Transportation Research Part B 100 (2017) 1–19

Contents lists available at ScienceDirect

Transportation Research Part B

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

Deficit function related to public transport: 50 year retrospective, new developments, and prospects

Tao Liu[a], Avishai (Avi) Ceder[a,b,*]

[a] Department of Civil and Environmental Engineering, The University of Auckland, Auckland 1142, New Zealand
[b] Technion-Israel Institute of Technology, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 24 September 2016
Revised 18 December 2016
Accepted 31 January 2017
Available online 3 February 2017

Keywords:
Deficit function
Public transport
Operations planning
Human-machine interactive
Graphical optimization method
History of work done

ABSTRACT

The Deficit Function (DF), with its graphical concept and modelling, was introduced 50 years ago by Linis and Maksim (1967) under the title of "On the problem of constructing routes." Since then, there have been many developments in the understanding of the theoretical, methodological and application aspects of the DF concept. This work, for the first time, makes a comprehensive and thorough retrospective examination of the major developments of DF modelling and applications in public transport (PT) planning and operations over the past 50 years, introduces some new developments, and offers future research directions. It is shown and proven that the graphical DF concept helps in creating efficient PT vehicle schedules, timetables, crew duties, networks of routes, bus rapid transit systems, and operational parking spaces. For instance, in one large bus company the total number of vehicles and crew duties were reduced by 6% to 12% and 8% to 15%, respectively. This work intends to stimulate further use of the DF concept as a bridge between the world of researchers and the world of practitioners.

1. Introduction

The public transport (PT) planning and operations process commonly includes five basic activities, usually performed in sequence: (1) network route design, (2) timetable development, (3) vehicle scheduling, (4) crew scheduling, and (5) real-time monitoring and control (Ceder, 2007a, 2016; Ibarra-Rojas et al., 2015; Liu et al., 2016). In order to make use of the system's capability to the greatest extent and maximize its productivity and efficiency, all five activities should be planned simultaneously. However, this integrated planning process, especially for medium and large scale PT agencies, is extremely cumbersome and complex. Therefore, separate treatment is required for each activity, with the outcome from one activity fed as an input into the next. As a result, this process needs to be supported by efficient and effective methods and tools to attain optimal resource allocation, e.g., vehicles and drivers, while maintaining an adequate level of service. To that end, a highly informative human-machine interactive graphical optimization technique has been demonstrated to be a very useful tool for PT planning and operations (Gertsbach and Gurevich, 1977; Ceder, 2007a, 2016). It is known as deficit function (DF), first introduced 50 years ago, by Linis and Maksim (1967) for estimating an optimal fleet size required for a fixed schedule.

The DF graphical optimization technique has been widely used in various PT planning and operations activities. The main advantage of the DF is its visual nature. The DF optimization process is performed in a conversational human-machine mode. It allows experienced schedulers to select computer-generated improvements or interject their own practical considerations,

* Corresponding author.
E-mail addresses: tliu773@aucklanduni.ac.nz (T. Liu), a.ceder@auckland.ac.nz (A. (Avi) Ceder).

http://dx.doi.org/10.1016/j.trb.2017.01.015

and to further advise themselves on evaluating the immediate effects on the final results. PT schedulers are in control of the steps included in the evaluation process and can feel the improvements, which improves their confidence and comfort in using the DF-based human-machine interactive approach.

The year 2017 marks the 50th anniversary of the first paper published on the DF. To commemorate this historical event in transportation science, in this work we have made a comprehensive review of the main contributions in the field throughout the history of the DF, and have cited some new developments. The purpose is to make a systematic, comprehensive and thorough retrospective examination of the major developments of the DF and their applications in PT planning and operations in the past 50 years and to suggest prospects for future research.

This work is comprised of six sections including this introduction. The second section provides a description of the DF theory, which includes the definitions, properties, theorems and a mirror image of the DF. A systematic overview of various applications of the DF in PT planning and operations activities is provided in the third section. The fourth section describes some DF-based computer-aided PT scheduling software packages and provides example of cost saving results in terms of saving vehicles and crew duties. Promising research directions are provided in the fifth section. Finally, the sixth section concludes this work.

2. Deficit function theory – background

In this section, a concise description of the definitions, properties, theorems and the mirror image of the DF theory is provided as explanatory background. They are based primarily on the works of Linis and Maksim (1967), Gertsbach and Gurevich (1977), Ceder and Stern (1981), Stern and Ceder (1981, 1983a), Ceder (1991a, 2002, 2003a, 2003b, 2007a, 2016) and Ceder and Israel (2009).

2.1. Definitions

The DF is a step function measured at a particular terminal in a multi-terminal PT system. The DF increases by one at the time of each trip departure and decreases by one at the time of each trip arrival.[1] Linis and Maksim (1967) and Gertsbach and Gurevich (1977) have called this step function a DF as its value represents the deficit number of vehicles required at a particular terminal in question in a multi-terminal PT system. To construct a set of DFs, the only information needed is a timetable of required trips. The DF has its appeal in its graphical nature and visual simplicity.

Let $I = \{i: i = 1, \ldots, n\}$ denote a set of required trips. The trips are conducted between a set of terminals $K = \{k: k = 1, \ldots, q\}$, each trip is serviced by a single vehicle, and each vehicle is able to service any trip. Each trip i can be represented as a 4-tuple (p^i, t_s^i, q^i, t_e^i), in which the ordered elements denote departure terminal, departure (start) time, arrival terminal, and arrival (end) time. It is assumed that each trip i lies within a schedule horizon $[T_1, T_2]$, i.e., $T_1 \le t_s^i \le t_e^i \le T_2$. The set of all trips $S = \{(p^i, t_s^i, q^i, t_e^i) : p^i, q^i \in K, i \in I\}$ constitutes the timetable. Two trips i, j may be serviced sequentially (feasibly joined) by the same vehicle if and only if (a) $t_e^i \le t_s^j$ and (b) $q^i = p^j$. If i is feasibly joined to j, then i is said to be the predecessor of j, and j the successor of i. A sequence of trips $i_1, i_2, \ldots, i_w$ ordered in such a way that each adjacent pair of trips satisfies (a) and (b) is called a vehicle chain or block. It follows that a chain is a set of trips that can be serviced by a single vehicle. A set of chains in which each trip i is included in I exactly once is said to constitute a vehicle schedule.

Let $d(k,t,S)$ denote the DF (the number of vehicles required) for terminal k at time t for schedule S. The value of $d(k,t,S)$ represents the total number of departures minus the total number of trip arrivals at terminal k, up to and including time t. The maximum value of $d(k,t,S)$ over the schedule horizon $[T_1, T_2]$, designated $D(k,S)$, depicts the deficit number of vehicles required at k. These DF notations are used for the example in Fig. 1, in which $[T_1, T_2] = [5{:}00, 8{:}30]$.

2.2. Properties and theorems

Note that S will be deleted when it is clear which underlying schedule is being considered. It is possible to partition the schedule horizon of $d(k,t)$ into a sequence of alternating hollow and maximum intervals $(H_0^k, M_1^k, H_1^k, .., H_j^k, M_{j+1}^k, \ldots, M_{n(k)}^k, H_{n(k)}^k)$. Maximum intervals $M_j^k = [s_j^k, e_j^k]$, $j = 1,2,\ldots,n(k)$ define the intervals of time over which $d(k,t)$ takes on its maximum value. Index j represents the jth maximum intervals from the left; $n(k)$ represents the total number of maximal intervals in $d(k,t)$, where s_j^k is the departure time for a trip leaving terminal k and e_j^k is the time of arrival at terminal k for this trip. The one exception occurs when the DF reaches its maximum value at $M_{n(k)}^k$ and is not followed by an arrival, in which case $e_{n(k)}^k = T_2$. A hollow interval H_j^k, $j = 0,1,2,\ldots,n(k)$ is defined as the interval between two maximum intervals. This includes the first hollow, from T_1 to the first maximum interval, $H_0^k = [T_1, s_1^k]$; and the last hollow, which is from the last interval to T_2, $H_{n(k)}^k = [e_{n(k)}^k, T_2]$. Hollows may contain only one point. If this case is not on the schedule horizon boundaries (T_1 or T_2), the graphical representation of $d(k,t)$ is emphasized by a clear dot.

The sum of all DFs over k is defined as the overall DF, $g(t) = \sum_{k\in K} d(k,t)$. This function g(t) represents the number of trips that are simultaneously in operation, i.e., a count, from a bird's-eye view at time t, of the number of transit vehicles in

[1] Note that the term 'trip' is used to denote vehicle trips rather than passenger trips.