2018 21st International Conference on Intelligent Transportation Systems (ITSC)
Maui, Hawaii, USA, November 4-7, 2018

Reliable bus dispatching times by coupling Monte Carlo evaluations with a Genetic Algorithm

Konstantinos Gkiotsalitis *Member IEEE*[1] and Francesco Alesiani *Member IEEE*[2]

***Abstract*— Bus operators plan the dispatching times of their daily trips based on the average values of their travel times. Given the trip travel time uncertainty though, the performance of the daily operations is different than expected impacting the service regularity and the expected waiting times of passengers at stops. To address this problem, this work develops a model that considers the travel time uncertainty when planning the dispatching times of trips. In addition, it introduces a minimax approach combining Monte Carlo evaluations with a Genetic Algorithm for computing dispatching times which are robust to travel time variations. This approach is tested in a circular bus line of a major bus operator in Asia Pacific (APAC) using 4 months of Automated Vehicle Location (AVL) and Automated Fare Collection (AFC) data for analyzing the travel time uncertainty and computing robust dispatching times. In addition, 1 month of data is used for validation purposes demonstrating a potential service regularity improvement of 5.5% in the average case and ≃22% in worst-case scenarios.**

Keywords: Tactical Planning; Timetable reliability; Travel time uncertainty; Robust bus scheduling.

I. INTRODUCTION

The availability of Automated Vehicle Location (AVL), Automated Fare Collection (AFC) and social media data has enabled transport authorities to monitor the adherence of bus operations to their intended schedules [1], [2], [3]. Using data from onboard units and fare payment systems, transport authorities provide incentives to bus operators to improve their performance [4]. For instance, bus operators in Singapore receive up to 2,000 Singaporean dollars per month for every 0.1-minute improvement of the service regularity [5].

These types of data can be used for improving the tactical planning stages of (i) frequency setting; (ii) timetable design (where the dispatching times of all daily trips are determined); and (iii) vehicle and crew scheduling [6], [7], [8]. Given that transport authorities incentivize the improvement of the service regularity, generating robust timetables which are robust to travel time variations is a significant step towards meeting the transport authorities' objectives and improving the daily operations.

The survey paper of [9] on public transport planning concluded that using the insights of historical AVL and AFC data is one of the highest potential areas for improving the planning of the daily operations. The problem of timetable unreliability was studied also in the work of [10] which showed that the daily bus operations deteriorate significantly when the actual travel times of buses vary by more than 30% from their expected values.

This work focuses on the robust bus timetabling[1] problem by introducing a minimax approach for addressing the performance degradation of the daily operations due to the travel time variations. The robust timetables can facilitate the bus operators and help them improve their performance levels without increasing the operational costs (i.e., increasing the fleet size or hiring more bus drivers).

II. RELATED WORKS

Several works have jointly addressed the bus frequency settings problem and the bus timetabling problem [11], [12], [13], [14], [15]. Nevertheless, in most cases the optimization of the timetable of a bus line is perceived as a standalone problem and is generally decoupled from the from the frequency settings problem [16], [17], [18].

When a bus operator plans the dispatching times of the daily trips of a bus line for determining its timetable, the expected trip travel times are considered as deterministic. This results in timetables that do not perform well in practice because the bus operations are unstable in nature given the stochasticity of the trip travel times [19]. For this reason, several works focus on defining optimal slack times for every daily trip [20], [21], [22]. These slack times can be used as buffer times that can absorb the unexpected delays of bus trips due to the road traffic without delaying the dispatching of future trips which are operated by the same buses.

Typical problem objectives when determining the dispatching times of trips in high-frequency services are the mitigation of bus headway variations for improving the service regularity [23] and the generation of practical timetables. For instance, [24], [25], [26] and [27] proposed to generate timetables where the dispatching times of all daily trips are as evenly-spaced as possible for improving the service regularity and addressing the practicability issue.

Closer to our work, [28], [19], [29] and [30] have considered the stochastic nature of the trip travel times during the daily operations. However, [28] was focused on the bus line synchronization problem, [29] on the reliable frequency settings problem and [30], [19] on the bus holding problem during the actual operations without addressing the dispatching time planning problem.

[1]Konstantinos Gkiotsalitis, PhD, is Assistant Professor at the University of Twente (Center for Transport Studies), 7522LW Enschede, The Netherlands k.gkiotsalitis@utwente.nl

[2]Francesco Alesiani, PhD, is Senior Researcher at the Data Science Applications Group of NEC Laboratories Europe, Kurfuerstenanlage 36, 69115 Heidelberg, Germany francesco.alesiani@neclab.eu

[1]since the timetable of a bus line is generated based on the dispatching times of all daily trips, in this work we use the terms "timetable optimization" and "dispatching time optimization of all daily trips" interchangeably

Focusing specifically on the problem of planning robust dispatching times for all daily trips, this work contributes to the state-of-the-art by: (i) incorporating the observed uncertainty of the bus travel times using historical AVL and AFC data; (ii) developing a model for timetable optimization which can handle uncertainty; (iii) introducing a Genetic Algorithm coupled with Monte Carlo evaluations for addressing the NP-hardness of the robust timetabling problem; and (iv) analyzing the performance of robust timetables in scenarios with low and high travel time variability.

III. PROBLEM FORMULATION

A. Modeling the movement of buses

The modeling part of this work relies on the following assumptions:

- Buses that serve the same line do not overtake[2] each other (first-in-first-out rule);
- Given the focus on high-frequency services, passenger arrivals at stops are random because the passengers cannot coordinate their arrivals with the arrival times of buses (see [31]);
- The capacity of buses can accommodate the passenger demand;
- The total travel time of a bus trip is affected by the travel times between stops and the dwell times at stops;
- Passengers use different door channels for boardings and alightings.

Before proceeding to the modeling, we introduce the following nomenclature:

NOMENCLATURE

N	Set of daily bus trips of a bus line, where each daily trip $n \in N$
S	Set of bus stops of a bus line, where each bus stop $s \in S$
δ_n	The originally planned dispatching time of a bus trip n
$\mathbf{x}$	Vector where each $x_n \in \mathbf{x}$ denotes the dispatching time deviation of a bus trip n from δ_n (decision variable of the robust dispatching time optimization problem)
Z	A set of all potential dispatching time deviation options (all options are in minutes)
$\theta_{n,s}(\mathbf{x})$	The departure time of bus trip n from stop s
$d_{n,s}(\mathbf{x})$	The dwell time of bus trip n at stop s
$t_{s,\tau}$	Uncertain parameter denoting the travel time between stop s and $s+1$ when a bus trip departs from stop s at time τ
$h_{n,s}(\mathbf{x})$	The headway between bus trip n and its preceding trip $n-1$ at stop s
$a_{n,s}(\mathbf{x})$	The arrival time of bus trip n at stop s
γ	The required time per passenger boarding
$b_{s,\varrho}$	The total number of passenger boardings at stop s during the ϱ^{th} hour of the day
$\mathcal{T}$	A pre-defined time before which the last trip of the day should have been already dispatched
ψ	The required layover time after completing a bus trip

The dispatching times, $(\delta_1, ..., \delta_n, ...)$, of the daily trips of a bus line, where $N = \{1, ..., n, ...\}$ is the set of all daily trips, can be initially considered as evenly-spaced across the day. The decision variable of the robust dispatching time planning problem is a vector of the dispatching time deviations $\mathbf{x} = \{x_1, ..., x_n, ...\}$ from the evenly-spaced dispatching times.

To avoid impractical dispatching times which cannot be implemented in practice, we do not allow significant dispatching time deviations from the evenly-spaced option. For this reason, each dispatching time deviation x_n , $n \in N$ can take values from a discrete set $Z = \{z_1, ..., z_i, ...\}$ which contains a limited number of dispatching time deviation options.

For a bus line that serves $S = \{1, ..., s, ...\}$ bus stops, the travel time for traversing a segment between any stop pair s and $s+1$ can be different from time to time. For instance, the travel time for traveling from one stop to another can be significantly higher during peak hours compared to the off-peaks. This effect is captured in this work by splitting the daytime into 1440 minutes and allowing the travel time for traversing a link[3] to vary based on the minute of the day the link is traversed. Therefore, the link travel times are modeled as a matrix $\mathbf{T} \in \mathbb{R}_+^{(|S|-1)\times 1440}$ where each $t_{s,\tau} \in \mathbf{T}$ denotes the expected link travel time for traversing a link that connects stops s and $s+1$ when a bus starts to traverse the link within the τ^{th} minute of the day.

The departure time, $\theta_{n,s}(\mathbf{x})$, of any bus trip $n \in N$ from any stop $s \in S$ can be calculated based on the link travel times and the dwell times at stops as:

$$\theta_{n,s}(\mathbf{x}) = \begin{cases} \delta_n + x_n \text{ if } s = 1 \\ \theta_{n,s-1}(\mathbf{x}) + t_{s-1,\tau} + d_{n,s}(\mathbf{x}), \forall s \in S \setminus \{1, |S|\} \end{cases} \tag{1}$$

where eq.1 is a recursive formula. In eq.1, if a bus trip departs from the first stop, $s = 1$, its departure time is $\delta_n + x_n$. If departs from another stop, $s \in S \setminus \{1, |S|\}$, then its departure time is equal to the departure time from the previous stop, $\theta_{n,s-1}(\mathbf{x})$, plus the link travel time for traveling from stop $s-1$ to stop s plus the dwell time at stops s which is denoted as $d_{n,s}(\mathbf{x})$. In eq.1, the link travel time from stop $s-1$ to stop s is $t_{s-1,\tau}$ and the minute τ within which the bus trip n starts to traverse link $s-1$ is:

$$\tau = \lfloor \theta_{n,s-1}(\mathbf{x}) \rfloor \tag{2}$$

where $\theta_{n,s-1}(\mathbf{x})$ is expressed in minutes.

The dwell time $d_{n,s}(\mathbf{x})$ of a bus trip n at stop s depends on the number of passenger boardings and alightings. In this study, the dwell time is not linked to the number of alightings for modeling simplification purposes since boarding times exceed alighting times in most cases [32].

[2]a common assumption used in related works (see [19])

[3]in this work, a link is the segment between two consecutive bus stops

At this point, we should note that in actual operations the alighting times might impact the dwell times at stops in the case where passengers use the same door channels for boardings and alightings [33]. Nevertheless, the main focus of this study is bus services that use the front door for boardings and all other doors for alightings. In such case, assuming that passenger arrivals at stops are random, the dwell time of a bus trip n at one stop s can be expressed as:

$$d_{n,s}(\mathbf{x}) = \frac{b_{s,\varrho}}{60(\text{min/hour})}\left(\theta_{n,s}(\mathbf{x}) - \theta_{n-1,s}(\mathbf{x})\right)\gamma \quad (3)$$

where γ is the required time per passenger boarding and $b_{s,\varrho}$ the hourly passenger boardings at stop s assuming that bus trips n and $n-1$ depart from stop n within the ϱ^{th} hour of the day.

The headway $h_{n,s}(\mathbf{x})$ between bus trip n and its preceding bus trip $n-1$ at stop s is:

$$h_{n,s}(\mathbf{x}) = a_{n,s}(\mathbf{x}) - a_{n-1,s}(\mathbf{x}) \quad (4)$$

where $a_{n,s}(\mathbf{x}), a_{n-1,s}(\mathbf{x})$ are the arrival times of bus trips n and $n-1$ at stop s. The arrival time $a_{n,s}(\mathbf{x})$ is equal to the departure time from the previous stop $s-1$ plus the link travel time from stop $s-1$ to stop s. For instance,

$$a_{n,s}(\mathbf{x}) = \theta_{n,s-1}(\mathbf{x}) + t_{s-1,\tau} \;\; \forall s \in S/\{1\} \quad (5)$$

Equations 1, 5 capture the movements of all bus trips of a bus line and can return the departure, arrival and dwell times at stops for all daily trips given the travel time values and the dispatching time deviations, **x**, from the originally planned dispatching times.

B. Modeling the constraints

1) Layover time constraints: The layover time for a bus that finished one bus trip is the minimum required time before starting its next trip. The layover time is equal to the required time for moving from the last stop of the finished trip to the first stop of the next trip (known as deadheading time) plus the recovery time for the bus driver (in most cases, bus drivers must take a short break after completing a bus trip).

If one bus operates trips n' and n one after the other and a layover time $\psi \geq 0$ is required after finishing trip n', then the layover constraint is:

$$a_{n',|S|}(\mathbf{x}) + d_{n',|S|}(\mathbf{x}) + \psi \leq \delta_n + x_n \quad \forall n \in N \setminus \{1\} \quad (6)$$

The constraint of eq.6 dictates that the dispatching time of a trip n, $\delta_n + x_n$, should be equal or greater than the arrival time of the previous trip n' which was operated by the same bus at the last bus stop $|S|$ plus the dwell time at stop $|S|$ plus the required layover time ψ. Evidently, the total number of layover time constraints can be up to $|N|-1$.

2) Schedule sliding constraint: In order to maintain the duration of the daily operations, the last trip of the day, $n = |N|$, should be dispatched before a pre-defined time, $\mathcal{T}$. This ensures that the daily operations will not be prolonged.

If the last trip of the day is operated after the completion of trip n^*, then:

$$a_{n^*,|S|}(\mathbf{x}) + d_{n^*,|S|}(\mathbf{x}) + \psi \leq \mathcal{T} \quad (7)$$

C. Objective function

As discussed in the introduction section, transport authorities provide monetary incentives to bus operators for improving the service regularity. As stated in [23], a typical key performance indicator (KPI) which measures the performance of high-frequency services is the excess waiting times of passengers at stops. In this KPI, the actual headways of buses should vary as little as possible from the desired headways to ensure that the actual passenger waiting times are close to the expected ones.

Following this KPI, the objective of our robust dispatching time optimization problem is to minimize the average squared deviation of the actual headways from their desired values at all bus stops:

$$f(\mathbf{x}) = \sqrt{\frac{1}{|S|(|N|-1)}\sum_{s\in S}\sum_{n\in N\setminus\{1\}}\left(h_{n,s}(\mathbf{x}) - h_s^*\right)^2} \; (\text{min}) \quad (8)$$

where h_s^* is the desired/ideal headway at any stop $s \in S$.

IV. ROBUST OPTIMIZATION OF TRIP DISPATCHING TIMES

A robust optimization of the dispatching times of all daily bus trips focuses on finding a solution of dispatching time deviations, **x**, which is robust to the inherent uncertainty of the link travel times **T**.

The minimax method [34] is one of the most commonly used methods in robust optimization. The minimax method is a decision-making method in which many problem solutions are ranked based on their worst-case performance in the presence of uncertainty and the optimal decision is the one with the slightest worst outcome.

If for a link travel time, $t_{s,\tau}$, we have a set of historical observations from AVL data which are stored in a set $\mathcal{B}_{s,\tau}$, then the link travel time parameter, $t_{s,\tau}$, is uncertain and can take any value from the set $\mathcal{B}_{s,\tau}$. By allowing the parameter $t_{s,\tau}$ to take any value from the set $\mathcal{B}_{s,\tau}$, one can examine the worst-case performance of a solution **x** that determines the dispatching times of the daily trips.

The objective of the robust optimization is to find the dispatching time modifications, **x**, that minimize the worst-case value of the objective function given the uncertainty of the link travel time parameters as it is expressed below:

$$\min_{\mathbf{x}}\max_{\mathbf{T}} f(\mathbf{x}, \mathbf{T}) \quad (9)$$

where each $t_{s,\tau} \in \mathcal{B}_{s,\tau}$ and

$$f(\mathbf{x},\mathbf{T})=\sqrt{\frac{1}{|S|(|N|-1)}\sum_{s\in S}\sum_{n\in N\setminus\{1\}}\Big(h_{n,s}(\mathbf{x},\mathbf{T})-h^*_s\Big)^2} \tag{10}$$

Therefore, the robust dispatching time optimization problem considering the physical constraints and the constraints related to the movement of buses can be formulated as:

$$\begin{aligned}
&\min_{\mathbf{x}}\max_{\mathbf{T}} f(\mathbf{x},\mathbf{T})\\
&\quad \text{s.t. Equations 1-5}\\
&\qquad a_{n',|S|}(\mathbf{x},\mathbf{T})+d_{n',|S|}(\mathbf{x},\mathbf{T})+\psi\le(\delta_n+x_n),\\
&\qquad\qquad\qquad\qquad\qquad \forall n\in N\setminus\{1\}\\
&\qquad a_{n^*,|S|}(\mathbf{x},\mathbf{T})+d_{n^*,|S|}(\mathbf{x},\mathbf{T})+\psi\le\mathcal{T}\\
&\quad \text{where } t_{s,\tau}\in\mathcal{B}_{s,\tau},\ \forall s,\tau\\
&\qquad x_n\in Z,\ \forall n\in N\\
&\qquad x_1=0
\end{aligned} \tag{11}$$

The robust optimization problem of eq.11 has a nonlinear objective function and the problem is combinatorial since the decision variables of the dispatching time deviations, $\mathbf{x}$, can take values from the discrete set Z. Given its combinatorial nature, computing robust dispatching times is an NP-hard problem and, for at least a class of instances, the problem cannot be solved with exact optimization methods due to its computational complexity.

V. SOLUTION METHOD

Given the computational intractability of the proposed robust dispatching time optimization problem, a metaheuristic solution method is introduced. A typical metaheuristic, such as a genetic algorithm (GA), can converge to a near optimal solution without scalability issues (however, it cannot be guaranteed that this solution is a globally optimal solution).

The work of [35] was one of the first on GAs. A typical GA contains a number of population members. Each population member is a GA chromosome and represents a solution of the optimization problem. In this work, each chromosome has $|N|$ genes and each one of the genes represents the dispatching time deviation of the n-th trip from the value δ_n.

The main stages of our proposed GA are: (1) encoding the initial population; (2) evaluating the fitness of each population member using Monte Carlo simulations; (3) parent selection for offspring generation; (4) crossover; and (5) mutation.

A. Encoding

For a GA population $\mathbf{P}=\{\mathbf{p}_1,...,\mathbf{p}_{|P|}\}$, each population member $\mathbf{p}_i\in\mathbf{P}$ is a potential solution of the robust dispatching time optimization problem. At the encoding stage, the genes of each population member $\mathbf{p}_i$ are: $(p_{i1},p_{i2},...,p_{i|N|})$. At the initialization stage, each gene can take a random value from the set Z that contains all dispatching time deviation options.

B. Fitness Evaluation and Population Member Selection

If a population member $\mathbf{p}_i\in\mathbf{P}$ satisfies the constraints of eq.11, then its fitness is evaluated; otherwise it is discarded and replaced by a new member. The fitness of a population member that satisfies all constraints is evaluated by calculating the cost of $\max_{\mathbf{T}} f(\mathbf{p}_i,\mathbf{T})$. The cost $\max_{\mathbf{T}} f(\mathbf{p}_i,\mathbf{T})$ is approximated by sampling link travel values $t_{s,\tau}$ from the corresponding uncertainty sets $\mathcal{B}_{s,\tau}$ and returning the maximum value of $f(\mathbf{p}_i,\mathbf{T})$ after many Monte-Carlo simulations that examine a large number of link travel time scenarios.

For a population member $\mathbf{p}_i\in\mathbf{P}$, each Monte Carlo simulation is an experimentation that evaluates the value of $f(\mathbf{p}_i,\mathbf{T})$ for link travel time values which are randomly sampled from the set $\mathbf{T}$. After performing a pre-defined number[4] λ of Monte Carlo simulations, the link travel times of the simulation that returned the highest $f(\mathbf{p}_i,\mathbf{T})$ cost are selected. This utilization of Monte Carlo simulations inside the GA for evaluating the fitness of each population member differentiates our GA and enables it to tackle the minimax problem.

One population member $\mathbf{p}_i\in\mathbf{P}$ is more fit for reproduction if its $\max_{\mathbf{T}} f(\mathbf{p}_i,\mathbf{T})$ value is low. Using the well-known roulette-wheel selection method [36], population members with better fitness have a higher probability of being selected for reproduction. In the roulette-wheel selection method, the probability of each population member $\mathbf{p}_i\in\mathbf{P}$ to be selected for reproduction is proportional to its fitness value divided by the sum of the fitness values of all other population members: $\frac{\max_{\mathbf{T}} f(\mathbf{p}_i,\mathbf{T})}{\sum_{\mathbf{p}_j\in\mathbf{P}}\max_{\mathbf{T}} f(\mathbf{p}_j,\mathbf{T})}$.

C. Crossover and Mutation

For each pair of parents which are selected from the initial population using the roulette-wheel selection method, a cross-over occurs at a randomly selected crossover point to produce two offsprings (recombination). The same process is repeated until the total number of generated offsprings is equal to the population size $|P|$.

In the mutation stage, a mutation can occur in any gene of an offspring allowing the exploration of new information that does not belong to the parents. In this work, each gene of an offspring has a very small probability, ϖ, to be replaced by a random value from the set Z of dispatching time deviation options.

D. Population Evolution and Termination

After completing the above stages, the initial population is replaced by the new generation. This procedure is repeated resulting in population evolutions until reaching a pre-determined limit of population generations, μ^{max}. This procedure is summarized in figure 1.

Theorem 1: For a pre-defined number of Monte Carlo simulations, λ, and a pre-defined limit of population evolutions, μ^{max}, the total number of computations until the

[4]the number of Monte Carlo simulations, λ, should be sufficiently high for increasing the probability of obtaining the worst-case link travel times for a given solution $\mathbf{p}_i$

termination of the algorithm is $\lambda|N||S||P|\mu^{max}$ and the computational complexity is polynomial.

Proof: For a population with $|P|$ population members the fitness of each population member $\mathbf{p}_i \in \mathbf{P}$ needs to be evaluated at each population generation. For evaluating the fitness of each population member we need to perform λ Monte Carlo simulations where at each simulation we evaluate the values of Eq.11 that return the value of $f(\mathbf{p}_i, \mathbf{T})$ for the randomly selected link travel times which are used in this simulation after performing $|N||S|$ computations for defining the values of the recursive functions of Eq.1-5. Therefore, the evaluating the fitness of each population member requires $\lambda|N||S|$ computations and the fitness of all population members in the population $\lambda|N||S||P|$. Thus, if the number of population generations until the termination of the algorithm is μ^{max}, then the total number of computations is $\lambda|N||S||P|\mu^{max}$. ∎

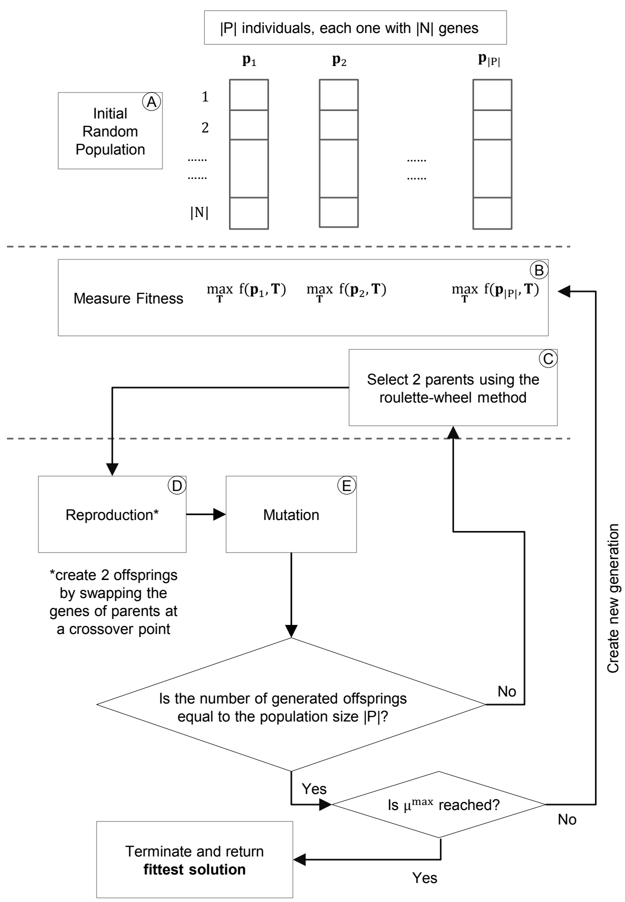


Fig. 1. GA-based robust dispatching time optimization for determining the dispatching times of all daily trips

VI. CASE STUDY

The effect of the robust optimization of the dispatching times of buses is tested using real AVL and AFC data from a high-frequency circular bus service in APAC. The examined bus line has a total number of $|N| = 132$ daily trips. It covers 7.5 km and serves 22 bus stops.

TABLE I

PARAMETER VALUES

ψ (layover times)	3 min
γ (extra time for each pass. boarding)	3 sec
$\mathcal{T}$ (latest dispatching time of last daily trip)	6:59 pm
h_s^* (desired headway at each stop $s \in S$)	5.5 min
λ (Monte Carlo simulations for each population member)	1,000

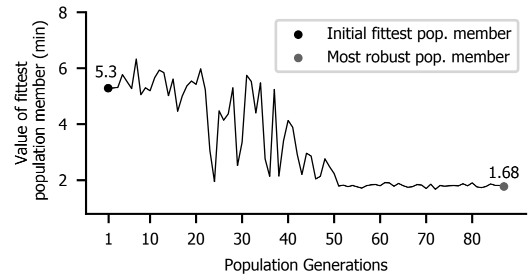


Fig. 2. Population generations of the GA until converging to a robust dispatching time solution

The available AVL and AFC data cover a five-month period. The dataset is split into two sets. The first set contains data from the first four months[5]. The second set contains data from the last month and is used for validation. In more detail, the robust dispatching times are computed based on historical data from the first four months and their robustness against the travel time variations is tested by evaluating their performance at each one of the 30 remaining days.

The values of the parameters for deriving a robust timetable for the circular bus line are summarized in table I.

Following the steps of the GA-based robust dispatching time optimization for all daily trips, the algorithm is programmed in Python 2.7 using the *Distributed Evolutionary Algorithms in Python* (Deap) package [37]. The fitness function for evaluating the performance of the population members is provided by the mathematical model of eq.11 and is programmed in Python 2.7.

The GA-based robust dispatching time optimization is initialized with a population size of $|P| = 100$ members. The fittest member of the initial population generation, named as $\mathbf{p}^{gen=1}$, has a worst-case performance of $\max_{\mathbf{T}} f(\mathbf{p}^{gen=1}, \mathbf{T}) = 5.3$ min as presented in figure 2. This means that when the dispatching times of all daily trips, $\mathbf{p}^{gen=1}$, are applied; then, the worst-case link travel times $\mathbf{T}$ will lead to an average squared headway deviation of 5.3 min from the desired headways.

As presented in fig.2, the fittest population members of the new generations of the GA do not have an improved worst-case performance until generation 21. After generation 22 and until generation 50 there is some instability since the fittest population members of some generations have a significantly improved performance while others do not. After generation 51 the GA stabilizes and the fittest population members of all generations until the termination

[5]the observed link travel times from the first four months are used for defining the sets $\mathcal{B}_{s,\tau},\ \forall s,\tau$

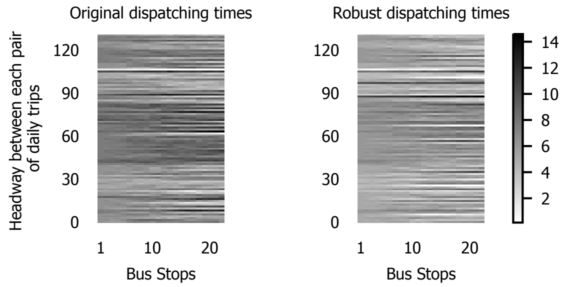


Fig. 3. Bus headways at each bus stop throughout the day at the worst-case scenario of link travel times

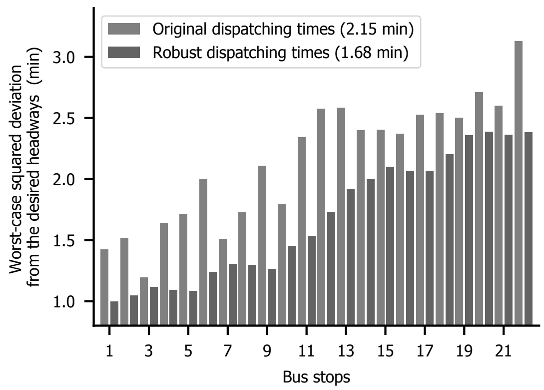


Fig. 4. Average squared deviation from the desired headways at each bus stop at the worst-case scenario of link travel times

of the algorithm exhibit a worst-case squared deviation from the desired headways of 1.68-1.82 min.

Before proceeding to the validation, in fig.3 we present the headways of all daily trips in the worst-case scenario of link travel times and in fig.4 the theoretical upper bound of improvement when using the robust dispatching times instead of the originally planned ones. For performing this task, we compute the worst-case performance of the squared headway deviation at all bus stops when the robust and the original dispatching times are applied. This worst-case performance is computed by using the worst possible link travel time values from the observed 4-month data. In such case, the average squared headway deviation from the desired headways is improved by $\simeq$22% when using the robust dispatching times as presented in fig.4.

Notwithstanding the above, the worst-case link travel times are not expected to occur in a typical day. Therefore, the actual improvement of the headway deviation can be much lower than 22%. For this reason, in the validation stage we use the observed link travel times from the last month of operations to evaluate the performance of the headway deviation when using the robust and the originally planned dispatching times.

The performance of the originally planned dispatching times and the robust ones at each day of the last month which is used for validation is presented in figure 5. From figure 5, one can observe that the average squared daily deviation of the actual headways from their desired values is improved

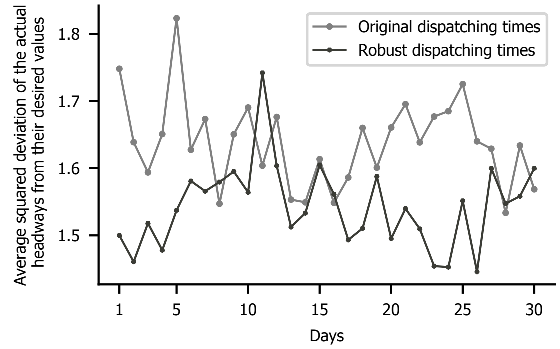
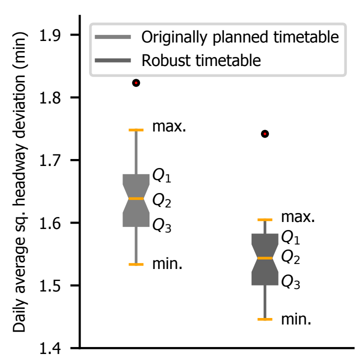


Fig. 5. Average squared deviation of the actual headways from their desired values for each one of the 30 days

when using the robust dispatching times in 25 days[6] out of the 30 days of the month.

This performance is summarized in figure 6 using the Tukey boxplot convention [38]. In this convention, Q_1 is the first quartile, Q_2 the mean and Q_3 the third quartile, min. the lowest datum still within 1.5 the interquartile range (IQR) of the first quartile and max. the highest datum still within 1.5 IQR of the third quartile. All other values outside the [min., max.] set are outliers.

In this month, the observed average improvement of the headway deviation when using the robust dispatching times is $\frac{1.54-1.63}{1.63} = 5.5\%$. More importantly, the max. value of the boxplot is improved by 8.5% when using robust dispatching times demonstrating an improved performance at extreme link travel time scenarios in the long run. In contrary, the interquartile range (IQR) of the robust dispatching times is IQR:= $Q_3 - Q_1 = 1.58 - 1.50 = 0.08$ min and is equal to the observed IQR when implementing the originally planned dispatching times. This finding shows that even if the deviation of the actual headways from their desired values is improved when using robust dispatching times, the robust dispatching times cannot guarantee the extinction of headway deviations when the link travel times vary from time to time.



	Original Disp. Times	Robust Disp. Times
max.	1.75	1.60
Q_3	1.67	1.58
Q_2	1.63	1.54
Q_1	1.59	1.50
min.	1.53	1.45
outliers	1.82	1.74

Fig. 6. Validation results

[6]in the remaining five (5) days, the originally planned dispatching times performed better. The actual link travel times of those days were very close to their expected values providing an advantage to the original dispatching times

VII. CONCLUSIONS

This work modeled the dispatching time optimization problem in the presence of travel time uncertainty and presented a GA-based approach for computing dispatching times which are robust to travel time variations. This approach was tested in a circular bus line of a major bus operator in APAC using 4-months of AVL and AFC data for computing dispatching times that are robust to the observed link travel time variations.

The last month of the operational data was used for validation demonstrating an improvement potential of 5.5% on the average squared deviation of the actual headways from their desired values. In addition, the analysis showed that, given the uncertainty of the link travel times, the daily operations of buses can vary significantly even if a robust schedule is applied.

In future research, the impact of applying robust dispatching times to the in-vehicle travel time of passengers can be examined. Finally, an interesting future research topic can be the consideration of the uncertainty of passenger demand by expanding the model of this study.

REFERENCES

[1] K. Jansson and R. Pyddoke, "Quality incentives and quality outcomes in procured public transport–case study stockholm," *Research in Transportation Economics*, vol. 29, no. 1, pp. 11–18, 2010.

[2] K. Gkiotsalitis and A. Stathopoulos, "Joint leisure travel optimization with user-generated data via perceived utility maximization," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 532–548, 2016.

[3] K. Gkiotsalitis and A. Stathopoulos, "Demand-responsive public transportation re-scheduling for adjusting to the joint leisure activity demand," *International Journal of Transportation Science and Technology*, vol. 5, no. 2, pp. 68–82, 2016.

[4] K. Gkiotsalitis and N. Maslekar, "Improving bus service reliability with stochastic optimization," in *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*. IEEE, 2015, pp. 2794–2799.

[5] W. Leong, K. Goh, S. Hess, and P. Murphy, "Improving bus service reliability: The singapore experience," *Research in Transportation Economics*, vol. 59, pp. 40–49, 2016.

[6] A. Ceder, "Public transit planning and operation: Theory," *Modeling and practice. Oxford: Elsevier*, 2007.

[7] R. Z. Farahani, E. Miandoabchi, W. Y. Szeto, and H. Rashidi, "A review of urban transportation network design problems," *European Journal of Operational Research*, vol. 229, no. 2, pp. 281–302, 2013.

[8] K. Gkiotsalitis and O. Cats, "Exact optimization of bus frequency settings considering demand and trip time variations," in *96th Annual Meeting of the Transportation Research Board*, no. 17-01871, 2017.

[9] O. Ibarra-Rojas, F. Delgado, R. Giesen, and J. Muñoz, "Planning, operation, and control of bus transport systems: A literature review," *Transportation Research Part B: Methodological*, vol. 77, pp. 38–75, 2015.

[10] K. Gkiotsalitis and R. Kumar, "Bus operations scheduling subject to resource constraints using evolutionary optimization," *Informatics*, vol. 5, no. 1, 2018.

[11] P. G. Furth and N. H. Wilson, "Setting frequencies on bus routes: Theory and practice," *Transportation Research Record*, vol. 818, no. 1981, pp. 1–7, 1981.

[12] G. Gallo and F. D. Miele, "Dispatching buses in parking depots," *Transportation Science*, vol. 35, no. 3, pp. 322–330, 2001.

[13] L. Peeters and L. Kroon, "A cycle based optimization model for the cyclic railway timetabling problem," *Computer-aided scheduling of public transport*, pp. 275–296, 2001.

[14] P. Ávila-Torres, F. López-Irarragorri, R. Caballero, and Y. Ríos-Solís, "The multimodal and multiperiod urban transportation integrated timetable construction problem with demand uncertainty," *Journal of Industrial & Management Optimization*, pp. 29–42, 2017.

[15] F. Zhao and X. Zeng, "Optimization of transit route network, vehicle headways and timetables for large-scale transit networks," *European Journal of Operational Research*, vol. 186, no. 2, pp. 841–855, 2008.

[16] D. J. Sun, Y. Xu, and Z.-R. Peng, "Timetable optimization for single bus line based on hybrid vehicle size model," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 2, no. 3, pp. 179–186, 2015.

[17] Y. Wang, D. Zhang, L. Hu, Y. Yang, and L. H. Lee, "A data-driven and optimal bus scheduling model with time-dependent traffic and demand," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2443–2452, 2017.

[18] H. Yu, H. Ma, H. Du, X. Li, R. Xiao, and Y. Du, "Bus scheduling timetable optimization based on hybrid bus sizes," in *UK Workshop on Computational Intelligence*. Springer, 2017, pp. 337–348.

[19] Y. Xuan, J. Argote, and C. F. Daganzo, "Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis," *Transportation Research Part B: Methodological*, vol. 45, no. 10, pp. 1831–1845, 2011.

[20] C. F. Daganzo, "A headway-based approach to eliminate bus bunching: Systematic analysis and comparisons," *Transportation Research Part B: Methodological*, vol. 43, no. 10, pp. 913–921, 2009.

[21] A. Adamski and A. Turnau, "Simulation support tool for real-time dispatching control in public transport," *Transportation Research Part A: Policy and Practice*, vol. 32, no. 2, pp. 73–87, 1998.

[22] J. Zhao, M. Dessouky, and S. Bukkapatnam, "Optimal slack time for schedule-based transit operations," *Transportation Science*, vol. 40, no. 4, pp. 529–539, 2006.

[23] M. Trompet, X. Liu, and D. Graham, "Development of key performance indicator to compare regularity of service between urban bus operators," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2216, pp. 33–41, 2011.

[24] Y. Shafahi and A. Khani, "A practical model for transfer optimization in a transit network: Model formulations and solutions," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 6, pp. 377–389, 2010.

[25] A. A. Ceder, S. Hassold, and B. Dano, "Approaching even-load and even-headway transit timetables using different bus sizes," *Public Transport*, vol. 5, no. 3, pp. 193–217, 2013.

[26] K. Gkiotsalitis and N. Maslekar, "Towards transfer synchronization of regularity-based bus operations with sequential hill-climbing," *Public Transport*, pp. 1–27, 2018.

[27] K. Gkiotsalitis and N. Maslekar, "Multiconstrained timetable optimization and performance evaluation in the presence of travel time noise," *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 9, p. 04018058, 2018.

[28] Y. Wu, J. Tang, Y. Yu, and Z. Pan, "A stochastic optimization model for transit network timetable design to mitigate the randomness of traveling time by adding slack time," *Transportation Research Part C: Emerging Technologies*, vol. 52, pp. 15–31, 2015.

[29] K. Gkiotsalitis and O. Cats, "Reliable frequency determination: Incorporating information on service uncertainty when setting dispatching headways," *Transportation Research Part C: Emerging Technologies*, vol. 88, pp. 187–207, 2018.

[30] M. D. Hickman, "An analytic stochastic model for the transit vehicle holding problem," *Transportation Science*, vol. 35, no. 3, pp. 215–237, 2001.

[31] P. Welding, "The instability of a close-interval service," *Journal of the operational research society*, vol. 8, no. 3, pp. 133–142, 1957.

[32] W. H. Kraft and T. F. Bergen, "Evaluation of passenger service times for street transit systems," *Transportation Research Record*, no. 505, 1974.

[33] Kittelson & Assoc., Parsons Brinckerhoff, KFH Group, Texam A&M Transportation Institute and Arup, "Transit capacity and quality of service manual," 2013.

[34] A. Wald, "Statistical decision functions which minimize the maximum risk," *Annals of Mathematics*, pp. 265–280, 1945.

[35] J. H. Holland, *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, Michigan, 1975.

[36] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of genetic algorithms*. Elsevier, 1991, vol. 1, pp. 69–93.

[37] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, "DEAP: Evolutionary algorithms made easy," *Journal of Machine Learning Research*, vol. 13, pp. 2171–2175, jul 2012.

[38] J. W. Tukey, *Exploratory data analysis*. Reading, Mass., 1977, vol. 2.