return to the node later. The above inequality ensures the service order relationships for customers, and it can prevent ineffective delivery of multiple drones.

$$t_j^d + s_j^d \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh} \geq \bar{t_j^d}, j \in C \tag{16}$$

$$t_j^d + s_j^d + M\left(1 - \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh}\right) \geq \bar{t_j^d}, j \in C \tag{17}$$

Eq.16 and Eq.17 require the completion time of the drone at node j. In order to meet the service time windows requirement for customers, the earliest and latest time for the drone to arrive at the delivery node need to be within the service interval. Meanwhile, Eq.17 makes the drone depart from node j immediately after completing the service, which guards against sticking for too long.

$$\bar{t_i^d} + \gamma_{ij}^{d} - M\left(1 - \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh}\right) \leq t_j^d, j \in C \tag{18}$$

Eq.18 indicates the arrival time at node j for the drone. If a drone travels from the node i to the node j, the upper and lower bounds of the time windows for node i can be considered. Eq.19 also imposes the same restrictions.

$$\bar{t_i^d} + \gamma_{ij}^{d} + M\left(1 - \sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} x_{ijh}\right) \geq t_j^d, j \in C \tag{19}$$

The above customer service time constraints ensure that the demand nodes in the logistics system are met within a reasonable time interval. The truck and a fleet of drones can assign the reasonable delivery sequence based on time windows.

5) Service sequence of the vehicles: According to the parameter notations in Table II, service sequence constraints can be set. If the drone is launched at node i, the truck should serve the node i. No matter which the truck or drone reaches node i first, the node must be serviced by the truck. Eq.20 is the mathematical expression in case of drone launch, and the drone launch meets the node sequence arrangement.

$$\mu_{0,d,i}^{l} + \mu_{d,0,i}^{l} = \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{ijh}, i \in \overrightarrow{N} \tag{20}$$

$$\mu_{d_1,d_2,i}^{l} + \mu_{d_2,d_1,i}^{l} \leq 1, d_2, d_1 \in D, i \in \overrightarrow{N} \tag{21}$$

Eq.21 prevents the multiple drones from being launched concurrently at node i. The truck launches the d_2 before the d_1, or the truck launches the d_1 before the d_2. In short, the launch sequence of the multiple drones is constrained.

$$\mu_{d_1,d_2,i}^{l} \leq \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{d_1 ijh}, d_2, d_1 \in D, i \in \overrightarrow{N} \tag{22}$$

$$\mu_{d_1,d_2,i}^{l} \leq \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{d_2 ijh}, d_2, d_1 \in D, i \in \overrightarrow{N} \tag{23}$$

Eq.22 and Eq.23 indicate that the launch numbers of d_1 or d_2 at node i is less than the total number of launches. They make regulations on the drone launch numbers and the orders of drone launch are constrained. Similarly, Eq.24-27 describe the service sequence when the drones are retrieved at network requirement nodes. When drones are recycled, the vehicle service sequence needs to be strictly restricted.

$$\mu_{0,d,h}^{r} + \mu_{d,0,h}^{r} = \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{ijh}, h \in \overleftarrow{N} \tag{24}$$

$$\mu_{d_1,d_2,h}^{r} + \mu_{d_2,d_1,k}^{r} \leq 1, d_2, d_1 \in D, h \in \overleftarrow{N} \tag{25}$$

Eq.24 states if the drone is retrieved at node h, the truck should serve the node h. Eq.25 guards the multiple drones against being retrieved at the same time. The truck retrieves the d_2 before the d_1, or the truck retrieves the d_1 before the d_2.

$$\mu_{d_1,d_2,h}^{r} \leq \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{d_1 ijh}, d_2, d_1 \in D, h \in \overleftarrow{N} \tag{26}$$

$$\mu_{d_1,d_2,h}^{r} \leq \sum_{j \in C} \sum_{h \in \overleftarrow{N}} x_{d_2 ijh}, d_2, d_1 \in D, h \in \overleftarrow{N} \tag{27}$$

Eq.26 and Eq.27 require the recovered sequence at node h. For example, the recovered times for d_1 at node h is less than the total number of the launch, delivery and recovery processing. This prevents drone excessive recovering.

C. Objective Function

The MILP objective is to find the minimum delivery time to complete all the delivery tasks. Node $c+1$ indicates the final destination visited by the truck carrying multiple drones. The arrival time t_{c+1} at node $c+1$ is the total operation duration, so the objective function can be expressed as Eq.28.

$$\min f = t_{c+1} \tag{28}$$

Besides getting the earliest arrival time at depot, total delivery cost is an auxiliary objective function. Total delivery cost is related to transportation cost and energy consumption cost, which can be formulated as

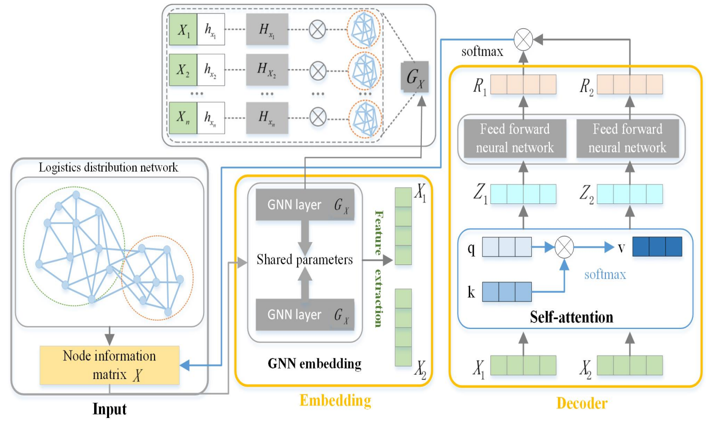


Fig. 3. The model framework of the constraint-based hybrid pointer network.

$$\begin{aligned} \min f' = \textstyle\sum_{i \neq j} \sum_{h \in \overleftarrow{N}} c_d \left(\delta'_{ij} + \delta'_{jh}\right) x_{ijh} \\ + \textstyle\sum_{i \in \overrightarrow{N}} \sum_{j \in \overleftarrow{N}} c_t \delta_{ij} y_{ij} + \sum_{i \neq j} \sum_{h \in \overleftarrow{N}} c_e p_i x_{ijh} \end{aligned} \tag{29}$$

Additional objective function f' minimizes the total delivery cost. It becomes a multi-objective MILP formulation, when minimizing delivery time and minimizing delivery cost are considered simultaneously.

Since two objectives are involved in the vehicle collaborative delivery problem, it cannot be solved directly. Firstly, we construct CH-Ptr-Net model to obtain a single objective optimization solution for delivery time f. Secondly, to obtain multi-objective solutions, we address the MOP as a weighted combination problem. Model adopts different weight coefficients β, satisfying $\sum_v \beta_v = 1$, as shown below:

$$\begin{gathered} \min : \beta_1 f + \beta_2 f' \\ s.t. \, (1) - (27) \end{gathered} \tag{30}$$

Ultimately, the multiple objectives can be integrated into a single objective formulation. Solving MOP is transformed into single objective optimization problem. We need to quantify the relative weights of total delivery time and cost to get the set of Pareto minimizers or Pareto front. According to the above built constraints Eq.1-27, optimization variables x_{ijh} and y_{ij} are continuously adjusted to change the truck-drone delivery plans. We consider the constraints of energy consumption, delivery routing availability, drone endurance, customer timing windows and service sequence. Optimization variables, i.e, the trip sequence of the truck and drones, determines the final time and cost.

IV. Constraint-Based Hybrid Pointer Network Model for Collaborative Delivery Optimization

A. Overall framework of the CH-Ptr-Net model

During the vehicle delivery, the customer needs will be continuously met with the transportation activities, making the model output dynamic and complex. The basic framework of pointer network is difficult to solve the model output. Fig. 3 puts a new encoder and decoder structure, called the CH-Ptr-Net model. The hybrid model consists of GNN embedding and attention decoder. We replace the RNN encoder with a simple network embedding. This is because the RNN encoder is only helpful for input order related tasks, but not necessary for input order unrelated tasks. So

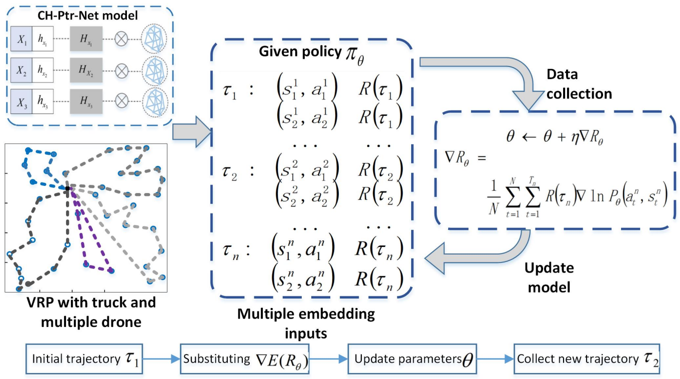


Fig. 4. Policy gradient training based on policy-based method.

the encoder part is removed and only the decoder part is retained. This can simplify the model decoding process.

Ptr-Net encoder is replaced by GNN embedding which is all inputs of the current step, so that the model output is independent of the input order. Each node of original Ptr-Net is only input once during the whole training. However, in the new CH-Ptr-Net model, each node is input once in each step, so it can handle dynamic delivery environment. The whole network mainly contains two parts, one is encoder, which consists of static node coordinates and dynamic demand states. The other part is the decoder, which points to the input node.

B. GNN Embedding and Attention Decoder

Data representation learning is essential matter for Ptr-Net to extract the characteristics of input data. At the same time, embedding learning can improve the generalization ability of the model and extract the effective data features. Embedding process can not only better express the internal relationship between data, but also simplify the model decoder. It uses low dimensional vectors to represent the nodes and topological structure for the graph network, which can be used as the input of the machine learning algorithm.

The GNN embedding is a kind of neural network directly acting on the graph structure. It can handle the initial data with irregular and unfixed structures. Each customer node is regarded as an individual object, and each edge is treated as a certain delivery connection between different nodes. The GNN embedding characterizes the relationship between nodes through the adjacency matrix. The adjacency matrix stores the source and target information. Then, adjacency matrix and original eigenvectors are used as model inputs. Through aggregation operation, the ambient nodes associated with node C_i are weighted on it, making a feature update dynamically. After feature updates for many times, the network delivery needs is extracted. Therefore, GNN implements dynamic embedding as encoder of the CH-Ptr-Net model.

The decoding process of Ptr-Net is to obtain the output element by calculating the largest weight input. Attention mechanism establishes a weight relationship for each position association between the output and the input. Self-attention and feed forward neural network help attention decoder output the node matrix. Data pass through the self-attention module to obtain a weighted feature vector Z. The algorithm steps are as follows:

Algorithm 1 the decoding process of the CH-Ptr-Net model

Parameters: q-query vector, k-key vector, v-value vector,
X-embedding vector, Z-weighted feature vector,
W^q, W^k and W^v-weight matrix of q, k and v.

Input: input node delivery information, then convert the node delivery into vectors q, k and v.

Output: percentage of each distribution node being selected;
service sequence of drone to each network node;
the feasible optimization paths;

Start: transform node delivery information into embedded vector X.

for each: 1) according to embedded vector X, getting the q, k and v vector;
2) calculate each vector score, $score = q \cdot k$;
3) normalize the gradient for convergence speed: $\frac{qk^T}{\sqrt{d_k}}$;
4) input score to soft-max activation function: $softmax\left(\frac{qk^T}{\sqrt{d_k}}\right)$;
5) calculate the score Z of the weighted input vector:
$$Z = softmax\left(\frac{qk^T}{\sqrt{d_k}}\right) \cdot v;$$
end for

repeat

until converge

Each delivery node information is composed of q, k and v vector. They are obtained by the embedding vector X_i multiplied by the weight matrix W^q, W^k and W^v respectively. Then, each node can normalize a score, and the soft-max activation function dot multiplies the value to obtain the weighted score v of each input vector. Above decoding process can be expressed as $Attention\,(Z) = softmax\left(\frac{qk^T}{\sqrt{d_k}}\right) v$. Feed forward neural network is a fully connected layer. After obtaining the weighted feature vector Z, and its activation function can be expressed as $Feed\,(R) = max\,(0, ZW_1 + b_1)\,W_2 + b_2$. Feed forward neural network helps to establish the relationship between the current delivery task and the encoded feature vector.

C. Training Method via Policy-Based Method

To train the Ptr-Net parameters and increase the convergence rate, the policy-based method is used. Policy-based method consists of actor, environment and reward function. The environment and reward functions are predetermined and cannot be changed. The only thing that can be adjusted is the actor policy, so that the actor can obtain the maximum reward. The policy is to give an external input, and then the neural network model will output the behavior that the actor should perform in the current moment.

For the collaborative delivery problem with the truck and multiple drones, model input is the logistics delivery network, which is composed of the depot and customer demand nodes. Model output is the route that the truck and drones can choose, and its neuron number is determined by the number of actions taken by the truck and drones. The agent, i.e., the single truck and multiple drones, travel over environment with the entire square grids area. Each grid cell is a state s_i and the vehicle can change its state by choosing one of four directional movement options, i.e., up, down, left and right that form the action set A. When the agent adopts the policy $\pi\,(s_t, a_t)$, it will receive the positive or negative rewards. Positive reward tends to minimize the vehicle delivery or service time. The total reward at each cell is the sum of its immediate reward plus all of the future rewards discounted using a discount factor less than 1.

Policy-based training first initialize the variables and grid, in which the state space S and action space A including four main direction movements are set. Method benefits from the greedy search by random action function using ε variable as a holder for the given threshold value. If the random variable epsilon is less than a threshold, then the random action function will randomly choose an action among all possible actions. Otherwise, the action that maximizes the total reward will be chosen. The policy gradient is the basis of policy-based method, and its function is to make the policy parameters θ move update in the rise direction. The expectation of cumulative reward is a value function, which can measure the quality of state. When the truck and drone go for the policy, the cumulative reward R_i follows mathematical distribution. The state value function v_π can be expressed as Eq.31.

$$v_\pi(s_i) = E_\pi\left[\sum_{k=0}^{\infty}\gamma^k R_{t+k+1}, S_t = s_i\right] \tag{31}$$

Policy parameters θ obeys the policy gradient rise. The steps of sample update are $\theta_{t+1} = \theta_t + \alpha\sum_\alpha \bigtriangledown_\theta \pi_\theta\left(s_t, a\right) R_\pi\left(s_t, a\right)$. The model goal is to maximize cumulative rewards. However, the drone is in a strange environment, it does not know the reward of the selected path at the beginning. Delivery actions are random, so the corresponding rewards are also random. Therefore, our goal is to estimate the cumulative reward, that is, to find its expectation. The expectation function can be expressed as $E\left(R_\theta\right) = \sum_\tau R\left(\tau\right) P_\theta\left(\tau\right)$, where τ represents a set of trajectories with drones and truck. The final derivation result is Eq.32.

$$\begin{aligned}&\bigtriangledown E\left(R_\theta\right) = E_{\tau\sim P_\theta}\left[R\left(\tau_n\right)\cdot \bigtriangledown lnP_\theta\left(\tau\right)\right]\\&= \tfrac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T_n} R\left(\tau_n\right)\bigtriangledown lnP_\theta\left(a_t^n, s_t^n\right)\end{aligned} \tag{32}$$

Fig. 4 shows the policy gradient training process. First, the actor interacts with the environment to obtain the initial trajectory $\tau_1 = \left\{s_1^1, a_1^1, s_2^1, a_2^1 ..., s_t^1, a_t^1\right\}$. The model brings the data into the gradient formula $\bigtriangledown E\left(R_\theta\right)$, and then updating the policy parameters θ. The new network continues to collect the next delivery trajectory, and then keeping on optimizing parameters θ. In the above steps, each trajectory is multiple embedding input to handle the dynamic delivery environment. Model processing node information is an important innovation of trajectory optimization.

V. Experimental Results and Analysis

To assess the performance of the proposed CH-Ptr-Net model for collaborative delivery optimization, it will be applied in the randomly generated instances. We will validate the drone energy consumption model, and the performance comparison between CH-Ptr-Net model and other approaches for a single objective and multi-objective optimization problem will be analyzed, respectively. Especially, we conduct an experiment to compare the proposed algorithm against an recent hybrid ILS-VND method in Gu *et al.*(2022). Moreover, it will discuss the impact of GNN embedding structure, compared with RNN encoder. Choosing the optimal service number of drones will be also an indispensable issue.

TABLE III
Parameter profiles for the MT-VRPD instances

Parameter	Description	Value
s_j	The service time for the truck at customer node (min)	2
s_j^d	The service time for the drones at customer node (min)	2
e_{ijh}	The endurance time of the drones during the route (min)	300
γ	The maximum battery capacity and payload weight (kg)	1.8×10^4
λ_t	The travel speed of the truck (m/s)	3
λ_d	The travel speed of the drone (m/s)	6
d	The vector space of GNN encoder-attention decoder	128
S	The state pace of the training network	$11 * 11$
ε	The greedy search through a random action (m/s)	0.9
α	The initial learning rate	10^{-3}

A. MT-VRPD Test Instances

As the collaborative delivery problem with the truck and multiple drones is a new issue, there are no existing benchmark problems. We derive a new piece of test instances from the randomly generated set. The instances are used to form scenarios, which include three types of customer distribution. In the uniform instances, customers are independently and uniformly distributed. The customer locations are randomly generated from the unit region within $[0, 10]\times[0, 10]$. Each customer has a certain service time windows, randomly generated from the unit $[1, 9] \times [1, 9]$. Delivery demand is also randomly generated from $\{1, 2, ...9\}$. A truck and a fleet of drones are employed, with each drone having the same limited battery capacity. The single truck starts from and returns to the given depot until the end of task.

Delivery distance and optimal gap are cited as two performance metrics, under a time limit of 30-min. Delivery distance indicates the total trip that the truck and drones complete all delivery tasks. Less delivery distance implies the less time cost spent. The best situation is to pursue the minimum delivery distance. Optimal gap is the percentage difference between the solutions for test instances. Solution quality is embodied by the gap between the best upper and lower bounds. Less value of the optimal gap means the higher training efficiency.

B. Scenario Description

To possess the overall behavior of the MT-VRPD, we average the calculations of a large amount of randomly generated instances of each scenario, which contains the small-scale, medium-scale, and large-scale delivery. Each type of scenario has 20 unique instances with uniform customer locations. Small-scale delivery scenarios have $[20, 50]$ delivery locations, while medium-scale and large-scale delivery scenarios have 100 and more locations. Each scenario has an area size of 0.5, 1 and 2 km^2. We generate 20 random instances for each type of scenario. In every instance, customer nodes are randomly distributed throughout the entire area. Each instance has a random demand of $\{1, 2, ...9\}$. The depot is located in the area of $[0, 0]$. The CH-Ptr-Net algorithm is run 100 times for per instance, and then the maximum, mean and minimum distance and optimal gap are calculated. We run the MILP implementation once per instance, and we provide the average experimental result for all 20 instances of a scenario.

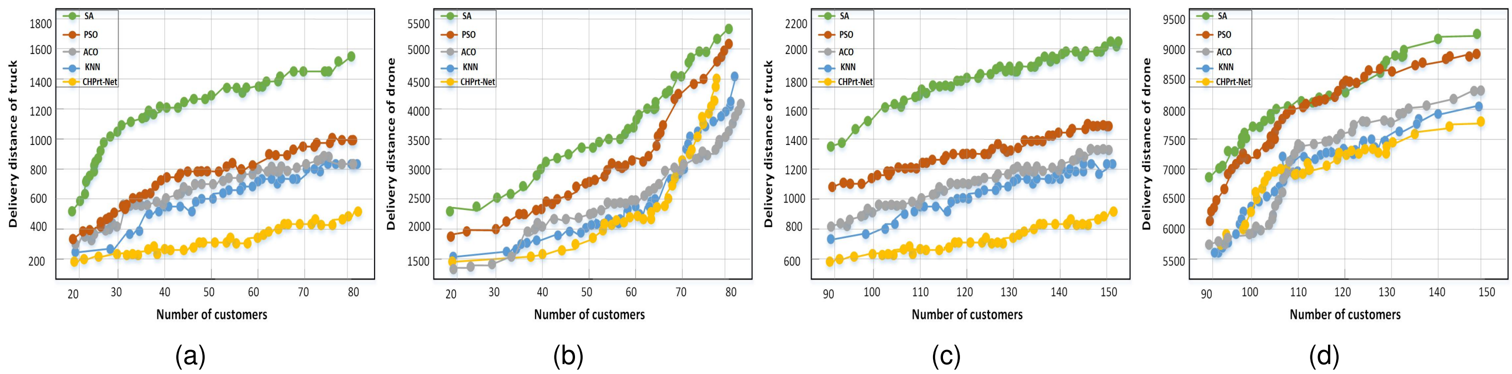


Fig. 5. The delivery distance under the different number of customers. Customer number interval of (a) and (b) is $[20, 80]$. Customer number interval of (c) and (d) is $[90, 150]$.

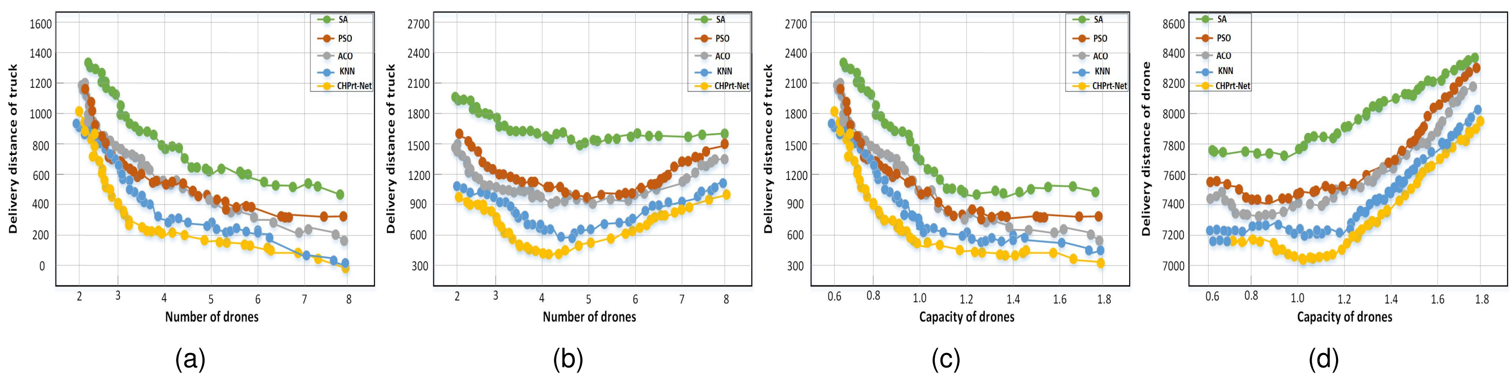


Fig. 6. The delivery distance under the different drone number and capacity. Customer number of (a) and (c) is 50. Customer number interval of (b) and (d) is 120.

C. Parameter Settings

According to each customer locations and demands, we set the maximum endurance time of drones as $e_{ijh} = 300$. Considering convenience, the configuration of each drone is uniform without distinction. The travel speed of truck is set to $\lambda_t = 3$, and the flight speed of drones is set twice as fast as the truck. Every drone has the maximum capacity including its battery and payload weight of 1.8×10^4. We assign the service time of drones $s = 2min$, while the truck has the same service time settings. Detailed parameter information is summarized in Table III.

For the CH-Ptr-Net algorithm parameters, we use one layer of GNN encoder-attention decoder with a size of 128. Every customer node is also embedded into a vector of size 128. The visiting nodes are mapped to a vector in a 128-dimensional vector space and used in the attention decoder layer. Attentive encoder consists of 3 stacks with $h = 16$ parallel heads and $d = 128$ hidden dimensions. For each head, we make use of $d_h = d/h = 8$. Training state space S is set as a $11 * 11$ grid network. We set the search step to 1, experimentally optimized destination grid to 1000, no-fly grid to -30. Random action for greedy search is set to $\varepsilon = 0.9/s$, and initial learning rate is set to $\alpha = 10^{-3}$.

The MT-VRPD is solved by the Python IDE-PyCharm 2022, and the programs are implemented in a Python version 3.10.9. All computational works are completed on a Windows 10 desktop PC server with Intel(R) Core(TM) i5-2.11 GHz processor.

D. Energy Consumption Measurements

We conduct key elements to validate our energy consumption model and evaluate different interrelated aspects of energy consumption per meter, speed, drone flight range and payload. It explores the specific situation of modeling batteries, drone operation, especially speed and payload mass, as well as energy adjustments such as wind power, avionics, and energy loss.

Parameter values used in drone energy consumption is listed in Table IV. Environmental parameters (e.g., air density, gravity, etc.) independent of drone design are shown. It includes a common value for the power transfer efficiency of the drone, and we also assume empty returns so the values of energy transfer rate are the average of the loaded and unloaded. We change either airspeed or payload mass to explore the impact on energy consumption rate. Thus, by using the same drone standards and flight conditions, we can get the changes in energy consumption rate or flight distance caused by different input structures and assumptions.

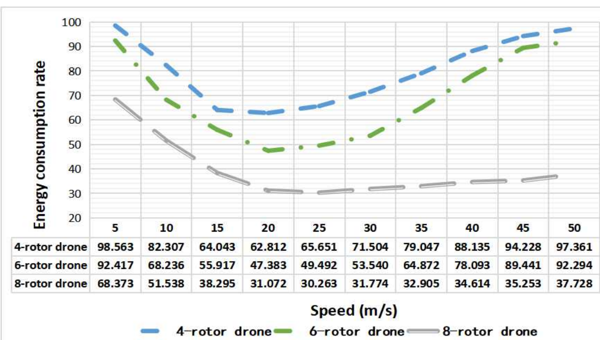
The impact of speed and payload on drone energy consumption is shown in Fig. 7, under the 4-rotor, 6-rotor and 8-rotor drone. Fig. 7 (a) and (b) display the tendency of energy consumption rate caused by speed and payload, respectively. The shapes of the curves for energy consumption rate are similar for the 4-rotor, 6-rotor and 8-rotor drone, though the curves are less steep for the 8-rotor drone. When the speed approaches $20m/s$ from $5m/s$, energy consumption rate continuously decreases and the 8-rotor drone remains lower than the other two types of drones.

For the impact of payload mass, energy consumption rate increases by up to 13.481% for 6-rotor drone as the payload rises from $76.203kg$ to $86.476kg$. On the whole, the payload range of the 4-rotor drone has the least impact on drone energy consumption rate.

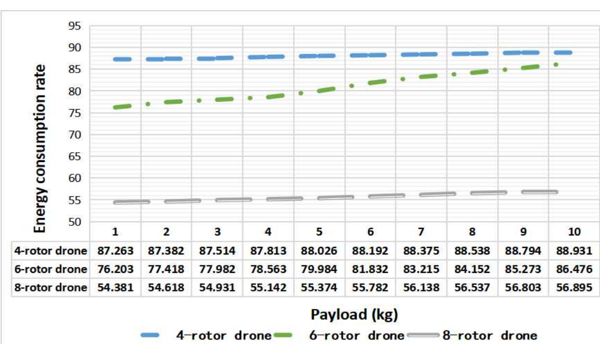
This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3318524

TABLE IV
PARAMETER PROFILES FOR DRONE ENERGY CONSUMPTION

Parameter	Description	Value
ϑ_1	The mass of drone body (kg)	4
ϑ_2	The mass of drone battery (kg)	2
ϑ_3	The mass of drone payload (kg)	8
g	Acceleration of gravity $\left(m/s^2\right)$	9.807
ρ	The air density $\left(kg/m^3\right)$	1.205
n	The number of drone rotors	4/6/8
ϱ	The disc area of the spinning blade $\left(m^2\right)$	0.035
ε_1	The drag coefficient of drone body	1.25
ε_2	The drag coefficient of drone battery	1
ε_3	The drag coefficient of drone payload	1.25
ς_1	The vertical projected area of drone body $\left(m^2\right)$	0.06
ς_2	The vertical projected area of drone battery $\left(m^2\right)$	0.005
ς_3	The vertical projected area of drone payload $\left(m^2\right)$	0.01
ℓ	Battery power transfer efficiency	0.8



(a)



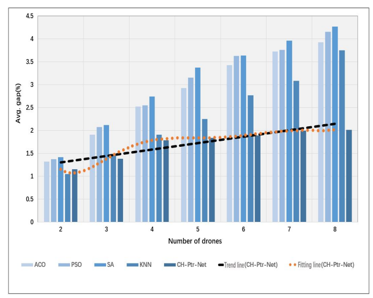
(b)

Fig. 7. The impact of speed and payload on drone energy consumption.

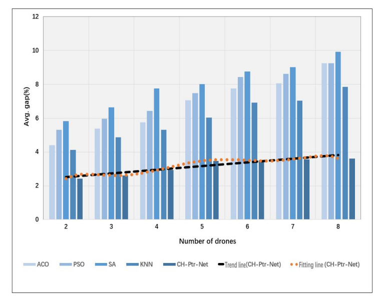
E. Performance Analysis Between CH-Ptr-Net and Exact and Heuristic Algorithms

1) Performance analysis for single objective: To compare the proposed CH-Ptr-Net algorithm performance on the single objective optimization problem for minimum duration, four candidate VRPD solution approaches are investigated. They are frequently-used exact and heuristic algorithms, including ant colony optimization(ACO), particle swarm optimization (PSO), simulated annealing (SA) and K-nearest neighbor (KNN). ACO uses a positive feedback mechanism to make the search process converge and finally approach the best route. PSO belongs to one of the evolutionary algorithms, similar to the simulated annealing algorithm. PSO starts from the random solution, searches for the optimal solution through iteration. SA is affected by the cooling rate. If the cooling rate is slow, the search time is longer. In this way, a better solution can be obtained, but it will take a lot of time. By calculating the distance between the test data and each training data, KNN is a neural network method that sorts according to the distance increasing relationship.

Firstly, we evaluate the delivery distance of the truck and drones by setting other factors unchanged. The change range of the customer number is $[20, 80]$ and $[90, 150]$. The change range of the drone capacity is $\left[0.6 \times 10^4, 1.8 \times 10^4\right]$, and the computing time of each solution is within 30 minutes. Fig. 5 presents the impact of the customer number on the route distance, with the number of drones being 3 and the drone capacity being 1.8×10^4, respectively. Moreover, the impacts of changes in the number of drones and capacity on delivery distance are analyzed in Fig. 6. The number of customers is determined to be 50 and 120, and the delivery distance of the truck and drones show a significant trend.



(a)



(b)

Fig. 8. Validation average optimal gap with different customers. (a) exists 50 customer nodes and (b) exists 120 customer nodes.

When setting 3 drones in the logistics network, Fig. 5 shows the visual representation of truck and drone delivery distance. When the customer number interval is $[20, 80]$, drone delivery distance is more sensitive to the customer number than the truck. Truck delivery distance is less affected by the number of customers. As the number of customers increases, the curve slope of drone delivery distance decreases. We propose CH-Ptr-Net method performs better than other algorithms for total delivery distances. Fig. 6 shows the intuitive impact of drone number and capacity on the truck and drone delivery distances. With the number of drone increasing, the delivery distance of truck decreases rapidly. However, when the number of drone increases to over 6, the driving distance of the truck and drone does not significantly decrease. In addition, when the drone capacity changes from 0.6×10^4 to 1.8×10^4, the sensitivity of truck delivery distance decreased significantly. This is due to the increase in the drone capacity, which enhances the drone delivery potential and reduces the truck operational pressure.

Optimal gap of different solutions reflect the model performance for the collaborative delivery problem. We compare the large-scale and small-scale node delivery problems with ACO, PSO, SA, KNN and CH-Ptr-Net algorithms. As shown in Table VIII, service nodes are set as 50 and 120 respectively, and the gap values reflect different characteristics. Fig. 8 is a more intuitive average optimal performance comparison. The CH-Ptr-Net has

TABLE V
IMPACT OF THE WEIGHT COEFFICIENT ON THE OBJECTIVES

Scenario	Weight coefficient		Delivery duration	Delivery cost	Optimal gap	Model runtime
	β_1	β_2				
1	0.1	0.9	1847	578	1.842	122.703
2	0.2	0.8	1829	635	2.218	124.262
3	0.3	0.7	1753	704	2.533	127.318
4	0.4	0.6	1701	762	2.782	129.294
5	0.5	0.5	1683	812	2.914	131.237
6	0.6	0.4	1572	851	2.721	127.115
7	0.7	0.3	1422	884	2.526	125.826
8	0.8	0.2	1285	912	2.185	122.630
9	0.9	0.1	1136	928	1.901	121.415

promising optimization performance, especially after the number of drone exceeds 2. When network service node number is 50, one truck and eight drones completing the delivery task, the average gap of the proposed algorithm is only 2.017. Compared to the SA with the 4.269 average gap, which of CH-Ptr-Net is increased by 52.752%. When the number of customers increases to 120, the overall trend of optimal performance remains consistent. For the fitting results about the CH-Ptr-Net, the model optimality ability gradually decreases, especially it has obvious decline for more than six drones combination.

2) Performance analysis for multi-objective: Based on the scenario description and parameter settings, we continue to explore the results of multi-objective optimization for minimum duration and delivery cost. To investigate the impact of the multi-objective consideration, different weight coefficient β are configured. Table V shows the impact of the weight coefficient on the objectives. When encountering the scenario with $\beta = 0$, we use the lexicographic ordering strategy, which ranks objectives by optimizing the latter objective at the optimal value of former one, so that the solutions are Pareto optimal.

Table V reveals the significant characteristics of the minimum value. When the weight coefficient β_1 of the delivery time is relatively large, a lower time will naturally be obtained, but the delivery cost is relatively high. On the contrary, when the weight coefficient β_2 of the delivery cost is relatively high, delivery cost becomes the lowest and the duration is not ideal. If we consider two objectives $\beta_1 = \beta_2 = 0.5$ at the same time, and the results of each objective are between those of a system dedicated to a single objective, average optimal gap has the worst performance.

Compared with single objective optimization problem, the proposed multi-objective optimization can better balance multiple objectives and develop more reasonable solutions. It can make ideal trade-offs based on the relative importance of weight coefficient. Decision makers can make adjustments based on actual needs, investment preferences, or expert advice.

F. Enhanced Comparison Between CH-Ptr-Net and Existing ILS-VND on the MT-VRPD

To further evaluate our CH-Ptr-Net model, we consider recent hybrid approaches developed by Gu *et al.* for a very similar problem [56]. Their method hybridize an iterative local search heuristic with a variable neighborhood descent procedure (ILS-VND) to solve the MT-VRPD. Two evaluation indexes contain the average optimal gap and run-time are used. Average optimal gap records the feasibility plan, and the run-time shows the calculation time required for the feasible plan. The size of test instances ranges from 50 to 150, and the recorded time is in seconds. For each instance, we run 100 times with different random seeds and retrieve the optimal vehicle delivery scheme.

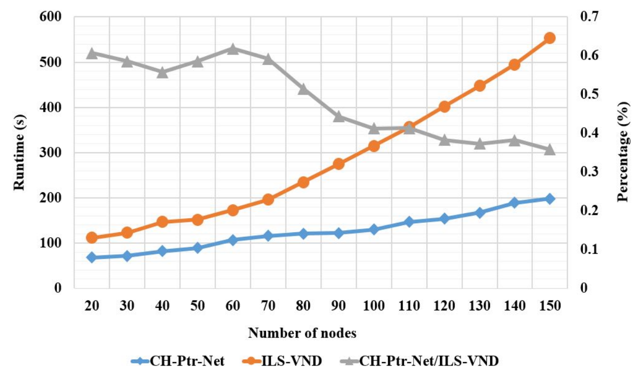


Fig. 9. The comparison about computational time.

Firstly, we measure the optimal gap of CH-Ptr-Net and ILS-VND with small-scale, medium-scale, and large-scale delivery instances, and each instance has 5 replication scenarios with one truck and three drones. Table VI summarizes the optimization results of the 15 scenarios for 3 type of delivery scales. In small-scale delivery scenarios, the minimum optimal gap of ILS-VND performs outstanding with 1.228. However, when the number of network nodes increases to 100 and 150, the superiority of ILS-VND is gradually losing. For example, the minimum, average and maximum optimal gaps of our proposed CH-Ptr-Net method are1.637,2.126 and 2.488 for scenario 11. It can be seen that proposed CH-Ptr-Net has advantages for handling medium and large-scale delivery.

Secondly, the run-time and the ratio of two algorithms are shown in Fig. 9. The number of consumer locations changes from 20 to 150, and the final run-time of CH-Ptr-Net gradually approach around $200s$. Intuitively, the run-time consumed by ILS-VND algorithm is always higher than CH-Ptr-Net. With the increase of the number of customers, the ratio of CH-Ptr-Net to ILS-VND has a significant decline trend, and this also proves the superiority of our proposed method in terms of computational time.

By comparing with existing effective algorithm, the proposed CH-Ptr-Net is demonstrated to have an advantage on optimal gap

TABLE VI
OPTIMAL GAP COMPARISON BETWEEN CH-PTR-NET AND ILS-VND

Scenario	Number of nodes	Minimum optimal gap		Average optimal gap		Maximum optimal gap	
		CH-Ptr-Net	ILS-VND	CH-Ptr-Net	ILS-VND	CH-Ptr-Net	ILS-VND
1	50	1.264	1.332	1.765	1.975	2.237	2.406
2	50	1.285	1.228	1.803	1.991	2.152	2.389
3	50	1.282	1.284	1.778	1.964	2.228	2.416
4	50	1.279	1.297	1.814	1.982	2.195	2.390
5	50	1.254	1.306	1.791	1.963	2.230	2.403
6	100	1.923	2.237	2.247	2.625	2.905	3.204
7	100	1.917	2.294	2.383	2.597	2.886	3.171
8	100	1.908	2.306	2.261	2.614	2.879	3.213
9	100	1.929	2.284	2.357	2.588	2.914	3.185
10	100	1.934	2.301	2.312	2.607	2.906	4.152
11	150	1.673	2.497	2.126	3.527	2.488	4.039
12	150	1.729	2.526	2.096	3.473	2.375	4.207
13	150	1.784	2.518	2.130	3.516	2.491	4.262
14	150	1.702	2.496	2.129	3.607	2.513	4.304
15	150	1.664	2.529	2.104	3.492	2.392	4.271

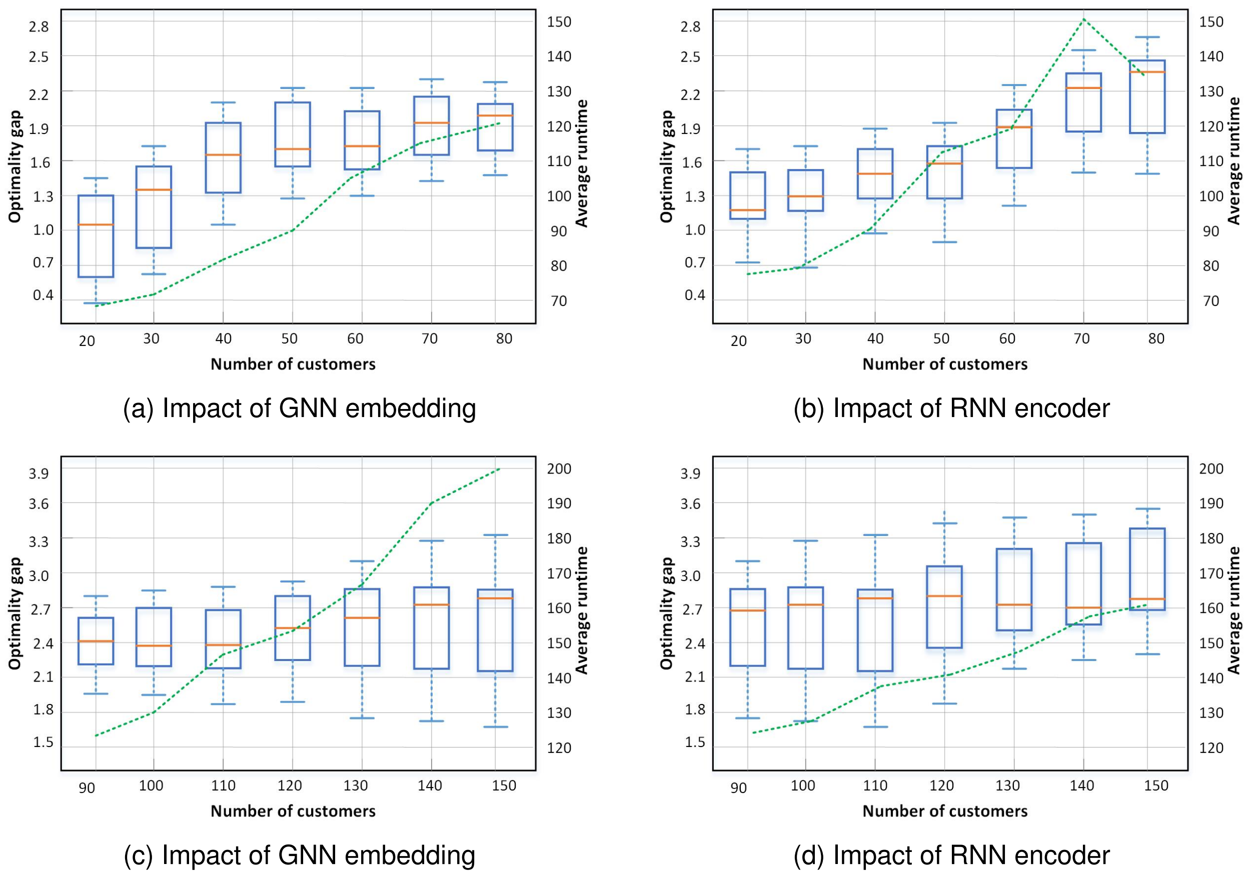


Fig. 10. Impact analysis of algorithm embedding structure. Customer number interval of (a) and (b) is $[20, 80]$. Customer number interval of (c) and (d) is $[90, 150]$.

and computational time. Therefore, the analysis results supports the claim on the effectiveness of reinforcement learning for collaborative delivery optimization with one truck and multiple drones.

G. Effect of Embedding Structure

Network encoding part uses a simplified version to handle irregular and unfixed data structures, which reflects the routing input with a dynamic process. The proposed hybrid model consists of GNN embedding instead of RNN or LSTM encoder. GNN embedding is a kind of neural network directly acting on the graph structure. It embeds all initial inputs with abnormal structures in the current step, including the static node coordinates and dynamic demand states. After that, model decoder points to input dynamic nodes. To compare the impact of GNN embedding structure for CH-Ptr-Net model, we reference the RNN encoder to perform optimization calculations.

Taking the collaborative delivery with single truck and five drones as an example, Table VII summarizes percentage gap involving combinations of GNN embedding and RNN encoder,

TABLE VII
A SUMMARY OF OPTIMAL GAP FOR DIFFERENT NETWORK STRUCTURE.

Metric	GNN embedding on CH-Ptr-Net						RNN encoder on CH-Ptr-Net					
	Optimal gap(%)			Run-time(s)			Optimal gap(%)			Run-time(s)		
Customer	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.
20	1.181	0.374	1.492	68.762	54.973	78.732	1.215	0.704	1.763	78.214	66.849	91.372
30	1.369	0.581	1.743	72.583	60.163	81.628	1.293	0.682	1.772	80.926	68.482	96.623
40	1.751	1.092	1.983	82.046	71.424	90.372	1.475	0.974	1.891	92.375	84.126	102.251
50	1.838	1.286	2.147	89.905	77.693	93.751	1.582	0.745	1.925	114.265	103.041	127.834
60	1.874	1.305	2.155	107.274	92.861	112.592	1.904	1.263	2.194	119.973	112.042	119.728
70	1.937	1.459	2.217	116.096	104.066	121.731	2.096	1.486	2.462	151.208	140.843	167.037
80	2.048	1.513	2.204	121.479	113.095	126.832	2.237	1.471	2.581	135.872	130.729	159.482
90	2.413	1.937	2.803	122.385	113.053	134.538	2.681	1.741	3.102	124.016	120.842	138.163
100	2.384	1.904	2.835	130.163	119.752	142.532	2.742	1.716	3.273	127.803	119.856	140.528
110	2.391	1.892	2.817	147.632	138.749	159.535	2.793	1.694	3.324	138.043	131.638	147.533
120	2.668	1.903	2.918	154.973	146.542	171.743	2.804	1.891	3.415	141.016	138.538	151.546
130	2.716	1.759	3.126	167.739	154.633	183.904	2.735	2.174	3.486	148.228	139.972	157.352
140	2.725	1.748	3.282	189.930	176.732	203.927	2.909	2.251	3.503	157.029	149.053	166.384
150	2.826	1.682	3.348	198.625	191.352	214.627	3.196	2.306	3.552	161.273	153.563	169.534

with different customer number interval. For optimal gap, GNN embedding has better competitiveness than RNN encoder on CH-Ptr-Net. Especially when the number of customers exceeds 80, this difference is even more pronounced. When the network service node increases to 150, the average gap of GNN embedding on CH-Ptr-Net is only 2.8, far lower than 3.196. While model run-time matters, GNN embedding enables the delivery routing sparingly.

Fig. 10 discusses the impact of network embedding structure intuitively, by box-whisker plot. In the Fig. 10(a) and Fig. 10(c), the box plot changes from long to short and then to long, meaning that optimal gaps of GNN embedding are first dispersed, then concentrated within a small range, and finally dispersed. It can be seen that GNN embedding has strong optimization ability for customer interval between 60 and 120, while RNN encoder shows the opposite trend as shown in Fig. 10(b) and Fig. 10(d). When the median is close to the bottom, it indicates that most of the gap values are relatively small. The median distribution also indicates the gap characteristics of GNN embedding and RNN encoder. For average run-time, when customer number interval is $[20, 80]$, GNN embedding outperforms the RNN encoder. As the number of customers increases, the run-time advantage of GNN embedding is lost.

From the above analysis, it can be seen that the proposed CH-Ptr-Net algorithm can improve optimal performance for the collaborative delivery problem with the truck and multiple drones. Especially, GNN embedding is a more simplified version than RNN encoder. When a certain customer is visited, the requirements may be reset to zero. The RNN encoder is only helpful for input order related tasks, but not necessary for input order unrelated sequences. So the encoder part is removed and only the decoder part is retained, which simplifies the model decoding process and improves the accuracy and optimization performance.

TABLE VIII
OPTIMAL PERFORMANCE COMPARISON BETWEEN DIFFERENT MODELS

Average gap	50 customers					120 customers				
Methods	ACO	PSO	SA	KNN	CH-Ptr-Net	ACO	PSO	SA	KNN	CH-Ptr-Net
2 *drones*	1.318	1.374	1.416	***1.047***	1.158	4.403	5.316	5.827	4.126	**2.437**
3 *drones*	1.906	2.074	2.122	1.483	**1.379**	5.384	5.958	6.637	4.856	**2.617**
4 *drones*	2.521	2.547	2.737	1.905	**1.784**	5.761	6.439	7.748	5.308	**2.625**
5 *drones*	2.927	3.153	3.369	2.248	**1.838**	7.057	7.465	8.018	6.041	**2.668**
6 *drones*	3.421	3.627	3.637	2.769	**1.891**	7.763	8.439	8.748	6.908	**2.725**
7 *drones*	3.727	3.756	3.962	3.082	**1.994**	8.057	8.617	9.018	7.041	**2.768**
8 *drones*	3.926	4.153	4.269	3.751	**2.017**	9.236	9.253	9.917	7.843	**2.861**

H. Discussions

Different from the traditional exact algorithm or a certain integrated heuristic algorithm, the proposed constraint-based hybrid pointer network model is more suitable for the collaborative delivery problem with the truck and multiple drones. Moreover, the proposed reinforcement learning model derives the reward calculation and adapts to the new delivery routing automatically, regardless of explicit distance matrix. GNN embedding helps the decoder suitable for the irregular and unstable data structures, and it uses graph structures to characterize the sequence internal connections.

VI. Conclusion

In this paper, we propose a novel self-driven CH-Ptr-Net reinforcement learning model to analyze and optimize the collaborative delivery problem with the truck and multiple drones. The CH-Ptr-Net model is constructed by a set of integral generated sequences, and the generated structure coupled with GNN embedding and attention mechanism helps to adapt to the new trajectory structure automatically. Specifically, we develop the mixed integer linear program formulation related to energy consumption, vehicle routing availability, drone endurance, customer service time windows and service sequence. After that, the proposed CH-Ptr-Net model learns from routing optimization matrix. It can handle the initial data with irregular and unfixed structures, and then model establishes a weight relationship for each position association between the output and the input by attention mechanism. Experimental results show that the proposed model achieves superior performance than other algorithms for the collaborative delivery optimization. Regarding to future work, we will further study the logistics delivery routing with multiple trucks and drones, so as to adapt to more complex delivery network.

REFERENCES

[1] M. A. Nguyen, G. T.-H. Dang, M. H. H, and M.-T. Pham, "The min-cost parallel drone scheduling vehicle routing problem," *Eur. J. Oper. Res.*, vol. 299, no. 3, pp. 910–930, Jul. 2022.

[2] P. Bachmann and D. Frutos-Bencze, "Rd and innovation efforts during the covid-19 pandemic: The role of universities," *J. Innov. Knowl.*, vol. 7, no. 4, pp. 100 238–100 250, Aug. 2022.

[3] C. C. Murray and A. G. Chu, "The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery," *Transp. Res. Part C*, vol. 54, no. 1, pp. 86–109, May. 2015.

[4] L. Peng and C. C. Murray, "Veroviz: A vehicle routing visualization toolkit," *INFORMS J. Comput.*, vol. 34, no. 4, pp. 1842–1848, Sep. 2022.

[5] C. Buck, J. Clarke, R. Torres de Oliveira, K. C. Desouza, and P. Maroufkhani, "Digital transformation in asset-intensive organisations: The light and the dark side," *J. Innov. Knowl.*, vol. 8, no. 2, pp. 100 335–100 346, Feb. 2023.

[6] A. Bogyrbayeva, T. Yoon, H. Ko, S. Lim, H. Yun, and C. Kwon, "A deep reinforcement learning approach for solving the traveling salesman problem with drone," *Transp. Res. Part C*, vol. 148, no. 1, pp. 103 981–103 998, Jan. 2023.

[7] S. Lee, D. Hong, J. Kim, D. Baek, and N. Chang, "Congestion-aware multi-drone delivery routing framework," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9384–9396, Sep. 2022.

[8] T. Zhou, M. L. Kris, D. Creighton, and C. Wu, "Gmix: Graph-based spatialtemporal multi-agent reinforcement learning for dynamic electric vehicle dispatching system," *Transp. Res. Part C*, vol. 144, no. 1, pp. 103 886–103 905, Oct. 2022.

[9] M. Chiu and C. Lin, "Developing supply chain open innovation capability: The mediating role of the knowledge creation process, governance mechanism and technology as a driver," *J. Innov. Knowl.*, vol. 7, no. 4, pp. 100 264–100 279, Sep. 2022.

[10] C. Archetti, L. Peirano, and M. G. Speranza, "Optimization in multimodal freight transportation problems: A survey," *Eur. J. Oper. Res.*, vol. 299, no. 1, pp. 1–20, Sep. 2022.

[11] H. Ensafian, Abdullah, M. G. Bell, D., P. Kilby, and J. Li, "Cost-optimal deployment of autonomous mobile lockers co-operating with couriers for simultaneous pickup and delivery operations," *Transp. Res. Part C*, vol. 146, no. 1, pp. 103 958–103 985, Dec. 2023.

[12] H. Huang, A. V. Savkin, and C. Huang, "Reliable path planning for drone delivery using a stochastic time-dependent public transportation network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 4941–4950, Aug. 2021.

[13] K. Chen, M. Xie, Y. Chen, T. Chu, and Y. Lin, "Dronetalk: An internet-of-things-based drone system for last-mile drone delivery," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15 204–15 217, Jun. 2022.

[14] J. P. Aurambout, K. Gkoumas, and B. Ciuffo, "A drone hop from the local shop? where could drone delivery as a service happen in europe and the usa, and how many people could benefit from it?" *Transp. Res. Interdp. Perspec.*, vol. 16, no. 1, pp. 100 708–100 716, Dec. 2022.

[15] P. Li, A. Bastone, T. A. Mohamad, and F. Schiavone, "How does artificial intelligence impact human resources performance. evidence from a healthcare institution in the united arab emirates," *J. Innov. Knowl.*, vol. 8, no. 2, pp. 100 340–100 347, Feb. 2023.

[16] D. Sun, X. Peng, R. Qiu, and Y. Huang, "The traveling salesman problem: Route planning of recharging station-assisted drone delivery," *Intl. Conf. Mgt. Sci. Egn. Mgt.*, vol. 22, no. 1, pp. 13–23, Jan. 2020.

[17] H. Huang, A. V. Savkin, and C. Huang, "Drone routing in a time-dependent network: Toward low-cost and large-range parcel delivery," *IEEE Trans. Ind. Inform.*, vol. 17, no. 2, pp. 1526–1534, Jul. 2021.

[18] C. C. Murray and R. Raj, "The multiple flying sidekicks traveling salesman problem: Parcel delivery with multiple drones," *Transp. Res. Part C*, vol. 110, no. 1, pp. 368–398, Dec. 2020.

[19] G. Wu, N. Mao, Q. Luo, B. Xu, J. Shi, and P. N. Suganthan, "Collaborative truck-drone routing for contactless parcel delivery during the epidemic," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 25 077–25 091, Feb. 2022.

[20] M. Y. Arafat and S. Moh, "Jrcs: Joint routing and charging strategy for logistics drones," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21 751–21 764, Jun. 2022.

[21] R. G. Mbiadou Saleu, L. Deroussi, D. Feillet, N. Grangeon, and A. Quilliot, "The parallel drone scheduling problem with multiple drones and vehicles," *Eur. J. Oper. Res.*, vol. 300, no. 2, pp. 571–589, Sep. 2022.

[22] M. Perreault and K. Behdinan, "Delivery drone driving cycle," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1146–1156, Jan. 2021.

[23] Z. Meng, Y. Zhou, E. Y. Li, X. Peng, and R. Qiu, "Environmental and economic impacts of drone-assisted truck delivery under the carbon market price," *J. Clean. Pr.*, vol. 401, no. 1, pp. 136 758–136 772, Mar. 2023.

[24] Q. Luo, G. Wu, B. Ji, L. Wang, and P. N. Suganthan, "Hybrid multi-objective optimization approach with pareto local search for collaborative truck-drone routing problems considering flexible time windows," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 13 011–13 025, Oct. 2022.

[25] N. Elsokkary, H. Otrok, S. Singh, R. Mizouni, H. Barada, and M. Omar, "Crowdsourced last mile delivery: Collaborative workforce assignment," *IEEE Internet Things J.*, vol. 22, pp. 100 692–100 709, Apr. 2023.

[26] X. Wang, Z. Liu, and X. Li, "Optimal delivery route planning for a fleet of heterogeneous drones: A rescheduling-based genetic algorithm approach," *Comput. Ind. Eng.*, vol. 179, no. 1, pp. 109 179–109 187, Mar. 2023.

[27] Y. Yang, C. Yan, Y. Cao, and R. Roberti, "Planning robust drone-truck delivery routes under road traffic uncertainty," *Eur. J. Oper. Res.*, vol. 315, no. 2, pp. 1–13, Feb. 2023.

[28] D. N. Das, R. Sewani, J. Wang, and M. K. Tiwari, "Synchronized truck and drone routing in package delivery logistics," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5772–5782, May. 2021.

[29] G. Xue, "A two-stage heuristic solution for multi-depot collaborative pickup and delivery network with transfers to reduce carbon emissions," *J. Clean. Pr.*, vol. 373, pp. 133 839–133 847, Aug. 2022.

[30] Y. Boxu, L. Xingguang, and K. Kou, "Research on the influence of network embeddedness on innovation performance: Evidence from china's listed firms," *J. Innov. Knowl.*, vol. 7, no. 3, pp. 100 210–100 221, Jun. 2022.

[31] F. B. Sorbelli, F. Cor, L. Palazzetti, C. M. Pinotti, and G. Rigoni, "How the wind can be leveraged for saving energy in a truck-drone delivery system," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4038–4049, Jan. 2023.

[32] J. Zhang, J. F. Campbell, D. C. Sweeney II, and A. C. Hupman, "Energy consumption models for delivery drones: A comparison and assessment," *Transp. Res. Part D*, vol. 90, no. 7, p. 102668, Jan. 2021.

[33] N. A. Kyriakakis, S. Aronis, M. Marinaki, and Y. Marinakis, "A grasp/vnd algorithm for the energy minimizing drone routing problem with pickups and deliveries," *Comput Ind Eng*, vol. 182, no. 1, p. 109340, Aug. 2023.

[34] H. Malapur, Y. Singh, M. Shendkar, A. Barve, and M. Bedekar, "Diffused casing of drone propeller for reduced operational noise and optimized energy consumption," *Mat. Tod: Proc.*, vol. 63, no. 1, pp. 136–140, Jun. 2022.

[35] F. Tamke and U. Buscher, "The vehicle routing problem with drones and drone speed selection," *Comput Oper Res*, vol. 152, no. 1, p. 106112, Jan. 2023.

[36] F. B. Sorbelli, F. Cor, S. K. Das, and C. M. Pinotti, "Energy-constrained delivery of goods with drones under varying wind conditions," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 6048–6060, Sep. 2021.

[37] O. Dukkanci, B. Y. Kara, and T. Bekta, "Minimizing energy and cost in range-limited drone deliveries with speed optimization," *Transp. Res. Part C*, vol. 125, no. 1, p. 102985, Apr. 2021.

[38] A. Nauman, M. Obayya, M. M. Asiri, K. Yadav, M. Maashi, M. Assiri, M. K. Ehsan, and S. W. Kim, "Minimizing energy consumption for noma multi-drone communications in automotive-industry 5.0," *J King Saud Univ-Com*, vol. 35, no. 6, p. 101547, Jun. 2023.

[39] B. Sun, Q. Zhang, L. Hu, C. Zou, N. Wei, Z. Jia, and X. Zhao, "A prediction-evaluation method for road network energy consumption: Fusion of vehicle energy flow principle and two-fluid theory," *Physica A: Stat. Mech. Appl.*, vol. 626, no. 1, p. 129077, Sep. 2023.

[40] K. Dorling, J. Heinrichs, M. Geoffrey G, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Trans.SMC*, vol. 47, no. 1, pp. 70–85, Jul. 2017.

[41] M. A. Figliozzi, "Lifecycle modeling and assessment of unmanned aerial vehicles (drones) co2e emissions," *Transp. Res. Part E*, vol. 57, no. 1, pp. 251–261, Dec. 2017.

[42] J. K. Stolaroff, S. Constantine, E. R, L. Alia, M. A. S, and C. Daniel, "Energy use and life cycle greenhouse gas emissions of drones for commercial package delivery," *Nat. Commun.*, vol. 9, no. 1, p. 409, Dec. 2018.

[43] Y. Yin, D. Li, D. Wang, J. Ignatius, T. Cheng, and S. Wang, "A branch-and-price-and-cut algorithm for the truck-based drone delivery routing problem with time windows," *Eur. J. Oper. Res.*, vol. 324, no. 3, pp. 269–287, Feb. 2023.

[44] H. N. Tariq and M. B. Bashir, "An optimized protocol for cost effective communication in a multi-agent environment," *J. Parallel Distr. Comput.*, vol. 169, no. 1, pp. 24–41, Jun. 2022.

[45] W. Najy, C. Archetti, and A. Diabat, "Collaborative truck-and-drone delivery for inventory-routing problems," *Transp. Res. Part C*, vol. 146, pp. 103 791–103 910, Aug. 2023.

[46] Y. Zhao, S. Wen, T. Zhou, W. Liu, H. Yu, and H. Xu, "Development and innovation of enterprise knowledge management strategies using big data neural networks technology," *J. Innov. Knowl.*, vol. 7, no. 4, pp. 100 273–100 281, Sep. 2022.

[47] I. S. Damoah, A. Ayakwah, and I. Tingbani, "Artificial intelligence (ai)-enhanced medical drones in the healthcare supply chain (hsc) for sustainability development: A case study," *J. Clean. Pr*, vol. 328, no. 1, p. 129598, Feb. 2021.

[48] A. Arishi, K. Krishnan, and M. Arishi, "Machine learning approach for truck-drones based last-mile delivery in the era of industry 4.0," *Eng. Appl. Artif. Intel.*, vol. 116, no. 1, p. 105439, Jan. 2022.

[49] M. Choi, M. Shin, and J. Kim, "Dynamic video delivery using deep reinforcement learning for device-to-device underlaid cache-enabled internet-of-vehicle networks," *J. Commun. Net.*, vol. 23, no. 2, pp. 117–128, Aug. 2021.

[50] T. Kim, Y.-W. Kim, D. Lee, and M. Kim, "Reinforcement learning approach to scheduling of precast concrete production," *J. Clean. Pr.*, vol. 336, no. 1, pp. 130 419–130 427, Jan. 2022.

[51] C. Fu, Z. Gao, W. Wu, V. Chau, J. Wang, X. Xu, and J. Luo, "A learning approach for multi-agent travelling problem with dynamic service requirement in mobile iot," *Comput. Electr. Eng.*, vol. 104, no. 1, pp. 108 397–108 408, Oct. 2022.

[52] K. Zhang, X. Lin, and M. Li, "Transformer-based reinforcement learning for pickup and delivery problems with late penalties," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24 649–24 661, Aug. 2022.

[53] J. Li, L. Xin, Z. Cao, A. Lim, W. Song, and J. Zhang, "Heterogeneous attentions for solving pickup and delivery problem via deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2306–2315, Feb. 2022.

[54] F. Kong, J. Li, B. Jiang, H. Wang, and H. Song, "Trajectory optimization for drone logistics delivery via attention-based pointer network," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 4, pp. 4519–4531, May. 2023.

[55] K. Ren, "Helicopter-mounted hf direction-finding antenna with rotor-blade modulation using the characteristic mode theory," *IEEE Antenn Wirel Pr*, vol. 22, no. 7, pp. 1597–1601, Mar. 2023.

[56] R. Gu, M. Poon, Z. Luo, Y. Liu, and Z. Liu, "A hierarchical solution evaluation method and a hybrid algorithm for the vehicle routing problem with drones and multiple visits," *Transp. Res. Part C*, vol. 141, p. 103733, Jun. 2022.

Fanhui Kong received the M.S. and Ph.D. degree in management science and engineering from Tianjin University of Technology, Tianjin, China, in 2016 and 2020. She was also a joint Ph.D student in the Davis College of Business, Jacksonville University, Jacksonville, FL, USA, from 2018 to 2019. She was a Postdoctoral Research Fellow in the College of Computer and Software Engineering, Shenzhen University, Shenzhen, China, from Aug.2020 to Jul.2022.

Dr.Kong is currently an Associate Professor in the School of Management Engineering, Qingdao University of Technology, Qingdao, China. Her research interests include Intelligent network optimization, complex system analysis and computation.

Bin Jiang (Member, IEEE) received the Ph.D. degree in information and communication engineering from Tianjin University, Tianjin, China, in 2020, where he also received the B.S. and M.S. degree in 2013 and 2016. He is currently an Associate Professor with the College of Oceanography and Space Informatics, China University of Petroleum (East China). Dr.Jiang was a joint Ph.D student in the Security and Optimization for Networked Globe Laboratory, Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA, from Nov. 2017 to Feb.2018, and from Oct.2018 to Oct.2019. He was a postdoctoral in College of Computer Science and Software Engineering, Shenzhen University, China, from Aug.2020 to July.2022. His major research interests include Internet of Things, cyber-physical security and privacy computing.

Jian Wang (Member, IEEE) is an assistant professor in the department of Computer Science, The University of Tennessee at Martin, Martin, Tennessee, and a senior member in the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). He received the Ph.D. degree in Computer Science and Electrical Engineering from Embry-riddle Aeronautical University in 2021. His major research interests include wireless networks, unmanned aircraft systems, and machine learning. Jian Wang is a recipient of 3 Best Paper Awards from CPSCom-2019, CBDCom 2020, and WASA 2020.

Huihui Wang (Senior Member, IEEE) is an Associate Professor and the Director of Cybersecurity program at St. Bonaventure University. Her current interests are Cybersecurity of Cyber Physical Systems/Internet of Things, and the cyber/engineering education. She has been awarded about 1.8M dollar from the external grant agencies including NSF, private foundations and industry. She has been actively serving as a TPC chair, member or reviewer of various panels, journals and conferences. She is a senior member of IEEE, a member and the Program Chair of ASEE.

Houbing Song (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in August 2012.

He is currently a Tenured Associate Professor, the Director of the NSF Center for Aviation Big Data Analytics (Planning), the Associate Director for Leadership of the DOT Transportation Cybersecurity Center for Advanced Research and Education (Tier 1 Center), and the Director of the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us), University of Maryland, Baltimore County (UMBC), Baltimore, MD. Prior to joining UMBC, he was a Tenured Associate Professor of Electrical Engineering and Computer Science at Embry-Riddle Aeronautical University, Daytona Beach, FL. He serves as an Associate Editor for IEEE Transactions on Artificial Intelligence (TAI) (2023-present), IEEE Internet of Things Journal (2020-present), and IEEE Transactions on Intelligent Transportation Systems (2021-present). He was an Associate Technical Editor for IEEE Communications Magazine (2017-2020). His research interests include cyber-physical systems/internet of things, and AI/machine learning/big data analytics. His research has been sponsored by federal agencies (including National Science Foundation, National Aeronautics and Space Administration, US Department of Transportation, and Federal Aviation Administration, among others) and industry. His research has been featured by popular news media outlets, including IEEE GlobalSpecs Engineering360, Association for Uncrewed Vehicle Systems International (AUVSI), Security Magazine, CXOTech Magazine, Fox News, U.S. News and World Report, The Washington Times, and New Atlas.

Dr. Song is an IEEE Fellow (for contributions to big data analytics and integration of AI with Internet of Things), and an ACM Distinguished Member (for outstanding scientific contributions to computing). He is an ACM Distinguished Speaker (2020-present), an IEEE Vehicular Technology Society (VTS) Distinguished Lecturer (2023-present) and an IEEE Systems Council Distinguished Lecturer (2023-present). Dr. Song has been a Highly Cited Researcher identified by Clarivate (2021, 2022) and a Top 1000 Computer Scientist identified by Research.com. He received Research.com Rising Star of Science Award in 2022 (World Ranking: 82; US Ranking: 16). In addition to 2021 Harry Rowe Mimno Award bestowed by IEEE Aerospace and Electronic Systems Society, Dr. Song was a recipient of 10+ Best Paper Awards from major international conferences, including IEEE CPSCom-2019, IEEE ICII 2019, IEEE/AIAA ICNS 2019, IEEE CBDCom 2020, WASA 2020, AIAA/IEEE DASC 2021, IEEE GLOBECOM 2021 and IEEE INFOCOM 2022.