International Journal of Transportation Science and Technology 5 (2016) 68–82

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

journal homepage: www.elsevier.com/locate/ijtst

Demand-responsive public transportation re-scheduling for adjusting to the joint leisure activity demand

Konstantinos Gkiotsalitis *, Antony Stathopoulos

National Technical University of Athens, Iroon Polytechniou 5, Zografou 15773, Athens, Greece

ARTICLE INFO

Article history:
Received 26 April 2016
Received in revised form 6 September 2016
Accepted 19 September 2016
Available online 22 September 2016

ABSTRACT

Fixed daily trips such as trips to work/school have fixed departure/arrival times and destination points. The recurrent nature of fixed activities facilitates individuals on making more well-informed decisions about the transport mode selection. On the contrary, selecting a transportation mode for non-recurrent leisure trips, which can account for up to 60% of trips in some cities (Transport for London, 2014), is a more complex task due to the fact that individuals have little knowledge about the alternative modal options. In this paper, we try to improve the operations of demand-responsive public transportation systems by increasing their service quality and their ridership related to joint-leisure-trips via timetable rescheduling. First, we model the public transport service re-scheduling problem considering operational regulations and the quality of service. Then, a sequential heuristic method is introduced for re-scheduling the timetables of demand-responsive public transport modes in near-real time and accommodating the joint leisure activity demand without deteriorating the quality of service. The public transport re-scheduling for increasing the joint leisure activity ridership was tested in a case study using user-generated data from social media in Stockholm and the General Transit Feed Specification (GTFS) data from Sweden focusing especially on central bus lines 1 and 4.

1. Introduction

Joint leisure activities can account for up to 60% of trips in some cities (Transport for London, 2014) and require the timely arrival of several individuals at one pre-defined place which can be the location of an event or the location of a leisure activity. Nowadays, non-traditional data sources such as user-generated Social Media (SM) data, Floating Car Data (FCD), Cellular Data (CD) and Automated Fair Collection (AFC) data can provide more insights on individuals' preferences and enhance the location selection process for a joint leisure activity. For instance, Musolesi and Mascolo, 2007 utilized CD logs for correlating the mobility patterns of an individual with the mobility patterns of his friends and acquaintances. Theoretical concepts were also developed by Carrasco et al. (2008), Arentze and Timmermans (2008) and Chen et al. (2014) on predicting agents' mobility based on their social networks. Predicting the trips of individuals and deriving the OD matrices in urban areas are the main objectives of user-generated CD research including the works of De Domenico et al. (2013), Calabrese et al. (2011a, b), Gonzalez et al. (2008), Zhang et al. (2010) and Sohn and Kim, 2008. SM data from social networks like Facebook, Twitter, Foursquare and the image sharing service, Flickr, have also been used for capturing users' activities at different locations and

Peer review under responsibility of Tongji University and Tongji University Press.
* Corresponding author.
E-mail address: konstantinos.gkiotsalitis11@alumni.imperial.ac.uk (K. Gkiotsalitis).

http://dx.doi.org/10.1016/j.ijtst.2016.09.004
2046-0430/

day times via advanced spatio-temporal analysis and educated rules (refer to Alesiani et al. (2014), Gkiotsalitis and Stathopoulos (2015, 2016) and Sun (2016).

Processing and analyzing continuously updated user-generated data over time provides a better estimate of individuals' preferences and enables the selection of a more efficient joint leisure activity location. Unlike fixed trips, the joint leisure activity participants have less knowledge of the availability and punctuality of transport mode alternatives. For instance, one individual is aware of the alternative mode options for traveling to his/her work place but he/she is less aware of the mode options, their travel cost and their punctuality (i.e., travel time variability) when it comes to leisure trips to locations that the individual is not familiar with. This lack of information can deteriorate the perceived utility of the activity participants due to: i) the excess waiting time that individuals can wait at the joint leisure activity location until all individuals have arrived there; and ii) the inefficient selection of transport modes that can increase the travel cost.

Works on activity-based models from Miller et al. (2005) and Janssens et al. (2004) described two different decision-making models focused either on perceived utility maximization or rule-based selection. The work of Miller et al. (2005) focused on the selection of transport modes for participating at activities from all members of a household by utilizing the concept of Random Utility Maximization (RUM) and adding constraints to the selection problem (i.e., private mode availability subject to the use from another member of the household). Although, the literature on disaggregate (e.g., Asensio, 2002) or tour-based (e.g., Bowman and Ben-Akiva, 2000) mode choice is vast, the work of Miller et al. (2005) is one of the closest prior arts since it tackles the problem of mode selection for a joint activity from several household members. However, the transport mode selection for a joint leisure activity has a higher complexity level than the household-based activity planning because: (i) joint leisure activity participants start their trip from different origins and not from a common location (i.e., not necessarily home-based), (ii) the arrival time punctuality at the location of the leisure activity matters since the arrival time variance of all activity participants should be low to avoid excess waiting before the activity starts.

Due to that, this paper focuses on modeling and improving the joint-leisure-activity ridership of demand-responsive transportation via rescheduling the timetables of public transport modes subject to a set of operational constraints. In Section 2, the problem of public transport service re-scheduling considering operational regulations and the quality of service is modeled. In Section 3, a computational efficient heuristic search method for dynamic re-scheduling the public transport schedules of demand-responsive systems and increasing their ridership related to joint leisure trips is introduced. Finally, a case study utilizing Social Media and the General Transit Feed Specification (GTFS) data from Stockholm is presented in Section 4, followed by the concluding remarks in Section 5.

2. Modeling the public transport rescheduling problem for increasing the passenger ridership related to joint leisure activity trips

To increase public transportation ridership for trips related to joint leisure activities, the planned schedules of public transport modes should be adapted the joint-leisure-activity travelers' requirements. Demand responsive public transportation (DRPT) has been viewed as a mean of passenger communication with the transport operator in a direct way allowing the transport operator to create custom routes based on a priori knowledge of the passengers' locations, destinations, and schedules (Mauri and Lorena, 2009; Castex et al., 2004).

In this paper however, we assume fixed DRPT routes where only their timeschedules (i.e., departure times) can change subject to spatio-temporal demand variations related to joint leisure activities. The key individual-based data for deriving Joint Leisure activities demand variation in space and time is obtained from Smartphone apps that can open one-way communication channels between the travelers and the public transport operator (i.e., social media apps). In the past, smartphone apps have utilized for improving dynamic travel decisions (Dickinson et al., 2014). In our work, Smartphone data from Social Media Apps is transmitted to the control center of the transport operator (either via direct transmission or via a data crawling campaign) and provides information of joint-leisure-activity demand for adapting the DRPT operations to it via changing the planned timeschedules.

Let us assume that over the time period of a day a number of public transport trips $\pi = \{\pi_1, \pi_2, \ldots, \pi_\pi\}$ can serve joint leisure activity passengers. Those public transport trips belong, in general, to different public transport services while, some of them, can belong to the same public transport service (i.e., the bus service line). Let for instance a trip $\pi_i \in \pi$ to be part of a public transport service p. This public transport service has a planned daily schedule (service timetable). If this service serves $S_p = \{S_{p,1}, S_{p,2}, \ldots, S_{p,|S_p|}\}$ public transport stations and has a number of n_p trips during the day, where the trips are denoted as $K_p = \{K_{p,1}, K_{p,2}, \ldots, K_{p,|K_p|}\}$, then the planned daily schedule of service p is represented by a D: $|K_p| \times |S_p|$ dimensional matrix with integer elements denoting the planned arrival time of each trip $K_{p,j} \in K_p$ to each station $S_{p,i} \in S_p$.

Let $S_{p,i} \in S_p$ be the station of trip $K_{p,i} \equiv \pi_i \in K_p$ where one or more boardings from joint leisure activity passengers can occur and $S_{p,j} \in S_p$ be the alighting station of the joint leisure activity passenger/s which is the closest station to the location of the joint leisure activity Λ. The joint leisure activity passenger will be more tempted to use this public transport service if the alighting at the destination station occurs close to the time of the joint leisure activity. Given the joint leisure activity location Λ and the starting time of this activity t, we introduce a utility score for using the public transport service:

$$z = (t - t^w - D_{K_{p,i},S_{p_j}})^2 \tag{1}$$

which is zero at the optimal case (alighting at station $S_{p,j}$ is at the same time with the starting time of the joint leisure activity t minus the walking time required for from the station to the location of the activity denoted as t^w) and it progressively increases its value if trip $K_{p,i}$ arrives much earlier or later to that station.

During the rescheduling phase of this public transport service, the departure times of the daily trips are subject to changed. Those changes are the "unknows" of the rescheduling problem and are denoted by $x = \{x_{K_{p,1}}, x_{K_{p,2}}, \ldots, x_{K_{p,|K_p|}}\}$ which is a $|K_p|$ -dimensional vector of the problem variables and represents how many minutes the departure time of each trip of public transport service p can deviate from the planned one. Vector x is not allowed to take values from the Euclidean space $\mathbb{R}^{|K_p|}$ but only from a discrete set of values ($q = \{-5, -4, \ldots, 0, \ldots, 4, 5\}$ minutes) because the re-scheduled timeplan should be expressed in minutes (decimal values are not allowed to planned schedules). Therefore, the utility score for the joint leisure activity participant towards the public transport service is not the static value $z = (t - t^w - D_{K_{p,i},S_{p_j}})^2$, but can be expressed by a scalar function $z : \mathbb{N} \to \mathbb{R}$:

$$z(x_{K_{p,i}}) = (t - t^w - (D_{K_{p,i},S_{p_j}} + x_{K_{p,i}}))^2 \tag{2}$$

Realizing that an entire re-scheduling of all daily trips is unnecessary and will create many changes without reason, variables $x_{K_{p,\phi}} \in x$ are allowed to change only if the planned arrival time $D_{K_{p,\phi},S_{p_j}}$ of the trip $K_{p,\phi}$ at station S_{p_j} is within the range of ±60 min from the joint activity starting time minus the expected walking time to the activity location:

$$x_{K_{p,\phi}} : \begin{cases} \in q = \{-5, -4, \ldots, 0, \ldots, 4, 5\} & \textit{if } |D_{K_{p,\phi},S_{p_j}} - t - t^w| \leqslant 60 \text{ min} \\ 0 & \textit{otherwise} \end{cases} \tag{3}$$

One might reasonably claim that rescheduling the public transport service timeplan for adapting to the joint leisure activity demand might penalize other regular passengers of the public transport service who are not involved in leisure activities. In practical terms, it in not feasible to monitor the effect of the rescheduling phase by continuously surveying the passengers of the public transport service. However, we can ensure that (i) the operations are better of, or at least not deteriorated, for all passengers of that public transport service at a system-wide level and (ii) no passengers are over-penalized from those changes.

Joint leisure activity spatio-temporal demand variations have greater impact to the services of dense metropolitan areas due to the volume of leisure demand changes that justifies the public transport rescheduling efforts. Those services in metropolitan areas are high frequency services and they are regularity-based (their Key Performance Index (KPI) is the service-wide Excess Waiting Time (EWT) score instead of punctuality-based On-Time-Adherence (OTA) of operations to the planned schedule). Regularity-based services dictate that the daily trips of one public transport service should keep a certain, pre-defined time deviation (headway) when passing by each station of the service which should not be too small (bunching) or too high leading to excess waiting times for passengers that wait at the station for that service.

The operational performance of the public transport services are assessed through this EWT scheme at high-frequency services, and, in some cities, such as London and Singapore, the public transport operators receive monetary penalties or bonuses depending on the daily EWT score of each service. In many other cities, the EWT score of the public transport service indicates the performance of the service and the service operator can have his contracts terminated if the daily EWT scores of that services are below average.

The service-wide EWT of one public transport service is a linear function of the EWT scores observed at several stations of the transport service during different periods of the day (morning peak, afternoon peak, etc.). Those time periods have starting and ending times: $((T_1^s, T_1^e), (T_2^s, T_2^e), (T_o^s, T_o^e))$. By definition, the EWT at each station, $S_{p,j} \in S_p$ is calculated from the planned arrival times of consecutive trips for every time period $((T_g^s, T_g^e))$ and since it is a variance from the expected waiting time for passengers cannot take a negative value:

$$\begin{aligned} EWT_{S_{p,j}}^{(T_g^s,T_g^e)} = max\Bigg[0; & \frac{\sum_{i=2}^{|K_p|-1} \theta_{K_{p,i},S_{p,j}} ((D_{K_{p,i},S_{p,j}} + x_{K_{p,i}}) - (D_{K_{p,i-1},S_{p,j}} + x_{K_{p,i-1}}))^2}{2\sum_{i=2}^{|K_p|-1} \theta_{K_{p,i},S_{p,j}} ((D_{K_{p,i},S_{p,j}} + x_{K_{p,i}}) - (D_{K_{p,i-1},S_{p,j}} + x_{K_{p,i-1}}))} \\ & - \frac{\sum_{i=2}^{|K_p|-1} \theta_{K_{p,i},S_{p,j}} ((D_{K_{p,i},S_{p,j}} + x_{K_{p,i}}) - (D_{K_{p,i-1},S_{p,j}} + x_{K_{p,i-1}}))}{2\sum_{i=2}^{|K_p|-1} \theta_{K_{p,i},S_{p,j}}}\Bigg] \end{aligned} \tag{4}$$

where

$$\theta_{K_{p,i},S_{p,j}} : \begin{cases} 1 & \textit{if } D_{K_{p,i},S_{p,j}} + x_{K_{p,i}} \quad \text{and} \quad D_{K_{p,i-1},S_{p,j}} + x_{K_{p,i-1}} \in (T_g^s, T_g^e) \\ 0 & \textit{otherwise} \end{cases} \tag{5}$$

The EWT score at station $S_{p,j}$ for public transport service p is then derived by aggregating the EWT scores from different time periods of the day:

$$EWT_{S_{p,j}} = \frac{\sum_{g=1}^{o} EWT_{S_{p,j}}^{(T_g^s,T_g^e)}}{\sum_{g=1}^{o} 1} \tag{6}$$

In a similar fashion, the service-level EWT which is the main KPI for every regularity-based services is then derived by adding the EWTs computed at different stations (in some cases some stations might have higher weights than others):

$$EWT_p = \frac{\sum_{j=1}^{|S_p|} EWT_{S_{p,j}}}{\sum_{j=1}^{|S_p|} 1} \tag{7}$$

For satisfying condition (i): the operations are better of, or at least not deteriorated, at the service-level after the rescheduling of the service, the first objective is to reduce the EWT score at the service level or at least ensure that it is not increased. In addition, for not increasing operational costs we do not allow the insertion of new, additional transport modes at each re-scheduled public transport service and we also keep the number of planned trips the same as before (no inclusion of additional trips). Changing however the departure times of some trips might lead to operational cost changes such as fuel consumption due to traffic condition changes but we omit those costs from our analysis because they are of minor importance especially since our re-scheduled trips change departure times within the range of $q = \{-5, +5\}$ minutes resulting to small changes (no morning trips are shifted to afternoon or from peak to off-peak conditions where the network traffic is expected to change significantly).

Finally, for satisfying condition (ii) and not allowing any regular passengers to get over-penalized because of the re-scheduling, we adopt the frequency constraint rule which is also used for the tactical schedule planning from public transport operators. This rule is the outcome of the tactical frequency setting phase and dictates the range of minutes where successive trips can depart from the departure station for covering adequately the passenger demand at every station of the planned service. The frequency range changes over different time periods of the day (morning peak, etc.) and if the time periods of the day are: $\{1, 2, \ldots, o\}$ =((T_1^s, T_1^e), (T_2^s, T_2^e), (T_o^s, T_o^e)), then for each time period $g \in \{1, 2, \ldots, o\}$ there is a minimal frequency value $l(g)$ and a maximal frequency value $h(g)$ which should not be violated by any couple of successive trips within that time period. This rule, indirectly ensures that any re-scheduled timeplan that does not violate the frequency range for every time period of the day will cover the regular passenger demand in a satisfactory manner and no regular passengers of the service will not be over-penalized.

For instance, if one trip $K_{p,\phi}$ is subject to re-scheduling and belongs to the time period g, then the departure time deviation $x_{K_{p,\phi}}$ should be such that:

$$l(g) \leqslant (D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}}) - (D_{K_{p,\phi-1},S_{p,1}} + x_{K_{p,\phi-1}}) \leqslant h(g) \tag{8}$$

While re-scheduling the planned public transport service p that contains a joint leisure activity trip $K_{p,i} \in K_p$ we try to: (a) minimize the utility score of joint leisure activity passengers $(t - t^w - (D_{K_{p,i},S_{p_j}} + x_{K_{p,i}}))^2$, (b) disturb only the planned trips which are at the close vicinity of trip $K_{p,i} \in K_p$: $x_{K_{p,\phi}} = 0$ if $|D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60$ min, (c) minimize the service-wide EWT score, (d) do not violate the planned frequency range of successive trips derived from the tactical planning phase, (e) ensure that the new EWT will at least not be worse than the service-level EWT of the original timeplan. It is important to note here that points (c) and (d) are also the objectives for the public transport authorities when they plan their daily timeplans (timeschedules). If the service-level EWT of service p before the re-scheduling was EWT_p^{before}, then the above objectives result to the following optimization program written in the standard form:

$$\begin{array}{ll} \text{minimize} & f(x) = (t - t^w - (D_{K_{p,i},S_{p_j}} + x_{K_{p,i}}))^2 + EWT_p(x) \\ \text{subject to} & EWT_p^{before} - EWT_p(x) \geqslant 0 \\ \forall g \in \{0, 1, 2, \ldots, o\} & (D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}}) - (D_{K_{p,\phi-1},S_{p,1}} + x_{K_{p,\phi-1}}) - l(g) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ \forall g \in \{0, 1, 2, \ldots, o\} & h(g) - (D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}}) + (D_{K_{p,\phi-1},S_{p,1}} + x_{K_{p,\phi-1}}) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ & x \in q = \{-5, -4, \ldots, 0, \ldots, 4, 5\} \text{ min} \\ & x_{K_{p,\phi}} = 0, \quad \forall \phi : |D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60 \text{ min} \end{array} \tag{9}$$

In this optimization problem, the objective function is nonlinear and the service-level EWT constraint is also non-linear. This problem is combinatorial and includes also a series of linear inequality constraints for satisfying the allowed frequency ranges. Item (a) is covered by the objective function. Item (c) is also covered by the objective function whereas item (e) is covered from the nonlinear EWT constraint. Item (d) is covered by the series of linear inequality constraints and item (b) is covered by setting $x_{K_{p,\phi}} = 0, \forall \phi : |D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60$ min. The computational complexity of this problem is exponential $O(|q|^\rho)$ where ρ is equal to the number of re-scheduled trips (i.e., trips for which $|D_{K_{p,\phi},S_{p_j}} - t - t^w| \leqslant 60$ min); therefore, this problem is computational intractable and an exact solution can be computed only at small-scale scenarios. For instance, if within the 2-hour time period of interest we have 20 planned trips from this public transport service a minimum number of 6.727E + 20 computations is required. In addition, if we have not one, but P public transport services that require rescheduling for adjusting to the leisure activity demand, then the computational complexity becomes $O(P|q|^\rho)$.

Finally, it is evident that if within the time range $D_{K_{p,i},S_{p,j}} \pm 60$ min there is another trip $K_{p,m}$ that belongs to the same public transport service and can serve another joint leisure activity demand with $S_{p,m}$ the closest station to that activity location and

t_1, t_1^w the starting time and the walking distance from station $S_{p,m}$ to the location of that activity respectively, then the re-scheduling problem of this public transport service is transformed to:

$$\begin{array}{ll} \text{minimize} & f(x) = (t - t^w - (D_{K_{p,i},S_{p_j}} + x_{K_{p,i}}))^2 + (t_1 - t_1^w - (D_{K_{p,m},S_{pm}} + x_{K_{p,m}}))^2 + EWT_p(x) \\ \text{subject to} & EWT_p^{before} - EWT_p(x) \geqslant 0 \\ \forall g \in \{0,1,2,\ldots,o\} & (D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}}) - (D_{K_{p,\phi-1},S_{p,1}} + x_{K_{p,\phi-1}}) - l(g) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ \forall g \in \{0,1,2,\ldots,o\} & h(g) - (D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}}) + (D_{K_{p,\phi-1},S_{p,1}} + x_{K_{p,\phi-1}}) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ & x \in q = \{-5,-4,\ldots,0,\ldots,4,5\} \text{ min} \\ & x_{K_{p,\phi}} = 0, \quad \forall \phi : |D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60 \text{ min } U |D_{K_{p,\phi}S_{pm}} - t_1 - t_1^w > 60 \text{ min} \end{array} \tag{10}$$

3. Heuristic search for demand responsive public transportation rescheduling

The proposed heuristic search method is a stochastic numerical optimization method which attempts to explore intelligently the solution space and converge to the *global* minimum of the multivariate scalar function f. The heuristic method is composed of the following features: i) formulate a penalty function $p(x)$ for including the constraints to the objective function and penalize the objective function score if any constraint is violated leading to the formation of an unconstrained optimization problem where $p(x) = f(x)$ when all constraints are satisfied ii) generate two random parents and proceed to next generations with the formulation of an iterative Genetic Algorithm (GA).

This global minimization method is stochastic (heuristic) and there is no way to determine if the true global minimum has actually been found. Instead, as a consistency check, the algorithm can be run from a number of different random starting points to ensure the lowest minimum found in each example has converged to the global minimum.

For notation simplification, let us assume that our problem constraints (the service-level EWT score nonlinear inequality constraint and the frequency range linear inequality constraints) are denoted by $c_i(x) \geqslant 0, \forall i$ where i represents every different constraint. For instance $c_1(x) = EWT_p^{before} - EWT_p(x) \geqslant 0$. Then, our re-scheduling problem of Eq. (9) for service p can be expressed in a simplified manner as:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c_1(x) \geqslant 0 \\ \forall g \in \{0,1,2,\ldots,o\} & c_2(x), c_3(x), \ldots, c_{1+m1}(x) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ \forall g \in \{0,1,2,\ldots,o\} & c_{m1+2}(x), c_{m1+3}(x), \ldots, c_{1+m1+m2}(x) \geqslant 0, \forall D_{K_{p,\phi},S_{p,1}} + x_{K_{p,\phi}} \in (T_g^s, T_g^e) \\ & x \in q = \{-5,-4,\ldots,0,\ldots,4,5\} \text{ min} \\ & x_{K_{p,\phi}} = 0, \quad \forall \phi : |D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60 \text{ min} \end{array} \tag{11}$$

Introducing a penalty function, $p(x)$, can transform the above constrained optimization problem to an unconstrained one:

$$\begin{array}{ll} \text{minimize} & p(x) = f(x) + \sum_{i=1}^{1+m1+m2} (\min[-c_i(x), 0])^2 \\ & x \in q = \{-5,-4,\ldots,0,\ldots,4,5\} \text{ min} \\ & x_{K_{p,\phi}} = 0, \quad \forall \phi : |D_{K_{p,\phi},S_{p_j}} - t - t^w| > 60 \text{ min} \end{array} \tag{12}$$

The expression $(min[-c_i(x), 0])^2$ is the main addition to the penalty function and dictates that if a constraint $c_i(x)$ is negative, then min $[-c_i(x), 0] = -c_i(x)$ and the constraint is violated for the current set of variables x; therefore, the objective function is penalized by the term $(min[-c_i(x), 0] = -c_i(x))^2$. Otherwise, if $c_i(x) \geqslant 0$ this constraint is satisfied for the current set of variables x and is not penalizing the objective function since min $[-c_i(x), 0] = 0$ and $(min[-c_i(x), 0] = 0)^2 = 0$.

The heuristic search has a dual scope. Starting from an initial solution guess x_0 where $p(x_0) \geqslant f(x_0)$ it should explore the solution space until at some iteration k we have $p(x_k) = f(x_k)$. At that iteration it is ensured that all constraints are satisfied. However, the search should continue to reduce further the score of $p(x)$ until a convergence test is satisfied at some solution x_* which is also the solution of our stochastic search of the global optimum.

At the initial stage of the heuristic search, let us assume that we have a public transport service p which can serve a joint leisure activity that starts close to station $S_{p,j}$ and a number of ρ trips belong to that service such that $|D_{K_{p,\phi},S_{p_j}} - t - t^w| \leqslant 60$minutes,$\forall \phi \in \{1,2,\ldots,\rho\}$. Then, we can re-schedule the departure times of those ρ trips by selecting different values from the discrete set $q = \{-5,-4,\ldots,0,\ldots,+4,+5\}$ minutes.

At the first stage we generate two random sets (parents) of problem variables. Let the first random set (parent 1) be denoted as Pa. Pa is a ρ -dimensional vector where each element of the vector contains a random value from the set q: $Pa = \{Pa_1, Pa_2, \ldots, Pa_\rho\}$ and $Pa_1, \ldots, Pa_\rho$ take random values from the set q. In addition, the second random set (parent 2) has exactly the same characteristics with Pa, namely, $Pb = \{Pb1, Pb2, \ldots, Pb\rho\}$ and $Pb_1, \ldots, Pb_\rho$ take random values from the set q. Finally, we initialize another random set (child) denoted as Ch for which also $Ch = \{Ch_1, Ch_2, \dots, Ch_\rho\}$ and $Ch_1, Ch_2, \dots, Ch_\rho$ take random values from the set q.

The heuristic search exploration is iterative and at the first stage we compute the penalty function performance of the parent 1 and the parent 2: $p(Pa)$, $p(Pb)$ by selecting the penalty function as our fitness function. The parent with the highest penalty function score is the 'weak' parent and is the selected parent for re-placement at future iterations. At this stage, we start a sequential crossover phase starting from the first element of both parents (Pa_1 and Pb_1) and choosing randomly one of those two values to replace the child value Ch_1. If Pa_1 is randomly chosen, then the child vector becomes $Ch = \{Pa_1, Ch_2, \dots, Ch_\rho\}$. In addition, mutation is allowed by introducing a mutation parameter $m_u \leftarrow 1$. Then, a number a is randomly chosen from the integer set $\{0, 1, \dots, 9, 10\}$ and if $a \leqslant m_u$, then the 1st child vector element is replaced again by a random value ra from the set q: $Ch = \{ra, Ch_2, \dots, Ch_\rho\}$. In other words, during the mutation phase there is a 10% probability to change the randomly chosen element Pa_1 from the crossover stage with another random value ra from the discrete set q.

The fitness of the new child vector is then assessed by computing the penalty function for $p(Ch)$ and (i) if $p(Ch) < p(Pa)$ and $p(Ch) < p(Pb)$, then if $p(Pa) <= p(Pb)$ we replace Pb with Ch: $Pb \leftarrow Ch$ (elitism), (ii) if $p(Ch) < p(Pa)$ and $p(Ch) > P(b)$ we replace Pa with Ch: $Pa \leftarrow Ch$ (the same holds also for the opposite, i.e., $p(Ch) < p(Pb)$ and $p(Ch) > P(a)$), (iii) if $p(Ch) \geqslant p(Pa)$ and $p(Ch) \geqslant p(Pb)$, then the child has inferior performance and no parent is replaced. In addition, the selected child value at this sequence is dropped. For instance, if $Ch = \{ra, Ch_2, \dots, Ch_\rho\}$ after the crossover and mutation phase, then Ch drops the change of this sequence and becomes again $Ch = \{Ch_1, Ch_2, \dots, Ch_\rho\}$.

The sequential crossover continuous then with the second elements of parent 1 and parent 2: (Pa_2, Pb_2) and the same crossover/mutation/elitism replacement procedure continuous for all trips $\{1, 2, \dots, \rho\}$. When we have reached the crossover/mutation/elitism replacement procedure for the departure time deviation of trips ρ we re-start again from the

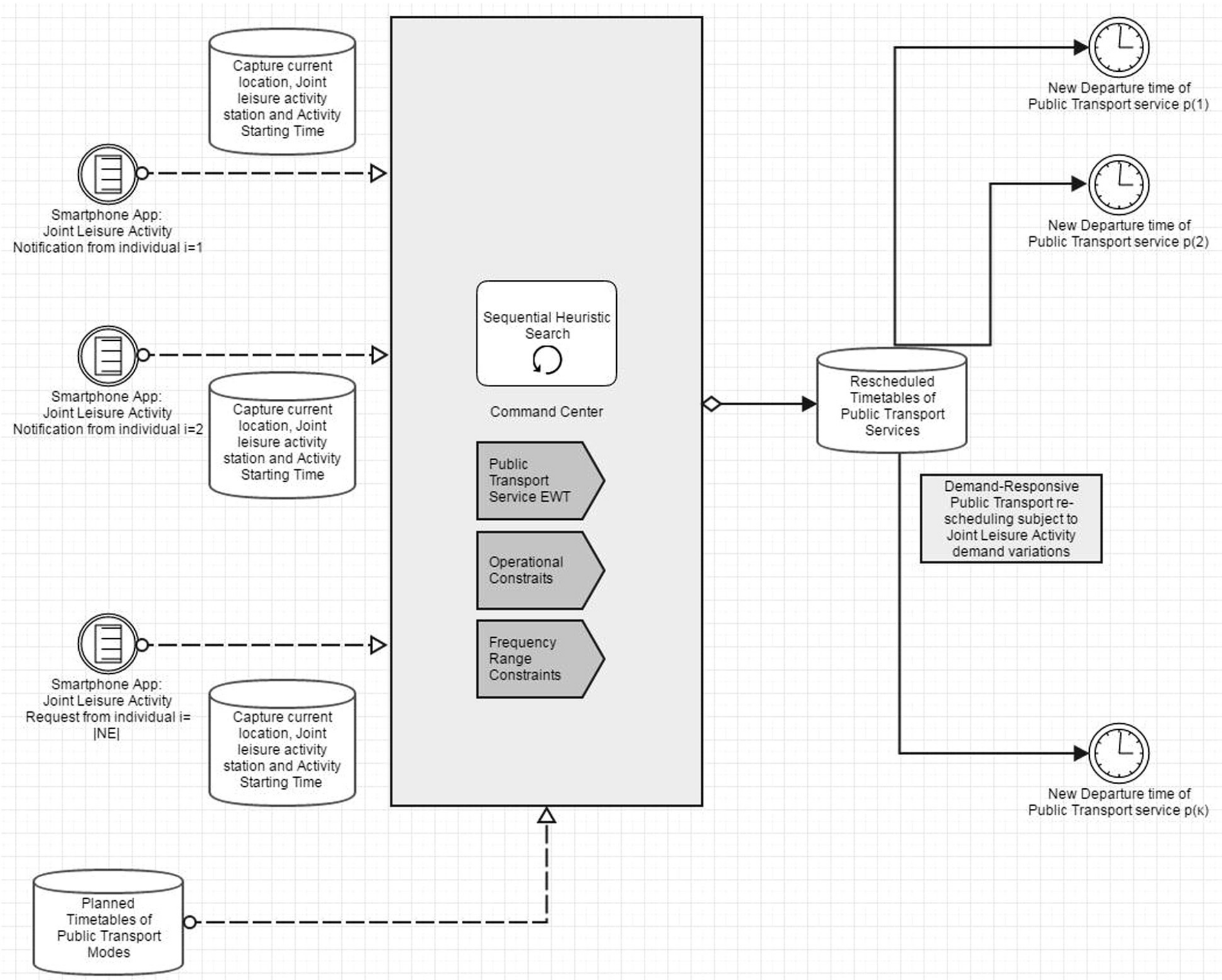

Fig. 1. Generic representation of the demand-responsive public transportation subject to operational KPIs and ridership maximization for joint leisure activities.

beginning performing another sequential GA search. Our first objective is to run this artificial evolutionary approach repeatedly until we find a child which has a penalty function score equal to the objective function $p(Ch) = f(Ch)$, something that guarantees that all constraints are met for the departure time deviations denoted by that child.

At some point, parent sets Pa, Pb might have become too similar after their continuous replacement with better-performing children. Then, the only way to generate a new child with different characteristics is via mutation but this might take too long if the mutation rate $m_u = 1$. For this reason, if at some point $b \in \{1, 2, \ldots, \rho\}$ we have $Pa_b = Pb_b$ we increase the mutation rate probability to $m_u = 7$. The same applies when for too many iterations (more than 500) we do not have a new child that performs better than its parents. When this new child is finally found, we turn again the mutation rate to $m_u = 1$ for avoiding a totally random generation of new children.

The final stage is the heuristic search termination criterion which is controlled by two elements. The first is the maximum allowed number of iterations *max_it* and the second is the convergence rate. For the convergence rate criterion, if for many iterations we do not have a new child that perfor *failed_changes_counter* $\leftarrow 0$ ms better than its parents even if the mutation rate is set to $m_u = 7$ and the previous children that replaced their parents had too close performance while also $p(Pa) = f(Pa)$ and $p(Pb) = f(Pb)$, we can assume that we have reached a minimum point (bottom of a valley) that can be any of the local optimums or the global optimum. Since we do not have any mechanism to justify that we have reached to the global optimum, we can terminate the algorithm assuming that we are close to the global optimum. The heuristic search method is summarized in the following algorithm:

---

function SEQUENTIAL HEURISTIC SEARCH $(p(x) = f(x))$
  Set $m_u \leftarrow 1$, $q = \{-5, -4, \ldots, 0, \ldots, +5, +5\}$ and iteration $\leftarrow 0$
  With an *initial random guess* choose a parent 1 set $Pa = \{Pa_1, Pa_2, \ldots, Pa_\rho\}$ and compute $p(Pa)$;
  With an *initial random guess* choose a parent 2 set $Pb = \{Pb_1, Pb_2, \ldots, Pb_\rho\}$ and compute $p(Pb)$;
  With an *initial random guess* choose a child set $Ch = \{Ch_1, Ch_2, \ldots, Ch_\rho\}$;
  while iterations $\leqslant$ max_*it* (1st termination criterion) **do**
    iteration $\leftarrow$ iteration $+ 1$
    for i in range $\{1,2,\ldots,\rho\}$ **do**
      Crossover: Set c $\leftarrow$ *random choice* between Pa_i and Pb_i;
      Mutation: if $a \leqslant m_u$ where a is a random choice number from the range $\{0, 1, \ldots, 9, 10\}$, then set c $\leftarrow$ *randomchoice* value from the set q;
      Replace the i^{th} element of the Ch vector with c;
      Compute $p(Ch)$;
      if $p(Ch) < p(Pa)$ or $p(Ch) < p(Pb)$ **then**
        Elitism: Replace the parent with the highest penalty function with Ch;
        Set $m_u \leftarrow 1$;
        Set failed_changes_counter $\leftarrow 0$;
      else
        Undo the replacement of the ith element of the Ch vector with c (set it back to its previous Ch_i value);
        Set *failed_changes_counter* $\leftarrow$ *failed_changes_counter* $+ 1$;
      endif
      if *failed_changes_counter* > 500 **then**
        Set $m_u \leftarrow 7$;
      endif
      if $p(Pa) = f(Pa), p(Pb) = f(Pb), p(Pa) \approx p(Pb)$ **then**
        If for many generations no child had better performance than its parents: *failed_changes_counter* > 10000
        Assume convergence and terminate (2nd termination criterion);
      endif
    endfor
  endwhile
  return optimal solution x^* where $x^* = Pa$ if $p(Pa) \leqslant p(Pb)$ and $x^* = Pb$ if $p(Pb) \leqslant p(Pa)$
end function

The sequential heuristic search has computational complexity $O(max\_it \times \rho)$ and if we have L public transport services that require rescheduling this complexity becomes $O(L \times max\_it \times \rho)$. With the sequential heuristic search the time complexity is reduced from exponential to polynomial and a solution search is feasible even for large values of ρ; however, our solution converges to a stochastic global optimal which cannot be guaranteed that is the real global optimal.

The re-scheduled departure times that improve the service-wide EWT and adjust the service to the joint leisure activity demand should be finally adopted by the Demand Responsive public transport service. For this, a seamless operation Command Center receives information from Smartphone Apps for deriving the current location of individuals on the transport network and the location and arrival times of their planned joint leisure activities. Later, the heuristic sequential search solution method is applied for computing the desired departure times of the public transport modes (Fig. 1).

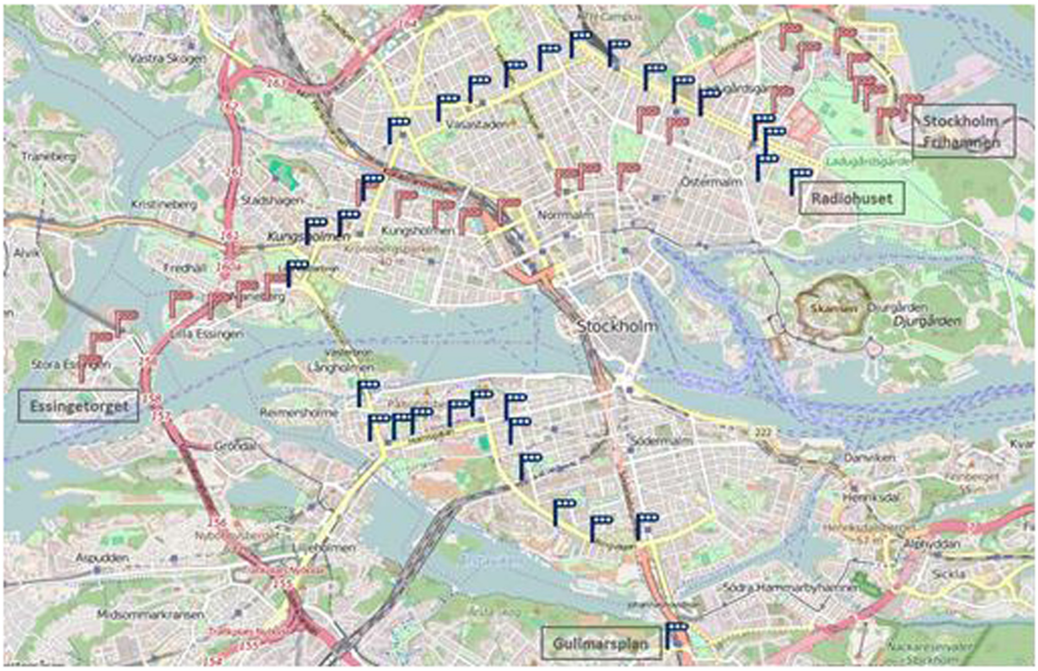

Fig. 2. Web-based visualization after querying bus services 1(bus stations in red), 4(bus stations in blue) from Sweden GTFS data with the use of Python GTFS library and OpenLayers.js (Fig. 3).

4. Case study

In this study, we utilized GTFS data from Sweden including the planned schedule of public transport modes for the period 13 February 2016–17 June 2016. The data includes the file presented at Table 1.

For deriving the planned schedules of public transport modes, a library was developed in Python 2.7. The library processes .txt files and converts/stores them to an sql database. This facilitates data queries and enables web-based visualization (Fig. 2) of the public transport operations with the use of OpenStreetMap (via OpenLayers, an open-source JavaScript library: http://www.openlayers.org/api/OpenLayers.js). The developed Python GTFS library: (i) converts .txt files to sql database tables, (ii) can query public transport routes from the database tables, (iii) creates new files containing the planned trips for each route in ascending order (starting from the earlier morning trip to the latest night trip).

After applying the Python GTFS library, we sorted the planned trips for every public transportation service and for each trip we have the planned arrival time at every station. In particular, we focused on two bi-directional central bus lines (1 and 4) in Stockholm which can be seen as 4 independent services because every line direction has another EWT score and another set of constraints. Those are bus service 1 (bus line 1, direction 1 (Essingetorget to Stockholm Frihamnen)), bus service 2 (bus line 1, direction 2 (Stockholm Frihamnen to Essingetorget)), service 3 (bus line 4, direction 1 (Gullmarsplan to Radiohuset)) and service 4 (bus line 4, direction 2 (Radiohuset to Gullmarsplan)). In Fig. 3 the stations of services 1, 2, 3, 4 for both directions are presented together with the EWT at each station calculated from the planned arrival times according to the Eq. (4).

Apart from public transportation data, the data of travelers that are heading to joint leisure activities is also required. One method for acquiring this sort of data is mining social media and especially posts from Smartphones that include geo-location information. For instance, in the work of Gkiotsalitis and Stathopoulos (2015), user-generated data from Smatrphones that post on Twitter was collected consistently over 14 months for capturing the mobility patterns of a portion of individuals and retrieving automatically users' willingness to travel certain distances for participating in different activity types over different days and times. The location and time of each joint leisure activity can also be optimized considering the preferences of participants as described at Gkiotsalitis and Stathopoulos (2016). Both works can provide valuable input regarding the expected starting times and locations of joint leisure activities which can be then used for adjusting the time-plans of public transport services to the joint leisure activity needs.

In this work, we performed a data crawling campaign in Stockholm. For the targeted data mining of Twitter, a platform in Python 2.7 was developed. Several libraries for authentication and data conversion were utilized (json, simplejson, oauth2, httplib2). In addition, we utilized Python Twitter, which is a Python wrapper around the Twitter API. After this campaign we retrieved geo-location information from $N_E^* = 62$ individuals from Stockholm who were traveling from their current location to other locations of joint leisure activities; therefore we did not have to rely on prediction models as those developed by

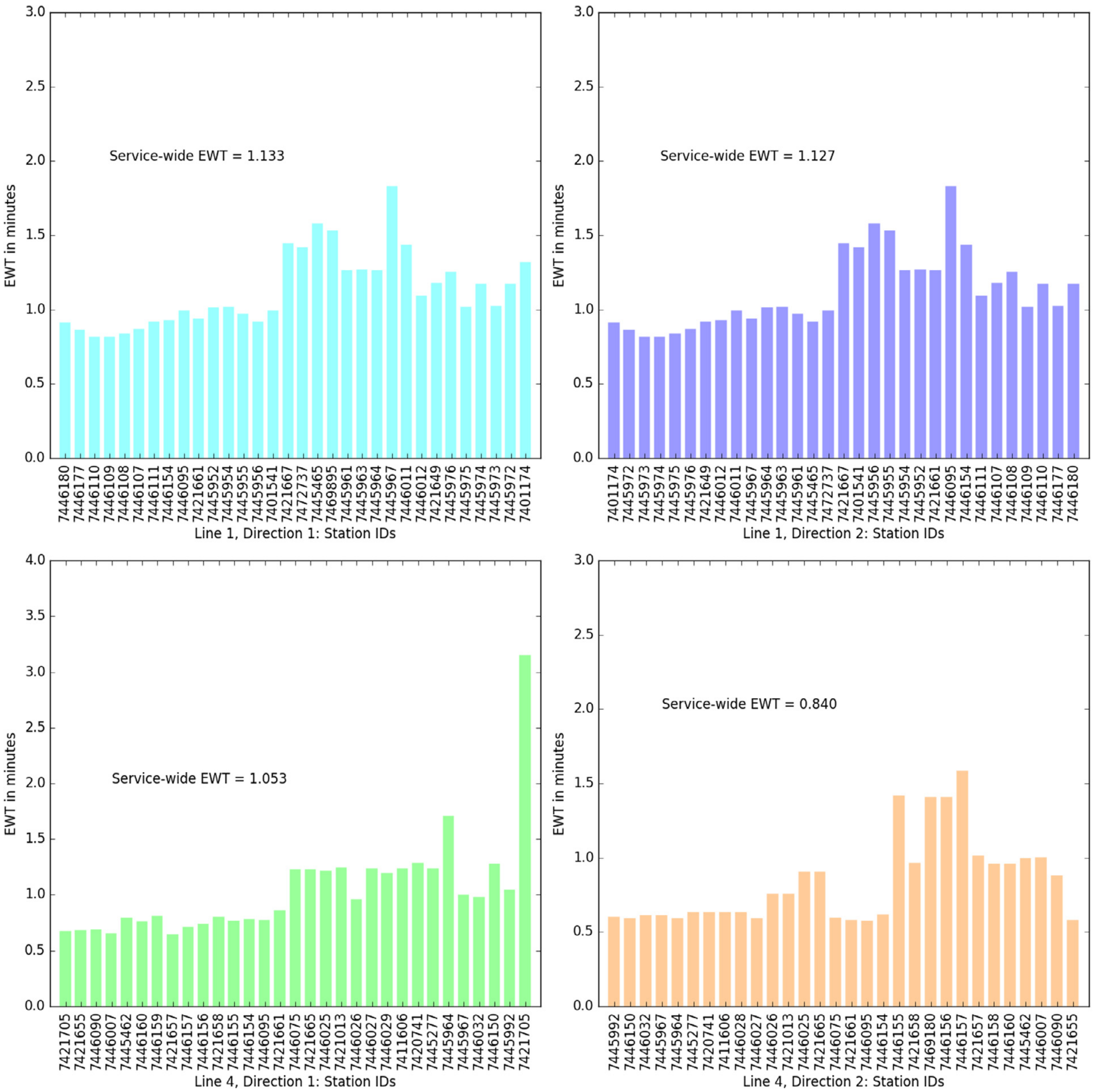

Fig. 3. EWT at every station and service-level EWT for every service according to the planned scheduled of daily trips for bus lines 1 and 4 in both directions for the afternoon EWT phase (time period 2:10 pm-7:30 pm).

Table 1
GTFS data files and size (total size: 0.25 GB).

File name	Size	File name	Size
agency.txt	5 KB	stop_times.txt	242,000 KB
calendar.txt	17 KB	stops.txt	6,384 KB
calendar_dates.txt	890 KB	transfers.txt	1,020 KB
routes.txt	218 KB	trips.txt	11,628 KB

Gkiotsalitis and Stathopoulos (2015, 2016). Those 62 individuals were selected according to the following criteria: (i) they change geo-location outside working hours (from 15:30–19:30); (ii) the location they are heading is not related to work/ home, but it is a leisure activity location for them; (iii) their final destination is close to one of the interchange stations of

Table 2
Nearest stations to individuals who are performing a joint leisure activity.

Individuals who are performing similar joint leisure activities	($N_E^* = 62$: individual trips from current station to Joint Leisure Activity Station)	Trip Destination Bus Station (Joint Leisure Activity Station)	Event Time
Trip Departure Bus Station	7445462, 7446157, 7421658, 7469180, 7401541	7446154	17:00
Trip Departure Bus Station	7445961, 7401541, 7445964, 7421665, 7421657, 7446075, 7446028, 7421661	7445967	16:30
Trip Departure Bus Station	7446095, 7446026, 7446025, 7445954, 7421649, 7421661, 7446177, 7445976, 7445974, 7446026, 7445955	7445964	16:00
Trip Departure Bus Station	7446108, 7446075, 7445992, 7421705	7445964	19:30
Trip Departure Bus Station	7446108, 7446177, 7446160, 7421657, 7446011, 7469180, 7446156, 7446109, 7446177	7446095	17:00
Trip Departure Bus Station	7446007, 7445972, 7445465, 7445952, 7446150	7446095	18:30
Trip Departure Bus Station	7446090, 7445964, 7421661, 7446154, 7445952, 7446160, 7446157	7421661	15:30
Trip Departure Bus Station	7446029, 7421705, 7420741, 7469180, 7421705, 7421665, 7445992	7421661	18:00
Trip Departure Bus Station	7421667, 7445976, 7446095, 7446095, 7445967, 7421655	7445952	17:00

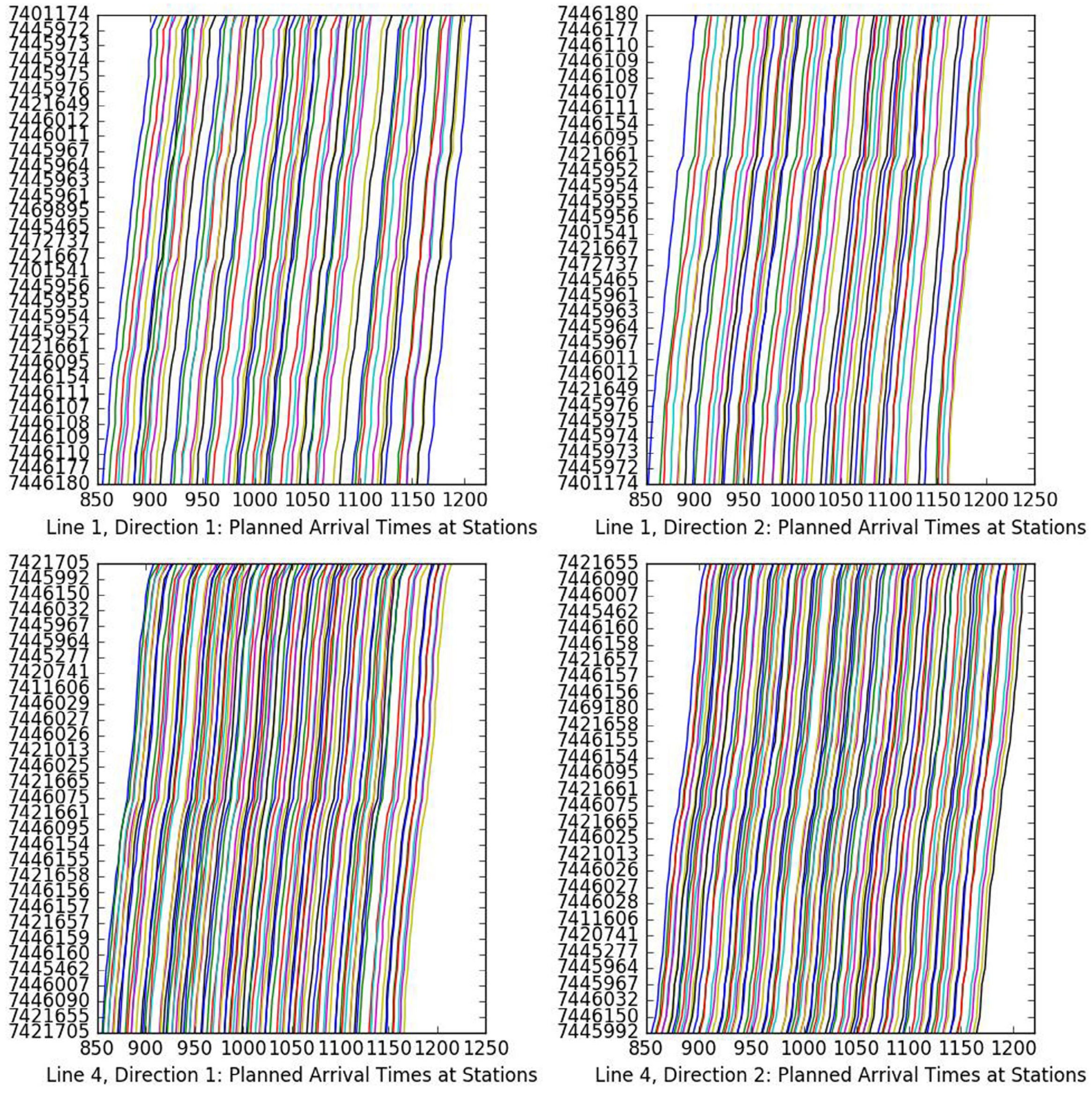

Fig. 4. Planned arrival times of trips at evert station for bus services 1, 2, 3, 4 from 2:10om (850 min.) until 7:30 pm (1170 min.).

Table 3
Trips IDs of π=97 trips that can be used from individuals for arriving to the locations of joint leisure activities[t3].

Trip IDs ($\pi_\kappa \in \{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$)	Line 1, Direction 1: 90656299, 90656481, 90656480, 90656293, 90656392, 90656601, 90656605, 90656408, 90656408, 90656613, 90656613, 90656607, 90656607, 90656409, 90656409, 90656615, 90656337, 90656725, 90656801, 90656271, 90656794 Line 1, Direction 2: 90655951, 90655947, 90655891, 90655953, 90655963, 90655963, 90655963, 90655886, 90655886, 90655886, 90655964, 90655824, 90655769, 90655903, 90655767, 90655938, 90655927, 90656128, 90656164 Line 4, Direction 1: 90661185, 90661816, 90661826, 90661821, 90660993, 90661189, 90661798, 90661808, 90661687, 90661832, 90661691, 90661835, 90661835, 90661519, 90661519, 90661519, 90661813, 90661693, 90661693, 90661428, 90660717, 90661426, 90660755, 90660756, 90661494, 90661651, 90661325 Line 4, Direction 4: 90661590, 90661059, 90660931, 90660962, 90661597, 90661585, 90661614, 90661616, 90661616, 90661037, 90661037, 90661588, 90661588, 90660956, 90660956, 90660956, 90661586, 90661586, 90661586, 90660939, 90661591, 90660946, 90660905, 90661035, 90660908, 90660977, 90661069, 90660918, 90661127, 90660920

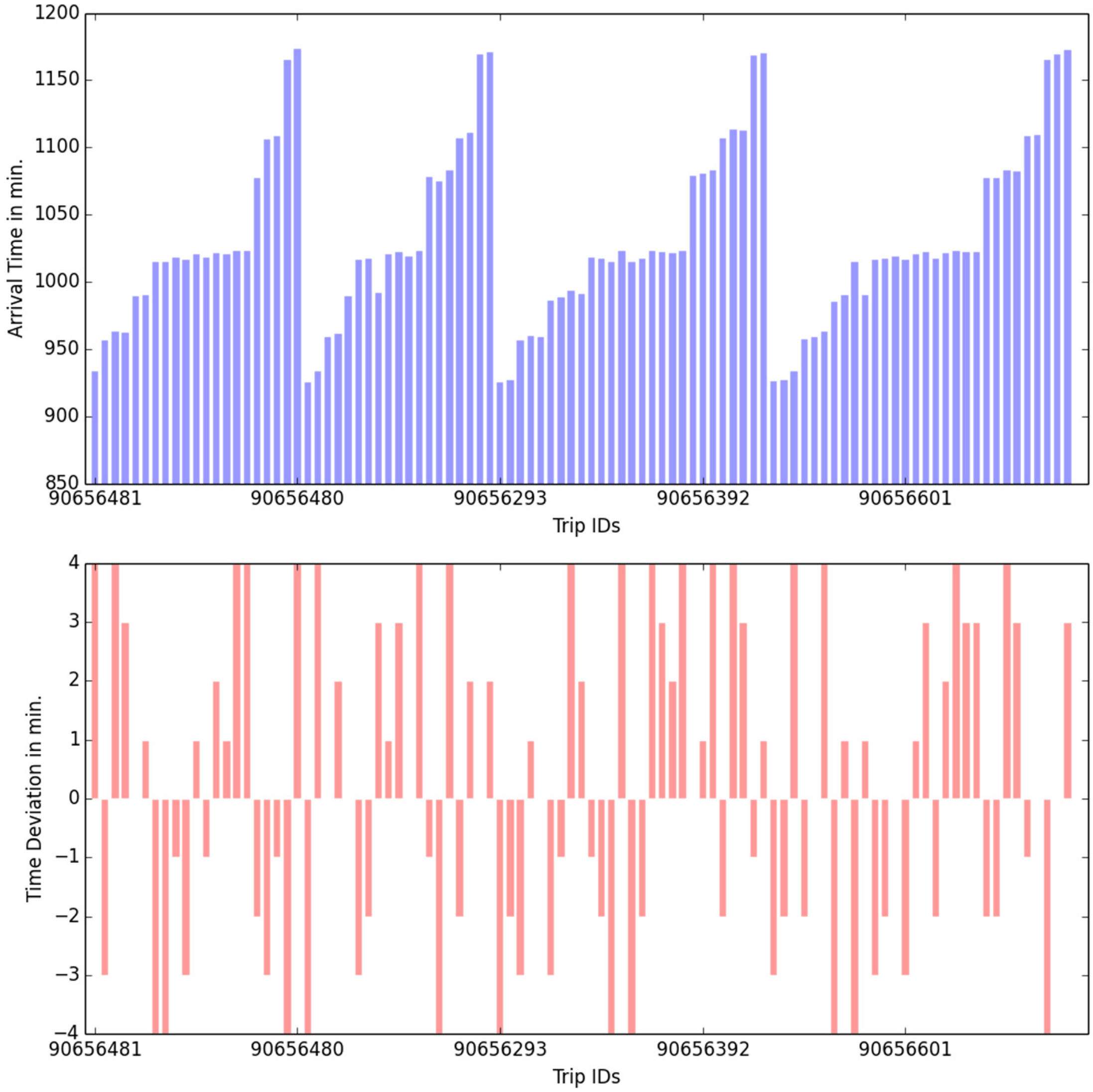

Fig. 5. Planned arrival times of trips $\{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$ at joint leisure activity stations $\{\Lambda_1, \Lambda_2, \ldots, \Lambda_{|\Lambda|}\}$ and their time deviation from the starting time of the activity in min. A positive deviation means that the bus arrived minutes after the activity started and a negative that arrived before.

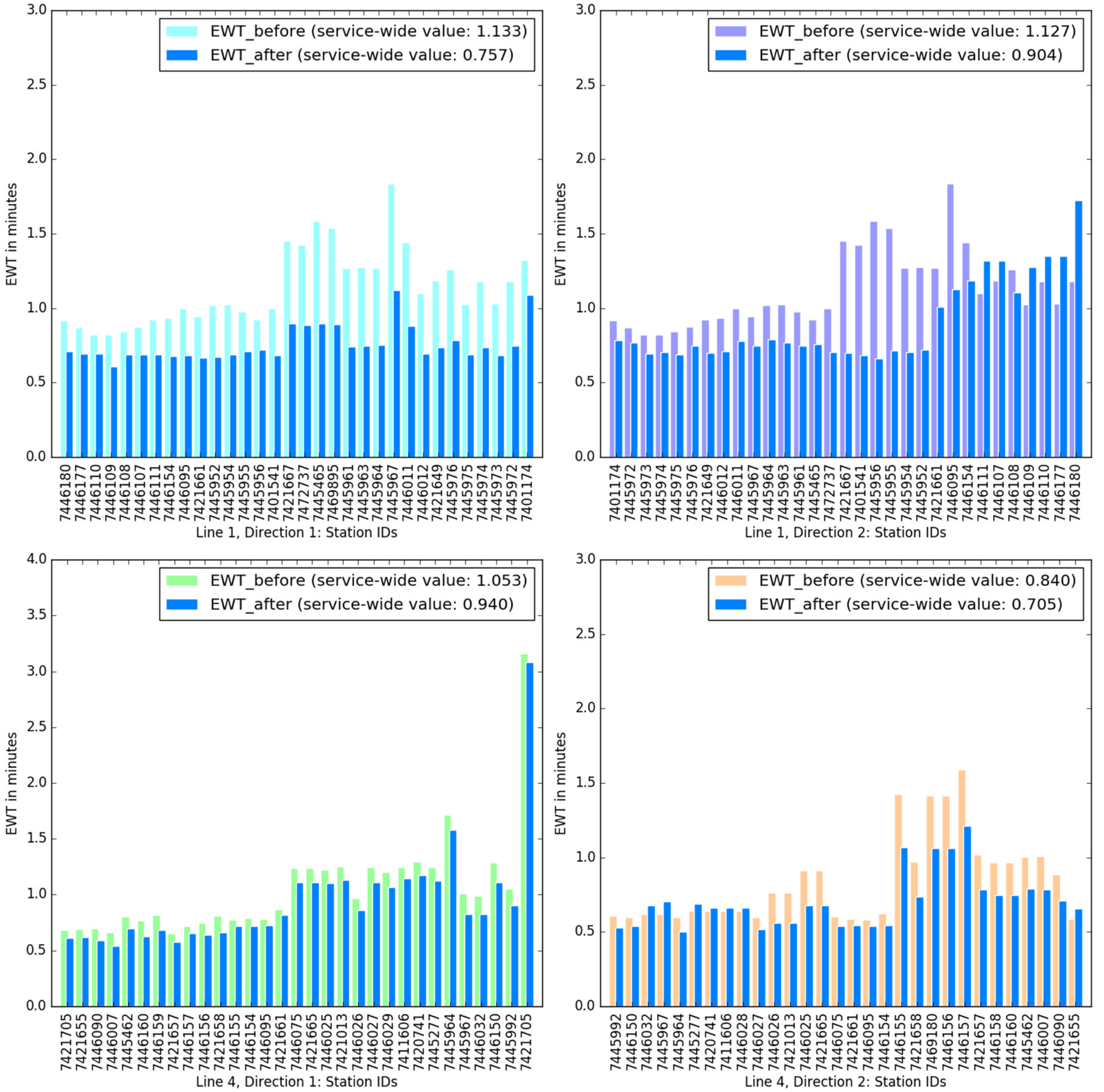

Fig. 6. EWT at every station and service-level EWT for every service before and after the heuristic search re-scheduling for the afternoon EWT phase (time period 2:10 pm–7:30 pm).

Table 4
Re-scheduling summary results for every service before and after the sequential heuristic search optimization.

		Service-wide EWT	Service-wide ASQ scores	Computational time (sec)
Line 1 Direction 1	Before	1.133273406	178	
	After	**0.75732**	**67**	129.44
Line 2 Direction 2	Before	1.127213956	129	
	After	**0.904360219**	**44**	112.981
Line 4 Direction 1	Before	1.05272815	210	
	After	**0.94029**	**97**	263.067
Line 4 Direction 2	Before	0.840175194	216	
	After	**0.704939879**	**105**	199.558

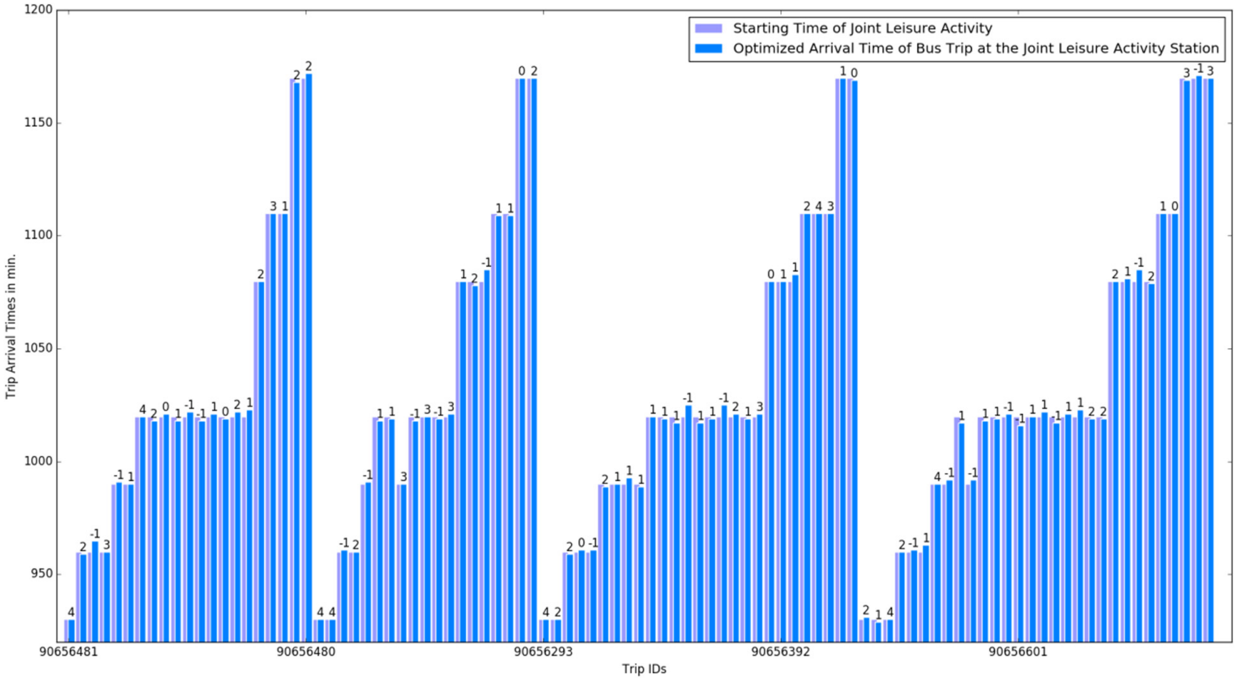

Fig. 7. Starting times of Joint Leisure activities and Re-scheduled arrival times of trips $\{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$ at joint leisure activity locations $\{\Lambda_1, \Lambda_2, \ldots, \Lambda_{|\Lambda|}\}$ in minutes. For every trip, the re-scheduling improvement in terms of adjustment closer to the starting activity time is also expressed in minutes (negative values indicate deterioration).

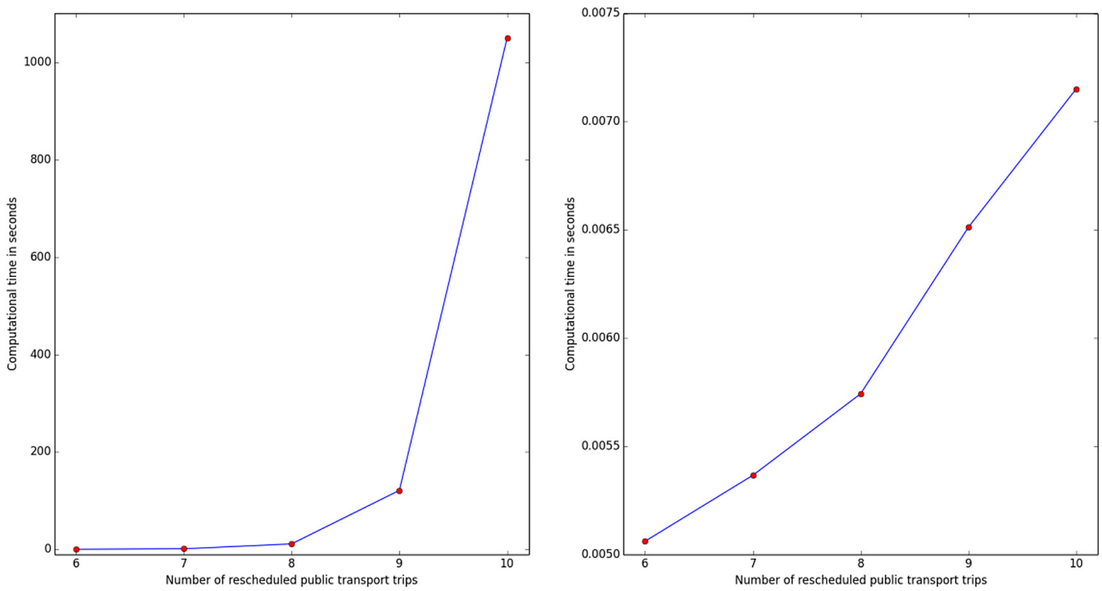

Fig. 8. Computing time for re-scheduling public transport operations related with joint leisure activities with brute force (left) and heuristic search (right). Only up to $\pi = 10$ trips were able to be tested due to the exponential computational cost of brute-force.

bus lines 1 and 4 (namely, bus stations 7446154, 7446095, 7421661, 7445964, 7445967 and 7445952); therefore, they have the option to use bus line 1 or bus line 4; (iv) their travel origin is within walking distance from at least one bus station of lines 1 and 4; (v) they arrive at the location of the joint leisure activity within a time variance of less than 25 min.; therefore, we assume that they participate at the same activity. Criterion (v) is our strongest assumption since multiple joint leisure activities can occur nearby concurrently. However, by depending solely on Twitter data we cannot justify the potential splitting of a joint leisure activity into multiple concurrent ones with a certain degree of confidence. The current locations and the final destinations of groups of individuals who are heading to the same leisure activity location are presented in Table 2.

From the above, we have $\Lambda = 9$ different joint leisure activity locations and each one has the following number of participants who are heading there from their previous location: $N_E^* = 62$, where $N_E^* = N_E(\Lambda_1) + N_E(\Lambda_2) + \ldots + N_E(\Lambda_9)$ and $N_E(\Lambda_1) = 5$, $N_E(\Lambda_2) = 8$, $N_E(\Lambda_3) = 11$, $N_E(\Lambda_4) = 4$, $N_E(\Lambda_5) = 9$, $N_E(\Lambda_6) = 5$, $N_E(\Lambda_7) = 7$, $N_E(\Lambda_8) = 7$, and $N_E(\Lambda_9) = 6$.

On another note, we utilized the Python GTFS library to query the database tables and select the number of public transport trips, π, that can serve those joint leisure activities. First, we present in Fig. 4 all planned trips from 14:10(850 min.)–19:30(1170 min.) of bus lines 1, 4 on both directions together with their planned arrival times at every bus station. Then, we select those trips, π, that can be used from the $N_E^* = 62$ individuals for arriving at the locations of joint leisure activities (refer to Table 3).

Those trips are planned to arrive at the joint leisure activity locations $\{\Lambda_1, \Lambda_2, \ldots, \Lambda_{|\Lambda|}\}$ at times which are close to the starting time of the joint leisure activity. In Fig. 5 we present the planned arrival times of trips $\{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$ at every joint leisure activity station and their deviation from the starting time of the joint leisure activities which should be as close as possible to zero after the re-scheduling phase.

After optimizing the schedule for the bus services with the sequential heuristic search method, we present the new EWTs at every station and the new service-wide EWT in Fig. 6 in comparison with the planned schedule EWTs at the do-nothing scenario. While re-scheduling each one of the public transport services that contain potential joint leisure activity trips $\{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$ we try to: (a) minimize the deviation between the arrival time of each trip at the joint leisure activity location and the starting time of the activity (presented at Fig. 5), (b) disturb only the planned trips which are at the close vicinity of trips $\{\pi_1, \pi_2, \ldots, \pi_{\pi=97}\}$ (less than 60minutes deviation), (c) minimize the service-wide EWT score, (d) do not violate the planned frequency range of successive trips derived from the tactical planning phase, (e) ensure that the new EWT will at least not be worse than the service-level EWT of the original timeplan.

In addition, the service-wide EWT score before and after the optimization together with the value of the penalty functions at the convergence state are presented in Table 4. In that table we show also how the arrival time of trips π at the joint leisure activity locations has been adjusted to the joint leisure activity starting times by presenting the time deviation scores (computed from Eq. (2)) before and after the rescheduling. Those scores are the aggregated values of the squared difference of the trips' arrival times to joint leisure activity locations and the starting times of those activities represented by the abbreviation ASQ scores.

For demonstration purposes also, we present the re-scheduled arrival times of bus trips $\pi_i \in \pi$ which can be potential joint leisure activity trips at the location of those activities. We present also the starting time of those activities. The re-scheduled trips have closer arrival times to the starting times of the activities after optimizing them with the sequential heuristic search method. How closer are the re-scheduled trip arrival times to the activity starting times compared to the planned arrival times is also presented in Fig. 7.

Finally, to show more detailed results on the computational costs, we present also different scenarios with different numbers of public transport mode trips $\pi \leqslant 10$ in Fig. 8, since a brute-force computation for $\pi = 97$ is not feasible due to the exponential time complexity.

5. Concluding remarks

In this work, we focused on the area of joint leisure activities. In particular, we modeled the public transport rescheduling for different public transport services that can cover joint leisure activity passenger demand subject to the no-deterioration of service quality and the adherence to a set of operational regulations, such as the adherence to the frequency setting ranges. Finally, we presented a sequential heuristic search algorithm for changing the public transport schedules in near real time (in a matter of minutes) for adjusting to the arrival time needs of joint leisure activity passengers without deteriorating the KPIs of public transport operations.

In the case study, we used GTFS data from Sweden, focusing on bus lines 1 and 4 in Stockholm and Twitter data for deriving individual trips to joint leisure activity locations in Stockholm. In Fig. 6 the re-scheduling changes on the bus line 1, 4 EWTs were presented. Due to the schedule changes, the operational performance of bus services demonstrated an EWT improvement at a service-wide level for all services while only some stations from line 1, direction 2 and line 4, direction 4 had a slight EWT deterioration(up to 0.6 min.) without affecting significantly the level of service of bus operations. At the same time, after re-scheduling joint leisure activity trips from all services enjoyed new arrival times to the joint activity stations which were closer to the starting times of those activities. Finally, the computational cost of the proposed heuristic algorithm for public transportation re-scheduling was presented in Table 4 demonstrating that a stochastic convergence to the global optimum requires from 2–6 min.

In future research, the public transport mode optimization method for covering joint leisure activity trips can be tested in more complex scenarios that include the entire public transport network of a city (bus/train services). Following this direction, the importance of muti-modality (i.e., mixed use of public and private transportation) for improving the public transport ridership related to joint leisure activities can be also studied.

References

Alesiani, F., Gkiotsalitis, K., Baldessari, R., 2014. A probabilistic activity model for predicting the mobility patterns of homogeneous social groups based on social network data. Transportation Research Board 93rd Annual Meeting. No. 14-1033.

Arentze, T., Timmermans, H., 2008. Social networks, social interactions and activity behavior: a framework for micro-simulation. Environ. Plann. B: Plann. Des. 35, 1012–1027.
Asensio, J., 2002. Transport mode choice by commuters to Barcelona's CBD. Urban Stud. 39 (10), 1881–1895.
Bowman, J., Ben-Akiva, M., 2000. Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. Part A 35, 1–28.
Calabrese, F., Lorenzo, G.D., Liu, L., Ratti, C., 2011a. Estimating origin-destination flows using mobile phone location data. IEEE Pervasive Comput. 10, 36–44.
Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., Ratti, C., 2011b. Real-time urban monitoring using cell phones: a case study in Rome. IEEE Trans. ITS 12, 141–151.
Carrasco, J.A., Hogan, B., Wellman, B., Miller, E.J., 2008. Collecting social network data to study social activity-travel behavior: an ego-centric approach. Environ. Plann. B: Plann. Des. 35, 961–980.
Castex, E., Houzet, S., Josselin, D., 2004. Prospective research in the technological and mobile society: new Demand Responsive Transports for new territories to serve. Proceedings of the 7th AGILE Conference, Heraklion, vol. 29.
Chen, Y., Frei, A., Mahmassani, H.S., 2014. From personal attitudes to public opinion: Information diffusion in social networks toward sustainable transportation. Transportation Research Board 93rd Annual Meeting.
De Domenico, M., Lima, A., Musolesi, M., 2013. Interdependence and predictability of human mobility interactions. Pervasive Mobile Comput. 9, 798–807.
Dickinson, J.E., Ghali, K., Cherrett, T., Speed, C., Davies, N., Norgate, S., 2014. Tourism and the smartphone app: Capabilities, emerging practice and scope in the travel domain. Curr. Issues Tourism 17 (1), 84–101.
Gkiotsalitis, K., Stathopoulos, A., 2015. A utility-maximization model for retrieving users' willingness to travel for participating in activities from big-data. Transp. Res. Part C: Emerging Technol. 58, 226–277.
Gkiotsalitis, K., Stathopoulos, A., 2016. Joint leisure travel optimization with user-generated data via perceived utility maximization. Transp. Res. Part C: Emerging Technol. 68, 532–548.
Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.L., 2008. Understanding individual human mobility patterns. Nature 453, 779–782.
Janssens, D., Wets, G., Brijs, T., Vanhoof, K., Arentze, T., Timmermans, H., 2004. Improving performance of multiagent rule-based model for activity pattern decisions with Bayesian networks. Transp. Res. Record: J. Transp. Res. Board, 75–83.
Mauri, G.R., Lorena, L.A.N., 2009. Customer's satisfaction in a dial-a-ride problem. IEEE ITS Mag. Fall 9, 6–14.
Miller, E.J., Roorda, M.J., Carrasco, J.A., 2005. A tour-based model of travel mode choice. Transportation 32 (4), 399–422.
Musolesi, M., Mascolo, C., 2007. Designing mobility models based on social network theory. ACM SIGMOBILE Mobile Comput. Commun. Rev., 59–70 11.3.
Sohn, K., Kim, D., 2008. Dynamic origin-destination flow estimation using cellular communication system. IEEE Trans. Veh. Technol. 57, 2703–2713.
Sun, Y., 2016. Investigating "Locality" of intra-urban spatial interactions in New York city using foursquare data. ISPRS Int. J. Geo-Inf. 5 (4), 43.
Transport for London, Travel in London, London Travel Demand Survey. http://www.tfl.gov.uk/cdn/static/cms/documents/london-travel-demand-survey.pdf, 2014, Accessed: 2014-09-30.
Zhang, Y., Qin, X., Dong, S., Ran, B., 2010. Daily O-D matrix estimation using cellular probe data. Transportation Research Board: 89th Annual Meeting.