IRSTI 73.01.11

Yernar Akhmetbek[1], Taukekhan Mustakhov[2]
[1,2]Suleyman Demirel University, Kaskelen, Kazakhstan

STOCHASTIC DYNAMIC VEHICLE ROUTING PROBLEM SURVEY

Abstract. The present article aims to offer an exhaustive and in-depth investigation of the Stochastic Dynamic Vehicle Routing Problem, which remains a significant challenge in the field of transportation logistics. To achieve this objective, we will undertake a meticulous analysis of the latest cutting-edge techniques and methodologies deployed to tackle this complex optimization problem. Furthermore, we will delve into the intricate and multifaceted stochastic and dynamic constraints that pose formidable obstacles to effective route planning and optimization. Through this survey paper, we seek to provide a comprehensive understanding of the current state-of-the-art approaches and highlight the potential avenues for future research in this critical area of transportation logistics. In addition, we will also analyze the application of advanced reinforcement learning methods and Markov decision processes to solve the problem of stochastic dynamic vehicle routing.

Keywords: transportation, combinatorial optimization, stochastic dynamic vehicle routing problem.

***

Аңдатпа. Бұл мақала көлік логистикасы саласындағы маңызды мәселе болып қала беретін стохастикалық динамикалық көлікті бағыттау мәселесін жан-жақты және терең зерттеуді ұсынуға бағытталған. Осы мақсатқа жету үшін біз осы күрделі оңтайландыру мәселесін шешу үшін қолданылатын соңғы озық әдістер мен әдістемелерге мұқият талдау жасаймыз. Бұдан басқа, біз маршрутты тиімді жоспарлау мен оңтайландыруға үлкен кедергілер тудыратын күрделі және көп қырлы стохастикалық және динамикалық шектеулерді қарастырамыз. Осы сауалнама арқылы біз қазіргі заманғы тәсілдер туралы жан-жақты түсінік беруге және көлік логистикасының осы маңызды саласындағы болашақ зерттеулердің әлеуетті жолдарын көрсетуге тырысамыз. Сонымен қатар, біз көлік құралын стохастикалық динамикалық бағыттау мәселесін шешу үшін кеңейтілген оқыту әдістерін және Марков шешім процестерін қолдануды талдаймыз.

Түйін сөздер: тасымалдау, комбинаторлық оңтайландыру, кездейсоқ динамикалық көлікті бағыттау мәселесі.

***

Аннотация. Цель данной статьи заключается в представлении комплексного и подробного исследования проблемы стохастической динамической маршрутизации транспортных средств. В работе будет произведен тщательный анализ современных методов и технологий, используемых для решения данной сложной задачи оптимизации. Особое внимание уделено изучению стохастических и динамических ограничений, которые создают значительные препятствия для эффективного планирования и оптимизации маршрутов транспортных средств. В результате данного обзора мы стремимся предоставить читателям всестороннее понимание современных методов решения этой проблемы и выделить потенциальные направления будущих исследований в данной важной области логистики транспорта. Помимо этого, мы также проанализируем применение передовых методов обучения с подкреплением и марковских процессов принятия решений для решения проблемы стохастической динамической маршрутизации транспортных средств.

Ключевые слова: транспорт, комбинаторная оптимизация, стохастическая динамическая задача маршрутизации транспортных средств.

Introduction

The pace of technological development is accelerating at an unprecedented rate, driven by a growing demand from consumers for products and services that meet their needs. In contrast to the past, when individuals would physically visit supermarkets to procure their required items, a significant proportion of the population now rely on online platforms to make purchases. Consequently, online orders are promptly fulfilled and delivered by courier services to their respective locations, this has changed the way people get the things they need.

The challenge of efficiently accommodating a high volume of customer orders has led to a pressing issue for many delivery companies. Although fulfilling orders for a limited number of clients within a set timeframe is relatively simple, scaling up to satisfy hundreds of customers while being constrained by the availability of only 3-4 vehicles becomes an arduous task. This complex problem is typically classified in computer science as a Vehicle Routing Problem, where the objective is to determine the optimal delivery route that can satisfy the customer's requests with a limited vehicle capacity load, within a specific timeframe. To complicate matters further, there may be additional sub-problems that can arise in such a scenario, such as traffic jams, deadlines for customer delivery, and new customer requests while delivering, which can all impact the solution. To clarify, the additional sub-problems that arise in the Vehicle Routing Problem can be described as a complication and are taken into account in the Stochastic Dynamic Vehicle Routing Problem (SDVRP), an extension of VRP.

One of the challenging combinatorial optimization problems is called the Stochastic Dynamic Vehicle Routing Problem (SDVRP) which requires the solver to find an optimal path of routes to serve customers within a short timeframe or at a minimum cost, depending on the problem's objective. This problem becomes even more challenging when new constraints are introduced, such as the use of multiple vehicles to serve customers or the incorporation of time windows for each customer. Moreover, vehicle availability and the possibility of breakdown while serving the customer add further complexity to the problem. Due to its complexity, the SDVRP falls under the class of NP-hard problems, which are impossible for existing algorithms to solve in a polynomial amount of time. Instead, the time required to tackle these issues increases dramatically as the input data amount increases. Thus, the primary objective of SDVRP researchers is to find solutions that can come as close as possible to solving the problem efficiently.

At present, there is no definitive state-of-the-art solution for the Stochastic Dynamic Vehicle Routing Problem (SDVRP). However, researchers have explored numerous methods, strategies, and techniques to address this challenging optimization problem. These include heuristic methods, exact methods, and reinforcement learning methods, among others. Computer scientists have introduced various approaches, tips, tricks, and formulas to solve the SDVRP. Our objective is to evaluate and compare these methods and analyze their effectiveness to determine the most viable ways of solving the SDVRP.

We shall discuss each of these sections separately in the following paragraphs. Section II introduces the subject and provides a literature evaluation. We shall pause on each approach to solving in Section III and provide succinct justifications for each. The best articles on our topic will be discussed and evaluated in Section IV. The conclusion on the study topic is then presented in section V.

Background

In this section, we will talk about the SDVRP problem in a more detailed way by giving additional information about the constraint and the variety of the problem. Also, we will provide a review of other works related to this topic.

Combinatorial Optimization

Finding the optimal option from a limited number of options falls within the research area of combinatorial optimization. This field has a wide range of practical applications, including logistics, transportation, manufacturing, and scheduling, among others.

The Traveling Salesman Problem is one of the most well-known combinatorial optimization issues (TSP). In order to complete the TSP, the shortest path that visits each of the supplied cities precisely once and returns to the beginning city must be found. The TSP is categorized as an NP-hard problem, which indicates that, for large problem instances, computing the exact solution to the issue is computationally intractable.

The Vehicle Routing Problem is another conjectural optimization issue (VRP). The VRP entails figuring out the best delivery routes for a fleet of vehicles to service a group of consumers while reducing expenditures like trip time, fuel use, and vehicle capacity. A further NP-hard issue is the VRP.

Stochastic Dynamic Vehicle Routing Problem

The Stochastic Dynamic Vehicle Routing Problem (SDVRP) is a combinatorial optimization problem that incorporates dynamic and stochastic elements into the traditional VRP. The SDVRP is a complex and challenging problem that requires new solution approaches to tackle its unique characteristics.

The SDVRP extends the traditional VRP by considering the uncertainty in travel times and demand, which can vary over time and across different regions. This uncertainty arises due to factors such as traffic congestion, road closures, and unexpected demand surges. As a result, the optimal solution to the SDVRP is not fixed but rather evolves as new information becomes available. This requires dynamic routing decisions that can adapt to changing conditions and stochastic elements to account for the uncertainty.

Solving the SDVRP is a challenging task, and traditional optimization techniques often fail to provide effective solutions. Researchers have developed various approaches to address the dynamic and stochastic nature of the SDVRP, including metaheuristics, simulation-based optimization, and robust optimization. These methods aim to find robust solutions to uncertain conditions and can adapt to changing circumstances in real time.

Among the various stochastic constraints inherent in the Stochastic Dynamic Vehicle Routing Problem (SDVRP), customer requests represent a particularly significant challenge. Vehicles must dynamically determine which customer requests to serve, based on which requests are most profitable and feasible to complete. The example presented in Figure 1 illustrates an instance of the SDVRP with stochastic customer requests. The black circle represents the depot, the white nodes denote known requests, and the yellow nodes depict requests that have not yet materialized and will appear along the route.

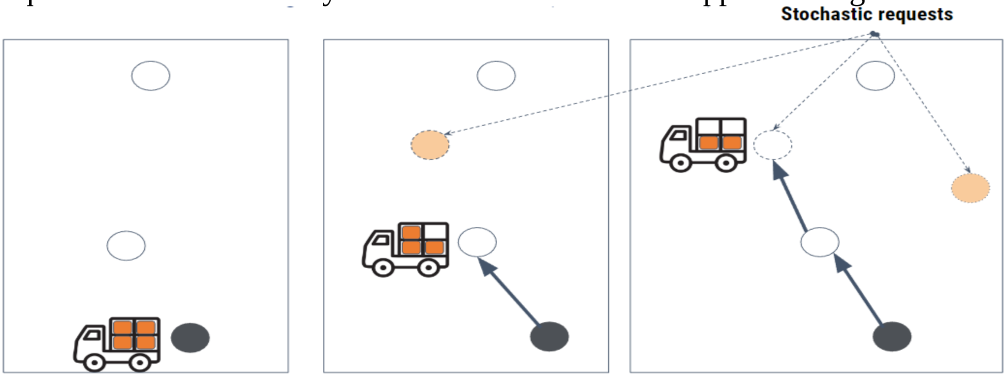

Fig. 1: SDVRP example

In the "Nature of Stochasticity" section, we will thoroughly examine each of the stochastic constraints that arise in the context of the Stochastic Dynamic Vehicle Routing Problem (SDVRP).

Markov decision process

A mathematical framework called a Markov Decision Process (MDP) is used to simulate decision-making in stochastic settings. It is a useful tool for solving the SDVRP because it allows us to model the dynamic and stochastic nature of the problem.

In an SDVRP MDP, the problem is modeled as a sequence of states, actions, and rewards. Each state represents the current situation of the system, each action represents a decision to be made, and each reward represents the immediate feedback for taking a certain action. In order to maximize the predicted cumulative reward over time, an optimum policy must be found, which is a mapping between states to actions.

The SDVRP problem is particularly well-suited for MDP modeling because of its dynamic and stochastic nature. The problem involves a changing set of customers with varying demands, and the optimal route must be updated in real-time to respond to new information. The MDP framework allows us to model this dynamic and stochastic nature and optimize the routing decisions accordingly.

Marlin W. Ulmer has conducted a comparative analysis of Markov Decision Process (MDP) approaches for solving the Stochastic Dynamic Vehicle Routing Problem (SDVRP). In [1], the author compared route-based MDP with conventional MDP and concluded that the route-based MDP is more effective in solving SDVRP than the conventional approach. The route-based MDP considers the sequence of customers to be visited on a route, and the author showed that this approach resulted in better solutions and reduced computation time compared to the conventional MDP approach. The study highlights the usefulness of MDP techniques in solving SDVRP and provides insights into the choice of MDP formulation for this problem.

Example

In summary, a mathematical description of each problem is necessary, and authors should determine the appropriate format for their respective analyses. As an example, Marlin W. Ulmer [2] endeavored to address an SDVRP problem with stochastic requests, formulating the relevant mathematical equations as depicted in Table 1 below.

Notation	Description
S_k	Decision state
t_k	Point of time
P_k	Position of the vehicle
C	Set of customers to serve
θ_k	Planned tour
C_k^{new}	New requests
x	Decision
$C_k^{x,\text{assign}}$	Set of newly assigned customers
C_k^x	Updated set of customers to serve
θ_k^x	Updated planned tour
$R(S_k, x)$	Reward of decision x in state S_k
S_k^x	Post-decision state
ω_k	Transition

Table 1: Notation of the Markov decision process[2]

Policy function: The policy function is defined as a means of maximizing the reward gained by the agent, as indicated by Equation (1).

$$\pi^* = \underset{\pi \in \Pi}{\operatorname{argmax}} \mathbb{E}\left[\sum_{k=0}^{K} R(S_k, X_k^{\pi}(S_k)) | S_0\right]. \tag{1}$$

Bellman equation: The Bellman equation function serves to determine which policy function is most likely to yield the most profitable rewards over all epochs as depicted in Equation(2).

$$X_k^{\pi^*}(S_k) = \underset{x \in X(S_k)}{\operatorname{argmax}} \left\{ R(S_k, x) + \mathbb{E}\left[\sum_{j=k+1}^{K} R(S_j, X_j^{\pi^*}(S_j)) | S_k\right] \right\}. \tag{2}$$

Nature of stochasticity

Several variations of SDVRP have been studied in the literature, including vehicle availability, travel time, customer demand, customer requests, and service time.

Customer request

Stochastic customer requests refer to the scenario where customer requests occur in real-time and are uncertain. These requests are typically characterized by their probabilistic distributions, which can be used to model the uncertainty in arrival times and demand levels. For example, in an online setting, customer requests can arrive randomly and unexpectedly, following a Poisson distribution, making it difficult to plan and optimize vehicle routes in advance. The uncertainty of stochastic customer requests can be due to several factors, such as customer behavior, traffic, and weather conditions.

To address this challenge, many strategies have been put out in the literature., such as using online algorithms to solve the SDVRP in real time based on incoming requests. These algorithms typically use heuristics or metaheuristics to quickly generate feasible solutions to the problem.[3][4] Another approach is to use reinforcement learning techniques to train agents to make routing decisions in an online setting.[5][6][7] These agents can learn from past experiences and adapt to new requests in real-time, taking into account the probabilistic nature of the customer requests.

Travel time

Travel time can be a significant source of uncertainty. The travel time between two points can vary depending on several aspects, such as traffic patterns, road conditions, and weather conditions. This variability can make it challenging to estimate accurate travel times for a given route, which can affect the efficiency of the routing plan.

One approach to addressing the uncertainty of travel time is to use stochastic modeling techniques.[8] These models can account for the variability in travel time and generate more accurate estimates of the expected travel time for a given route. For example, probabilistic models can be used to estimate the probability distribution of travel time based on historical data or real-time traffic information.[9] Similarly, other probability distributions such as Poisson or exponential distributions may be utilized to simulate stochastic client requests.

Another approach is to use online optimization algorithms that can adapt to changing travel times in real-time. These algorithms can use heuristics or metaheuristics to generate feasible solutions to the routing problem quickly.[10] Then, when new information about trip time becomes available, the solutions may be constantly updated.

Service time

The length of time a vehicle spends at a customer's location to provide products or services is referred to as service time. Many stochastic elements might have an impact on the service time, such as the type of service being provided, the number of items being delivered, and the customer’s availability.

Probability distributions can be used to model the variability in service time when dealing with stochastic service time in SDVRP. For instance, a normal distribution can represent the average service time with a certain variance to account for the variability in the service time. [11]

To solve the SDVRP with stochastic service time, a number of optimization techniques have been presented across the literature. One approach is to use metaheuristics such as genetic algorithms, simulated annealing, or tabu search to find an achievable solution to the issue. [12][13]

Vehicle availability

Vehicle availability refers to the ability of vehicles to carry out deliveries or services promptly. The availability of vehicles can be affected by various stochastic factors, such as traffic congestion, vehicle breakdowns, or unexpected delays.

To model stochastic vehicle availability, different approaches can be used depending on the level of detail required. For example, a simple approach is to assume that the probability of a vehicle being available at a particular time follows a certain probability distribution. More complex models can consider multiple factors, such as vehicle capacity, routing constraints, and scheduling constraints. [14] [15]

Customer demand

Stochastic customer demand is another important aspect of the SDVRP. In traditional vehicle routing problems, the demand for goods or services is assumed to be known in advance and constant. However, in real-world scenarios, customer demand can be highly variable and uncertain due to various factors, such as seasonality, weather conditions, and unexpected events.

To model stochastic customer demand, different probability distributions can be used to capture the variability in demand. For example, the demand for a particular customer may follow a normal distribution, a Poisson distribution, or a gamma distribution. In addition, a correlation between demand at different locations can also be taken into account.

The presence of stochastic customer demand poses significant challenges to the design of efficient routing plans. [16] Traditional routing algorithms that assume constant demand may lead to overestimation or underestimation of vehicle requirements, resulting in inefficient routing plans.[17]

Solution methods

This section continues to provide a summary of the suggested SDVRP solution approaches, focusing on both exact and heuristic approaches. The section highlights the main features, advantages, and limitations of each method, and discusses their applicability in different contexts.

Exact methods

Exact methods are algorithms that guarantee to find the optimal solution to an SDVRP problem, within a finite amount of time. These methods are based on mathematical programming formulations that can provide an optimal solution or an optimal bound on the solution.

Exact methods for the SDVRP typically involve formulating the problem as a mathematical programming model, such as a Mixed Integer Programming (MIP) model [18], and using specialized optimization software to solve the model. Alternatively, some exact methods use dynamic programming [7], branch-and-bound, cutting-plane, or column-generation techniques to solve the problem.

Exact methods are often computationally intensive and require significant computational resources, especially for large and complex instances of the problem. However, they offer the advantage of providing provably optimal solutions, which can be crucial in applications where the cost of suboptimal solutions is high.

Heuristic methods

The Stochastic Dynamic Vehicle Routing Problem (SDVRP) can have high-quality solutions found quickly using heuristic approaches, a family of optimization algorithms that employ problem-specific methodologies. While heuristic approaches are more effective than accurate methods in finding near-optimal solutions, they do not guarantee to find the ideal answer.

Ng et al. [9] proposed a heuristic method for solving an SDVRP with rerouting strategies under a dynamic travel time constraint, by implementing a Multiple Colonies Artificial Bee Colony Algorithm (MCABC). This method is effective in finding high-quality solutions to the SDVRP in a variety of contexts.

Heuristic methods can be particularly useful for large instances of the SDVRP or instances where the problem parameters change frequently, as exact methods may be computationally infeasible or impractical to use. However, heuristic methods may require careful tuning of their parameters and settings to achieve good performance, and their solutions may not always be reproducible. Therefore, it is important to carefully evaluate the performance of heuristic methods and compare them against other approaches fairly and rigorously.

Metaheuristic methods

Meta-heuristic methods are a class of optimization techniques that have been increasingly used to solve the SDVRP. These methods are based on intelligent search strategies that can efficiently explore a large search space to find good-quality solutions, without relying on explicit problem-specific knowledge. Alinaghian et al. [4] proposed a metaheuristic method that combines Genetic Algorithms (GA) and Variable Neighborhood Search (VNS) to solve an SDVRP problem with dynamic customer requests and demands.

The key advantage of meta-heuristic methods is their ability to handle complex optimization problems with a high degree of uncertainty, such as the SDVRP. They are also flexible and can be adapted to different problem settings, which makes them suitable for a wide range of applications. However, one drawback is that they do not guarantee the optimality of the solutions found and that the settings for certain problems and factors might affect how well they function.

Reinforcement learning

A subset of machine learning called reinforcement learning (RL) involves teaching an agent to make decisions in a given environment through trial and error. A vehicle routing agent may be trained to apply RL in the context of stochastic dynamic vehicle routing problems (SDVRP) to make routing decisions in a dynamic and unpredictable environment. The agent learns by receiving feedback in the form of rewards or penalties based on the quality of its routing decisions. The goal of the agent is to maximize its cumulative reward over time by learning a policy that maps the state of the environment to the optimal routing decision.

The field of reinforcement learning (RL) for solving the stochastic dynamic vehicle routing problem (SDVRP) is rapidly advancing, with numerous articles contributing to the development of SDVRP problem-solving. One such contribution is the work of Florentin D. Hildebrandt, which includes the articles [19] and [20]. In these works, the author provides a comprehensive explanation of SDVRP and RL, highlighting useful methods and discussing future works that can be undertaken to advance the field.

Approximate dynamic programming

A set of reinforcement learning techniques called approximate dynamic programming (ADP) is used to address challenging decision-making issues in dynamic systems. A sequence of decisions that maximizes a long-term goal function, such as anticipated reward or expected cost, must be made by the decision-maker in the ADP system, which is described as a Markov Decision Process (MDP).

The key idea behind ADP is to approximate the optimal value function or policy function of the MDP using a function approximator, such as a neural network, rather than explicitly computing the value or policy functions. This allows ADP to handle large-scale, high-dimensional state and action spaces that are difficult or impossible to solve using traditional dynamic programming methods.

The two primary phases of the ADP algorithm are (1) policy evaluation and (2) policy improvement. Using the present policy and the approximative value function, the value function or Q-function is approximated for the purpose of evaluating policies. The existing policy is updated for policy enhancement to be greedy with regard to the estimated value function or Q-function.

VFA

Value Function Approximation (VFA) is a technique used in reinforcement learning to estimate the value function in large state spaces. The value function estimates the expected cumulative reward that can be obtained by following a given policy.

VFA works by approximating the value function using a function approximator, such as a neural network or a linear model, which takes the state as input and outputs an estimate of the value function. This approach is useful in large state spaces where computing the exact value function is infeasible.

In order to tackle a Multi-Period Dynamic Vehicle Routing Problem with Stochastic Customer Request, Ulmer's study [21] presented a VFA-based limited horizon rolling algorithm (V-LHRA) (MDRPSR). In order to estimate the value function and change the policy based on in-the-moment observations, the V-

LHRA technique combines offline and online optimization methods. The suggested approach beat current state-of-the-art approaches when tested on a set of benchmark examples, according to the results.

PFA

In reinforcement learning, the Policy Function Approximation (PFA) method is used to directly estimate the best policy without calculating the value function. The action that maximizes the anticipated cumulative reward is mapped from the state by the policy function.

PFA operates by utilizing a function approximator, such as a neural network or a linear model, to approximate the policy function. Based on current observations and the projected policy function, the policy is changed.

In the work of M. Ulmer [22], the authors aimed to solve a unique case of the autonomous vehicle and station location problem, where customers can pick up goods and vehicles can deliver them, while also considering the capacities of both vehicles and stations. The location of the stations was based on a real case in the city of Braunschweig, Germany. The problem was solved by combining Value Function Approximation (VFA) and PFA methods. The authors proposed an algorithm that uses VFA to estimate the value function and PFA to optimize the policy function based on real-time observations.

Discussion

Through this survey paper, it is apparent that SDVRP has been an active area of research, and various approaches have been proposed to solve the problem. The reviewed literature suggests that researchers have primarily focused on addressing the problem under the constraint of stochastic customer requests, which is a significant challenge in this domain. Value function approximation, policy function approximation, and reinforcement learning have also been explored as effective methods for solving SDVRP.

In contrast, exact methods, heuristics, and metaheuristics have been used less frequently in recent years due to their limited scalability and performance issues. However, they have shown promising results in smaller-scale problems and should not be dismissed completely.

Moreover, this survey paper highlights that MDP is a crucial component of solving SDVRP by RL, and its implementation must be studied effectively to obtain better results. Although there has been significant research on RL, there is still much scope for improvement and further exploration in this area.

Conclusion

In conclusion, this survey paper provides an overview of SDVRP and its complexities, such as stochastic vehicle availability, travel time, customer demand, customer requests, and service time. The literature review demonstrates that many effective methods, such as VFA, PFA, RL, exact methods, heuristics, and metaheuristics, have been proposed to solve SDVRP.

However, there is still considerable potential for further improvement and research in this area. The findings of this survey paper could be a valuable reference for future researchers working on SDVRP and related problems. It is expected that future research in this field will contribute to the development of new and more efficient methods to solve the SDVRP problem, which has practical implications in various industries, such as logistics and transportation.

References

1 M. W. Ulmer, J. C. Goodson, D. C. Mattfeld, and B. W. Thomas, "On modeling stochastic dynamic vehicle routing problems," EURO Journal on Transportation and Logistics, vol. 9, no. 2, p. 100008, 2020.
2 M. W. Ulmer, "Anticipation versus reactive reoptimization for dynamic vehicle routing with stochastic requests," Networks, vol. 73, no. 3, pp. 277–291, 2019.
3 M. Ulmer, B. W. Thomas, A. M. Campbell, and N. Woyak, "Instances for" the restaurant meal delivery problem: Dynamic pick-up and delivery with deadlines and random ready times"," 2017.
4 M. Alinaghian, M. Aghaie, and M. S. Sabbagh, "A mathematical model for location of temporary relief centers and dynamic routing of aerial rescue vehicles," Computers & Industrial Engineering, vol. 131, pp. 227–241, 2019.
5 W. Joe and H. C. Lau, "Deep reinforcement learning approach to solve dynamic vehicle routing problem with stochastic customers," in Proceedings of the international conference on automated planning and scheduling, vol. 30, 2020, pp. 394–402.
6 M. W. Ulmer, N. Soeffker, and D. C. Mattfeld, "Value function approximation for dynamic multi-period vehicle routing," European Journal of Operational Research, vol. 269, no. 3, pp. 883–899, 2018.
7 M. W. Ulmer, J. C. Goodson, D. C. Mattfeld, and M. Hennig, "Offline–online approximate dynamic programming for dynamic vehicle routing with stochastic requests," Transportation Science, vol. 53, no. 1, pp. 185–202, 2019.
8 J. F. Ehmke, A. M. Campbell, and T. L. Urban, "Ensuring service levels in routing problems with time windows and stochastic travel times," European Journal of Operational Research, vol. 240, no. 2, pp. 539–550, 2015.
9 K. Ng, C. K. Lee, S. Zhang, K. Wu, and W. Ho, "A multiple colonies artificial bee colony algorithm for a capacitated vehicle routing problem and re-routing strategies under time-dependent traffic congestion," Computers & Industrial Engineering, vol. 109, pp. 151–168, 2017.
10 F. K¨oster, M. W. Ulmer, D. C. Mattfeld, and G. Hasle, "Anticipating emission-sensitive traffic management strategies for dynamic delivery rout-

ing," Transportation Research Part D: Transport and Environment, vol. 62, pp. 345–361, 2018.

11 G. Yu and Y. Yang, "Dynamic routing with real-time traffic information," Operational Research, vol. 19, pp. 1033–1058, 2019.

12 Z. Bian and X. Liu, "A real-time adjustment strategy for the operational level stochastic orienteering problem: A simulation-aided optimization approach," Transportation Research Part E: Logistics and Transportation Review, vol. 115, pp. 246–266, 2018.

13 M. Hyland and H. S. Mahmassani, "Dynamic autonomous vehicle fleet operations: Optimization-based strategies to assign avs to immediate traveler demand requests," Transportation Research Part C: Emerging Technologies, vol. 92, pp. 278–297, 2018.

14 Z. Wang, P. Jochem, and W. Fichtner, "A scenario-based stochastic optimization model for charging scheduling of electric vehicles under uncertainties of vehicle availability and charging demand," Journal of Cleaner Production, vol. 254, p. 119886, 2020.

15 N. R. Sabar, A. Bhaskar, E. Chung, A. Turky, and A. Song, "A self-adaptive evolutionary algorithm for dynamic vehicle routing problems with traffic congestion," Swarm and evolutionary computation, vol. 44, pp. 1018–1027, 2019.

16 J. C. Goodson, B. W. Thomas, and J. W. Ohlmann, "Restocking-based rollout policies for the vehicle routing problem with stochastic demand and duration limits," Transportation Science, vol. 50, no. 2, pp. 591–607, 2016.

17 C. E. Gounaris, P. P. Repoussis, C. D. Tarantilis, W. Wiesemann, and C. A. Floudas, "An adaptive memory programming framework for the robust capacitated vehicle routing problem," Transportation Science, vol. 50, no. 4, pp. 1239–1260, 2016.

18 M. Monroy-Licht, C. A. Amaya, A. Langevin, and L.-M. Rousseau, "The rescheduling arc routing problem," International Transactions in Operational Research, vol. 24, no. 6, pp. 1325–1346, 2017.

19 F. D. Hildebrandt, B. Thomas, and M. W. Ulmer, "Where the action is: Let's make reinforcement learning for stochastic dynamic vehicle routing problems work!" arXiv preprint arXiv:2103.00507, 2021.

20 F. D. Hildebrandt, B. W. Thomas, and M. W. Ulmer, "Opportunities for reinforcement learning in stochastic dynamic vehicle routing," Computers & Operations Research, p. 106071, 2022.

21 M. W. Ulmer, "Horizontal combinations of online and offline approximate dynamic programming for stochastic dynamic vehicle routing," Central European Journal of Operations Research, vol. 28, no. 1, pp. 279–308, 2020.

22 M. W. Ulmer and S. Streng, "Same-day delivery with pickup stations and autonomous vehicles," Computers & Operations Research, vol. 108, pp. 1-19, 2019.