

A Multiobjective Single Bus Corridor Scheduling Using Machine Learning Based Predictive Models

Chen, B., Bai, R., Li, J., Liu, Y., Xue, N., Ren, J.



**University of
Nottingham**
UK | CHINA | MALAYSIA

University of Nottingham Ningbo China, 199 Taikang East Road, Ningbo,
315100, China

First published 2020

This work is made available under the terms of the Creative Commons
Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

The work is licenced to the University of Nottingham Ningbo China
under the Global University Publication Licence:

[https://www.nottingham.edu.cn/en/library/documents/research-
support/global-university-publications-licence.pdf](https://www.nottingham.edu.cn/en/library/documents/research-support/global-university-publications-licence.pdf)



**University of
Nottingham**

UK | CHINA | MALAYSIA

A Multiobjective Single Bus Corridor Scheduling Using Machine Learning Based Predictive Models

Bing Chen^a, Ruibin Bai^b, Jiawei Li^b, Yueni Liu^b, Ning Xue^c and Jianfeng Ren^b

^a School of Mechanical Engineering, Northwestern Polytechnical University, 710072, Xi'an, China;

^b School of Computer Science, The University of Nottingham Ningbo China, 315100, China.

^c Faculty of Medicine and Health Sciences, The University of Nottingham, Nottingham, NG8 1BB, UK.

ARTICLE HISTORY

Compiled May 4, 2020

ABSTRACT

Many real-life optimisation problems, including those in production and logistics, have uncertainties that pose considerable challenges for practitioners. In spite of considerable efforts, the current methods are still not satisfactory. This is primarily caused by a lack of effective methods to deal with various uncertainties. Existing literature comes from two isolated research communities, namely the operations research community and the machine learning community. In the operations research community, uncertainties are often modelled and solved through techniques like stochastic programming or robust optimisation, which are often criticised for their over conservativeness. In the machine learning community, the problem is formulated as a dynamic control problem and solved through techniques like supervised learning and/or reinforcement learning, which could suffer from being myopic and unstable. In this paper, we aim to fill this research gap and develop a novel framework that takes advantages of both short-term accuracy from mathematical models and high-quality future forecasts from machine learning modules. We demonstrate the practicality and feasibility of our approach for a real-life bus scheduling problem and two controlled bus scheduling instances that are generated artificially. To our knowledge, the proposed framework represents the first multi-objective bus-

headway-optimisation method for non-timetabled bus schedule with major practical constraints being considered. The advantages of our proposed methods are also discussed, along with factors that need to be carefully considered for practical applications.

KEYWORDS

Bus Scheduling; Multi-Objective Optimisation; Combinatorial Optimisation; Machine Learning

1. Background and motivation

Production and logistics nowadays is a truly multi-disciplinary research field. In the past few years we have witnessed significant advances using various optimisation approaches as well as various data analytics approaches to address various problems occurred in production and logistics. Traditionally, there exist two independent groups of research with very different problem solving strategies and data requirements. The first group is to use operations research methods which often formulate problems with mathematical models embedded with innate problem structures and characteristics. Such models, however, require the values of a large number of parameters to be available (i.e. in the form of either their deterministic values or their stochastic distributions) before they can be set up and solved. However, real-life data are often messy, have a lot of uncertainty, and change over time. Optimisation methods are often criticised for its inflexibility or ineffectiveness to deal with complex problems involving a large amount of data or a high degree of data uncertainties. On the other hand, analytical approaches are entirely driven by data and do not rely much on rigid optimisation models. Although such methods are more flexible than optimisation methods, the resulting models and solutions (e.g., in the form of neural network models) have poor interpretability and hence lack of insights that can be easily explained and understood by human users.

This research aims to bridge the above described gap between optimisation and analytics by proposing a general framework that can integrate optimisation and machine learning in solving challenging real-life problems. Because of lack of access to large amount of real-life data in other areas, we demonstrated the practicality and

feasibility of the proposed framework through a bus scheduling problem which shares same characteristics with a production and logistics system, especially the natural of uncertainties and conflicting objectives from different stakeholders. We firmly believe the methodologies used in this paper for bus scheduling are transferable to production, logistics and other fields that encounter similar situations.

Optimisation of vehicle schedule is fundamentally important in freight transportations (Bai, Xue, Chen, & Roberts, 2015) and passenger transit services (Bai, Li, Atkin, & Kendall, 2014). Buses are probably the second most important urban public transport behind subways. However, the innovations in technologies and services related to bus transport lag behind other transport systems considerably. The bus timetables and schedules are still very much created from traditional planning models and approaches based on simple rough data estimation of travel demands and time. Various bus holding policies at control points are then utilised to execute the plan with minimum deviations and to improve the reliability of bus services.

In many countries, bus timetables are given as a priori, which assumes repetitive customer demand patterns for weekdays and weekends. These predefined timetables can be too rigid for volatile passenger demands, often affected by factors like rain, temperature, events, etc. In some other places, the bus timetables are not fixed. In this regard, the operating companies often adjust the dispatching density (or dispatching headways) dynamically for different traffic and travel demand scenarios. This method is particularly useful when the travel demand is very high. Similar mechanisms are also used in some metro-line timetables, which are defined by the start time, the finish time and the dispatching density or headways (e.g. every 5 min per dispatch). Practically, the determination of the dispatching density is primarily based on years of experience. In most cases, for ease of implementation and management, only two different dispatching densities are used for peak time and off-peak time respectively. However, because of some dynamic events (e.g. weather changes, road accidents, social events), the classic dispatching density models (for example, those by Sun and Zhang (2016)) are too slow to respond to real-time scenarios. Some of the issues can be illustrated by a plot of bus distributions along bus route stops based on real-life data in Figure 1.

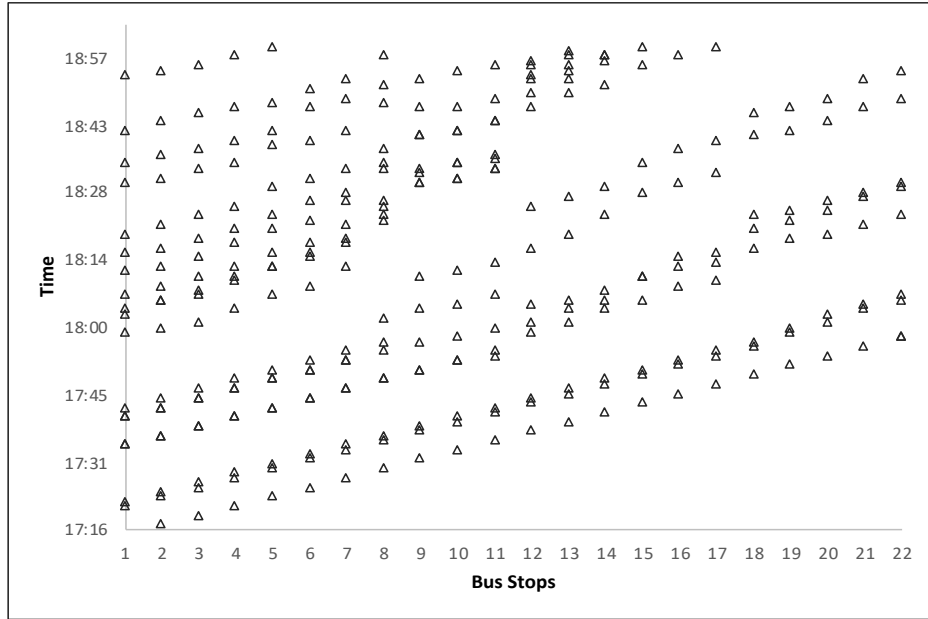


Figure 1.: A real-life example of bus distributions along the bus corridor.

One obvious issue is that the distribution of buses along the bus route is highly uneven, creating an unreliable bus service across different bus stops (e.g. bus users at some stops have to wait much longer than the expected waiting time). Additionally, the delay of buses would have negative impact on future bus dispatches at terminals. The root of the issues is the assumption of constant travel time between bus stops over time in this traditional bus dispatching model. Such an assumption often does not hold in urban areas in the event of accidents and traffic flow fluctuations. These kinds of events may introduce uncertainties to the bus dispatching model, e.g. the delay of buses. Furthermore, the uncertainties at one stop may be easily accumulated to the next stop so that towards the end of the bus route, the uncertainties might be so large that the model became infeasible to execute and could not work at all. Those methods focus on dynamic dispatching strategies on real time may be myopic in many cases (Berrebi, Watkins, & Laval, 2015; Chow, Li, & Zhong, 2017). Some research efforts have been made to analyze and address this issue. Habibi, Battaa, Cung, Dolgui, and Tiwari (2019) addressed a stochastic multi-vehicle collection-disassembly problem with uncertainties in the quality and quantity of products at collection centres. In their stochastic programming model, uncertainties were handled by two heuristic-based

approximation methods which were combined with a framework of sample average approximation in order to solve problems with large number of scenarios.

Inspired by multi-objective modelling presented in (Guo et al., 2019; Xu, He, & Zhu, 2019), in this research, we develop a novel multi-objective optimisation model to resolve conflicting objectives of the multiple stakeholders of the bus scheduling. Additionally, the proposed framework can handle uncertainties by exploiting both the long-term travel patterns in passenger demand and trip time and the short-term traffic scenarios. More specifically, we make use of machine-learning techniques such as support vector regression on both the historical data and the real-time data monitoring the bus services to estimate the long-term travel patterns such as passenger demand and trip time. As discussed early, the optimisation based on the real-time data alone may lead to accumulation of uncertainties to the subsequent bus stops. With the help of the historical data, we could effectively reduce uncertainties by digging out the long-term travel patterns through machine-learning techniques.

With these key parameters being estimated reliably, we then plug them in the multi-objective hybrid optimisation framework, aiming at minimising both the total operation cost and the total passenger waiting time at the same time. This multi-objective optimisation framework is particularly important when it is difficult for the decision makers to decide the weight of each objective function, whereas a valid bus schedule has to be planned by considering all kinds of constraints. Two multi-objective optimisation methods, NSGA-II (Deb, Pratap, Agarwal, & Meyarivan, 2002) and MOEA/D (H. Li & Zhang, 2009), are utilized to solve the problem and provide a set of valid bus schedules with direct estimation of these objective functions. The decision makers can then choose the best suitable schedule from the set of valid bus schedules. Such an optimisation framework could respond to the short-term traffic scenarios in a timely manner using the real-time monitoring data.

The main contributions of this paper are four-fold: 1) We propose a practical bus dispatching model and a solution framework that harnesses the strengths of both machine learning and optimisation-based planning. 2) The proposed framework could discover the long-term travel patterns by applying machine-learning algorithms on the historical data, and use the discovered patterns to reduce uncertainties in the optimisa-

tion framework. 3) The proposed multi-objective optimisation framework could better handle various real-life constraints and the trade-off among multiple objectives, and provide the decision makers an intuitive decision space. 4) The proposed framework is compared with a static model where no machine-learning techniques are utilized and the travel patterns are assumed constant over time without uncertainties. The proposed framework significantly outperforms the static model, which demonstrates the effectiveness of the proposed model.

The remainder of this paper is organised as follow: Section 2 provides a literature review of the relevant work related to bus scheduling, in particular those techniques focusing on dealing with uncertainties. Section 3 describes the problem and its mathematical formulation. The proposed multi-objective hybrid optimisation framework that is used to solve this problem is given in Section 4, followed by a real-life case-study and two artificially generated case-studies in Sections 5 and 6 respectively. Finally Section 7 concludes the work.

2. Literature review

There are numerous studies on optimising bus scheduling at different levels (strategic, tactical and operational). For a thorough review of bus planning and scheduling, one can refer to Bowman (1981), Daganzo (2009) and Ibarra-Rojas, Delgado, Giesen, and Munoz (2015). In this study, we use the single-bus-route-dispatching problem to demonstrate the novelties and benefits of our proposed models and solution philosophy (i.e. the integration of model-driven optimisation with machine learning). The same approach can also be used for multiple-bus problems.

Most of early bus scheduling studies adopted deterministic models but more and more recent studies have been focusing on dynamic strategies under an environment with uncertainties. Cortés, Jara-Díaz, and Tirachini (2011) proposed an integrated model that permits the options of short-turning strategy and deadheading into a normal bus dispatching model. The benefits of different options were evaluated. Sun and Zhang (2016) studied a bus headway optimisation which assumes a constant bus travel time and passenger arrival rate within the planning horizon. Mokhtari and Ghezavati

(2018) modelled a school bus routing problem as a multi-objective optimisation problem targeting at minimizing the total number of buses used and the average traveling time. Both traffic conditions and passenger demands are assumed constant over time. However, it is well accepted that bus scheduling problem has multiple random variables and hence the robustness of bus services is also important. Feng, Saito, and Liu (2016) addressed the urban passenger transport management by combining prediction with optimisation. A Bayesian network was used to predict the probability of the overall traffic congestion in an urban road network, which was showed effective in improving transport service. Chen, Yu, Zhang, and Guo (2009) analysed bus-service-reliability issues at different levels and defined three reliability measurements for bus routes and stops. The study found that the reliability is highly correlated to the length of bus routes. However, it fails to analyse how the reliability might change over time.

Two types of research efforts have been made in the literature to address the uncertainties in bus scheduling. The first type of strategy is to adopt myopic optimisation to address real-time dynamic events. Eberlein, Wilson, Barnhart, and Bernstein (1998) studied a real-time deadheading scheme in order to skip some stops in an “optimal” way in the event of disruptive incidents, so that the remaining schedules are not affected. The decisions to be optimised include the time at which a deadheading should take place and the number of stops that are skipped in a bus trip. Daganzo (2009) investigated a dynamic bus holding scheme at pre-defined control points to reduce in-vehicle passenger delay. However, this is a typical reactive approach and may lead to sub-optimal solutions over a long run. Liu, Yan, Qu, and Zhang (2013) proposed a genetic algorithm to solve a similar problem under random travel time for near-optimal solutions. S. Zhang and Lo (2018) proposed a dynamic control method that introduces additional headways ahead and behind a control point so that headways at different stops are distributed as evenly as possible. This scheme is useful for off-peak periods when the demand is relatively low and the variances of headways are independent and moderate. Although these dynamic scheduling strategies (e.g. short-turning and dead-heading) are useful to address some interruptive incidents and improve bus service reliability, the solutions may be too myopic because the decisions are made by only considering real-time data for the current bus service under optimisation. Repetitive

patterns and hidden structures in uncertain variables (e.g. travel time and passenger demand) in historical data are overlooked.

The second type of approaches try to introduce some redundancies in traditional deterministic models based on statistical information retrieved from historical data. Shen, Xu, and Li (2016) studied a transit-vehicle-scheduling problem with a novel probabilistic-delay-propagation model. Gkiotsalitis and Kumar (2018) highlighted the importance of reducing the fluctuation of expected waiting time of passengers at different bus stops and used a genetic algorithm to optimise the bus dispatching time subject to various regulatory constraints.

Andres and Nair (2017) investigated several time-series-analysis methods to directly predict the headways, instead of the bus arrival time. The prediction methods that were tested in their study include linear regression and extrapolation, kernel regression and extrapolation, artificial neural networks and autoregressive models. The prediction results were then used as inputs for a dynamic bus holding strategy to reduce the headway deviations. This predictive-control framework is suitable for fixed-timetable-bus-scheduling problem that is different from the problem studied in this paper. In our problem, rather than predict headways, we explicitly declare terminal station headways as the decision variables in our mathematical model (see Section 3) and headways at intermediate stops as functions of the terminal station headways and travel time, with the latter being predicted via a machine learning module. The work studied by J. Zhang et al. (2017) also modelled bus scheduling problem as a headway control problem at the terminal station in an attempt to maximize the fulfillment of a predefined target schedule or target headways. The main contribution of their work is the convergence properties of their two-way-looking (upstream and downstream) control scheme under both deterministic and stochastic travel time settings. However, their two-way-looking control scheme assumes static passenger demands and the bus capacity is not considered as a binding constraint. For practical use, this could be a major limitation because the method would not be suitable for bus dispatching during rush hours, when the bus capacity is very much a major factor to be taken into account in the optimisation. Another closely related work by Huang, Li, Li, and Xia (2019) extended the previous work with passenger flow and arrival time predictions based on

richer historical data and a more complex setting that supported passenger transfers at stops. The optimisation targets at the minimisation of total passenger waiting time in one decision epoch. Hence this framework is also a typical dynamical bus dispatching method.

3. The single-bus-route-dispatching problem and model formulation

3.1. Problem description

In this paper, we are concerned with a single-bus-route-dispatching problem, in which bus operators use terminal station headways as decision variables to balance the tradeoff between the service quality and the operation costs. The headway at a bus stop is the service gap between two successive bus visits to the stop. A smaller headway indicates a more frequent bus service but often leads to higher operation costs. This problem is different from traditional timetable-based bus planning where a pre-defined target timetable is given and the operators try to use different control strategies to minimise the deviation of the actual bus service time from the target timetable.

Formally, our problem can be described as follow. Given a bus route with a dispatching terminal and two bus movement directions (outbound or direction 1 and incoming or direction 2), let S and K be the list of stops in direction 1 and direction 2, respectively. Denote V and W be the list of bus trips to be made over the planning horizon for the two bus route directions, respectively. At any moment of decision making, CT , we want to determine the optimal dispatch headways (i.e. the time gap between two consecutive bus trips at a stop) at the terminal stop for outbound bus trips within the planning horizon. The objective factors for optimisation include minimisation of the total passenger waiting time, minimisation of the bus overload and minimisation of the total bus operation costs. The model contains three types of parameters: user-controlled parameters, real-time parameters and forecast parameters obtained by machine learning modules. Their definitions are given in the next section.

3.2. Parameters and notations

3.2.1. User controlled parameters

- *BusRoute*: including types (cyclic, or symmetric dual control for symmetric single control), number of stops, total distance, travel time under normal conditions in both directions (outbound and inbound).
- T_s : planning horizon start time. By default this is set to the current time CT .
- T_f : finish time of the planning horizon.
- P : the discretised planning horizon consisting of a set of continuous time periods of identical length τ , indexed by p .
- τ : length of each period $p \in P$.
- n : the total number of buses available.
- B : the list of buses available with properties including capacity, per trip running cost, available time window (or unavailable time window), vehicle types (normal bus or emergency back-ups). Broken-down vehicles should not be included in this list.
- g_{\max} : maximum bus dispatching gap permitted at different periods p .
- g_{\min} : minimum bus dispatching gap.
- *MinRestTime*: minimum rest time for a bus at a terminal.
- η : maximum bus load rate and $\eta \geq 1$.

3.2.2. Parameters obtained from real-time monitoring module

- The status of all buses defined in B , including their positions.
- The boarding and alighting data at each station in the most recent m trips. This can be empty if no bus trip has been made yet during that period.
- The GPS trajectory data of the most recent m bus trips.

3.2.3. Parameters estimated by machine learning modules

- $r(p, k)$: the passenger arrival rate at time period p and stop k . Unit: person/min.
- $a(p, k)$: the proportion of the bus load alighting at stop k during time period p .
 $0 \leq a(p, k) \leq 1$.

- $RT(d_0, i)$: the time for a given scheduled bus trip i to reach final destination (i.e. final terminal). d_0 is the departure time of the bus trip i . If $d_0 < CT$, the trip is on-going. The time should be estimated based on a combination of the vehicle's current position and historical data.
- $T(k, k + 1, p)$: the bus trip time from stop k to $k + 1$ at period p .

3.2.4. Decision variables and auxiliary variables

The solution of the bus schedule problem is encoded into two bus trip queues, V and W , for control direction 1 and direction 2, respectively. Both V and W are sorted by the planned departure time. The length of V and W should cover the entire planning horizon and can be estimated based on the average headways and average travel time for each direction.

- g_i : the dispatching headway for a scheduled bus trip $i \in V$ in control direction 1.
- h_j : the dispatching headway for a scheduled bus trip $j \in W$ in control direction 2.
- d_i : the departure time of a scheduled bus trip $i \in V$ in control direction 1. Hence $g_i = d_i - d_{i-1}$.
- \bar{d}_i : the departure time of the return trip in control direction 2 for the bus that starts the i -th trip in control direction 1.
- \bar{d}_j : the departure time of the return trip in control direction 1 for the bus that starts the j -th trip in control direction 2.
- e_j : the departure time of bus trip $j \in W$ in control direction 2, and $h_i = e_i - e_{i-1}$.
- d_i^k : the departure time of bus trip $i \in V$ from station/stop k .
- e_j^s : the departure time of bus trip $j \in W$ from stop $s \in S$.
- g_i^k : the headway for bus trip $i \in V$ at bus stop $k \in K$. $g_i^k = d_i^k - d_{i-1}^k$.
- h_j^s : the headway for bus trip $j \in W$ at bus stop $s \in S$ and $h_j^s = e_j^s - e_{j-1}^s$.
- L_i^k : the passenger load for bus trip $i \in V$ at stop $k \in K$ for direction 1.
- L_j^s : the passenger load for bus trip $j \in W$ at stop $s \in S$ for direction 2.

3.2.5. Other notations

- CT : current time.
- K : the set of successive stops in control direction 1, including start and finish terminals, indexed by k .
- S : the set of successive stops in control direction 2, including start and finish terminals, indexed by s .
- C_i, C_j : the capacity of vehicles used for trip i and j , respectively.
- F_i, F_j : the fixed operation cost for running trip i and j , respectively.

The scheduling optimisation procedure can be activated by a returning vehicle or a number of returning vehicles within the rolling planning horizon. For a current rescheduling point/time CT and a planning horizon P , comprising successive time periods of identical length, auxiliary variables can be calculated through the following recursive functions:

$$g_i^k = d_{i-1}^k - d_i^k \quad \forall i \in V, k \in K, \quad (1)$$

$$d_i^k = d_i^{k-1} + T(k-1, k, p) \quad \forall i \in V, k \in K, \quad (2)$$

$$g_i^0 = g_i \quad \forall i \in V, \quad (3)$$

$$d_i^0 = d_i \quad \forall i \in V, \quad (4)$$

$$h_j^s = h_{j-1}^s - h_j^s \quad \forall j \in W, s \in S, \quad (5)$$

$$h_j^k = h_j^{s-1} + T(s-1, s, p) \quad \forall j \in W, s \in S, \quad (6)$$

$$h_j^0 = h_j \quad \forall j \in W, \quad (7)$$

$$h_j^0 = h_j \quad \forall j \in W. \quad (8)$$

For most applications, we can approximate $T(k-1, k, p)$ by $T(k-1, k)$ for fast computation. If this does not meet practical real-life requirements, $T(k-1, k, p)$ should also be estimated through machine learning modules with real-time input features, where $p = (d_{i-1}^k + d_i^k)/(2 \times \tau)$ for direction 1 and $p = (e_{j-1}^s + d_j^s)/(2 \times \tau)$ for direction 2.

3.3. Constraints

As in the case for many real-world applications, our bus scheduling should satisfy several constraints. In the following, we list the most fundamental constraints that exist in almost all bus scheduling problems. It is possible for practitioners to extend these constraints further based on their actual requirements, while the main model and the solutions introduced in this paper are still valid.

$$g_{\min} \leq g_i \leq g_{\max} \quad \forall i \in V, \quad (9)$$

$$g_{\min} \leq h_j \leq g_{\max} \quad \forall j \in W, \quad (10)$$

$$d_i + RT(d_0, i) - \bar{d}_i \geq \text{MinRestTime} \quad \forall i \in V, \quad (11)$$

$$e_j + RT(d_0, j) - \bar{d}_j \geq \text{MinRestTime} \quad \forall j \in W, \quad (12)$$

$$\sum_{i \in V} g_i \geq T_f - T_s. \quad (13)$$

Constraint (9) and (10) ensure that the headways for each bus trip in both directions are between a pre-specified minimum and maximum. Constraint (11) and (12) make sure that drivers have a minimum rest time before their next trips start. Constraint (13) guarantees the full coverage of the planning horizon in any feasible list of bus trips V . The constraint of the maximum number of vehicles used should be automatically satisfied while initialising the trip list V and W .

3.4. Objectives

The objectives of the problem need to take account of preferences from different stakeholders. In this paper, we consider the following three factors, namely the bus service operation cost O_1 , the total passenger waiting time O_2 , and the total bus overload O_3 . The total operation cost can be estimated by the cost of bus trips in both control directions:

$$O_1 = \sum_{i \in V} F_i + \sum_{j \in W} F_j. \quad (14)$$

The total passenger waiting time at all stops can be calculated as follows:

$$O_2 = \sum_{i \in V} \sum_{k \in K} r(p, k) \times (g_i^k)^2 / 2 + \sum_{j \in W} \sum_{s \in S} r(p, s) \times (g_j^s)^2 / 2. \quad (15)$$

The first term is for bus route direction 1 and the second term is for direction 2. As stated earlier, period p can be estimated as $p = (d_{i-1}^k + d_i^k) / (2 \times \tau)$ for direction 1 and $p = (e_{j-1}^s + d_j^s) / (2 \times \tau)$ for direction 2. The arrival rate within each period p is assumed unchanged, resulting in a quadratic waiting time function. Alternatively, one can interpret it as an average waiting time of $g_i^k / 2$ multiplied by the total number of passengers accumulated at stop k during that period, which equals to $r(p, k) \times g_i^k$.

The final objective is to minimise the total bus overload, which takes into account not only the level of the overload (i.e. the number of passengers exceeding the stated bus capacity), but also the duration during which this overload occurs. The function is therefore defined as follows:

$$O_3 = \sum_{i \in V} \sum_{k \in K} [T(k, k+1, p) \times \max\{0, r(p, k)g_i^k + (1 - a(p, k))L_i^k - C_i\}] \\ + \sum_{j \in W} \sum_{s \in S} [T(s, s+1, p) \times \max\{0, r(p, s)g_j^s + (1 - a(p, s))L_j^k - C_j\}],$$

where p can be estimated similarly as in the previous equation. Note here the overload objective is defined mostly like a penalty function. If during a bus trip, overload never happens and the load at any time is no more than its capacity, this term equals zero.

The above three objectives measure different aspects of a bus schedule. O_1 considers the costing of the bus operation. O_2 and O_3 measure the quality of the bus service in terms of the waiting time by passengers at stops and possible overloading penalties. Although one can use a weighting function to combine these three functions into a single objective, it is often very difficult to set the weights because they measure different things with different units and are not directly comparable. To address this problem and make the model more friendly for practical applications, a multi-objective optimisation model is developed in this study. It is not difficult to understand that the overall cost of operation O_1 is in conflict with both O_2 and O_3 and should be treated as a standalone objective. However, instead of making both O_2 and O_3 as separate

objectives, which would make the problem unnecessarily harder to solve, we decide to move the bus overload, O_3 , into the constraints so that the maximum bus overload is restricted. In this case, we set the maximum bus load at any time be less than η ($\eta \geq 1$) times of the bus normal capacity. That is, the following conditions should be satisfied for all stops and all bus trips.

$$r(p, k)g_i^k + (1 - a(p, k))L_i^k \leq \eta C_i \quad \forall i \in V, \forall k \in K, \quad (16)$$

$$r(p, s)h_j^s + (1 - a(p, s))L_j^k \leq \eta C_j \quad \forall j \in W, \forall s \in S. \quad (17)$$

Our final multi-objective optimisation formulation becomes

$$\min O_1, O_2, \quad (18)$$

subject to constraint (9)-(13) and (16)-(17).

4. The proposed hybrid multi-objective optimisation methods

The model developed in the previous section is a typical non-linear formulation and it can be in large-scale if the granularity of periods P is significantly smaller than the planning horizon or if the dispatching density is high. Therefore, heuristic solutions are proposed for this problem. The primary decision variables are the headways (g_i, h_j) for bus route direction 1 and 2. Other variables can be computed through these two sets of decision variables. In this research, we propose to use two popular multi-objective optimisation methods, namely NSGA-II (Deb et al., 2002) and MOEA/D (H. Li & Zhang, 2009). Both methods are inspired by the natural evolutionary process and the principle of “survival of the fittest”. While NSGA-II evolves a set of evenly spread Pareto-frontier solutions via non-dominated sorting, MOEA/D achieves such a goal via decompositions of search into specific directions along the Pareto frontier. It is acknowledged that there have been some new developments in the multi-objective optimisation methods, particularly for problems with many objectives. However, the primary focus of this paper is a novel model for bus scheduling

which integrates machine-learning-based forecasts into a multi-objective optimisation framework to enhance its applicability and flexibility.

The overall framework of the bus dispatching system can be illustrated by Figure 2. The system makes use of both historical data gathered from passengers and bus GPS modules to produce high-quality, time-dependent parameters of our proposed model before it is solved by the multi-objective optimisation methods. During the execution stage of the solution, an execution monitoring module is used to collect the status of the plan execution and other information. In an event of a major disruption, the optimisation can be re-called to adapt to the new scenario if deemed necessary.

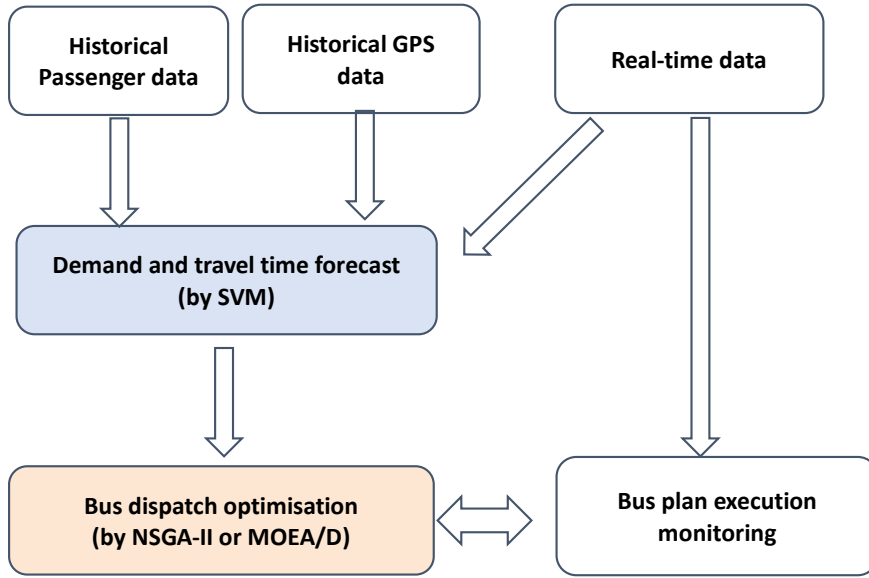


Figure 2.: The overall framework of the proposed bus dispatching system.

4.1. Solution encoding and fitness evaluation

In this application, a chromosome is encoded as a vector, consisting of all dispatch headway variables for both direction 1 and direction 2: $[g_0, g_1, \dots, g_i, \dots, g_{|V|}, h_0, h_1, \dots, h_j, \dots, h_{|W|}]$. Each allele in the chromosome takes positive integers from a domain of $[g_{\min}, g_{\max}]$. In addition, to speed up the objective evaluation, a number of arrays are used to store auxiliary variables. The list of

auxiliary variables are given in Section 3.2.4 and their values can be computed through equations (1)-(8).

Initial solutions are generated randomly. The variables are set to random integer between g_{\min} and g_{\max} . The length of the chromosome is set to a value so that the entire planning horizon is covered sufficiently (i.e. constraint (13) is satisfied), considering the possibility that most dispatching headways may take values close to their minimum. To achieve this, during the initialisation, we generate chromosomes with a few alleles longer than the actual required scheme to make sure that the planning horizon is guaranteed to be covered entirely. During the solution evaluation, however, the fitness will be calculated up to a point i of the chromosome so that its departure time d_i matches or exceeds the finish time of the planning horizon T_f . The remaining part of the chromosome beyond point i will not be evaluated because these trips shall not be executed.

4.2. Genetic operators

We adopt two-point crossover operation at probability of 1 and a uniform mutation (by either increasing or decreasing the headways by 1 unit) at a mutation rate of $1/R$ where R is the chromosome length. In practice, a local search procedure may be added at the end of each reproduction phase (i.e. crossover and mutation) so that a local optimum is reached at each generation of the evolution. The neighbourhood operator can be similar to the mutation operator (i.e. increase or decrease a dispatch headway by 1 minute each time), which is sufficient in this case because the constraints for this problem are not very tight and are separable. In our implementation in this study, we do not include the local search phase and rely on the evolution to converge naturally because the focus of the paper is the evaluation of practicality of a new model that requires the integration of machine learning and multi-objective optimisation.

4.3. Parameter settings

Like other applications, a number of parameters need to be tuned to the bus headways optimisation problem concerned in this paper. Some practical constraints should be

considered while setting them.

4.3.1. Computational time

In an ideal scenario, the formulation presented in the previous section is solved only once to generate a priori solution for execution. Most often, this can be done in the early morning before the first trip of bus services. However, for mega cities like London and Shanghai, there could be hundreds of bus lines, a batch run of the algorithm for all bus lines will take a long time. Therefore, we can't afford hours of computation for each bus line, even if the execution is carried out in parallel through multi-threading. Secondly, bus operations are subject to various dynamic events and uncertain demands from passengers. Given the dynamic nature of road traffic and passenger demands, the initial "optimal" schedule generated in the early morning may have well deviated from the optimality, or the solution may even become infeasible. If this happens, resolving the problem is needed in real-time, in which case, the time permitted to resolve the formulation would be even more limited. A lot efforts shall be required to make the optimisation method more efficient.

4.3.2. Data preprocessing

A real-life problem, one-terminal circular bus line, is used to investigate the feasibility and practicality of our proposed model and solution. The bus line has 22 stops, i.e. $|K| = 22$ and $|S| = 0$. The data are drawn from bus log files from 2016/3/11 to 2016/4/11. The number of records is over 100,000. The data contain at each stop (including the terminal) the bus arrival time, the number of boarding passengers, the number of passengers alighting, bus load and other information. A snapshot of the data is given in Table 1.

As shown in Section 3, there are four estimated parameters, namely the arrival rate, alighting ratio, travel time from stop k to stop $k + 1$ at different time periods, and the time for a bus trip between terminals. We can see that these parameters are not directly available from historical data. Furthermore, the data contain lots of noise and errors, which are largely due to the mechanism of data collection. More specifically, each time when the bus door opens, the system triggers a procedure to add a new piece

Table 1.: A snapshot of bus log data.

busID	stopID	boarding	alighting	load	datetime	day of week
1763	1	0	0	0	5:41:00 PM	5
1768	1	0	0	0	5:43:00 PM	5
1758	1	6	0	0	5:59:00 PM	5
1760	1	5	0	0	6:03:00 PM	5
1764	1	4	0	0	6:04:00 PM	5
1765	1	0	0	0	6:07:00 PM	5
1769	1	3	0	0	6:12:00 PM	5

of record to the database. Sometimes, the door might open several times at a single stop to let passengers arriving slightly late to get on the bus. This leads to some stops having several records in a very short time. It is also possible that the bus skips certain stops without opening the door since there are no people boarding, nor alighting, in which case no record shall be generated at these stops. Therefore, appropriate methods should be used to correct the data records to reflect real-life scenarios. In addition, the data for the first stop and the final stop tend to contain more errors. Fortunately, records of these two stops can be recalculated with data from other stops. For the starting stop, both alighting and load should be zero. (Notice that load means the number of people on the bus before passengers get on or get off at this stop.) For the final stop, the number of people boarding at the final stop is set to 0. And since all of the passengers still on the bus will get off at the final stop, alighting ratio is set to 100%. Finally, outliers (i.e. extreme values that unlikely happen) are replaced with the median value for that period.

4.4. Problem parameter estimation by support vector regression

One of the main contributions of this paper is the introduction of machine learning methods into a traditional optimisation problem. Through machine learning modules, high-quality parameters are estimated by utilising both the real-time traffic data and the historical data at a much finer time granularity (defined by the size of the period τ). The parameters estimated by the machine learning modules include passenger demand data such as $r(p, k)$ and $a(p, k)$, and travel time data such as $RT(d_0, i)$, $DT(k, k + 1, p)$

and $DT(p, k)$. In this research, we use a popular SVR (support vector regression) method (Vapnik, 1998) from the LIBSVM library (Chang, 2016) for the estimation of both passenger demand data and travel time data. Here, the adoption of SVR is mainly to demonstrate the feasibility of the system. It is not the focus of this paper to investigate the best forecast methods for travel time and passenger demands. We suggest interested readers to refer to (X. Li, Bai, Siebers, & Wagner, 2019) for more details.

Support Vector Regression is one of the most used methods for regression problems $y = f(x) + \epsilon$ where x is the input vector and ϵ is the noise. SVR estimates y by $\bar{y} = \sum_{i=1}^k w_i \phi_i(x)$ where $\phi(x) = [\phi_1(x), \phi_2(x), \dots, \phi_k(x)]$ is a vector of kernel functions (support vectors) and w_i are weights. SVR solves the regression problem by minimising a so-called ϵ -insensitive loss function over the input data points (Chang, 2016).

Once the data are cleaned, we then use around 10000 records to estimate the 4 sets of parameters that we mentioned earlier. In our study, SVR is tuned and evaluated using 10-fold cross-validation grid search to find the best parameter settings. The features used include the time of the day, day of the week, and passenger demand and travel time data of the most recent two trips. Once training is completed, a lookup table should be created for each set of estimated variables (i.e. $r(k, p)$, $a(p, k)$, $T(k, k + 1, p)$ and $RT(d_0, i)$) to speed up the search during the optimisation stage. Note that when the data size is significantly bigger, some of the latest learning methods, like advanced tree-ensemble methods or deep-learning methods, should be used instead if higher quality of forecasts is required.

4.5. Experimental environment

The proposed algorithms are implemented in Java. Both the parameter forecast and the optimisation modules are implemented on an iMac with Intel i7 CPU and 16G RAM. The system is supported by LibSVM Library (Chang, 2016) and MOEA Framework 2.12 (MOEA.Org, 2017) with default parameters.

5. A real-life case study

As mentioned previously, a case study (referred as **r-bus-01**) is conducted to evaluate the practicality and feasibility of the model and the algorithm. A circular bus line with 22 stops is selected. The time to make a full circular trip ranges from 38 minutes to 68 minutes, indicating high level of uncertainties in travel time. The problem instance we test is a 4-hour long schedule from 4:30pm to 8:30pm on 2016/03/26. Since the length of the period is set to $\tau = 30$ mins, the planning horizon, hence, contains 8 periods (i.e. $p = 1, 2, \dots, 8$). The boundaries of the headways are set to $[5, 20]$ minutes, giving plenty of room for optimisation. The total number of buses is set to 15 in our experiments. The bus capacity is set to 40 and the maximum bus load rate is $\eta = 1.2$, which means, at any time during the operation, the maximum number of passengers on a bus should not exceed $40 \times 1.2 = 48$. The operation cost of each bus trip is set to 250. For comparison reasons, a benchmark static instance is also created based on this real-life instance, in which the uncertain parameters (travel time and arrival rates) are estimated by their mean values. The benchmark instance is then solved by the NSGAI approach and the resulting solution is then evaluated in the dynamic setting in which the uncertain parameters take their true values. Figure 3 depicts Pareto frontier solutions for both instances.

Firstly, it is clear that the two objectives adopted in the model are in conflict with each other. The total passenger waiting time is almost inversely proportional to the bus service operation cost. Compared with the benchmark setting (i.e. a static model), the proposed method is able to produce about 5% improvement over the total waiting time with the same operation cost. Note that at the cost level 2640, the static model leads to an infeasible solution because of violations of bus capacity constraint. This is a major issue for many statically generated solutions, which are often over-optimised without taking account of the uncertainties in a dynamically changing environment. The solutions either become not executable after some time or their quality drops substantially, as illustrated in Figure 3. It should be noted that the potential users should read the results with caution. According to No Free Lunch Theorem, it is possible to artificially design specific problem instances for which static algorithm

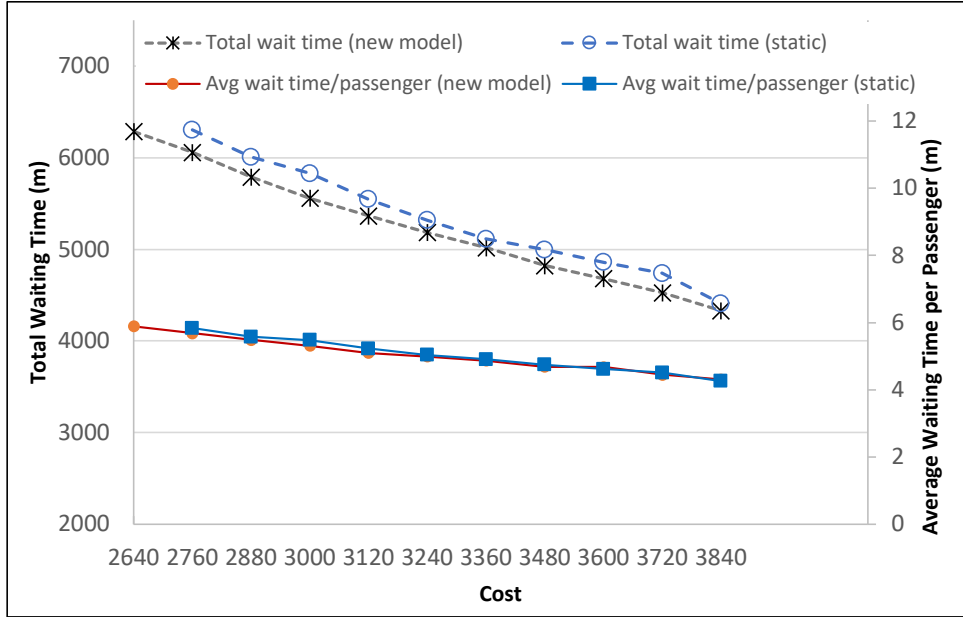


Figure 3.: The pareto frontier of the solutions returned by the proposed new model for a real-life instance. The results from the static model are also included for comparison purposes.

would perform better. Our approach, however, is aimed to address bus scheduling scenarios with stochastic demand and travel time in real-life, like the one in this paper.

One of the obvious advantages of the proposed multi-objective optimisation model is the flexibility in selecting one of the non-dominated solutions by the decision makers. In addition, with additional investment (by running more frequent bus services), the average passenger waiting time is also reduced, with the largest average waiting time 5.8 mins and the smallest average waiting time of 4.4 mins. These values are significantly smaller than the maximum permitted headways (set to 20 mins in this case). This clearly shows that the maximum and minimum dispatch headways at the terminal station are not accurate indicators of the true quality of the bus services at different stops. In contrast, the total waiting time (or the average waiting time) is a much better alternative as a good trade-off can be achieved between the service quality and the cost.

Figure 4 illustrates the capability of the model in adapting different levels of pas-

senger arrival rates at different time by adjusting the headways at stop 16. Note that, except the first few stops and the last few stops, similar patterns can be observed at other stops. Although the control variables in our model are the headways at the terminal station, through accurate estimations of travel time at different periods and passenger arrival rates, the optimisation through the model sets smaller headways at a stop when its arrival rate is high and vice versa. The headways are between 5 mins and 13 mins. Note that different stops may experience different demand characteristics (in terms of arrival rates and alighting ratios), and the best headways at different stops are inter-correlated. Setting the best headways at a particular stop should take account of both the demands at the current stop and the nature of the demands in the subsequent stops along the bus line. This is one of the advantages of the proposed method compared to dynamic-control methods existing in literature.

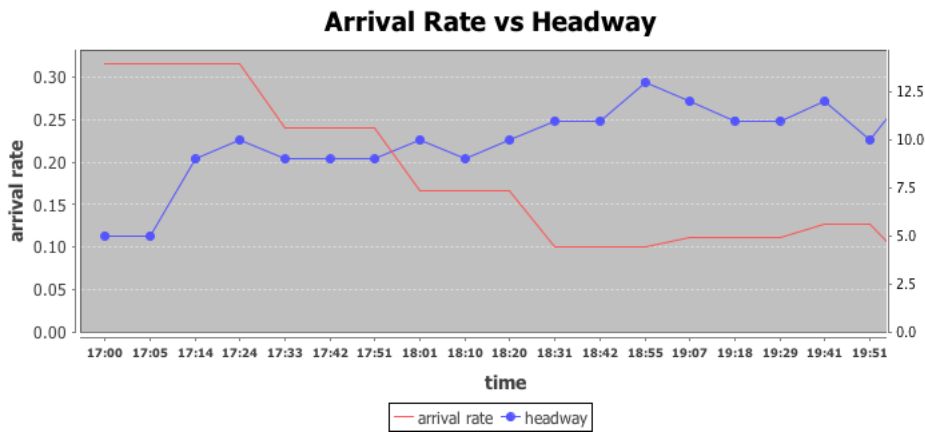


Figure 4.: The adaptation of headways to different arrival rates for stop 16.

Figure 5 shows the relationship between the bus load and the headways at stop 16. In most cases, they follow the same trend because higher headways will in general lead to more passengers arriving at the stop and hence more passengers boarding the bus. However, once the bus load approaches its maximum capacity, the headway should be reduced so that the capacity constraints are ensured. Of course, as mentioned early, other factors in both the current stop and the subsequent stops should be considered for the determination of headways.

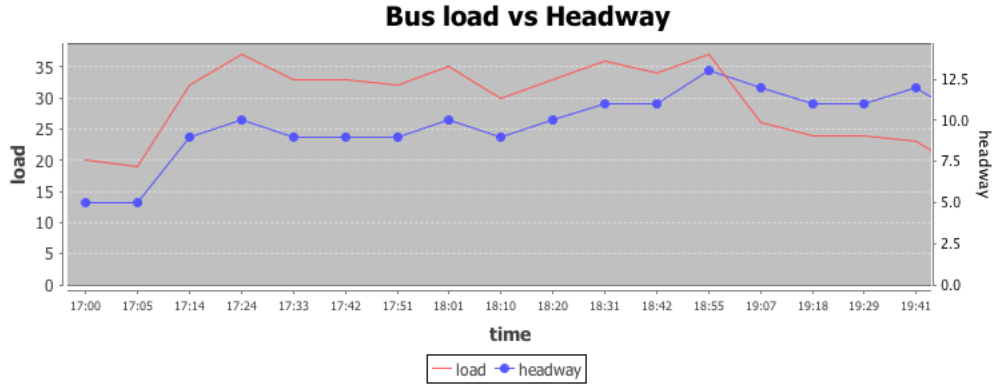


Figure 5.: Bus load vs. headways for stop 16.

6. Empirical experiments on simulated data

Although the real-life problem instances help check the feasibility of the model and the solution from a practical point of view, further evaluations for problem instances with specially controlled scenarios are required.

6.1. Data instances

To further verify the effectiveness of the proposed approach, another two problem instances are generated to simulate two controlled-passenger-demand patterns. The basic settings are the same as the real-life instance, i.e. the bus route has 22 stops in total and the scheduling time window starts at 4:30pm ($p = 1$) and ends at 8:30pm. For a period of 30 minutes, the planning horizon hence covers 8 periods. The travel time between any two successive stops remains the same as before but the passenger demands (arrivals and alighting) are generated in a controlled manner.

In the first generated instance (referred to as **g-bus-01** in the subsequent sections), both passenger arrival rates and alighting proportions are assumed time-independent but vary among stops. The passenger arrival rate $r(p, k) = \bar{r}(k)$ is set to 0.4 person/min at stop 1 and increases linearly until stop 11 where its value peaks at 1.0 person/min. The arrival rate then decreases linearly over the next 11 stops and reduces to 0 at stop 22. The alighting proportion $a(p, k) = \bar{a}(k)$ starts at 0 at stop 1 and then increases linearly and reaches maximum of 0.5 at stop 21 (the penultimate stop). The alighting

rate at terminal stop $\bar{a}(22)$ is set to 1, ensuring all passengers get off at the terminal stop.

In the second generated instance (denoted as **g-bus-02**), the arrival rates and alighting ratios are time-dependent. To model peak demand during rush hours, a time dependent coefficient function $\beta(p)$ is used. The starting period has a normal demand with a coefficient value $\beta(1) = 1.0$ and the last planning period has a below-than-normal demand $\beta(8) = 0.8$. A small peak demand is generated at period 3 with $\beta(3) = 1.3$. The coefficients in other periods are linearly interpolated. Finally, a random noise term $N(0, 0.15)$ is added to represent the fluctuations of demand over time where $N(0, 0.15)$ is a normal distribution with the mean of 0 and the variance of 0.15. Therefore, the final arrival rate function becomes $r(p, k) = \bar{r}(k) \times \beta(p) + N(0, 0.15)$. Similarly, the alighting ratio function is defined as $a(p, k) = \bar{a}(k) \times \beta(p) + N(0, 0.05)$ except for the first stop and the last stop where the alighting rate is set to 0 and 1, respectively.

In summary, the first instance aims to simulate a very stable passenger demand across the entire planning horizon. Although some stops (e.g. stop 11) are busier than others, the demands are not time-dependent. In the second instance, however, the demand at each stop changes over time and is also subject to random fluctuations. Our intention is to see how our proposed approach adapts itself to these two different scenarios.

6.2. Simulation results

Let us first focus on a slightly simpler problem case where the bus passenger demands are time-independent (**g-bus-01**). We pick two stops, stop 12 and stop 17, to observe and monitor the solutions by our multi-objective methods. Recall that demands tend to be much higher at stop 12 than stop 17. We want to observe how our approach handles busy and less-busy stops respectively. Out of all the Pareto front solutions with respect to our objective O_1 and O_2 , we select two solutions. The first solution emphasizes better quality of services (QoS) by prioritising objective O_1 while the second solution seeks to reduce the service operation cost O_2 more. Figure 6 illustrates bus schedule details across the planning horizon (4:30pm to 8:30pm). Recall that the bus load reflects the accumulated passengers from the current and all preceding stops

while the headway is the time gap between two successive bus trips. Therefore, a smaller headway implies shorter waiting time and better quality of services. On the other hand, maintaining a relatively high bus load (still within its capacity) can be considered as a reliable indicator of economical operations of bus services. We can see that our multi-objective approach can indeed provide good alternative solutions for decision makers. When the quality of the service is considered more important (Figure 6(a) and (b)), the headways tend to be smaller, implying shorter waiting time at stops for passengers. If the optimisation focuses more on cost-effective delivery of services because of tight budget, headways are getting higher (see Figure 6(c) and (d)). Since the bus load is generally proportional to the headways, the average bus load in Figure 6(c) and (d) is much higher (but still within the limit of 48 people at any time). This shows the benefits of a multi-objective optimisation approach in this work, which gives decision makers better insights about the conflicting objectives and constraints in bus scheduling problem and good solutions with different trade-offs between these objectives.

Note that because of time-dependent travel time between stops, the headways at two different stops (12 and 17) follow a similar pattern but are not identical, reflecting a fundamental challenge in this problem. Sometimes, a more frequent dispatch at the departure terminal may be an effort to address the suddenly increased travel time at some segments of the bus route, so that headway spikes can be avoided and buses are distributed relatively evenly along the bus route.

In the case of fluctuating passenger demands over time (instance g-bus-02), we are interested in how our multi-objective approach adapts to these changes. Figure 7 shows the relationship between the arrival rate/bus load and headways at stop 12. Again two possible solutions with different trade-offs are considered. Figure 7(a) and Figure 7(b) show a solution that focuses more on the quality of the services and Figure 7(c) and (d) plot a solution that emphasises more on the minimisation of the operation cost.

In general, it can be seen from Figure 7(a) and (c) that smaller headways are used for periods with higher arrival rates. Once the arrival rate declines, the headways increase from the range of [5, 8] to the range of [7, 9], so that the system becomes more cost-efficient. When the budget is tight (Figure 7(c)), this trend is more evident, i.e. the

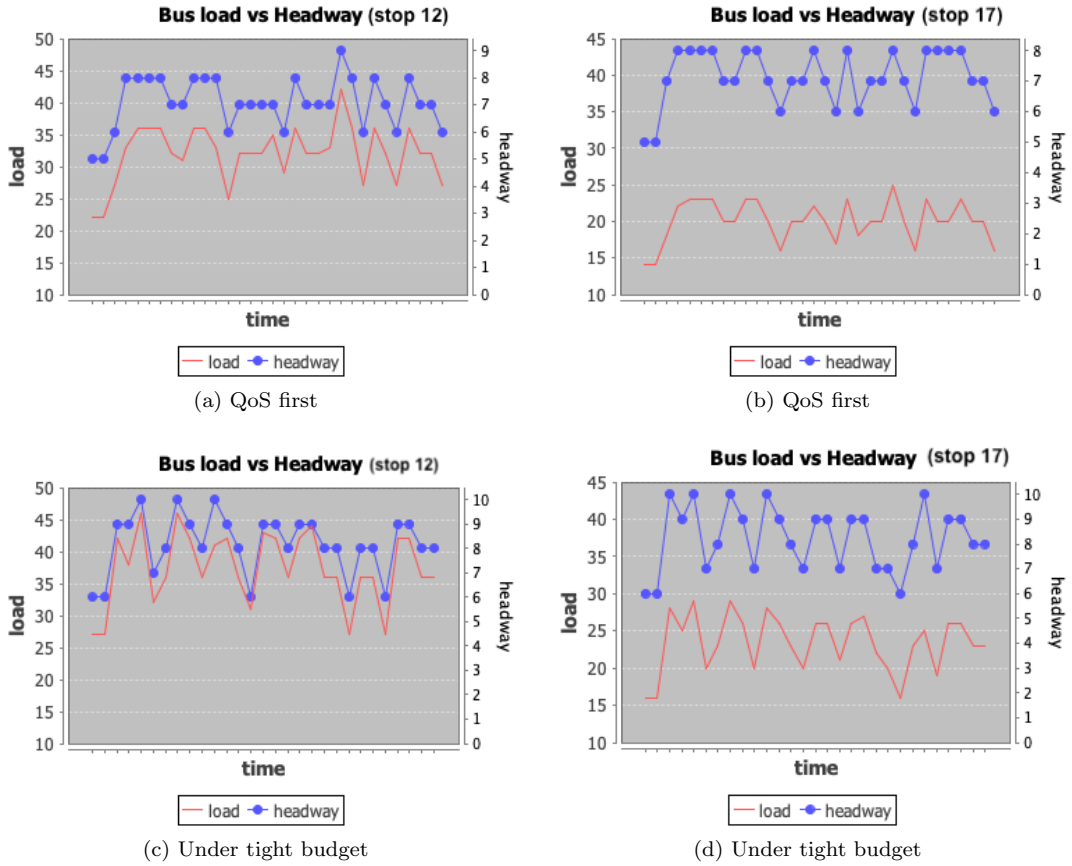


Figure 6.: An illustration of the solutions under time-independent parameters with different optimisation priorities (for instance g-bus-01). (a) and (b) focus on better quality of services (QoS) while solutions in (c) and (d) try to minimise the operation cost.

headways rise to the range of [9, 11]. In terms of the bus load, we can observe patterns similar to those in Figure 6. For all cases, we can see the optimisation method is able to produce solutions that satisfy the constraints (e.g. capacity) while giving decision makers flexibility to choose among solutions of different trade-offs.

6.3. Managerial insights for practioners

Uncertainties exist in many complex systems, regardless whether large scale IoT devices are deployed or not for monitoring and data capturing purposes. It is a common misconception that the provision of more data would readily lead to better planning and scheduling in production and transportation. This is because real-life data arrives in huge volume and in high dimensions. Therefore, without advanced tools, captured

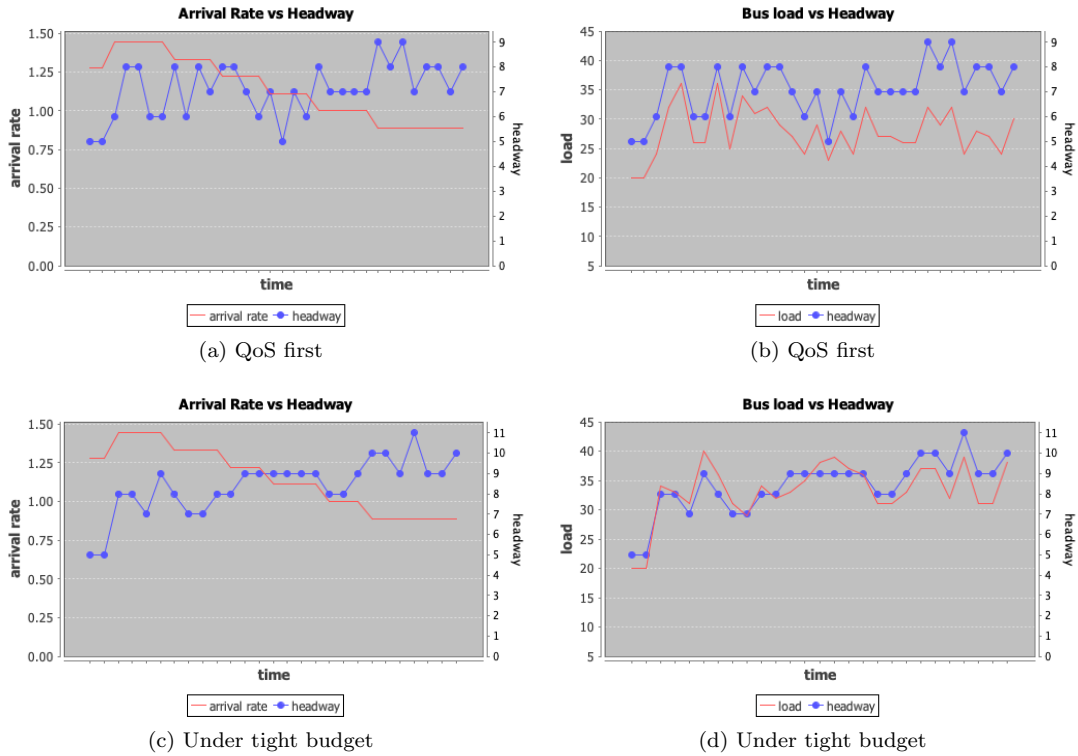


Figure 7.: An illustration of the solutions at a busy stop (stop 12) under time-dependent parameters with different optimisation priorities. (a) and (b) focus on better quality of services (QoS) while solutions in (c) and (d) try to minimise the operation cost.

real-life data is often beyond human’s capability to handle directly. In reality, when the system is complex and uncertainties exist, managers and production planners should embrace advanced decision support systems, especially those that integrate traditional model based optimisation with data mining techniques for better decision making. The goal is to respond to different scenarios appropriately by both avoiding myopic decisions and achieving satisfactory level of robustness. When the optimisation involves multiple stakeholders, it is worth considering the multi-objective optimisation.

7. Discussions and future research

This paper is concerned with a novel framework to integrate analytics and machine learning in addressing uncertainties and conflicting objectives in challenging real-life optimisation problems. The research take bus scheduling as a case study but the

methodologies can be adapted to other problems, like those in production and logistics. For example, similar to bus services, one of the most challenging issues in production scheduling is the variations related to product orders, and job processing times. Our investigation found that most of the existing approaches formulate the bus scheduling problem either as a static one solved by traditional optimisation methods, or as a dynamic one solved by some machine-learning methods. While solutions obtained from the static formulations are too rigid and brittle in the presence of uncertainties, the solutions from dynamic models are also problematic because of myopic decisions and lack of exploitation of problem structures. In this paper, we present a novel multi-objective bus dispatching optimisation model that can be repeatedly resolved under a rolling scheduling horizon. The machine learning modules take inputs of both the historical bus trip data and the current traffic conditions that may be obtained through sensors and probing vehicles/buses currently in the road network. A real-life case is then used to evaluate the practicality and effectiveness of the proposed system. The proposed model is further evaluated for two artificially generated instances. The experimental results demonstrate both the feasibility and the flexibility of our proposed method.

We believe the proposed framework for handling uncertainties and conflicting objectives can be adapted to other scheduling problems of similar characteristics. In future, if sufficient real-life data become available, it will be interesting to extend this framework to various optimisation problems in production and logistics that are constantly subject to changes.

Another interesting direction for future research would be extension of our bus scheduling model with joint optimisation of headways, deadheading and express bus lines to deal with extreme uncertainties when the variations become significantly large and optimising headways alone may not be sufficient.

Funding

This work is supported by the National Natural Science Foundation of China (Grant No. 71471092), China Scholarship Council (Grant No. 201706295033), Natural Science

Foundation of Zhejiang Province (Grant No. LR17G010001) and Ningbo Municipal Bureau of Science and Technology (Grant No. 2017D10034).

References

- Andres, M., & Nair, R. (2017, OCT). A predictive-control framework to address bus bunching. *Transportation Research Part B-Methodological*, *104*, 123-148.
- Bai, R., Li, J., Atkin, J. A., & Kendall, G. (2014). A game theoretic approach for the taxi scheduling problem with street hailing. *Journal of the Operational Research Society*, *65*(10), 1501-1510.
- Bai, R., Xue, N., Chen, J., & Roberts, G. W. (2015). A set-covering model for a bidirectional multi-shift full truckload vehicle routing problem. *Transportation Research Part B-Methodological*, *79*, 134-148.
- Berrebi, S. J., Watkins, K. E., & Laval, J. A. (2015). A real-time bus dispatching policy to minimize passenger wait on a high frequency route. *Transportation Research Part B-Methodological*, *81*(2, SI), 377-389.
- Bowman, L. (1981). Service Frequency, Schedule Reliability And Passenger Wait Times At Transit Stops. *Transportation Research Part A-Policy and Practice*, *15*(6), 465-471.
- Chang, C.-C. (2016). *LIBSVM: A library for support vector machines*. Retrieved 2017-08-26, from <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Chen, X., Yu, L., Zhang, Y., & Guo, J. (2009). Analyzing urban bus service reliability at the stop, route, and network levels. *Transportation research part A: policy and practice*, *43*(8), 722-734.
- Chow, A. H. F., Li, S., & Zhong, R. (2017). Multi-objective optimal control formulations for bus service reliability with traffic signals. *Transportation Research Part B-Methodological*, *103*(SI), 248-268.
- Cortés, C. E., Jara-Díaz, S., & Tirachini, A. (2011). Integrating short turning and deadheading in the optimization of transit services. *Transportation Research Part A: Policy and Practice*, *45*(5), 419-434.
- Daganzo, C. F. (2009). A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons. *Transportation Research Part B: Methodological*, *43*(10), 913-921.
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective

- genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2), 182-197.
- Eberlein, X. J., Wilson, N. H., Barnhart, C., & Bernstein, D. (1998). The real-time deadheading problem in transit operations control. *Transportation Research Part B: Methodological*, 32(2), 77-100.
- Feng, X., Saito, M., & Liu, Y. (2016). Improve urban passenger transport management by rationally forecasting traffic congestion probability. *International Journal of Production Research*, 54, 3465-3474.
- Gkiotsalitis, K., & Kumar, R. (2018). Bus Operations Scheduling Subject to Resource Constraints Using Evolutionary Optimization. *Informatics-Basel*, 5(1).
- Guo, X., Wu, J., Zhou, J., Yang, X., Wu, D., & Gao, Z. (2019). First-train timing synchronisation using multi-objective optimisation in urban transit networks. *International Journal of Production Research*, 57, 3522-3537.
- Habibi, M. K. K., Battaa, O., Cung, V.-D., Dolgui, A., & Tiwari, M. K. (2019). Sample average approximation for multi-vehicle collection/disassembly problem under uncertainty. *International Journal of Production Research*, 57, 2409-2428.
- Huang, Z., Li, Q., Li, F., & Xia, J. (2019). A novel bus-dispatching model based on passenger flow and arrival time prediction. *IEEE Access*, 7, 106453-106465.
- Ibarra-Rojas, O. J., Delgado, F., Giesen, R., & Munoz, J. C. (2015). Planning, operation, and control of bus transport systems: A literature review. *Transportation Research Part B-Methodological*, 77, 38-75.
- Li, H., & Zhang, Q. (2009). Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II. *IEEE Transactions on Evolutionary Computation*, 13(2), 284-302.
- Li, X., Bai, R., Siebers, P.-O., & Wagner, C. (2019). Travel time prediction in transport and logistics: Towards more efficient vehicle gps data management using tree ensemble methods. *VINE Journal of Information and Knowledge Management Systems*, 49(3), 277-306.
- Liu, Z., Yan, Y., Qu, X., & Zhang, Y. (2013). Bus stop-skipping scheme with random travel time. *Transportation Research Part C: Emerging Technologies*, 35, 46-56.
- MOEA.Org. (2017). *Moea framework: A free and open source java framework for multiobjective optimization*. <http://moeaframework.org>.
- Mokhtari, N., & Ghezavati, V. (2018). Integration of efficient multi-objective ant-colony and a heuristic method to solve a novel multi-objective mixed load school bus routing model.

Applied Soft Computing, 68, 92 - 109.

Shen, Y., Xu, J., & Li, J. (2016). A probabilistic model for vehicle scheduling based on stochastic trip times. *Transportation Research Part B-Methodological*, 85, 19-31.

Sun, Q., & Zhang, X. (2016). Optimization model and algorithm design of bus lines non-fixed headways problem based on passenger arrival rates. *Open Journal of Transportation Technologies*, 5(1), 7-16.

Vapnik, V. (1998). *Statistical learning theory*. John Wiley and Sons, Inc., New York.

Xu, W.-J., He, L.-J., & Zhu, G.-Y. (2019). Many-objective flow shop scheduling optimisation with genetic algorithm based on fuzzy sets. *International Journal of Production Research*, 57, 1-12.

Zhang, J., Shen, D., Tu, L., Zhang, F., Xu, C., Wang, Y., . . . Li, Z. (2017, NOV). A Real-Time Passenger Flow Estimation and Prediction Method for Urban Bus Transit Systems. *IEEE Transactions On Intelligent Transportation Systems*, 18(11), 3168-3178.

Zhang, S., & Lo, H. K. (2018). Two-way-looking self-equalizing headway control for bus operations. *Transportation Research Part B: Methodological*, 110, 280-301.