The Historical Development and Future Outlook of Public Transit Vehicle and Crew Scheduling: A Literature Review

Authors: Ruikun CHEN, Xuewu CHEN, Xuewu CHEN

Date: 2026-04-03T09:31:54+00:00

Abstract

This paper overviews the historical developments in public transit vehicle and crew scheduling using various modelling frameworks. The three major modelling frameworks are operation research-based models, heuristic models and reinforcement learning models, each have a unique mathematical foundation and a series of resolution techniques. And each framework has its strengths and weaknesses as well. There are no permanently perfect models, for which the basic models have been adjusted and extended in not only smaller, simpler modules like the objective function and the constrain conditions, but also larger, complex parts like the model structures. These extensions fit the realistic operation requirements of public transit corporations. This work also proposes some potential future research directions for public transit vehicle and crew scheduling based on the up-to-date development trends in this field.

Full Text

Preamble

The Historical Development and Future Outlook of Public Transit Vehicle and Crew Scheduling:

Literature Review Ruikun C ,Xuewu School of Transportation, Southeast University, Nanjing, People' s Republic of China

Abstract

This paper overviews the historical developments in public transit vehicle and crew scheduling using various modelling frameworks. The three major modelling frameworks are operation research based models, heuristic models and reinforcement learni models, each have a unique mathematical foundation and a series

of resolution techniques. And each framework has its strengths and weaknesses as well. There are no permanently perfect models, for which the basic models have been adjusted and extended in not only smaller, simpler modules like the objective function and the constrain conditions, but also larger, complex parts like the model structures.

These extensions fit the realistic operation requirements of public transit corporations. This work also proposes some potential future research directions for public transit vehicle and crew scheduling based on the development trends in this field

Keywords

Public Transit; Vehicle Scheduling Problem; Scheduling Problem; Modelling; Optimization

1. Introduction

In recent years, with the acceleration of urbanization and rising travel demand, the scale and complexity of public transit systems in major cities have expanded significantly, particularly in densely populated developing regions Modern public transportation has become an indispensable part of urban mobility, providing accessible and affordable travel options for a growing number of residents.

However, the increasing reliance on public transit also brings challenges such as oper ational inefficiency, overcrowding, service delays, and financial strain.

These issues are further compounded by environmental concerns and the pressing need for sustainable urban development. An effective public transit scheduling system is not only cruci al for minimizing operational costs and enhancing passenger service efficiency, but also serves as a vital tool for promoting green transportation, reducing carbon emissions, and building a

more sustainable and resilient urban mobility framework.

To sche dule a public transit system is to legally and effectively use existing vehicles and crews to cover a determined service plan.

A service plan describes the expected departure times of vehicles and their routes if the system has a fixed route.

Be aware that how to generate an appropriate service plan is NOT the focus of this literature.

Also, when executing the plan, incidents suc h as traffic jams may occur and countermeasures may be applied.

This literature recognizes some incidents as random inputs of the scheduling problem , but do NOT take into account how to properly conduct countermeasures when vehicles are on the road, i.e. t operation control measures.

In short, in this literature, ublic ransit cheduling roble is about: a service mission is needed to be executed, how choose the best vehicle and crew(s) for this mission.

According to the object being scheduled (i.e. the decision variables) , public transit scheduling problems can be divided into two categories: the vehicle scheduling problems (VSPs in short) and the crew scheduling problems (CSPs in short) . Generally, the two categories share the same modelling framework and often the same o ptimization objective However, they are somehow distinct in terms of constrain conditions. This is mainly because in reality operation, vehicles and crews have different physical and legally limits need to be considered.

For example, it' s usually acceptable to keep a vehicle running for 12 hours or longer as long as it' s gas or battery does not run out, but in most countries it' s forbidden to require a driver drive all day long.

To now, there are three major modelling frameworks for VSPs and CSPs. first is operation research based models. Most articles using this framework apply a row decomposition (i.e. column generation) method named Dantzig Wolfe Decompos tion.

The second is heuristic models. Although these heuristic algorithms may vary , they are all heuristic and mostly inexact (i.e. a theorical optimized solution is not guaranteed) , thus this literature recognize algorithms following a "guided random search" modelling procedure as a single framework: the heuristic models. The third is reinforcement learning. This framework constructs one or more agents. Agents interacts with the environment, for example the di fferent service plan modes, and learn how to make best scheduling decisions.

In Chapter 2, the technical details, development history, strengths and weaknesses of each modelling framework discussed.

Since public transit scheduling has a long developm ent history, lots of scholars have tried to extend the basic modelling frameworks for specific realistic operational requirements or to integrate more than one framework for better performance . This literature dive into these trials in Chapter 3. By anal yzing a large number of articles on VSPs and CSPs, this literature also propose some future outlook of this research field in Chapter 4.

2. Major

odelling ramework of VSPs and CSPs

In early years, scheduling is implemented manually by operators hired by transit corporations.

During , scholars started systematic theorical and experimental research on CSPs in urban rail crew scheduling (Desrochers & Soumis, 1989) , and these methodologies were soon introduced into public transit The first mature modelling framework is operation research based models. Then in heuristic models emerged and thrived (Haghani & Banihashemi, 2002; Valouxis & Housos, 2002; Wang & Shen, 2007) Finally after the year , reinforcement learning methods were introduced large scale (Wang et al., 2020; Wang et al., 2023) Following a chronological order, this literature will demonstrate these three frameworks one by one.

The Operation Research based Models Manual scheduling is flexible, and the complexity of the problem acceptable when the bus fleet small.

So most corporations rarely cared about systematic and automatic scheduling methods at that time.

But as the urbanization went on, the city grew larger, and bus fleets greatly expanded.

Operators soon found it too hard to maintain high scheduling efficien cy at such complexity.

Propelled by the realistic needs, scholars proposed the primary models.

Since similar problems have been researched in operation research and applied mathematics, some experiences and results were directly brought into the scheduli ng problem In operation research based models, the scheduling problems were considered as a set covering or set partitioning model (Desrochers & Soumis, 1989; Feng et al., 2024) . Here we introduce the former since it is more generalized for scenes where deadheading is allowed.

$= 1$

$= 1$

$= 1$

$, \sum \geq 1$

$= 1$

Where indicates the driver number, indicates the mission number, and indicates the time period number. are usually provided in local labor laws This a simplest model a CSP, in which only one departure depot is considered.

The objective function is the total salary of all drivers. If a driver works too long, he will be paid an extra salary for the exceeded work time.

third strain formula is to ensure one driver can only execute one mission at the same time.

The fifth formula is to ensure all missions are executed. we add more depots or legal requirements (e.g. drivers should not continuous execute missions for more than 4 hours without a 20 minute or longer break), we can also create additional constrain formulas for them Apparently this is a integer programming problem, and operation research provide various methods to tackle it. Some commercial solvers such as Gurobi are also feasible.

Early literature like (Fügenschuh, 2009) solved these original forms of VSPs and CSPs in smaller bus fleets directly However when bus fleets became large, the decision variables exponentially. ven commercial may fail to solve the original problems in acceptable time.

Thus some deco mposition methods were proposed. Decomposition methods were specific modelling skills. They can transform the original problems into other forms, decrease the number of decision variables or the scale of feasible domain significantly without changing the o ptimality. operation research, Dantzig Wolfe Decomposition and Benders Decomposition are the major decomposition methods. For VSPs and CSPs, the former is usually easier to operate , for which most articles choose Dantzig Wolfe Decomposition.

The core of Dantzig Wolfe Decomposition is to break the problem into one master problem and subproblems. Master problem deals with all feasible one driver mission plans, and it's objective is to choose the best combination of them. Of cours e, the feasible plans may be numerous, so in effect master problem usually treat only some promis plans, named restricted master problem (RMP) And the goal of subproblems is to generate these "promising plans" . (Dantzig & Wolfe, 1960) Each time a RMP is settled, it also provides a dual value for each mission. The dual value is the shadow pri ce of each mission, i.e. how much money or resources are spent for this mission in current optimal combination. Using dual value, we can find the best one driver plans using tailored shortest path algorithms. these subproblems are sometimes recognized as shortest path problems with resources constrains (SPPRC). (Dantzig & Wolfe, 1960; Morabit et al., 2022) In each iteration, RMP and SPPRC are solved alternatively, and the objective value gradually converges to optimal.

Note that when the model converges, it usually get a non integer solution, which violate the physical meanings of decision variables. So a branch bound (B&B) or branch price (B&P) procedure should be applied to get a n integer optimal solution.

Early typical literature using this decomposition method include (Desrochers & Soumis, 1989) Operation research models are elegant and theorical complete. It has absolute optimality and great interpretability (Dantzig & Wolfe, 1960) , and do work in early practices.

The most fatal weakness is decomposition models are still compute intensive when the bus fleet went larger. This is mainly because the diagram of SPPRC subproblems still grew exponentially as more missions

were added RMP also became harder to solve when a lot of one driver plans were considered, however it usually plays a less important part than SPPRC.

To settle this, scholars further extended this framework. These extensions will be discussed in Chapter (Feng et al., 2024; Morabit et al., 2021) Another major weakness of operation research based models is that these models are usually static, meaning they are calculated only once before the corporation executes the actual schedule. (Amberg & Amberg, 2023; Li et al., 2009) This can bring troubl incidents happen and the schedule is forced to change, which is always the case in practice. To make the model robust and flexible is thus another important extension direction of this framework. These extensions will also be discussed in Cha The Heuristic Models Heuristic models are not the first models being proposed. It' s only when the bus fleet grew larger and operation research based models could hardly give a enough solution in relatively short calculation time that scholars tried to import heuristic mod This is mainly because nearly all heuristic models are incomplete in theory and do not guarantee optimality. In fact, these heuristic models are actually guided random search models. However, they are still useful if the calculation resources are intense and optimality is not a necessity Before demonstrating specific algorithms, this paragraph will examine VSPs and CSPs were modelled in a heuristic framework. First, the schedule plans were encoded. Encoding me ans transforming concrete plans into abstract expressions. A commonly used way is use a set of arrays to represent a full schedule plan (Ma et al., 2016; Mertens et al., 2024) . Each array means the mission executed by one particular vehicle or driver.

To search better solutions is to change the arrays partly, and different algorithms vary in how to change them to get better search efficiency.

The next paragraphs will show the details of some major he uristic algorithms.

The first batch of heuristic algorithms were relatively simpler models.

Initially they were not designed for scheduling problems, but scholars imported and tailored them to fit th scene.

A typical example of early algorithms can be Generation Algorithm is a population based metaheuristic optimization technique inspired by the principles of natural selection and genetics, first established by (Holland, 1962) . It is designed to efficiently search e, complex, and non differentiable solution spaces for optimal or near optimal solutions. The algorithm operates by iteratively evolving a population of candidate solutions, referred to as individuals chromosomes , through the application of stochastic operators.

In the scheduling context, each individual represent one full schedule plan, i.e. the set of arrays mentioned before. Individual has multi chromosomes, and each chromosome represents a single vehicle or single driver plan, i.e. an ssion array.

The first population is generated randomly, or following a greedy

algorithm. In each iteration, individuals will change some of their chromosomes and some chromosomes may mutate, thus expand the scale and diversity of the population.

Then the algorithm select some best individuals to start the next iteration. Usually the selection is implemented by killing the unpromising individuals, which violate some constrains or have worst fitness (i.e. objective values) Many articles including (Mertens et al., 2024) used GA like evolutionary algorithms in VSPs and CPSs.

As the time went, more heuristic algorithms were imported, such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and neighborhood search based algorithms , each following a distinct rule of changing current schedule plan and searching for better one.

Generally, the effectiveness of a heuristic algorithm vary in different physical scenes. To our best knowledge, the most successful algorithm in VSPs and CSPs might be neighborhood search based algorithms. The next paragraph will focus on the details of t algorithms.

Neighborhood search based algorithms use the same encoding mode as GA.

However unlike GA, NS do not try to imitate the behavior of natural evolution.

The philosophy of NS is to make the tiniest , feasible change each time, and improve the objective value through small steps.

This is implemented by destruct and repair operators Destruct operators are functions that accept current schedule plan as an input, and output a partly broken plan. Common methods to do this include randomly one mission from one array (i.e. an one vehicle or one driver or randomly remov one array from the plan Because the removal of missions, the plan becomes unfeasible after destruction.

Once destruct operators finish their jobs repair operators are called to fix the plan, making it at least feasible. The removal missions may be randomly inserted into arrays, or following a greedy rule in repair operators.

After these procedures, the plan will either get better or otherwise. Usua lly only when the plan get better will it be kept as the start of the next iteration. Hopefully, these iterations will improve the plan over time. (Ma et al., 2016) used a tailored NS, namely VNS, to properly schedule electric buses. (Wen et al., 2016) used a NS mixed with an adaptive operator selection mechanism, namely ALNS for VSPs The advantages of using heuristic models are obvious. They do not have a consuming convergence process, and require less efforts in transforming different objectives and constrains into mathematical expressions compared with operation research based odels. Most heuristic models, especially those have a population mechanism or use multi search units, are naturally suitable parallel computing.

This means they will perform much better in modern processors with multi physical or logical cores.

However, heuristic models were far from perfect. Although some heuristic models do have a theorical proof that it can reach optimality (Stutzle & Dorigo,

2002) gap between the lower bound of the model optimal objective value can be estimated in specific conditions (for example, the

Simulated Annealing is proven to have optimality when initial temperature is high enough , the number of search units is large enough and the annealing e is small enough (D. Mitra & Sangiovanni Vincentelli, 1985) , most heuristic models lack rigorous proof of optimality Due to this theorical incompleteness , using heuristic models is somehow similar to guessing the answer. How good the answer is depends on the parameters set before calling the models (namely hyper parameters), and randomness.

The expression of feasible plans can become complex when the bus fleet is large and a lot of missions exist. This may lead to unacceptable long computing time.

Like operation research based models, heuristic models are usually static.

They are good for planning the schedule of the next day, but when incidents occur and situations change, they may be forced to re allocate the following mission roughly , and this process is time consuming So they may not be helpful in case of an emergency when fast and reliable re schedule is needed.

Heuristic models also have numerous extensions. Some typical extensions will be highlighted in Chapter 3.

The Reinforcement Learning Models Reinforcement learning models are the latest modelling framework of VSPs and CSPs.

Compared with the two static frameworks above, reinforcement learning adopts a different, dynamic modelling path.

Before it was imported into VSPs and CSPs, RL had proven itself to be a stronger m than both manual and traditional models to find optimal solution in many fields, including playing chess like games.

RL is based on the mathematical ground of Markov Decision Process (MDP).

A random process is MDP only when the potential future development of the system depends solely on the system' s current configurations state If the bus fleet' s configurations are expressed properly and completely, then scheduling can be a MDP.

That is to say, when we know the location, on mission status, cumulated and continuous driving time and other necessary details of each driver and the c oming missions , which drivers can be chosen mission and the following outcome for choosing each driver is determined. et al., 2023) Given the fact that one or more optimal schedule plan(s) exist(s), there must at least one optimal driver or vehicle to execute each mission.

Since scheduling is a MDP, it can be inferred that step's optimal driver or vehicle can be decided by assessing the system' s current state , and this decision needs no extra information The philosophy of RL is to automatically find the optimal step schedule orders (i.e. actions , and gradually make an op timal schedule plan. (Liu et al., 2023) Nearly all RL algorithms reward function as a signal for assessing and adjusting the model.

Reward is the measurement of goodness of a schedule plan.

It can be the total generalized cost of the plan, and is nearly same as the

fitness function in heuristic models or the objective function in operation research based models.

However using th measurement of the whole plan can be problematic, since this reward is not revealed until the last action is made. This is called sparse rewards in RL, and will confuse the model as it cannot know immediately whether an action is appropriate right after it makes the decision.

In practice, reward is designed based on specific problems . Reward is usually different from the final objective function , but can lead the model to approximate optimal For example, we may use the current average mission as a reward to optimize the generalized total cost.

Optimizing reward can work, but model may choose some actions that good in this step but leads to bad outcomes later to make the model more visionary, another concept named return is proposed to handle this. Return is sum of current reward and potential future rewards (these future undetermined parts are usually discounted) Finding the decision to optimize return can result in a high reward , and thus making the final plan better.

In terms of finding the best actions, different RL algorithms may utilize different methods.

Before diving into these algorithms, the environment model concept should be introduced.

The environment model is a model that describes how one schedule order can differentiate the state , and thus differentiate the final reward . In scheduling problems , the environment model always exists, but the RL model may or may not try to explicitly rn this environment model (Konda & Tsitsiklis, 1999; Sutton, 1991; Watkins & Dayan, 1992; Williams, 1992) Model based RL are algorithms that tries to directly learn the environment model. After that, a optimal step step scheduling plan can be made to reach the final state with best reward value. (Sutton, 1991) Then comes the model free RL.

Learning the environment model can be hard, and environment model are usually problem specific (this means a trained model may deteriorate even only the salary of drivers or battery capacities of buses change , and needs to be re adjusted with the new environment model free RL algorithms proposed.

They do not try to learn the environment. model free RL algorithms may construct a value function, which determine the goodness of actions (namely Q values) , and the actions with highest Q values may be chosen.

The value functions vary, from a simple state action value table to a complex neural network.

Also some algorithms choose to learn a best state action mapping, or policy function.

There are also algorithms which learn a value function and a policy function simultaneously , and the two function may cooperate to optimize each other the training process.

This trick is called actor critic (Konda & Tsitsiklis, 1999; Watkins & Dayan, 1992; Williams, 1992) Table shows some typical RL algorithms and their categories. . Summary of influence factors for MFDs with data sources and models

Category Typical lgorithms Literature in VSPs and Model based Dynamic Programming, Monte Carlo Tree Search (Wang et al., 2020) Value based SARSA, (Liu et al., 2023) Policy based REINFORCE Actor Critic (Wang et al., 2023) Reinforcement learning is effective. (Liu et al., 2023) (Wang et al., 2023) proven this fact by numerical experiments. Also, RL agents are highly adaptative, meaning they can handle different objectives, different constrains without much difficulty. The most amazing feature of RL may be its dynamic essence. RL makes decision step, so unlike static frameworks, RL can easily adjust the following schedule plan when incidents and emergencies happe Despite many strengths RL has, it is not a widely deployed modelling framework in practice. One reason is it is relatively novel, but the most important reason might be it is usually not so interpretable as simpler operation research based or heuristic mod , especially for those model free algorithms with a large neural network.

If drivers and corporation staffs know little about how a scheduling system generate its orders , they hardly choose to trust it. nderstand how this black framework actually works in a scheduling context in order to make it stronger and safer might be a future direction of VSPs and CSPs.

3. Extension

of basic frameworks Public transit scheduling is a well studied field, and many extensions have been applied to the simple basic frameworks so far. In this chapter, some key extensions will be examined.

General extension of objectives and constrains One important extension direction for all frameworks is to gradually import diverse objectives and constrains This is in fact thrusted by industry: different corporations in different countries and regions have different rule s in scheduling vehicles and crew.

The difference in laws can extend the constrains and implements of models.

A driver is not allowed to drive 4 hour or longer without a 20 minute or longer k in China , but in some countries th e threshold may be longe constrain may cause trouble in some traditional modelling frameworks since a constrain such complexity can only be dealt with by largely tailor resolution implement, for example revising the graph structure or path finding mechanism in the subproblems of operation research based model The need for rostering can extend crew scheduling models as well.

In a longer

term, drivers are usually not allowed to work without complete rest periods, so they should alternatively work. A driver may work from Monday to Friday, and rest from Saturday to Sunday. any scholars adopted a separated or sequential method to handle ros tering constrains in crew scheduling.

In a separated method, daily missions are first assigned to anonymous drivers, and these unnamed mission arrays are called shifts duties . Then after generating all shifts in a roster period, say one week or one month, these shifts can be allocated to specific drivers, satisfying their work and rest requirements during the process. (Breugem et al., 2023) separated method breaks the long term problem down and thus is much less compute intense, but may result in a deterioration in optimality (Feng et al., 2024) For instance, when generating shifts, the model usually minimize the number of shifts , ignoring the distribution work time or other factors Imagine there are hour shifts and hour shifts every day in a week , and a driver can work no more than 6 days a week with the total work time no than 40 hours.

In this case, drivers can only drive as long as 7*5=35 hours because the shifts are determined , and 5 hours is "wasted" . hour and 5 shifts may have better performance in a weekly scale, even if more shifts are generated for each day. s computers' ability grow, some scholars have tried to integrate the generation of shifts and rostering process. This means their mod els will output the weekly or monthly mission arrays (with rest days, of course) for each vehicle or driver directly without generating anonymous shifts first and may get more global optimized solutions. (Feng et al., 2024) The operation mode of bus fleets significantly influence the structure models. Traditionally one bus fleet is in charge of one or a small number of bus lines in only one departure depot, however more and more corporations try to schedule buses and crew throughout the system with multi depots.

Imagine there are two near depots, one with a line whose peak hour is 8:00 a.m. and another with a line whose peak hour is 10:00 a.m.

Apparently we can schedule resources between the two depots, using the excess vehicles and crew from the non peak line t support the peak line in different time periods.

These schedule practice between depots can be modelled and optimized in the extension of single depot scheduling problems, namely multi depot scheduling problems VSPs or MD CSPs).

Specifically, in operation research based models, a tailored special temporal graph is needed to express multi depot mission chain ; in heuristic models, scholars mainly revise the encoding of mission arrays to this extension; and for RL models, adding action which move a vehicle or driver between depots is suitable (Liu et al., 2023) The diversity of vehicles will also lead to model extensions.

New York and London may use different types of buses with different gas or battery capacities and may run different hours before returning for supplies These extensions are usually smaller changes in parameters and not influence the original framework.

However as people focus on energy and environmental protection scholars are trying to find a best scheduling strategy considering battery usage for electric buses nowadays.

For instance, (Wang et al., 2020) researched the battery aging mechanism and proposed a scheduling method to slow down the aging of vehicle batteries These efforts can not only reduce the energy cost of the bus fleets, but also help protecting the environment. (Cong et al., 2024) studied a scheduling methodology for heterogeneous fleets with both fuel buses and electric buses This is helpful for regions transitional period where both types of buses are in s ervice.

Other ethical considerations also extend the models. Fairness is getting more and more highlighted in modern transit operations. Drivers want to have a balanced driving time compared with their colleges, especially for those working in fleets giving salaries not by actual driving time but by days at work Thus, the model should be both cost efficient and fair in time. (Breugem et al., 2023) proposed a method to model fairness by the normalized work time.

And in optimization process, this literature minimized the total cost satisfying the constrain that the fairness of work time should be higher than a threshold.

Extensions to counter uncertainty As mentioned in Chapter 1, in public transit operation, uncertainty is unavoidable. Traffic jams, crowded passengers, long red lights, unstable traffic flows and unexpected emergencies are all reasons causing uncertainty. chapter will study the influence of uncertainty to scheduling prob lems, and how scholars tried to counter it.

Definitely, uncertainty can break a well scheduled plan by disabling it from being executed accurately.

If a vehicle breaks down or a driver quits, the original plan must be revised , or some of the following missions would not be covered.

Even there are no bigger incidents or accidents, a long delay on the road can also make the plan unfeasible, which is a common scene for those long bus routes a lot of stops (Amberg & Amberg, 2023; Li et al., 2009) One general and straightforward method to counter uncertainty is re scheduling. Each time uncertainty makes the plan not feasible, the model is called again to make a new schedule plan for the following missions , based on new initial conditions after the incident happens Adopting a re scheduling method can be convenience in modelling since no adjustment is applied to the model itself , but can be catastrophic in practice because every re heduling process require much calculation while the incidents should be responded as quickly as possible. (Li et al., 2009) Another method is to make the model itself more robustness. That is to making the model able to consider and counter uncertainty.

This can be achieved in different interesting ways. (Poullet & Parmentier, 2020) used a probabilistic robust optimization method to model and counter uncertainty in ground staff scheduling of airports, and might be learnt from by transit corporations In this literature, uncertainty expressed by "cases" , and each case ha d a chance to happen, and the chances

were calculated based on data and some hypotheses.

A case was a incident like flight No.22 is delayed for 10 minutes" .

Since the cases and the chances were calculable, the performance of a particular scheduling plan i n each case and the chances of happening of them can be measured. So a weighted performance of this plan can be got. Then re scheduling and outsourcing methods were applied to decrease the influence of uncertainty, and a optimal solution was provided. (Amberg & Amberg, 2023) proposed a method named controlled trip shifting to counter uncertainty.

In this literature, the departure and arrival time of missions were not a fixed value, but a flexible value which can be adjusted in a time window.

The adjustment of departure times called trip shifting . Then a multi objective optimization problem was constructed, minimizing the cost while minimizing the delay propagation, numbers of trip shifting and deviations from planned headways ailored Dantzig Wolfe method utilized to calculate a optimal solution in both articles above uncertainty is made by design in non traditional operation modes , and heuristic models may work in such situations. (An et al., 2026) studied a robust and flexible method to schedule customized buses in mega events like Olympics.

In this literature, "customized" actually meant "flexible" or "demand response" .

The buses may cho se different routes and stops depending on current passenger OD demands between sites , and this is the reason for uncertainty since each run may have different time interval.

To counter this designed uncertainty, a NS based heuristic model, namely SWK-LUCB@VNS. This algorithm based on VNS, but was augmented by R It simultaneously optimized the scheduling and routing of bus fleets serving a mega event RL models are a great tool in terms of handling uncertainty. As mentioned in Chapter 2, RL algorithms have a dynamic essence, and make scheduling decision for missions one by one. hen incidents happen, RL needs no re scheduling. It can immediately update the state, and make new decisions for the following missions without much extra calculation.

But that doesn' t mean adjustment is not needed when using RL in a uncertain scene.

A RL trained static , determined environment may try to make decisions assuming the travel speed or time is fixed and vehicles or drivers are always available.

This can sometimes be silly. Imagine there are 5 missions during the peak hour, and each run usually cost 20 minutes. A static optimal solution may be two vehicles running alternatively . However during peak hours, each run may cost 22 28 minutes.

Thus, his static plan will frequently fail, and the model is forced to calculate again and again.

This is not a problem technically as long as each calculation is done fast enough, but this "plan" is obviously unacceptable for drivers: no one wants to be reallo cated every 30 minutes.

Generally, scholars would train the RL agent(s) in a stochastic, uncertain environment in order to make the plan more stable in practice. (Liu et al., 2023)

3.3 Extensions to augment

model abilities Besides improving robustness here are also plenty of extensions whose goal is to enhance the models' solution goodness, calculation effectiveness, and other abilities. feasible method is the fusion of frameworks Given the fact that each framework has its strengths and weaknesses, using the wisdom from two or more frameworks may create a combined model which is stronger than both. (Gerbaux et al., 2025; Morabit et al., 2021, 2022) tried to import machine learning methodologies into operation research based models (Dantzig Wolfe Decomposition or column generation models to be exact). In Chapter 2, technical details , strengths and weaknesses of this basic model is examined.

It' s known that this model optimizes the scheduling plan iteratively , solving RMP and subproblems alternatively.

It guarantees optimality, but both RMP and subproblems can be time consuming when number of missions is large.

Theref these article s used some algorithms to accelerate the convergence of column generation was applied to either subproblems or In subproblems, ML helped selecting arcs. An arc meant the linkage between missions.

Apparently if we have lot of missions, we have numerous arcs.

Numerous arcs mean numerous feasible one vehicle or one driver plan (i.e. feasible paths in SPPRC). And this is the reason why subproblems is time consuming. (Morabit et al., 2021) found only very few of arcs were used in an optimal solution, and proposed a ML algorithm to select arcs in advance, so as to save the time solving subproblems.

In RMP, ML helped selecting columns. RMP is to find a optimal combination of some one vehicle or one driver plans, covering all missions and minimizing total cost.

This is a simple linear programming problem, but when a lot of columns exist, it can also be compute intensive.

Similarly, (Morabit e t al., 2022) found only very few of columns were used in an optimal solution.

And they proposed a ML algorithm to identify the most promising columns before solving the RMP.

The ML models in both articles are classifiers, but the refined features layer structure is problem specific.

The ML empowered column generation had great enhancement in calculation efficiency, but loses the theorical optimality of operation models. Although the probability is small, ML selection may miss some arcs or columns which actually is a part of optimal solution.

It was advised that at least one iteration without ML selection should be conducted to examine if the solution s optimal and further improve the solution. (Morabit et al., 2021) Heuristic models can also be augmented by Heuristic models are faster and more convenience, but hyper parameters and randomness matter when it comes to the goodness of solutions.

For GA, the hyper parameters are population size, mutation probability, e And for NS, the hyper parameters

can be which destruct repair operators to use, and the weights of operators being selected.

Traditionally scholars manually adjusted these hyper parameters, or used some simple methods like grid search and adaptive search. (An et al., 2026) proposed a stronger, RL based way to automatically and effectively adjust the weights of VNS operators.

A non stationary KL UCB algorithm estimate expected rewards of different operators through repeated interaction with the environment.

One important improvement of augmented VNS was that it accounted for delayed rewards . Compared with traditional Adaptive Large Neighborhood Search (ALNS) , it might be more visionary.

Some operators might perform better at first but perform poor in the f ollowing process . In ALNS, these operators become dominant because the initial great performance , and other better operators may not be selected frequently enough after that But this RL augmented VNS can deal with this situation.

4. Future directions of scheduling models

The future directions of VSP and CSP models can be stated in five aspects.

The first aspect is the extension of objective functions. In addition to operational cost, valley energy ratio , battery usage and fairness mentioned in Chapter 3, less deadheading and less carbon emission (Wang et al., 2023) also be considered.

To achieve multi objective optimization , the comparison between weighted, hierarchical or pareto models can also be discussed.

The second aspect is more operational scenes. Nowadays we have demand responsive buses whose travel time is not fixed, and some emergencies where resources are very limited and efficiency is life. Scheduling models can make a big difference in th novel practical scenes.

The third aspect is the fusion of modelling frameworks. Operation research based models are old, but they may not be out. Some philosophy, like the dual theory, may still be used in modern heuristic or RL frameworks to provide valuable and useful guidan ce information.

Reversely, using ML and RL as a tool to augment operation research based models are proven a feasible will continue to develop. (Gerbaux et al., 2025; Morabit et al., 2021, 2022; Xu et al., 2025) The fourth aspect is the re import of human wisdom.

Operators send less and less scheduling orders nowadays, but their experience and wisdom may be inherited in a technical way. For example, by the model structure or reward tion.

The last but not least aspect is improving the interpretability of RL. RL is not commonly deployed in real life systems. One important reason is we cannot trust a black box. tailored policy function or other tricks an agent explain itself

Disclosure Statement No potential conflict of interest was reported by the author(s).

ORCID Ruikun CHEN:

References

Amberg, B., & Amberg, B. (2023). Robust and cost efficient integrated multiple depot vehicle and crew scheduling with controlled trip shifting.

Transportation Science (1), 82 An, X., Li, X., & Zhang, B. (2026). Flexible scheduling of customized bus for green events: a distributionally robust optimization approach.

Computers & Operations Research .cor.2025.107249 Breugem, T., Schlechte, T., Schulz, C., & Borndörfer, R. (2023). A three phase heuristic for the fairness oriented crew rostering problem.

Computers & Operations Research Cong, Y., Bi e, Y., Liu, Z., & Zhu, A. (2024). Collaborative vehicle crew scheduling for multiple routes with a mixed fleet of electric and fuel buses.

Energy D. Mitra, F. R., & Sangiovanni Vincentelli, A. (198 Convergence and finite behavior of simulated annealing . 1985 24th IEEE Conference on Decision and Control, Fort Lauderdale, FL, USA.

Dantzig, G. B., & Wolfe, P. (1960). Decomposition principle for linear programs.

Operations Research (1), 101 Desrochers, M., & Soumis, F. (1989). A column generation approach to the urban transit scheduling problem.

Transportation Science Feng, T., Lusby, R. M., Zhang, Y., Tao, S., Zhang, B., & Peng, Q. (20 24). A branch price algorithm for integrating urban rail crew scheduling and rostering problems.

Transportation Research Part B: Methodological , 102941.

Fügenschuh, A. (2009). Solving a school bus scheduli ng problem with integer programming.

European Journal of Operational Research (3), 867 Gerbaux, J., Desaulniers, G., & Cappart, Q. (2025). A machine learning based column generation heuristic for electric bus scheduling.

Computers & Operations Research Haghani, A., & Banihashemi, M. (2002). Heuristic approaches for solving large scale bus transit vehicle scheduling problem with route time constraints.

Transportation Research. Part a, Policy and Practice (4), 309 8564(01)00004

Holland, J. H. (1962). Outline for a logical theory of adaptive systems.

Journal of the Acm (3), 297 Konda, V. R., & Tsitsiklis, J. N. (1999).

Actor critic algorithms . NeurIPS , Denver, , USA.

Li, J., Mirchandani, P. B., & Borenstein, D. (2009). A lagrangian heuristic for the time vehicle rescheduling problem.

Transportation Research Part E: Logistics Transportation Review Y., Zuo, X., Ai, G., & Liu, Y. (2023). A reinforcement learning based approach online scheduling.

Knowledge Based Systems Ma, J., Ceder, A. A., Yang, Y., Liu, T., & Guan, W. (2016). A case study of beijing bus crew scheduling: a variable neighborhood based approach.

Journal of Advanced Transportation (4), 434 Mertens, L., Amberg, B., & Kliewer, N. (2024). Integrated bus timetabling, vehicle scheduling , and crew scheduling with a mutation based evolutionary scheme.

Transportation Research Procedia Morabit, M., Desaulniers, G., & Lodi, A. (2021). Machine learning based column selection for column gen eration.

Transportation Science (4), 815 Morabit, M., Desaulniers, G., & Lodi, A. (2022). Machine learning based arc selection for constrained shortest path problems in column generation.

Informs Journal timization Poullet, J., & Parmentier, A. (2020). Shift planning under delay uncertainty at air france: a vehicle scheduling problem with outsourcing.

Transportation Science Stutzle, T., & Dorigo, M. (2002). A short convergence proof for a class of ant colony optimization algorithms.

Ieee Transactions On Evolutionary Computation Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting.

Sigart Bulletin Valouxis, C., & Housos, E. (2002). Combined bus and driver scheduli Computers Operations Research 0548(00)00067 Wang, H., & Shen, J. (2007). Heuristic approaches for solving transit vehicle scheduling problem with route and fueling time constraints.

Applied Mathematics Computation Wang, J., Kang, L., & Liu, Y. (2020). Optimal scheduling for electric bus fleets based on dynamic programming approach by considering battery capacity fade.

Renewable Sustai nable Energy Reviews

Wang, Y., Qiu, D., He, Y., Zhou, Q., & Strbac, G. (2023). Multi agent reinforcement learning for electric vehicle decarbonized routing and scheduling.

Energy 129335. https: Watkins, C. J. C. H., & Dayan, P. (1992). Technical note: q learning.

Machine Learning 4), 279 Wen, M., Linde, E., Ropke, S., Mirchandani, P., & Larsen, A. (2016). An adaptive large neighborhood search heuristic for the electric vehicle scheduling problem.

Computers Operations Research Williams , R. J. (1992). Simple statistical gradient following algorithms for connectionist reinforcement learning.

Machine Learning , 229 Xu, K., Shen, L., & Liu, L. (2025). Enhancing column generation by reinforcement learning based hyper heuristic for vehicle routing and scheduling problems.

Computers Industrial Engineering

Note: Figure translations are in progress. See original paper for figures.