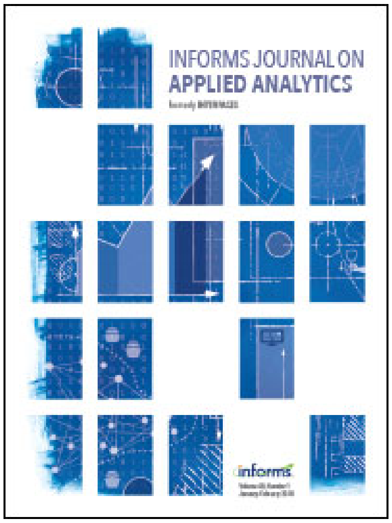


INFORMS Journal on Applied Analytics

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Alibaba Vehicle Routing Algorithms Enable Rapid Pick and Delivery

Haoyuan Hu, Ying Zhang, Jiangwen Wei, Yang Zhan, Xinhui Zhang, Shaojian Huang, Guangrui Ma, Yuming Deng, Siwei Jiang

informs
http://pubsonline.informs.org/journal/inte
INFORMS JOURNAL ON APPLIED ANALYTICS
Vol. 52, No. 1, January–February 2022, pp. 27–41
ISSN 2644-0865 (print), ISSN 2644-0873 (online)

THE FRANZ EDELMAN AWARD
Achievement in Operations Research

Alibaba Vehicle Routing Algorithms Enable Rapid Pick and Delivery

Haoyuan Hu,[a] Ying Zhang,[a] Jiangwen Wei,[a] Yang Zhan,[a] Xinhui Zhang,[b,*] Shaojian Huang,[b] Guangrui Ma,[b] Yuming Deng,[b] Siwei Jiang[c]

[a] Cainiao Network, Hangzhou, Zhejiang 311100, China; [b] Alibaba Group, Hangzhou, Zhejiang 310052, China; [c] Lazada Group, Singapore 068811
*Corresponding author

Contact: haoyuan.huhy@cainiao.com (HH); youzhu.zy@cainiao.com (YiZ); jiangwen.wjw@alibaba-inc.com (JW); zhanyang.zy@cainiao.com (YaZ); maisui.zxh@alibaba-inc.com (XZ); shaojian.hsj@alibaba-inc.com (SH); guangrui.ma@sjtu.edu.cn (GM); yuming.dym@alibaba-inc.com (YD); siwei.jiangsiwei@lazada.com (SJ)

Accepted: September 23, 2021

https://doi.org/10.1287/inte.2021.1108

Abstract. Alibaba Group pioneered integrated online and offline retail models to allow customers to place online orders of e-commerce and grocery products at its participating stores or restaurants for rapid delivery—in some cases, in as little as 30 minutes after an order has been placed. To meet these service commitments, quick online routing decisions must be made to optimize order picking routes at warehouses and delivery routes for drivers. The solutions to these routing problems are complicated by stringent service commitments, uncertainties, and complex operations in warehouses with limited space. Alibaba has developed a set of algorithms for vehicle routing problems (VRPs), which include an open-architecture adaptive large neighborhood search to support the solution of variants of routing problems and a deep learning-based approach that trains neural network models offline to generate almost instantaneous solutions online. These algorithms have been implemented to solve VRPs in several Alibaba subsidiaries, have generated more than $50 million in annual financial savings, and are applicable to the broader logistics industry. The success of these algorithms has fermented an inner-source community of operations researchers within Alibaba, boosted the confidence of the company's executives in operations research, and made operations research one of the core competencies of Alibaba Group.

Keywords: vehicle routing problems • last-mile delivery • adaptive large neighborhood search • deep reinforcement learning • inner source • Edelman Award

Introduction

E-Business, Supply Chain, and Logistics Endeavors at Alibaba

Alibaba Group (Alibaba in the remainder of this paper), which was founded in 1999 and operates one of the world's largest e-commerce platforms, is a multinational technology company in China. Alibaba recorded 757 million annual active customers (i.e., consumers who make at least one purchase during the year) and 818 million monthly active users (i.e., consumers who visit Alibaba's website at least once during the month) on its mobile and desktop retail platforms within China (Alibaba 2020). Since 2009, Alibaba has successfully organized the Chinese Singles' Day shopping festival: the 2020 festival generated a record-high $74.1 billion in gross merchandise volume in an 11-day period from November 1, 2020, to November 11, 2020. In 2020, Alibaba accounted for 50.8% of all online retail sales in China.

To help its merchants, such as manufacturers and wholesalers, to connect with customers worldwide and to facilitate their capabilities to do business anywhere in the world, Alibaba has actively driven technology innovation in the digital transformation of businesses from commerce to logistics. In recent years, it has pioneered new retail models that integrate online and offline shopping, enable customers to order almost any product available at its participating stores or restaurants, and have the product delivered through its logistics services within a short time after placing the order (e.g., in as little as 30 minutes). Alibaba subsidiaries implementing these retail models include Freshippo, Ele.me Taoxianda (Ele.me and Taoxianda were merged to create Local Services in 2020), Lazada, and Cainiao Network.

Freshippo, established in 2016, is a chain of more than 300 retail stores in metropolitan areas. It is an exemplary, and perhaps the first, omnichannel online and offline supermarket; a store comprises two sections: a semiautomated back-end warehouse and a front-end store. Freshippo handles more than 1.5 million orders per day, half of which are digital online orders. Because

Table 1. Summary of the Business Services that Alibaba's Subsidiaries Provide

Business (subsidiary)	Products	Delivery radius	Service commitment
Freshippo	Grocery	3 km	30 minutes
Ele.me	Food, drug, miscellaneous	5 km	45 minutes
Taoxianda	Grocery	10 km	Two to four hours
Lazada	General merchandise	10 km	Next day or two days
Cainiao Network	General merchandise	Large cities and most small cities in China	Next day or two days

of space limitations in its stores, a suspension chain conveyor system is custom designed for each store to expedite the movement of orders, and delivery is guaranteed in 30 minutes or less if the customer lives within a three-kilometer radius of the store.

Taoxianda, launched in 2019, is an e-commerce platform that enables Alibaba's customers to shop for groceries at traditional brick-and-mortar supermarkets; it has expanded these stores to integrate online shopping. Customers within 10 kilometers of a participating supermarket can place orders and schedule home delivery within four hours at the same price they would pay in the store. The aggregate volume of online orders over many stores makes efficient logistic operations possible. Taoxianda is available in 5,000 grocery stores in more than 100 cities and delivers half a million orders per day.

Ele.me, established in 2008 and acquired by Alibaba in 2018, is a catering and food delivery platform that provides immediate pickup and delivery services (i.e., delivery in less than 45 minutes) of customers' time-sensitive orders at restaurants, drug stores, and other stores. A high-density population and relatively affordable deliveries have stimulated the rapid development of the industry and of Ele.me. Ele.me operates in more than 670 cities, has more than 3 million registered delivery drivers, connects over 4 million merchants and 260 million users, and delivers more than 15 million orders per day.

Lazada, founded in 2012 and acquired by Alibaba in 2016, is South Asia's leading e-commerce platform with a presence in six countries—Singapore, Indonesia, Malaysia, Vietnam, the Philippines, and Thailand. Lazada connects customers to thousands of international and local brands, provides an end-to-end logistics capability with more than 380 warehouses across 17 cities, serves customers living within 10 kilometers of a warehouse with guaranteed next-day and two-day services, and delivers more than two million orders per day.

Cainiao Network, established in 2013, is Alibaba's official logistics platform and has an objective of delivering to and from major cities and most small suburban cities in China in one to two days. Cainiao has built more than 1,000 warehouses and 40,000 last-mile delivery stations and provides logistics services to the merchants on Alibaba's various e-commerce platforms. It handles more than 42 million packages per day originating from more than 2,800 districts to millions of customers in China.

In 2020, Taoxianda and Ele.me were merged into Alibaba's newly founded Local Services, which has an objective of collaborating with diverse local stores and has new retail models under development. These models will provide a variety of service options to meet customers' demands for almost any product, such as fresh food, groceries, and over-the-counter drugs. Table 1 summarizes these subsidiaries and their service characteristics.

Logistic Operations in the New Businesses

The success of these businesses requires the creative and careful design of business models and efficient and effective supply chain operations. These operations require critical yet complex business decisions; examples include assortment planning, demand forecast, inventory management, and routing decisions, where operations research plays a key role in making these decisions. The focus of our research is the optimal routing decisions that arise as part of the warehouse and logistic operations. These decisions affect customer service, logistics costs, and the efficiency of the entire supply chain, and they are of vital importance to the bottom lines of these businesses.

Routing decisions in our business models mainly address (1) the order-picking operations at stores and warehouses and (2) route planning, typically for last-mile delivery. Both of these decisions are often modeled as (multiple) vehicle routing problems (VRPs). A multiple VRP consists of determining a set of routes, one for each vehicle, that starts and ends at the depot. Each vehicle visits a subset of nodes exactly once and travels along the edges (connecting one node to another). The multiple VRPs involve both partition and routing decisions. For example, in last-mile delivery routing, drop-off sites (i.e., the locations to which customer orders will be delivered) are partitioned such that they are associated with vehicles, and the nodes in a delivery route represent the subset of drop-off sites assigned to the vehicle; in the order-picking operation, however, customer orders are partitioned such that they are assigned to pickers—who are modeled as vehicles in the VRP—but the nodes in a picker route represent the storage locations of items within the warehouse that are associated with the orders assigned to that picker. Many complicated forms of VRPs exist, as we discuss in the *Problems and Challenges* section, and effective and efficient solutions to these routing problems are the focus of this paper.

The literature includes many studies to solve practical routing problems in various logistics operations. Examples include optimizing delivery routes at Waste Management (Sahoo et al. 2005), UPS (Holland et al. 2017), and Coca-Cola (Kant et al. 2008) and routing school buses for Boston's public schools (Bertsimas et al. 2020). In contrast to our problem, the first three studies involve the delivery of future orders where all orders are known. In these examples, sufficient time (e.g., one to two hours) is typically available to solve the corresponding problem.

In contrast to many of these traditional planning problems, because of Alibaba's stringent expectation for rapid and on-time delivery, the routing problems in many of our problems are online optimization problems where decisions must be made quickly (i.e., in seconds, in many cases) after an order has been placed, to leave sufficient time for subsequent offline order-picking and delivery operations. Longer computational times would cause a delivery delay and impair customer service.

The remainder of the paper is organized as follows. In the *Problems and Challenges* section, we provide some examples of routing problems in Alibaba's subsidiaries, their characteristics, and the challenges encountered. In *Technical Solutions*, we describe our technical solutions and computational results. In *Inner-Source Development Leads to Broad Use*, we discuss our inner-source framework for software development and how it led to broad use of the work and continuous model and algorithm improvement. In *Portability to Related Problems*, we give examples of models developed for related problems under our inner-source framework to foster synergies among operations researchers and to illustrate the portability of the algorithms. In *Benefits*, we summarize the benefits accrued. Finally, in the *Summary* section, we conclude and summarize this paper.

Problems and Challenges

Practical VRPs are complex, and the rapid development of these business models has introduced additional routing problems, which must be modeled promptly and solved effectively. In the following subsections, we present examples of routing problems to illustrate the variety of problems that Alibaba faces and their characteristics—a set of three complex order-picking and delivery-routing problems at Freshippo, the open-routing problem with pickup and delivery and due-time requirements at Ele.me, the two-echelon routing problem with synchronization and stochastic travel times at Taoxianda, and the multiple-trip routing problems at Lazada.

Example 1: The Complex Order-Picking and Delivery-Routing Problem at Freshippo

Figure 1 depicts the order-picking and delivery-routing problem at Freshippo. As the figure shows, online orders, upon their arrival, are assigned such that batches are created, as we show in panel (a). Each batch is then split into several picking tasks or routes to be picked in different areas in the back-end warehouse or the front-end store, as we show in the top portion in (b). Once each picking task is completed, the item(s) picked are transported through a suspension chain to be sorted at the packing area, as we see in the bottom portion in panel (b). Because of space limitations on the suspension chain and at the packing area, the picking tasks must be synchronized so that all items of a batch arrive at the packing area at about the same time; consequently, the customer orders in a batch, such as orders 2, 4, 8, and 9 in batch A, must be delivered together on one route by a delivery driver, deliverer B, as we show in panel (c).

This results in three VRPs that must be solved: the order-batching problem as a VRP with both warehouse and delivery objectives, the picker routing problem as a VRP with a synchronization requirement, and delivery routing as a VRP with batch restrictions:

1. The order-batching problem must consider warehouse objectives (e.g., the number of distinct items in customer orders, picking routes in the warehouse), the due times (i.e., the times by which customer orders must be delivered), and delivery routes by delivery drivers, as panel (a) in Figure 1 shows.

2. The picker routing problem must consider the available workforce and the synchronization of the arrival of picking tours at the suspension drop point (i.e., intermediate point) at about the same time to avoid congestion, as panel (b) in Figure 1 shows.

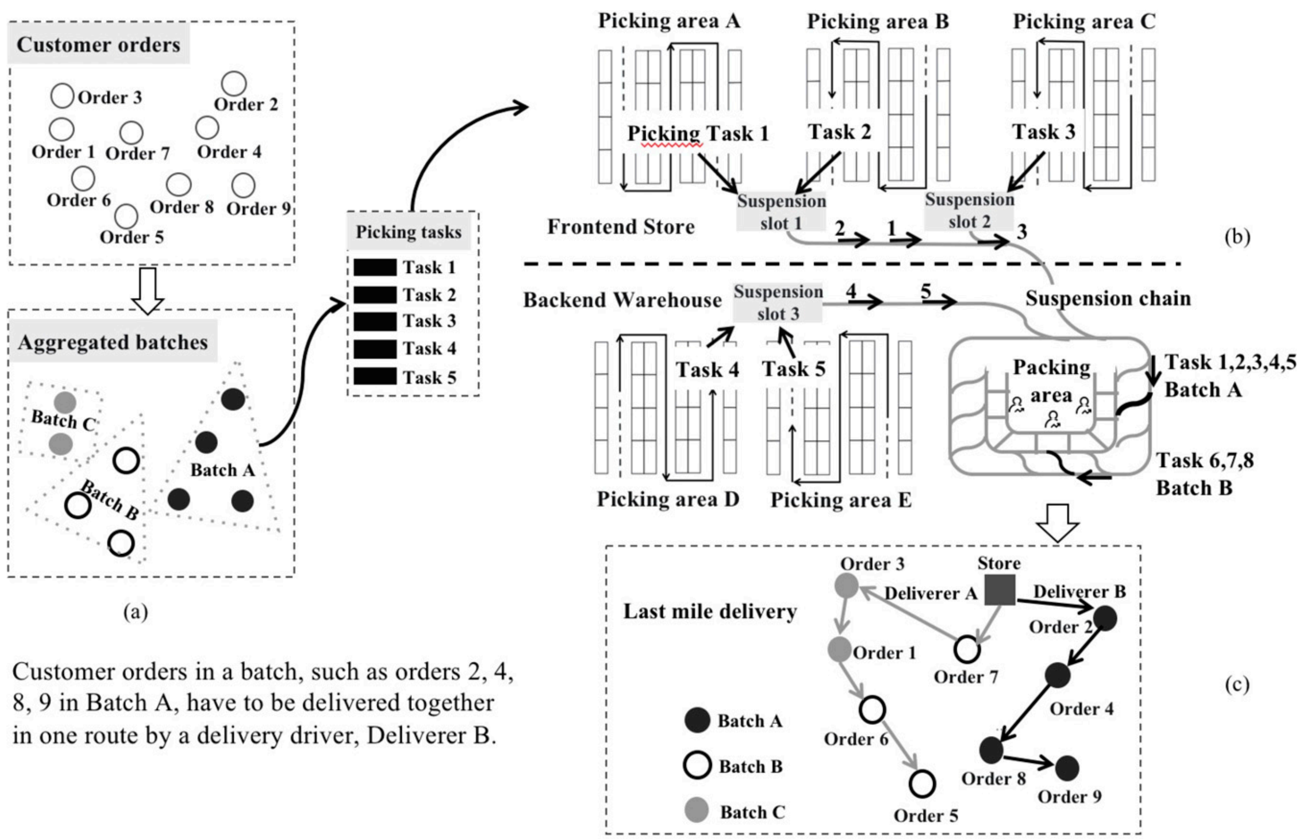
3. The last-mile delivery problem must determine the assignment of batches from the order-batching solution to drivers and the sequence of the orders in these batch assignments to form routes such that a single route is assigned to one driver, as panel (c) in Figure 1 shows.

These optimization problems are difficult to model and solve, and they must be solved online in seconds because orders arrive dynamically. The speed and quality of the solutions directly affect customer satisfaction and delivery efficiency.

Example 2: Open VRP with Pickup and Delivery at Ele.me

Ele.me provides almost instantaneous pickup and delivery services to customers for food catering services,

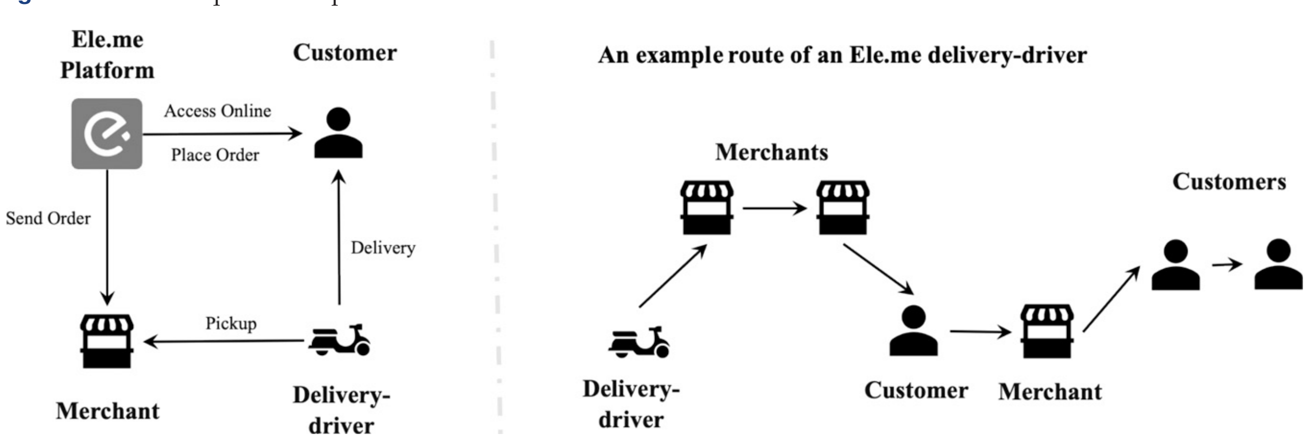
Figure 1. Because of Space Limitations, Order Picking and Last-Mile Delivery at Freshippo Result in Complex VRPs That Must Consider Both Warehouse and Delivery Objectives and Synchronization to Avoid Warehouse Congestion



over-the-counter drugs, flowers, and miscellaneous items. Upon an order's arrival, the merchant, typically a restaurant, is notified immediately to start its preparations; decisions then must be made on assigning a delivery driver and optimizing that driver's route to pick up and deliver the order. The driver may pick up from multiple merchants to satisfy a single customer or pick up from a single merchant to satisfy multiple customers (Figure 2).

The problem is a large-scale, dynamic, open VRP (i.e., a VRP in which the vehicle does not have to return to the depot) with pickup and delivery and due-time requirements. The problem must consider a driver's current orders; prospective future orders; delivery capacity; and multiple objectives such as minimizing the total travel distance (or travel time) to fulfill the orders, providing fairness in the assignment of orders to drivers, avoiding delays to customer delivery times, and

Figure 2. Ele.me Requires the Open VRPPD to Be Solved in Seconds



minimizing the time between when an order is ready (e.g., at a restaurant) and the time it is delivered to the customer to ensure the product's freshness is preserved (Ulmer et al. 2021). The number of pickup and delivery catering orders at Ele.me, which is in an urban city with one of the largest populations in the world, is staggering and is an example of a large industry problem, which must be solved in seconds nearly one and half million times per day.

Example 3: Two-Echelon Stochastic VRP at Taoxianda

Taoxianda, now part of Alibaba's Local Services, provides customers with an option to place online orders for items in brick-and-mortar grocery stores and have them delivered through a two-echelon distribution system (Figure 3). In the top portion of the figure, customer orders are picked up from the stores and then transported by small trucks to intermediate facilities (i.e., drop points). In the bottom portion of the figure, these orders are delivered from these drop points to customers by drivers on motorcycles. Because of limited capacity at drop points, it is often necessary to synchronize the trucks arriving at the drop points with the motorcycles departing from them. A truck arriving at a drop point unloads its orders, which must then be moved quickly to a motorcycle. Highly variable travel times are the norm in urban transportation, and the problem can be modeled as a two-echelon VRP with synchronization and stochastic travel times.

Example 4: Multitrip Heterogeneous VRP with Time Windows at Lazada

Lazada's logistics operations are composed of three stages: first-mile collection, mixed-mode transportation, and last-mile delivery (Figure 4). In the first-mile operation, products are collected from merchants or retail hubs and delivered to sortation hubs, where mixed-mode transportation is used to transport them to the last-mile hubs for final delivery to the customers.

In metropolitan areas in South Asia, small vehicles are preferred because of accessibility limitations on some roads, and the travel times between pickup points are short relative to the planning horizon. As a result, vehicles often return to the depot several times to reload, thus requiring a multitrip VRP (MTVRP). Last-mile delivery operations could involve 100 or more vehicles of various types and as many as 15,000 orders that must be delivered each day via a last-mile hub. MTVRPs are hard to solve, and MTVRPs with heterogeneous vehicles, which include such large numbers of orders and vehicles, are even more difficult to solve; see the example in Francois et al. (2019).

Complexities of the Routing Problems

We encountered three challenges in our search for a solution to these routing problems. Although these problems share similarities, operational constraints resulted in many variants of the VRP and are complicated by uncertainties.

First, we are not solving one specific version of the VRP; we are trying to solve multiple VRPs. These problems include multitrip VRPs at Lazada, dynamic open VRPs with pickup and delivery and with due-time requirements at Ele.me, multiechelon VRPs with synchronization constraints at Taoxianda, and the three dynamic VRPs at Freshippo (e.g., the VRP with both warehouse and delivery objectives and time synchronization to avoid congestion). We also encountered many other VRPs that we do not discuss in this

Figure 3. Taoxianda Requires the Solution of a Two-Echelon Stochastic VRP with the Synchronization of Trucks Arriving and Motorcycles Departing at Drop Points Because of Space Limitations

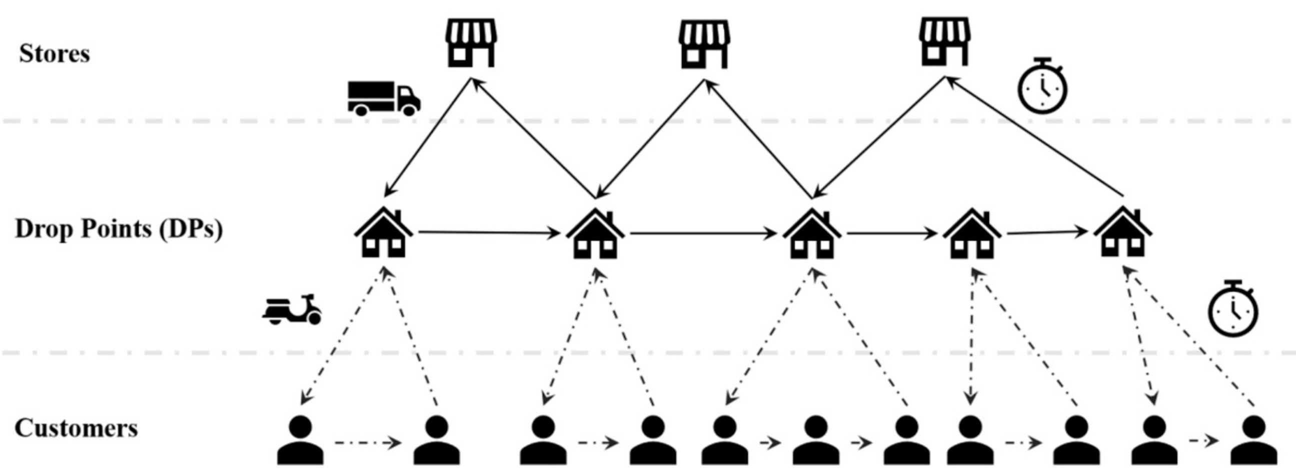
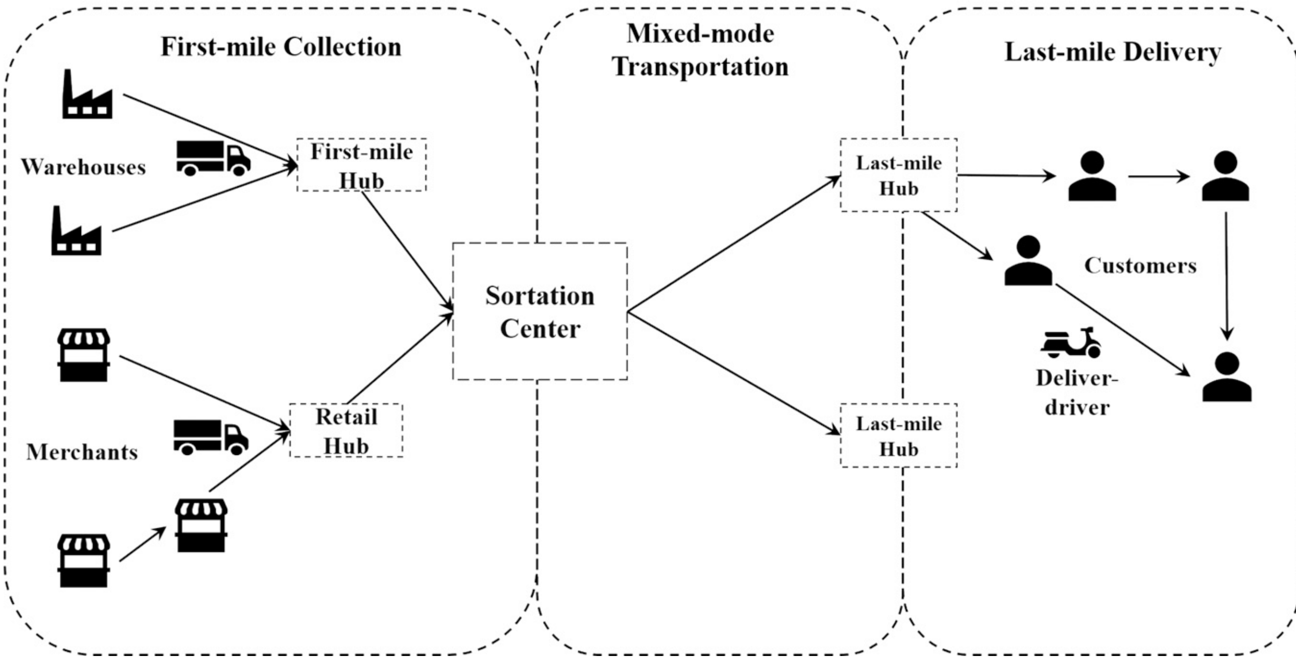


Figure 4. Lazada Logistics Operations Require Large-Scale Multitrip VRPs



paper because of space limitations. Examples include VRPs with driver-customer consistency requirements, which restrict the drivers' routes to customers over a planning horizon; split-delivery VRPs where a customer can be visited by more than one truck; and VRPs with various requirements imposed by businesses, such as geographic compactness and workload balancing.

Second, uncertainties such as highly variable travel times in congested cities, unpredictable service times in high-rise apartments, prospective future demands, variable meal-preparation times, and operational constraints in confined metropolitan warehouses introduce more complexities, some of which are difficult to model and significantly complicate the solutions to these problems.

Third, many of these problems are dynamic online optimization problems, not static offline planning problems, and the stringent expectation for rapid and on-time delivery requires many routing decisions to be made in a short time. We often must solve them in minutes or seconds to allow sufficient time for downstream picking and delivery operations. See Bertsimas et al. (2019) for an example of online routing—a dial-a-ride problem involving taxi drivers.

The variety of VRPs and their sophisticated characteristics required us to develop an open and versatile solution framework that can handle these varieties and provide computationally effective and efficient algorithms capable of solving online VRPs rapidly and solving offline planning problems to near optimality. In Table 2, we summarize these characteristics.

Technical Solutions

Cainiao Network, the logistics arm of Alibaba, began research on routing algorithms in October 2016 when it was attempting to solve the problem of delivering packages to rural villages in China. The resulting algorithm resulted in reducing transportation costs by up to 15%. Since then, as Alibaba's businesses have evolved, the algorithm has evolved on two frontiers: an adaptive large neighborhood search (ALNS) to support the solution of variants of routing problems and deep reinforcement learning (DRL), which trains neural network models offline to make instantaneous online predictions of optimal solutions.

Table 2. Problem Characteristics of the VRPs in the Selected Subsidiaries

Subsidiary	Problem characteristics
Freshippo	Online VRP with synchronization to avoid congestion or with batching restrictions
Ele.me	Online open VRPPD with due times and prospective future orders
Taoxianda	Online two-echelon VRP with synchronization and stochastic travel times
Cainiao Network	Online/offline CVRP with geographic compactness and consistency requirements
Lazada	Offline multitrip VRPTW with heterogeneous vehicles

ALNS

The literature includes many heuristics and exact algorithms to solve VRPs. However, as the size of a problem increases, exact algorithms quickly become intractable, and heuristics and metaheuristics become the method of choice. Although many researchers have their preferred heuristic algorithms, most heuristics are designed on a case-by-case basis and require significant manual tuning; however, this is unacceptable in a rapidly evolving business environment. Hence, in 2016, Alibaba adopted an ALNS framework because of its excellent results on varying real-life routing problems (Ropke and Pisinger 2006; Pisinger and Ropke 2007, 2010; Laporte et al. 2010; Francois et al. 2019).

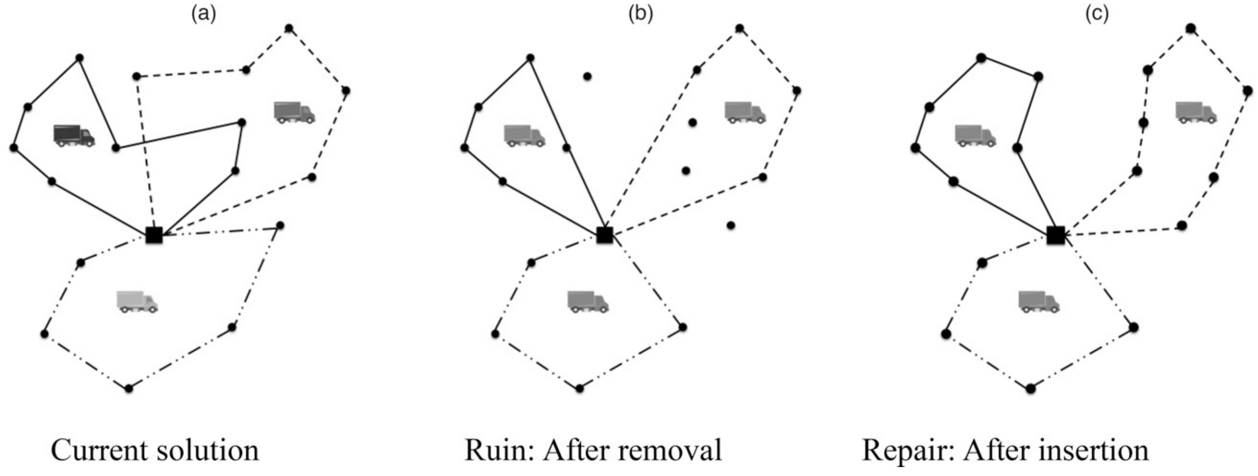
General Framework. The classic ALNS framework is an iterative process during which, at each iteration, part of the solution is destroyed or ruined and then repaired in an attempt to find a better solution (Ropke and Pisinger 2006). In the case of a VRP, given the solution in the example in Figure 5(a), the ruin phase would simply drop or remove some customers from their current routes, as Figure 5(b) illustrates, to form an unassigned customer pool; the repair phase would then insert these customers into routes to form a different solution, as Figure 5(c) illustrates. Ruin-and-repair processes are performed by appropriate heuristics, selected at each iteration from a set of predefined procedures, also known as operators (e.g., via biased random mechanisms, such as roulette wheel selection).

Our ALNS algorithm follows the scheme in Figure 5 but also includes enhancements adopted from the latest developments in modern heuristics. These enhancements include (1) the development of sophisticated ruin-and-repair and pairs of ruin-and-repair schemes, as well as procedures to improve computational effectiveness and to support various VRPs; (2) the development of an adaptive selection of procedures based on specific problem types; (3) the design of diversification and intensification strategies through solution combinations, as implemented in memetic algorithms; and (4) the development of various acceptance methods—examples include greedy acceptance, probability acceptance, and threshold acceptance. Because of space limitations, we present only selected enhancements in what follows.

Ruin-and-Repair Procedures. We believe that local search is the key to a neighborhood search and, in the case of ALNS, the ruin-and-repair procedures (i.e., operators). In addition to common operators, such as worst removal, cluster removal, and route removal in the case of ruin operators, regret insertion, and best-first insertion in the case of repair operators, our algorithm maintains a large pool of ruin-and-repair operators. We designed many of these operators in the process of solving variants of VRPs encountered in our research. For example, in solving the last-mile delivery VRP at Freshippo, we developed a *batch removal* ruin operator where the orders in the batch are removed from a route and *a batch insertion* repair operator that inserts the batch of orders into a route.

ALNS currently has more than 20 ruin operators and 10 repair operators. Because one operator may work well for one problem instance but not for another, we allow ruin-and-repair operators to be configured; therefore, we developed an adaptive strategy to choose operators or pairs of operators that allow application developers to select the ones that best suit the problem instances based on past performance. These operators have greatly improved the computational efficiency in many complex problems that Alibaba encountered.

Figure 5. The Ruin-and-Repair Scheme in an ALNS Iteration Allows Solutions to a Large Variety of VRP Problems



Diversification and Intensification Strategy Through a Memetic Algorithm. Once the search process falls into local optimality, the memetic algorithm is applied by crossing the candidate individuals stored in the parent solution pool through the edge assembly crossover operator to generate new offspring. A similarity measure, defined by a sigmoid-type function and Jaccard index (i.e., a statistic used to understand the similarities between sample sets), is introduced to quantify the similarity between any two solutions. Exploration refers to the selection of parents with low similarity (i.e., solutions that are far distant from each other), whereas exploitation refers to the selection of high-similarity parents; see Zhang et al. (2020) for details.

Computational Results. Our ALNS algorithm, which is based on some of the above-mentioned enhancements, was refactored in 2017 to take advantage of the parallel processing capability of multiprocessor computer chips. We designed this version of the algorithm to challenge the well-known VRP benchmarks, maintained by the Transportation Optimization Portal of SINTEF Applied Mathematics (https://www.sintef.no/projectweb/top/, accessed November 20, 2020). As of November 2020, our ALNS algorithm included 34 of the best-known solutions for VRP with time windows (VRPTW) instances and 23 of the best-known solutions for VRP with pickup and delivery and time windows (VRPPDTW) instances. For example, the benchmark solutions for 400 nodes are shown at https://www.sintef.no/projectweb/top/vrptw/homberger-benchmark/400-customers/ (accessed November 20, 2020) where our solutions are listed under CAINIAO.

Modeling Schema and Adoption. In the middle of 2018, we named the algorithm GreedSolver, and we developed a VRP modeling schema—essentially, a modeling language similar to GAMS or AMPL—to enable programmers to quickly develop their own VRP solutions to the problems they encounter. The language is composed of variables, constraints, and objectives to allow application developers to formulate variants of VRPs and applications. In late 2018, GreedSolver was promoted to all subsidiaries within Alibaba, and their successes thereafter form the basis of this application.

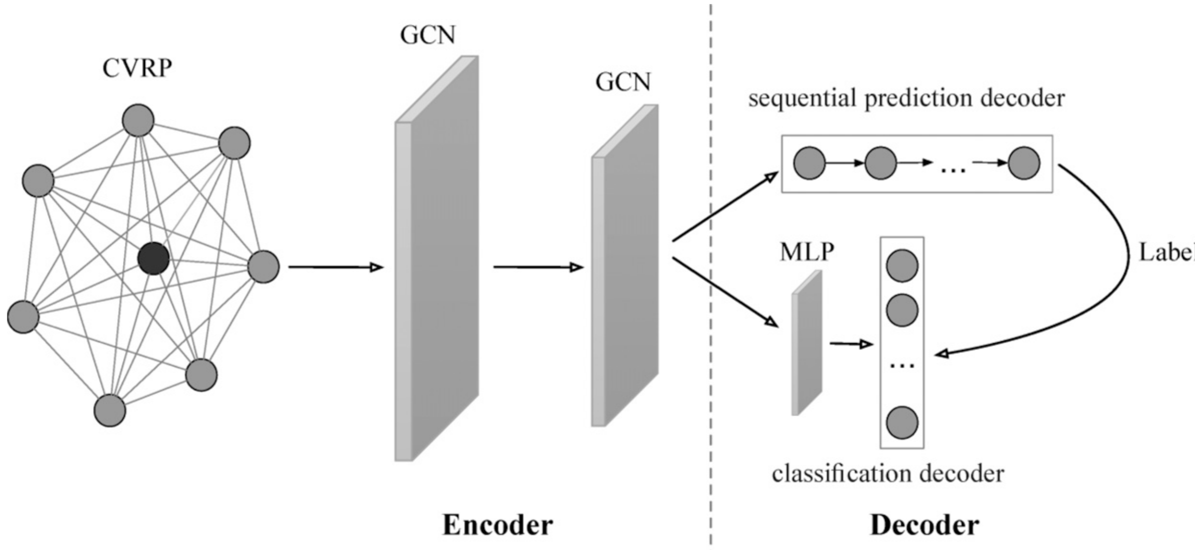
Learning-Based Framework

The ALNS algorithm is flexible; however, for some online optimization problems, sufficient time is not available for an ALNS algorithm to do its full explorations of the solution landscape. As a result, the quality of the solutions suffers, especially in the case of large problems. To resolve this issue, in 2018, we began to evaluate DRL to solve VRPs. DRL has been shown to solve some combinatorial problems (Bello et al. 2016), and we hoped to train DRL models offline with a large number of cases and use the pretrained models online to immediately create online solutions.

DRL Model Development. Several of our initial attempts to develop DRL models were not as successful. Our first attempt was to train a machine learning model using supervised learning with optimal tours, hoping to capture the structures within these tours. This attempt failed apparently because "learning from optimal tours is harder … due to subtle features that the model cannot figure out only by looking at given supervised targets" (Bello et al. 2016, p. 7). Our second attempt was to follow many of the studies of DRL on VRPs, such as those that Nazari et al. (2018), James et al. (2019), and Kool et al. (2018) proposed; however, we soon realized that these models were unable to provide comparable good solutions in many practical cases. The reason we believe this is that in many practical instances, the actual travel distance (or travel time) of edges (between nodes) could dramatically differ from inferred calculations of distance (or time) between nodes, yet these models use only features of nodes (e.g., customer and demands) but not the edge features, which would reflect complexities such as one-way streets or highways. To overcome these problems, more complex models needed to be developed. To this end, we have developed a graph convolutional network (GCN) model with both node and edge features and adopted a joint learning strategy. In early 2020, after much exploration, we were pleased to realize that our DRL model, presented in the following subsection, was able to outperform the ALNS algorithm in both solution quality and computation time.

DRL Models. Our DRL model is based on the GCNs and follows the general encoder-decoder architecture (Figure 6). The encoder (on the left side of the picture, as indicated by the dashed line) takes as input the VRP instances, represented by the node features (coordinates and demand) and the edge features (distances or time between nodes), and it produces the node and edge representations through two GCNs—specifically, a series of aggregation (such as max pooling) and combinations layers. The decoder (see the right side of Figure 6) differs from most existing models in that it incorporates two decoders, a sequential prediction decoder and an edge classification decoder. The sequential prediction decoder, using a recurrent neural network, takes the node representations as input and produces a sequence as the solution for a VRP instance. The classification decoder, using a multilayer perceptron (MLP), takes as input the sequence produced by the other decoder as labels (i.e., solutions from which to learn through supervised learning) and generates the probability matrix that denotes the probabilities

Figure 6. Our Deep Reinforcement Learning Model Differs from Most Models in That It Uses Two Decoders



of the edges being present for vehicle routes. As the method converges, the sequence generated from the sequential prediction decoder and the probability matrix from the classification decoder should be consistent and hopefully approach optimal solutions. For details on the model, please see Duan et al. (2020).

Compared with other learning-based methods used to solve VRPs, our approach explicitly considers both the node and edge features of the networks and the joint learning strategy, which combines reinforcement learning with supervised learning to train the model, and it optimizes simultaneously the parameters of the GCNs to improve the solutions of the sequential prediction decoder and the supervised learning in the classification decoder. Our numerical experiments on real-world data show that this approach can accelerate training convergence and improve the solution quality. It significantly outperforms several well-known algorithms discussed in the literature, especially when the problem size is large, as Tables 3 and 4 illustrate.

Computational Results. Tables 3 and 4 present comparisons of computational results for the DRL model and the ALNS algorithm for the capacitated vehicle routing problem (CVRP) and open VRPPD, respectively. The objective for the CVRP is the total travel distance in the benchmark data set, and the objective for the open VRPPD is a weighted summation of travel distance and the penalty for violation of due times, as specified by Ele.me. The DRL improvement percentage is calculated as

$$100\% \times (\text{the average cost of ALNS} - \text{the average cost of DRL})/\text{the average cost of ALNS}.$$

As Table 3 shows for the CVRP, the most current DRL provided better solutions than ALNS in much less computation time. For example, the DRL-based algorithm is approximately 10,000 times faster than ALNS when solving the CVRP with 100 nodes, and its solution quality is 4.32% better than that of ALNS. As the problem scale increases, DRL's advantages over ALNS also increase. When solving a CVRP with 5,000 nodes, DRL's solution is 40.13% better than the ALNS solution.

To solve an open VRPPD at Ele.me (Table 4), DRL is approximately 37 times faster than ALNS, and its solution quality is 5.50% better than that of ALNS for problems with approximately 15–20 tasks or future locations to visit per driver. Similarly, as the scale of the

Table 3. The DLR-Based Algorithm Outperforms ALNS for Solving the CVRP

		Average cost (travel distance)			Average running time (seconds)	
Number of nodes	Number of problems	DRL	ALNS	DRL improvement (%)	DRL	ALNS
100	10,000	16.60	17.35	4.32	0.0002	2.52
500	2,000	67.10	84.05	20.17	0.0051	12.50
1,000	1,000	128.62	174.12	26.13	0.0221	26.09
2,000	500	243.55	365.28	33.33	0.1506	35.64
5,000	200	608.38	1,016.18	40.13	2.5511	87.47

Table 4. The DRL-Based Algorithm Outperforms ALNS for Solving an Open VRPPD

Number of tasks (i.e., locations to visit per driver)	Number of problems	Average objective value (sum of travel distances and due-time violation penalty)			Average run time (milliseconds)	
		DRL	ALNS	DRL improvement (%)	DRL	ALNS
1–5	293,484	29.44	29.73	0.98	0.152	0.224
5–10	406,698	36.86	38.28	3.71	0.225	1.328
10–15	140,550	44.24	46.77	5.41	0.349	5.384
15–20	13,304	49.08	51.94	5.51	0.681	25.16
20–30	7,018	50.77	57.05	11.01	0.903	42.12

problem increases, DRL's improvement over ALNS also increases.

Our DRL models have been successfully applied to the solution of several variants of the VRP, such as split-delivery VRPs, VRPTWs, and VRPPDTWs, and large-scale VRPs (e.g., up to 5,000 nodes); nevertheless, compared with various local search algorithms, such as ALNS, that have been studied for more than three decades, DRL for VRP is still at an early stage of development. Although computational results for DRL are promising for several VRP variants, considerable development is necessary to solve more complex VRPs.

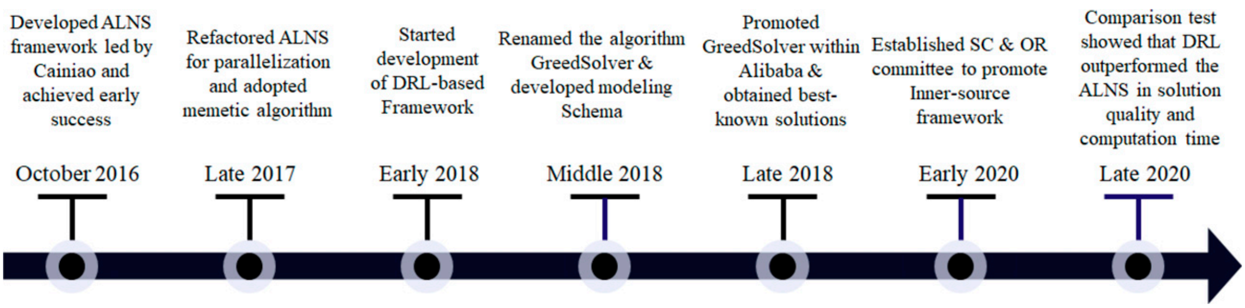
Inner-Source Development Leads to Broad Use

In early 2020, Alibaba formed the Supply Chain and Operations Research (SC&OR) committee to promote the development and application of operation research within Alibaba, and it selected our VRP algorithm as one of the key areas of research. Inner-source software (Stol and Fitzgerald 2015) is similar to open-source software; however, the software is proprietary, and users must reside within the company—in this case, within Alibaba. To expand the logistics operations within Alibaba, the VRP algorithm—specifically, GreedSolver—was inner sourced within the company in August 2020 (DRL is under development and is presently in the process of being inner sourced because of its complexity) and is managed by a committee of operations research developers from Cainiao Network, Freshippo, Taoxianda, Lazada, and Ele.me.

Alibaba's inner-source community, supervised by the SC&OR committee, encourages sharing among the company's business subsidiaries to avoid the necessity for the time-consuming and error-prone reimplementation of algorithms. This approach has facilitated continuous model and algorithm improvement throughout Alibaba. The open architecture allows the extension of the algorithm to solve new variants of the VRP and has, in turn, enriched the inner-source algorithm repository, as we discuss in the *Portability to Related Problems* section. In Figure 7, we summarize the history and evolvement of the algorithms.

To implement the algorithm to address varying business scenarios, we developed a user-friendly graphical interface and a customization-friendly schema to allow developers to solve routing problems. The algorithm can be invoked either by an application programming interface (API) service or by a Java package. The API service is a standard in-house service application running on Alibaba's cloud service. A developer can call the API using specific cases as input, and the API will return a solution. The Java package provides in-depth functionalities and direct integration into individual applications, allowing algorithm developers to manipulate the internal classes and methods to customize various components of the model, such as decision variables, objective functions, and constraints, and to make modifications to the algorithm (e.g., neighborhood definition, seed solution construction, adaptive scoring).

Figure 7. The Timeline of Our Algorithm Development Spans More than Five Years



Portability to Related Problems

Alibaba has developed several practical models since our VRP algorithm (GreedSolver) was inner sourced in August 2020 and has used these algorithms to solve complex problems, including stochastic VRPs with stochastic travel times, VRPs with geographic compactness requirements, and VRPs with driver-customer consistency requirements. ALNS provides an extremely fast prototype capability to model these variants, as we illustrate in the following two examples.

VRP with Stochastic Travel and Service Times

Stochastic travel or service times can be incorporated into a VRP commonly through either (1) the addition of the constraint, referred to as a chance constraint, which ensures that the probability of having a route duration in excess of a specific time does not exceed a given threshold, or (2) the change of objective function to include an expected penalty cost associated with duration that does not exceed a specific criterion. Algorithms for either version of the stochastic VRP can be developed easily based on modeling schema in GreedSolver. In our implementation, we modified the objective function to include the expected penalty costs of multiple scenarios and designed special ruin and insertion operators to target routes with time violations to the chance constraints to speed up the computation. The algorithm *to incorporate stochastic travel and service times was fully functional within only two weeks from the beginning of the algorithm design and is being shared with developers within Alibaba through our inner-source community.*

Incorporating uncertainty can significantly improve the performance of routing plans. To prove this, we generated two solutions, one using a deterministic VRP with mean travel time and the other using a stochastic VRP with chance constraints. We set a total travel time (e.g., 75 minutes) for each vehicle. In the stochastic VRP, we set the chance constraint to ensure that the probability of overtime delivery for each trip was less than 10%. These solutions were implemented in practice on a different, yet similar, set of days. Computational results showed that exceeding the vehicle time limitation (i.e., overtime cost) could be significantly reduced from 15% using a deterministic VRP to 6.7% using a stochastic VRP, and the total cost could be reduced by as much as 10% because of reduced penalty (i.e., overtime) costs.

Geographic Compactness

In practical routing problems, implementing solutions where routes are geographically compact and do not overlap is usually preferable. These solutions visually appeal to human operators and quickly earn their trust. For details on examples of such applications, see Kant et al. (2008) and Hollis and Green (2012).

Although the exact definition of geographic compactness is difficult to state, studies suggest that such a subjective concept of a route exhibits certain characteristics such as compactness of the routes, minimum overlapping between the convex hulls of the routes, and reductions in the number of crossings between routes. For details, see Rossit et al. (2019). These characteristics can be modeled easily through the ALNS framework either by designing problem-specific operators to avoid route overlaps or by incorporating additional objective terms or measures. We adopted both approaches in our implementation in approximately one week by incorporating them into the ALNS framework.

Figure 8 shows the design of the solutions to (a) the traditional CVRP and (b) a CVRP with geographic compactness for the largest instances, A-n80-k10 from "set A" of the CVRP benchmarks (NEO Research Group 2013). As Figure 8 illustrates, the solution with geographic compactness is visually appealing and simpler to understand and execute, and drivers and operational management usually prefer it.

Summary of Framework Portability

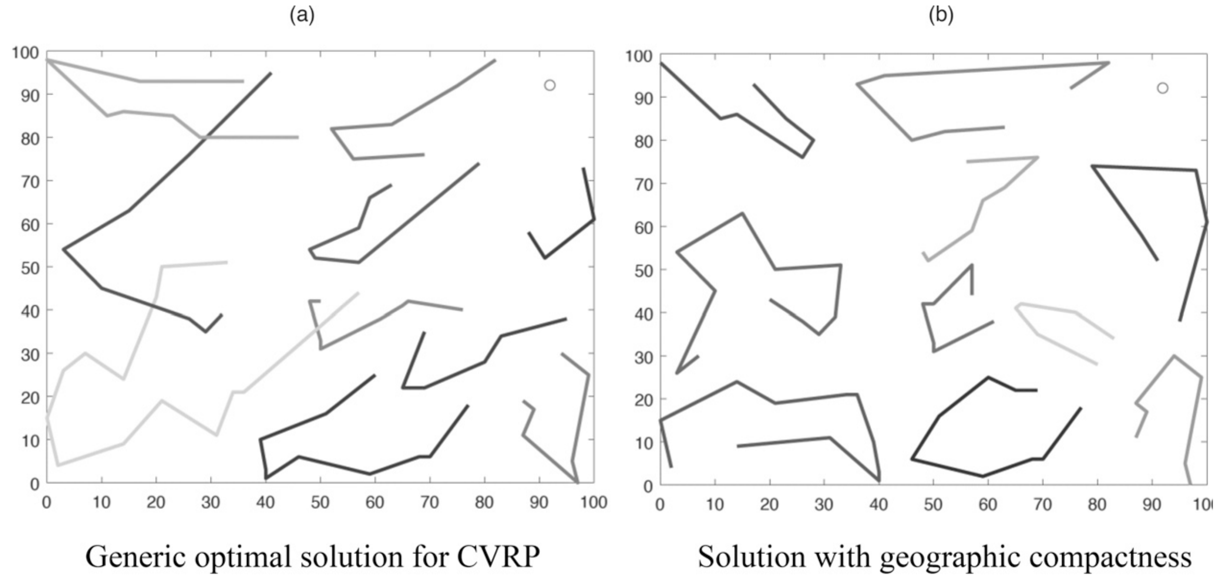
The above-mentioned extensions demonstrate the portability of the ALNS algorithm (and to some extent, the DRL framework) and their applications to similar scenarios in the supply chain and logistics business. We summarize the portability of the algorithms in what follows.

First, the ALNS algorithm is designed to solve a large variety of VRPs through (1) an extendable domain-specific language that facilitates user-defined constraints and objectives and (2) an inner-source framework, which allows the customization of neighborhood search operators and evaluation functions that suit specific business practices. These characteristics of the algorithm have made it widely accepted for solving a variety of VRPs.

Second, the DRL framework provides a set of data-driven, self-learning algorithms that can generate offline optimization training to derive models that are used for online real-time decision making. When we applied these algorithms to selected online routing problems, they yielded high-quality solutions in milliseconds. We believe such a methodology can apply to a set of online optimization problems.

The development of domain-specific language for VRPs has significantly reduced the development time of our algorithms. To implement a new scenario, developers may use the modeling schema to model their own problem—that is, define decision variables, constraints, and objective functions—which they can send to GreedSolver to find desirable solutions. Multiple

Figure 8. The Solution of (b) a CVRP with Geographic Compactness Is Visually Appealing Compared with (a) the Solution of a CVRP with a Generic Solution



predefined ruin-and-repair operators can be configured and problem-specific operators can be customized to expedite the time required to find high-quality solutions. Finally, search strategies can be tuned and stopping criteria modified to meet the specific needs of customers.

Benefits

The models and algorithms have been implemented to solve various VRPs in Freshippo, Taoxianda, Ele.me, Cainiao Network in China, and Lazada. They are embedded in many Alibaba warehouses or logistics systems, run millions of route optimizations per day, and have produced immediate and significant financial results. In the following subsections, we present a summary of the implementation and financial savings.

Freshippo

Since 2017, Freshippo has applied the ALNS algorithms in the solution of its three VRPs: the order-batching problem, the picker routing problem, and the last-mile delivery problem. These problems, as we described previously, are complex VRPs with both warehouse and delivery objectives, VRPs with synchronization to avoid congestions, or VRPs with order-batching restrictions as a result of warehouse operational constraints. The problem cases, which are solved in parallel on multithread machines (one store per thread), involve approximately 100–400 customer orders, and the solution time limit is four seconds. The algorithm is invoked almost half a million times per day. Our numerical experiments show that the algorithms can achieve an average optimality of 90% compared with the best-known results obtained offline with typical orders of magnitude of computation time.

Ele.me

The routing problem in Ele.me is an open VRPPD with heterogeneous vehicles and capacity constraints. The problem is challenging in peak hours such as lunchtime and dinner time when about 12 million orders could be assigned to 0.5 million active drivers. We divide the orders into small batches; each batch typically has 10–40 orders (including pickup and delivery points). The problem in each is an open VRP with pickup and delivery and due-time requirements, and the batches are solved in parallel. The time allowed for one batch is 0.5 seconds. In our offline testing, the ALNS algorithm can achieve 80% optimality (compared with exact solutions obtained by using mixed-integer programming) for small problems within the required response time. However, as we show in Table 4, the DRL algorithm outperforms the ALNS algorithm by achieving better performance in a shorter time, and plans are in place to upgrade the algorithm from ALNS to DRL.

Cainiao Network

We first applied the ALNS algorithm to solve the order-batching problem in Cainiao's warehouse operations (a large-scale CVRP that must be solved in seconds). The typical size of a problem is about 1,000

Table 5. Summary of the Benefits Each Business Subsidiary Accrued

Business	Year of launch	Current annual financial savings ($ million)	Savings on transportation costs (%)	Expected annual financial savings ($ million)
Freshippo	2017	29	5.3	33
Ele.me	2020	9	5.0	11
Taoxianda	2019	3	2.0	9
Cainiao Network	2017	7	5.1	7
Lazada	2020	7	3.9	9
Total		55	4.3[a]	69

[a]Average.

nodes once a request for order batching has been triggered. Because of its large size, we shifted to the DRL algorithm to solve these problems because it can handle as many as 5,000 nodes in seconds. We usually train the neural network model once a month using the current month's data. Although the offline training time lasts several days, the trained model can produce online solutions within milliseconds, thus saving time for order warehouse operations.

Taoxianda

The problem solved for Taoxianda's four-hour delivery service is a stochastic two-echelon VRP model. We developed a stochastic version on top of the ALNS framework to incorporate the uncertainties of travel and service times where the objective functions to minimize the total cost—a summation of travel distance and expected penalty costs with respect to the duration in excess of vehicle travel time limitations. The algorithm was launched in late 2019 and is invoked daily to generate routes for trucks and drivers. Taoxianda has approximately 3,000 vehicles and employs 15,000 drivers per day.

Lazada

The VRP problem we solved is a multitrip VRPTW with heterogeneous vehicles. We launched the algorithm in January 2020 and achieved great success. The VRP, which runs daily and generates solutions within 15 minutes, handles about 2 million orders and 12,000 vehicles per day. We extended the ALNS framework with new ruin operators that produce shorter routes. We then developed a bin-packing approach to progressively assemble them into feasible working hours before evaluating the objective value. From the flexible ALNS framework, within one week, we were able to develop the new multiple-trip feature to handle the delivery work in Lazada's first-mile collection operations.

The total annual financial savings for all of the above-mentioned business subsidiaries of Alibaba are approximately $55 million. We believe that the expected annual savings can reach about $69 million, assuming business growth and broader adoption of the solver. These savings are calculated as the summation of cost reduction through field experiments in all subsidiaries. These financial savings are shown in Table 5, and details of these calculations are shown in supplemental materials that are available from the authors upon request.

In addition to the above-mentioned cost-saving benefits, the algorithm has provided several other intangible benefits. It significantly freed up time for the algorithm developers and became the key to the success of the business subsidiaries discussed in this paper. By providing quick online decisions, the algorithms have allowed sufficient time for offline order picking and delivery operations, which has maintained—and, in many cases, improved—on-time delivery. Improved on-time and fast delivery services at even lower costs are critical to the success and development of these businesses in a competitive market. In addition, decreasing the number of miles vehicles travel reduces carbon dioxide emissions, which significantly and positively impacts China's environmental objectives.

Summary

Alibaba has pioneered a new retail model that enables it to deliver e-commerce and grocery products to customers through its logistics networks in a period of between 30 minutes and 4 hours after order placement. The stringent service commitment requires timely online decisions to be made immediately after orders are received. To meet this objective, Alibaba developed vehicle routing algorithms to optimize warehouse order-picking operations and generate delivery routes for its drivers. One algorithm framework is built on an adaptive large neighborhood search that can model various scenarios by adding customized move operators, constraints, and objectives. Another framework is based on the deep reinforcement learning-based approach, which trains neural network models offline to make almost instantaneous online generations of solutions. Our algorithms have been implemented to solve various VRPs in five Alibaba subsidiaries and generated $55 million in annual financial savings.

The methodology can be easily applied to solve more static routing and dynamic online routing problems in the logistics service industry and public sectors. The success of these algorithms has fermented an inner-source community of operations researchers within Alibaba, increased the confidence of the company's executives toward operations research, and established optimization as one of the core components of Alibaba's supply chain and logistics operations. As Daniel Zhang, chairman and chief executive officer of Alibaba Group, stated in the Edelman competition video, "Supply chain and operations research is very important to our e-commerce business, and Alibaba will continue to pursue innovations that will help the digital transformation of enterprises across the sector" (Zhang 2021).

References

Alibaba Group (2020) Alibaba group anounces September quarter 2020 results. Retrieved https://www.alibabagroup.com/en/news/press_pdf/p201105.pdf.

Bello I, Pham H, Le QV, Norouzi M, Bengio S (2016) Neural combinatorial optimization with reinforcement learning. Preprint, submitted November 29, https://arxiv.org/abs/1611.09940.

Bertsimas D, Jaillet P, Sebastien M (2019) Online vehicle routing: The edge of optimization in large-scale applications. *Oper. Res.* 67(1):143–162.

Bertsimas D, Delarue A, Eger W, Hanlon J, Martin S (2020) Bus routing optimization helps Boston public schools design better policies. *INFORMS J. Appl. Anal.* 50(1):37–49.

Duan L, Zhan Y, Hu H, Gong Y, Wei J, Zhang X, Xu Y (2020) Efficiently solving the practical vehicle routing problem: A novel joint learning approach. *Proc. 26th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (Association for Computing Machinery, New York), 3054–3063.

Francois V, Arda Y, Crama Y (2019) Adaptive large neighborhood search for multi-trip vehicle routing with time windows. *Transportation Sci.* 53(6):1706–1720.

Holland C, Levis J, Nuggehalli R, Santilli B, Winters J (2017) UPS optimizes delivery routes. *Interfaces* 47(1):8–23.

Hollis BL, Green PJ (2012) Real-life vehicle routing with time windows for visual attractiveness and operational robustness. *Asia-Pacific J. Oper. Res.* 29(4):1250017.

James JQ, Yu W, Gu J (2019) Online vehicle routing with neural combinatorial optimization and deep reinforcement learning. *IEEE Trans. Intelligent Transportation Systems* 20(10):3806–3817.

Kant G, Jacks M, Aantjes C (2008) Coca-Cola enterprises optimizes vehicle routes for efficient product delivery. *Interfaces* 38(1):40–50.

Kool W, Van Hoof H, Welling M (2018) Attention, learn to solve routing problems! Preprint, submitted March 22, https://arxiv.org/abs/1803.08475.

Laporte G, Musmanno R, Vocaturo F (2010) An adaptive large neighborhood search heuristic for the capacitated vehicle arc-routing problem with stochastic demands. *Transportation Sci.* 44(1):125–135.

Nazari M, Oroojlooy A, Takáč M, Snyder L (2018) Reinforcement learning for solving the vehicle routing problem. Bengio S, Wallach HM, Larochelle H, Grauman K, Cesa-Bianchi N, eds. *Proc. 32nd Conf. Neural Inform. Processing Systems* (Curran Associates, Red Hook, NY), 9861–9871.

NEO Research Group (2013) Capacitated VRP instances. Accessed July 7, 2020, https://neo.lcc.uma.es/vrp/vrp-instances/capacitated-vrp-instances/.

Pisinger D, Ropke S (2007) A general heuristic for vehicle routing problems. *Comput. Oper. Res.* 34(8):2403–2435.

Pisinger D, Ropke S (2010) Large neighborhood search. Gendreau M, Potvin J-Y, eds. *Handbook of Metaheuristics, International Series in Operations Research and Management Science*, Vol. 146 (Springer, Boston), 399–419.

Ropke S, Pisinger D (2006) An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transportation Sci.* 40(4):455–472.

Rossit DG, Vigo D, Tohme F, Frutos M (2019) Visual attractiveness in routing problems: A review. *Comput. Oper. Res.* 103(March): 13–34.

Sahoo S, Kim S, Kim BI, Kraas B, Popov A Jr (2005) Routing optimization for waste management. *Interfaces* 35(1):24–36.

Stol KJ, Fitzgerald B (2015) Inner source—Adopting open source development practices in organizations: A tutorial. *IEEE Software* 32(4):60–67.

Ulmer MW, Thomas BW, Campbell AM, Woyak N (2021) The restaurant meal delivery problem: Dynamic pickup and delivery with deadlines and random ready time. *Transportation Sci.* 55(1):75–100.

Zhang D (2021) 2021 Edelman finalist: Alibaba. INFORMS, April 13, 2021. YouTube video, 39:12, https://www.youtube.com/watch?v=dOBvgzHoQb8.

Zhang Y, Zhang D, Wang L, He Z, Hu H (2020) Balancing exploration and exploitation in the memetic algorithm via a switching mechanism for the large-scale VRPTW. Nah Y, Cui B, Lee SW, Yu JX, Moon YS, Whang SE, eds. *Database Systems for Advanced Applications, Lecture Notes in Computer Science*, Vol. 12112 (Springer, Cham, Switzerland), 324–331.

Haoyuan Hu is a director at Alibaba Group, leading the Artificial Intelligence Department of Cainiao Network. Under his leadership, the team developed the GreedSolver, which is broadly used to solve various routing problems in Alibaba. His research interests include the integration of operation research and machine learning, large-scale optimization in logistics, financial engineering, and server scheduling.

Ying Zhang is a senior algorithm expert in the Artificial Intelligence Department of Cainiao Network. His research interests include supply chain design and logistics optimization, especially routing algorithms. He obtained a PhD degree from the Department of Industrial Engineering, Tsinghua University.

Jiangwen Wei is a senior engineer at Cainiao Network. He leads the research and development of a learning-based framework to solve vehicle routing problems. He obtained a bachelor's degree in mathematics from Shanghai University.

Yang Zhan is an algorithm engineer in the Artificial Intelligence Department of Cainiao Network. Her research interests include stochastic optimization, reinforcement learning, and their applications in logistics and service operations management. She graduated from the joint PhD program of Shanghai Jiao Tong University and the University of Michigan, Ann Arbor. She was a visiting scholar at the University of Minnesota, Twin Cities.

Xinhui Zhang is a senior director and head of supply chain and operations research committee at Alibaba Group. As a tenured professor at Wright State University, he was a finalist for the 2013 Franz Edelman Award. His research interests include large-scale optimization, inventory management, vehicle routing, and various aspects of supply chain management. He obtained his PhD in the Operations Research

and Industrial Engineering program at the University of Texas at Austin.

Shaojian Huang is a senior algorithm expert leading the logistics algorithm team in the Freshippo Division at Alibaba Group. His research interests include warehouse operations and logistics optimization—specifically, warehouse design and optimization, line-haul optimization, and last-mile delivery optimization. He obtained his master's degree in computer science from Huazhong University of Science and Technology.

Guangrui Ma is a staff research engineer in the Digital Supply Chain Division at Alibaba Group. Prior to Alibaba, he was a faculty member in the College of Management and Economics, Tianjin University. His research interests include operations research, machine learning, supply chain and logistics systems, and retailing operations. He obtained his PhD degree from the Hong Kong University of Science and Technology.

Yuming Deng is a director at Alibaba Group, leading the digital supply chain department. Prior to Alibaba, he was a supply chain optimization expert for e-commerce at Amazon. His research interests are assortment planning, network planning, pricing strategy, forecasting, inventory optimization, and simulation-based optimization and its applications. He obtained his PhD degree in the Operations Research and Industrial Engineering program at the University of Texas at Austin.

Siwei Jiang is the senior vice president and team leader for the logistics algorithm department in Lazada. His research interests include multiagent evolutionary algorithms, vehicle routing problems, workforce scheduling and rostering, and logistics optimization. He obtained his PhD degree in computer science from the School of Computer Engineering, Nanyang Technological University, Singapore, in 2014.