WORKING PAPER

ITLS-WP-15-17

Frequency-based transit assignment revisited

By
Michael G. H. Bell; Michiel Bliemer and Mark Raadsen

September 2015

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER:	Working Paper ITLS-WP-15-17
TITLE:	**Frequency-based transit assignment revisited**
ABSTRACT:	This working paper reformulates the Spiess and Florian frequency-based transit assignment method in matrix algebra revealing a new solution method. It is shown that the number of destination-specific passenger wait times at stops is equal to the number of flow conservation constraints (Proposition 1). The frequency-based transit assignment is found by matrix manipulation and when there are line capacity constraints the equilibrium effective frequencies are obtained iteratively. The existence of equilibrium effective frequencies is proven (Proposition 2). It is shown that a wider range of fare schemes, for example flat fares, can be modeled by the use of legs in the network representation. Numerical examples are presented and solved by R code.
KEY WORDS:	***public transport; transit; assignment***
AUTHORS:	**Bell, Bliemer and Raadsen**
ACKNOWLEDGEMENTS:	This paper has been informed and enriched by discussions with former colleagues at Imperial College London, in particular Valentina Trozzi.
CONTACT:	INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (C13) The Australian Key Centre in Transport and Logistics Management The University of Sydney NSW 2006 Australia Telephone: +612 9114 1824 E-mail: business.itlsinfo@sydney.edu.au Internet: http://sydney.edu.au/business/itls
DATE:	September 2015

1. Previous work

Liu et al (2010) provide a comprehensive review of the route choice aspect of transit assignment. The review presented here, which loosely follows Liu et al, is couched in terms of bus networks, so refers to buses and stops. However, the methods covered are equally applicable to rail-based transit systems or indeed maritime networks, when buses are replaced by ships and passengers by containers (see Bell et al., 2011, 2013).

Early work on transit assignment assumes deterministic vehicle running times and a passenger waiting time dependent on the frequencies of lines serving a given OD pair. Dial (1967) and le Clercq (1972) look at exponentially distributed headways and random passenger arrivals. In the case of *common lines*, where a number of lines share a bus stop, passengers may have a choice of lines. Typically some lines lie on potentially optimal paths while others do not. For a given destination, Chriqui and Robillard (1975) define the optimal subset of common lines (referred to as the *attractive* lines) as that which minimizes the expected travel time to the destination when the first arriving bus from this subset is chosen. They proposed a greedy heuristic to find this choice set. When all lines have independently and exponentially distributed headways, this heuristic finds the optimal choice set. The concept of an optimal choice set has proved very influential for later work.

Nguyen and Pallotino (1988) apply the concept of the *hyperpath* in graph theory to represent the set of attractive paths, namely those defined by the optimal choice sets at the first boarding stop and subsequent interchange stops. Spiess and Florian (1989) formulate a linear program to determine for a given destination the optimal probability of choosing each line at every boarding stop under the Dial and le Clercq assumptions of Poisson passenger and bus arrivals. They also present a modification to Dijkstra's shortest path algorithm to solve the problem efficiently.

Congestion, which can be an important feature of transit networks, takes the form of queuing at stops and crowding on board. Attempts to include the effects of congestion on transit assignment in a static framework include Last and Leak (1976), Nguyen and Pallottino (1988), Spiess and Florian (1989), de Cea and Fernandez (1993), and others. Wu et al (1994) extend the work of Spiess and Florian (1989) and de Cea and Fernandez (1993) by including empirically derived link travel time functions of flow, resulting in a user equilibrium (UE) transit assignment model. Cominetti and Correa (2001) obtain the waiting time function of flow from a bulk queueing model.

There have also been stochastic user equilibrium (SUE) transit assignment models. Following the approach of Bell (1995) for traffic assignment, Lam et al. (1999) present a logit transit assignment model formulated as an entropy maximizing problem with bottleneck constraints. The dual variables for the bottleneck constraints provide estimates of equilibrium queuing times at boarding stops. Nielsen and Frederiksen (2006) propose a nested logit SUE transit assignment model while Nielsen (2000) present a multinominal probit SUE transit assignment model.

Reliability is an important concept in transit assignment and is frequently associated with congestion. Yang and Lam (2006) propose a reliability-based SUE transit assignment model with normally distributed in-vehicle travel times and concave disutility functions to represent risk averse behaviour. Schmoecker et al (2008) consider boarding as a source of unreliability in congested conditions. A failure-to-board probability is estimated for congested lines and taken into account in the utility function. Passengers who fail to board in one time slice are moved to the next time slice, so a series of static assignments linked by overflow queues at stops are considered (this approach is referred to elsewhere as quasi-dynamic).

While the *frequency-based* approach to transit assignment has captivated many because of its simplicity, power and elegance, it has limitations. Where frequencies are low and services are reliable, the schedule becomes more important. There have also been criticisms of the frequency share approach of assigning passengers to attractive common lines (Marguier and Ceder, 1984; Jansson and Ridderstolpe, 1992). Link service frequencies perform two roles; the first is the determination of delay at a boarding node or link and the second is the sharing of boarding passenger flows bound for the same destination between attractive lines. Sharing by frequency leads to a discontinuity; while a line remains attractive, the share is proportional to the frequency, however the share drops to zero if a line becomes unattractive (Nökel and Wekeck, 2009).

Another limitation relates to the Poisson assumption for bus arrivals. Bus headway distributions may deviate from exponential (the consequence of Poisson arrivals) because of bus bunching, which would tend to increase headway variability, or timing points, which would reduce headway variability. Gentile et al (2005) propose the Erlang distribution for headways as it can represent cases where a number of headways must lapse before a passenger can board and includes exponential headways as a special case. When services have an element of regularity, elapsed waiting time influences the set of attractive lines. Billi et al (2004) have analysed the impact of regularity on the set of attractive lines and shown that as waiting time elapses lines drop out of the attractive set in order of the remaining expected travel time, until a core set of lines that are always attractive remains.

This paper presents the Spiess and Florian (1989) model in matrix notation, revealing some important dimensional properties of the method. Proposition 1 proves that the number of destination-specific node delays is equal to the number of linearly independent flow conservation constraints. Making use of the matrix notation, a new way of finding a Spiess and Florian assignment is presented together with a way of calculating effective frequencies. Proposition 2 shows that a set of effective frequencies exists. In order to represent a greater range of fare structures and coordination between services, the concept of a leg is introduced. Illustrative numerical examples are presented and solved by R.

2. Definitions

The basic components of a transit network are *lines*, *stops*, *links* and *legs*. A *line* is a route on which the vehicles are operated. It consists of a series of *stops* where passengers can board or alight. A *link* is a consecutive pair of stops on a given line, so a line consists of a series of links. Boarding and alighting are represented by additional links. A transfer consists of one alighting movement followed by one boarding movement at the same stop. A *leg* is a trip by one line, beginning with the boarding link and ending with the alighting link. A trip involving more than one line would consist of a chain of legs, each leg being specific to one line. A leg uses a consecutive series of links of one line while a link may be used by a number of legs. For each boarding link there is a finite service frequency; for all other links or legs service is continuous, so the service frequency is infinite (in practice, a large number is used).

A	is the set of links and legs.
N	is the set of stops.
L	is the set of links that make up the lines (boarding and alighting links are not included).
R	is the set of lines.
S	is the set of destination stops.
A_{rl}	is the set of legs for line $r \in R$ using link $l \in L$.
$\mathbf{x}$	is a vector with elements x_{as}, which is the flow of passengers on link or leg $a \in A$ bound for destination $s \in S$.

c is a vector with elements c_a, which is the travel time on link or leg $a \in A$.

w is a vector with elements w_{is}, which is the waiting time at stop $i \in N$ for passengers destined for $s \in S$.

θ is a vector with elements θ_{rl}, which is the carrying capacity of line $r \in R$ on link $l \in L$.

z is a vector of destination-specific flow conservation constraints with elements z_{is}, which is equal to the difference between the in- and out-flow destined for s at node i.

D is the flow conservation matrix for link or leg flows with elements

$$d_{(is)(as')} = \begin{cases} 1 \text{ if } s = s' \text{ and link or leg } a \in A \text{ leads into node } i \in N \\ -1 \text{ if } s = s' \text{ and link or leg } a \in A \text{ leads out of node } i \in N \\ 0 \text{ otherwise} \end{cases}$$

F is a diagonal matrix of link and leg frequencies of service. Hence:

$$\mathbf{F} = \text{Diag}\{f_{as}\}$$

where $f_{as} = f_a, \forall a \in A, \forall s \in S$ is the frequency of service associated with link or leg flow x_{as}. This has an infinite (or very large) value unless a is a boarding link. While the frequency of service is not destination-dependent, the diagonal matrix is expanded to match the dimensions of **D** and **E**.

E is an incidence matrix linking destination-specific node delays to destination-specific link or leg flows with elements

$$e_{(as')(is)} = \begin{cases} 1 \text{ if } s = s' \text{ and link or leg } a \in A \text{ leads out of node } i \in N \\ 0 \text{ otherwise} \end{cases}$$

3. Model formulation

The uncapacitated frequency-based transit assignment model may be formulated as P_0:

(1) P_0: $$\min_{\mathbf{x},\mathbf{w} \geq \mathbf{0}} \mathbf{c}^{\mathrm{T}}\mathbf{x} + \mathbf{1}^{\mathrm{T}}\mathbf{w}$$

where **1** is a vector of appropriate size all of whose elements are 1. Pre-multiplication by this vector sums the elements of the following vector, in this case **w**. Spiess and Florian (1989) showed that the solution to this problem is a frequency-based transit assignment.

The flow conservation constraints are:

(2) $$\mathbf{z} = \mathbf{D}\mathbf{x}$$

These constraints include the origin-destination matrix. The node dwell times are:

(3) $$\mathbf{E}\mathbf{w} \geq \mathbf{F}^{-1}\mathbf{x}$$

Pre-multiplying by the diagonal matrix of service frequencies and introducing a vector of slack variables, we obtain

(4) $\mathbf{FEw} = \mathbf{x} + \boldsymbol{\xi}$ where $\boldsymbol{\xi} \geq \mathbf{0}$

Hence

(5) $\mathbf{DFEw} = \mathbf{z} + \mathbf{D}\boldsymbol{\xi}$

Proposition 1 below proves that $(\mathbf{DFE})^{-1}$ exists. Hence

(6) $\mathbf{w} = (\mathbf{DFE})^{-1}\mathbf{z} + (\mathbf{DFE})^{-1}\mathbf{D}\boldsymbol{\xi}$

Moreover

(7) $\mathbf{FEw} = \mathbf{x} + \boldsymbol{\xi} = \mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{z} + \mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{D}\boldsymbol{\xi}$

Hence

(8) $\mathbf{x} = \mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{z} + (\mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{D} - \mathbf{I})\boldsymbol{\xi}$

where **I** is an appropriately dimensioned identity matrix.

Problem P_0 can now be reformulated as P_1:

(9) P_1: $\min_{\boldsymbol{\xi} \geq \mathbf{0}} \mathbf{h}^{\mathrm{T}} \boldsymbol{\xi}$

subject to

(10) $(\mathbf{DFE})^{-1}\mathbf{z} + (\mathbf{DFE})^{-1}\mathbf{D}\boldsymbol{\xi} \geq \mathbf{0}$

where

(11) $\mathbf{h}^{\mathrm{T}} = \mathbf{c}^{\mathrm{T}}(\mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{D} - \mathbf{I}) + \mathbf{1}^{\mathrm{T}}(\mathbf{DFE})^{-1}\mathbf{D}$

Proposition 1: There are as many linearly independent flow conservation constraints as there are destination-specific node waiting times.

Proof 1: Each entry node for a link has both a flow conservation equation for each destination and a waiting time for each destination.

□

Let:

(12) $\mathbf{k} = \mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{z}$

and

(13) $\mathbf{B} = \mathbf{I} - \mathbf{FE}(\mathbf{DFE})^{-1}\mathbf{D}$

To solve P_1 execute the following two-step algorithm A_0:

Algorithm A_0

1. *Identify hyperpaths*: Note the sign of each element of $\mathbf{h}$. When $h_a > 0, \xi_a = 0$ and $x_a \geq 0$. Conversely, when $h_a < 0, \xi_a > 0$ and $x_a = 0$. Rearrange the order of the links and legs so that

(14) $$\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{k}_1 \\ \mathbf{k}_2\end{bmatrix} - \begin{bmatrix}\mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{0} \\ \boldsymbol{\xi}_2\end{bmatrix}$$

2. *Load hyperpaths*: Solve

(15) $$\boldsymbol{\xi}_2 = (\mathbf{B}_{22})^{-1}\mathbf{k}_2 \text{ and } \mathbf{x}_1 = \mathbf{k}_1 - \mathbf{B}_{12}(\mathbf{B}_{22})^{-1}\mathbf{k}_2$$

□

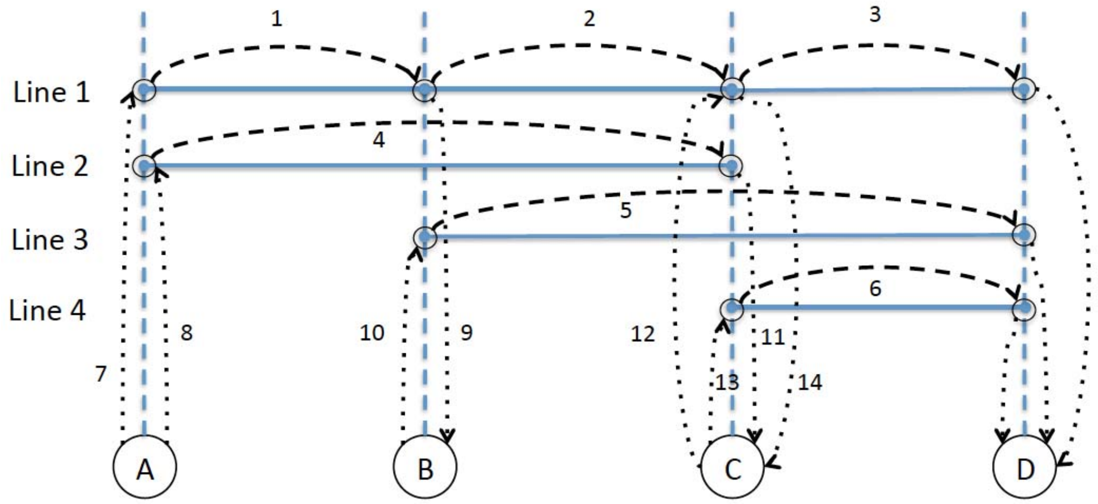

Fig. 1: Four bus line example (dashed arrows are the onboard links while the dotted arrows are the boarding and alighting links; the numbered links are those included in set A)

Example 1: Fig. 1 shows the conventional representation of a transit system without legs. There are four lines and four stops (A, B, C and D). 100 passengers travel from A to D. They can either do so directly on line 1, or start on line 2 then change at C to either line 1 or line 4, or they can start on line 1 then change at B to line 3. We assume that it takes 0.5 mins to board and alight. We do not need to include the final alighting links at the destination D in set *A* as the alighting time is the same in each case. We assume that waiting is associated only with boarding and that the bus dwell time at stops is built into the onboard link travel times.

The assumed link travel times and service frequencies are given in Table 1 followed by the transit assignment results. Boarding links have finite service frequencies while all other links are served continuously. The results are given in Table 1. The *h*-values indicate that only the slack variables for links 9 and 14 will be non-zero at the solution. As a consequence, links 5, 9, 10 and 14 are unused. Flow is divided between the used links according to the service frequencies.

The value of the objective function at the solution is 2283.33 passenger-mins. Note that if the passengers use only line 1 the expected journey time is 25.50 mins (15 mins in the bus, 0.5 mins boarding and 10 mins waiting at stop A) leading to an objective function value of 2550 passenger-mins, so although the other paths are no faster in terms of onboard time, the reduction in waiting time at stop A by including line 2 in the choice set makes paths using line 2 attractive. Line 3 is unattractive because the time involved in changing from line 1 to line 3 is not compensated by any reduction in waiting time.

□

Table 1: Link travel time (mins), frequency (bus/min), h-, xi- and x-vectors

a	c_a	f_a	h_a	ξ_a	x_a
1	5	1e10	0.0000	0.0000	33.3333
2	5	1e10	1.0000	0.0000	33.3333
3	5	1e10	4.0000	0.0000	66.6667
4	10	1e10	0.0000	0.0000	66.6667
5	10	1e10	0.0000	0.0000	0.0000
6	5	1e10	0.0000	0.0000	33.3333
7	0.5	0.1	5.3333	0.0000	33.3333
8	0.5	0.2	2.3333	0.0000	66.6667
9	0.5	1e10	-1.0000	33.3333	0.0000
10	0.5	0.2	5.0000	0.0000	0.0000
11	0.5	1e10	0.0000	0.0000	66.6667
12	0.5	0.1	3.0000	0.0000	33.3333
13	0.5	0.1	7.0000	0.0000	33.3333
14	0.5	1e10	-4.0000	66.6667	0.0000

Table 2: Constraints and waiting times (passenger-mins) on the boarding links

i	*Flows**	z_i	w_i
1	7,8	100	333.3330
2	-7,1	0	0.0000
3	-1,2,9	0	0.0000
4	-2,12,3,14	0	0.0000
5	-8,4	0	0.0000
6	-4,11	0	0.0000
7	-9,10	0	0.0000
8	-10,5	0	0.0000
9	-11,-14,12,13	0	333.3330
10	-13,6	0	0.0000

*Flows included in each constraint; negative refers to an inflow and positive to an outflow. Hence the first constraint is $x_7 + x_8 = 100$, the second is $x_1 - x_7 = 0$, etc.

The link representation of transit networks works well when travel time is the only criterion for assignment and when there is no coordination between services. However, fares are also an important factor and fare schemes are usually not link-based. Where flat fare schemes are found, for example with London buses, a leg can be used to connect the boarding stop with the alighting stop and the disutility of the flat fare added to the disutility of the travel time for the leg. Furthermore, certain interchanges may be coordinated. Where two lines are coordinated, a leg could connect the boarding stop on one line with the alighting stop on another.

Fig. 2 represents the same network as Fig. 1 but uses legs in place of onboard links. Note that this requires an increase in the number of variables to be estimated from 14 to 15 and the number of constraints from 10 to 11. However, flat fare schemes can in principle be accommodated in this representation. If additionally lines 1 and 4 were coordinated, a leg could be added to connect A on line 1 with D on line 4.

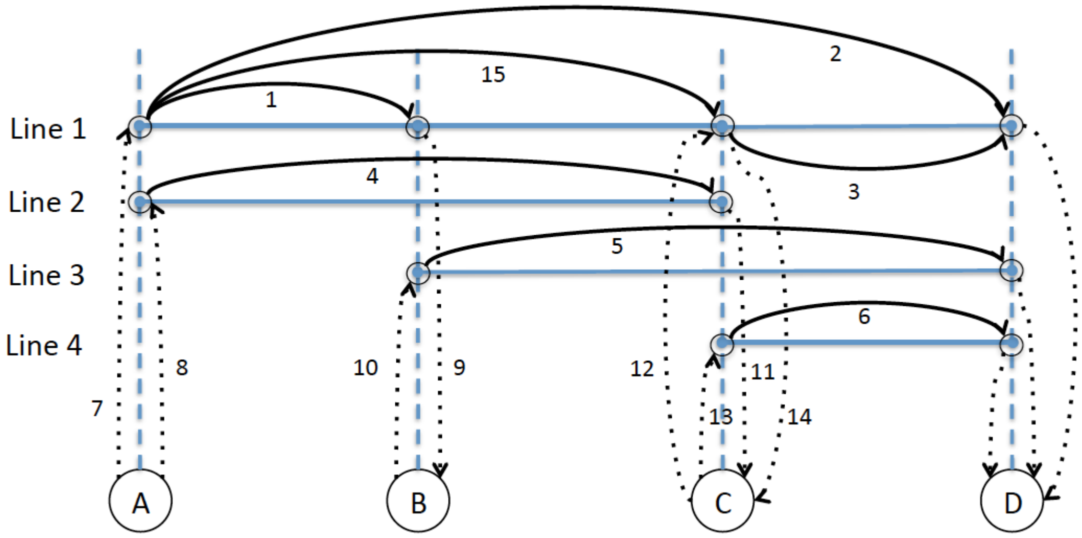

Fig. 2: Same four bus route example (continuous arrows represent legs, dotted arrows represent boarding and alighting links)

Example 2: Consider Example 1 but replace onboard links by legs. The link and leg parameters are given in Table 3. Note that the predicted passenger flows in Table 3 are exactly the same as those in Table 1, demonstrating the equivalence of the two network representations in this case.

Table 3: Link travel time (mins), frequency (bus/min), h-, xi- and x-vectors

a	c_a	f_a	h_a	ξ_a	x_a
1	5	1e10	-1.3333	33.3333	0.0000
2	15	1e10	4.6667	0.0000	33.3333
3	5	1e10	0.0000	0.0000	33.3333
4	10	1e10	0.0000	0.0000	66.6667
5	10	1e10	0.0000	0.0000	0.0000
6	5	1e10	0.0000	0.0000	33.3333
7	0.5	0.1	4.2222	0.0000	33.3333
8	0.5	0.2	2.8889	0.0000	66.6667
9	0.5	1e10	0.0000	0.0000	0.0000
10	0.5	0.2	5.0000	0.0000	0.0000
11	0.5	1e10	0.0000	0.0000	66.6667
12	0.5	0.1	5.0000	0.0000	33.3333
13	0.5	0.1	5.0000	0.0000	33.3333

14	0.5	1e10	0.0000	0.0000	0.0000
15	10	1e10	-3.3333	33.3333	0.0000

Table 4: Constraints and waiting times (passenger-mins) on the boarding links

i	*Flows**	z_i	w_i
1	7,8	100	333.3330
2	-7,1,2,15	0	0.0000
3	-1,9	0	0.0000
4	-12,3	0	0.0000
5	-8,4	0	0.0000
6	-4,9,11	0	0.0000
7	-9,10	0	0.0000
8	-10,5	0	0.0000
9	-11,-14,12,13	0	333.3330
10	-13,6	0	0.0000
11	-15,14	0	0.0000

*Flows included in each constraint; negative refers to an inflow and positive to an outflow. Hence the first constraint is $x_7 + x_8 = 100$, the second is $x_1 + x_2 + x_{15} - x_7 = 0$, etc.

Example 3: In the previous two examples, there was only one origin and one destination. As the model formulation makes clear, the flows in **x** are destination-specific. In Fig. 3 we present a network with 6 feasible origin-destination combinations, $\mathrm{A} \rightarrow \mathrm{D}, \mathrm{B} \rightarrow \mathrm{D}, \mathrm{C} \rightarrow \mathrm{D}, \mathrm{A} \rightarrow \mathrm{C}, \mathrm{B} \rightarrow \mathrm{C}, \mathrm{A} \rightarrow \mathrm{B}$. In Fig. 3 the flows for destination D are labeled. Figs. 4 and 5 show the flow labels for destinations C and B respectively. Matrix **D** is shown in Fig. 6. There are 26 destination-specific flow variables and 19 constraints. When passenger flows per minute are 50, 50, 50, 20, 20, 10 for $\mathrm{A} \rightarrow \mathrm{D}, \mathrm{B} \rightarrow \mathrm{D}, \mathrm{C} \rightarrow \mathrm{D}, \mathrm{A} \rightarrow \mathrm{C}, \mathrm{B} \rightarrow \mathrm{C}, \mathrm{A} \rightarrow \mathrm{B}$ respectively, the network loadings shown in Table 5 and the waiting times shown in Table 6 are obtained.

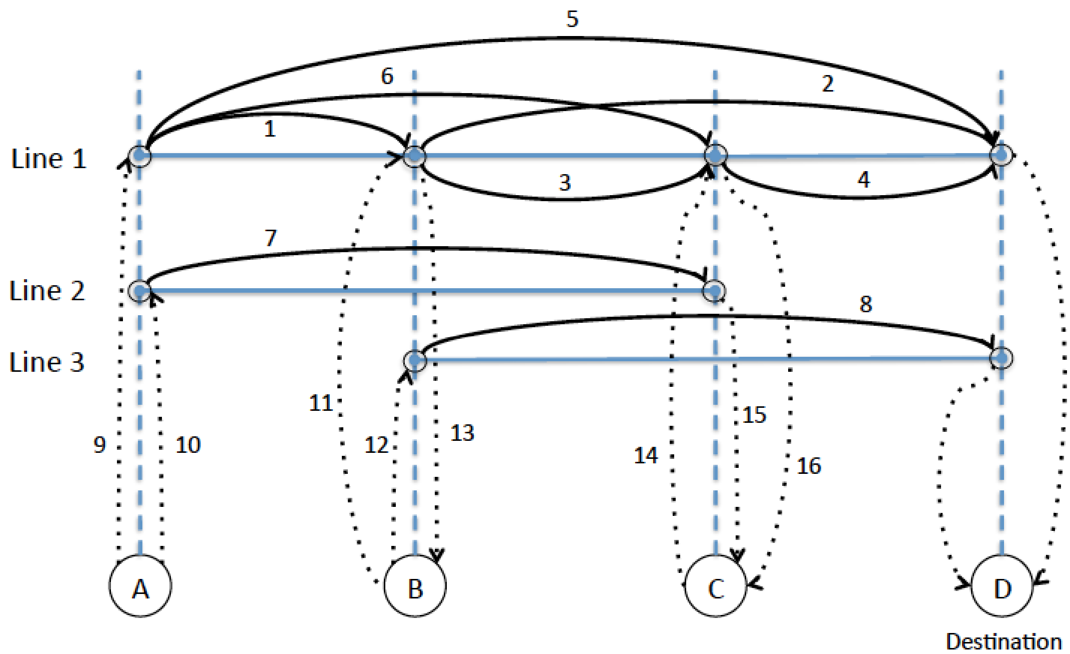

Fig. 3: Three-line network with 6 origin-destination pairs (flow labels refer to destination D)

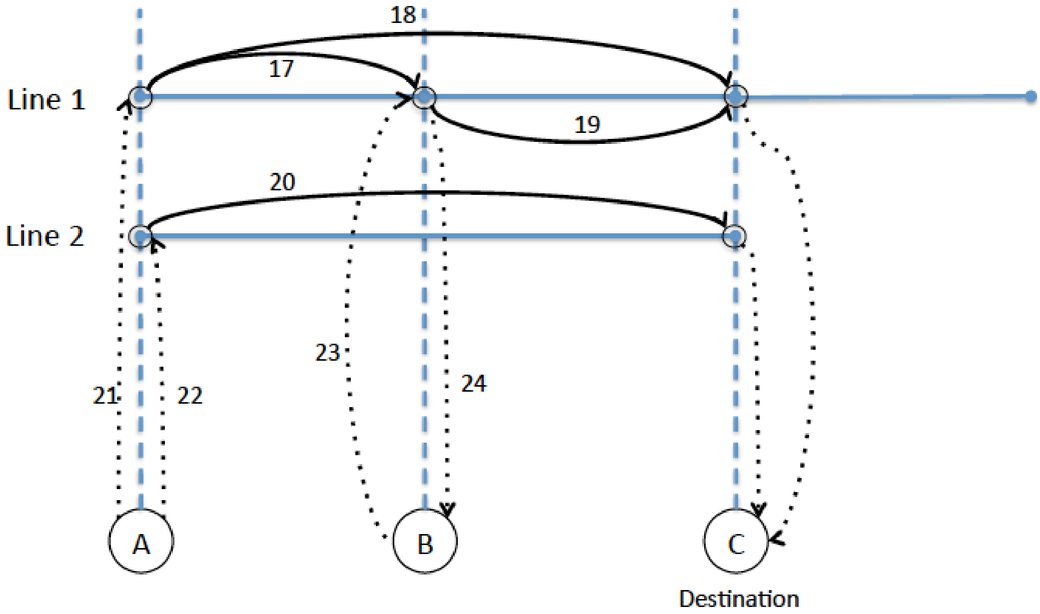

Fig. 4: Same network as Fig. 3 but showing only the leg and link labels for flows to destination C (Line 3 is removed as it could not be used to reach C)

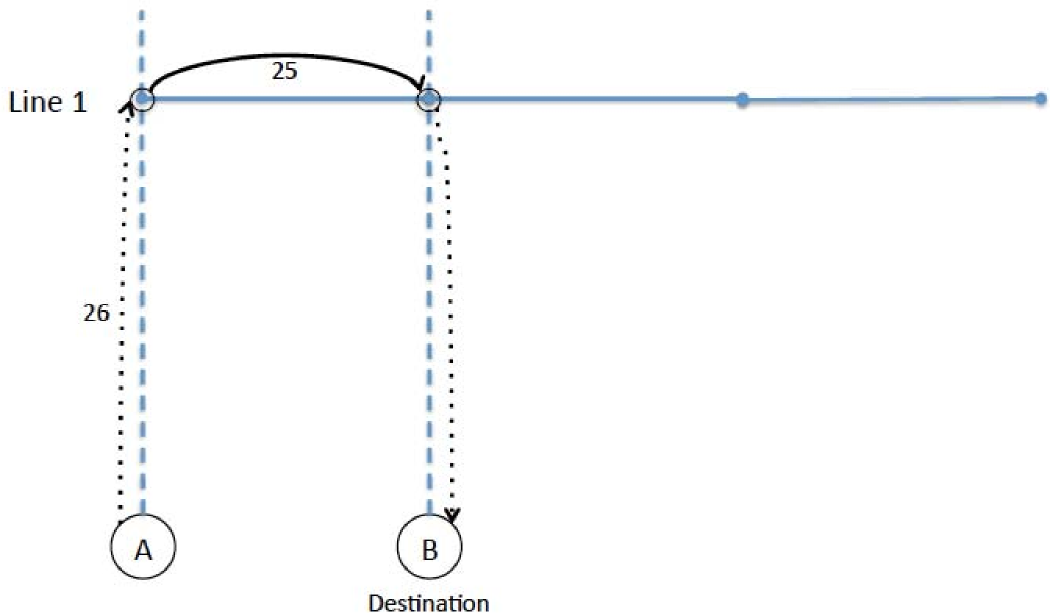

Fig. 5: Same network as Fig. 3 but showing only the leg and link labels for flows to destination B (Lines 2 and 3 are removed as they could not be used to reach B)

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]	[,21]	[,22]	[,23]	[,24]	[,25]	[,26]
[1,]	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[2,]	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[3,]	1	0	0	0	1	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[4,]	-1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
[5,]	0	1	1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[6,]	0	0	0	0	0	0	0	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[7,]	0	0	0	0	0	0	0	0	0	0	1	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
[8,]	0	0	-1	0	0	-1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
[9,]	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
[10,]	0	0	0	1	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
[11,]	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0
[12,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
[13,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	-1	0	0	0	0	0
[14,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0	0
[15,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	1	0	0
[16,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0
[17,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	-1	0	0	0
[18,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
[19,]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1

Fig. 6: Matrix D for Example 3

Table 5: Link travel time (mins), frequency (bus/min), h-, xi- and x-vectors

a	c_a	f_a	h_a	ξ_a	x_a
1	5	1e10	-0.4444	16.6667	0.0000
2	10	1e10	5.5000	0.0000	0.0000
3	5	1e10	-5.5000	0.0000	0.0000
4	5	1e10	0.0000	0.0000	83.3333
5	15	1e10	5.7222	0.0000	16.6667
6	10	1e10	-5.2778	16.6667	0.0000
7	10	1e10	0.0000	0.0000	33.3333
8	10	1e10	0.0000	0.0000	50.0000
9	0.5	0.1	6.8519	0.0000	16.6667
10	0.5	0.2	1.5741	0.0000	33.3333
11	0.5	0.1	-0.3333	25.0000	0.0000
12	0.5	0.2	5.1667	0.0000	50.0000
13	0.5	1e10	0.0000	0.0000	0.0000
14	0.5	0.1	10.0000	0.0000	83.3333
15	0.5	1e10	0.0000	0.0000	33.3333
16	0.5	1e10	0.0000	0.0000	0.0000
17	5	1e10	-5.5000	0.0000	0.0000
18	10	1e10	5.5000	0.0000	0.0000
19	5	1e10	0.0000	0.0000	20.0000
20	10	1e10	0.0000	0.0000	20.0000
21	0.5	0.1	-0.3333	10.0000	0.0000
22	0.5	0.2	5.1667	0.0000	20.0000
23	0.5	0.1	10.0000	0.0000	20.0000
24	0.5	1e10	0.0000	0.0000	0.0000
25	5	1e10	0.0000	0.0000	10.0000
26	0.5	0.1	10.0000	0.0000	10.0000

The constraints and the associated waiting times are given in Table 6.

Table 6: Constraints and waiting times (passenger-mins) on the boarding links

i	*Flows**	z_i	w_i
1	9,10	50	166.6670
2	-10,7	0	0.0000

3	-9,1,5,6	0	0.0000
4	-1,13	0	0.0000
5	-11,2,3	0	0.0000
6	-12,8	0	0.0000
7	-13,11,12	50	250.0000
8	-3,-6,16	0	0.0000
9	-7,15	0	0.0000
10	-14,4	0	0.0000
11	-15,-16,14,	50	833.3330
12	21,22	20	100.0000
13	-21,17,18	0	0.0000
14	-22,20	0	0.0000
15	-17,24	0	0.0000
16	-24,23	20	200.0000
17	-23,19	0	0.0000
18	26	10	100.0000
19	-26,25	0	0.0000

*Flows included in each constraint; negative refers to an inflow and positive to an outflow. Hence the first constraint is $x_9 + x_{10} = 50$, the second is $x_7 - x_{10} = 0$, etc.

Table 5 shows that for destination D, only legs 4, 5, 7 and 8 are used. For destination C, only legs 19 and 20 are used. Finally, for destination B leg 25 must be used. Table 6 shows that at A there are three destination-specific waiting times, at B there are two and at C there is just one.

4. Congested transit networks

As mentioned in the literature review, attention has been given to various aspects of congestion in transit networks. Spiess and Florian (1989) introduced the concept of *effective frequency* in recognition of the fact that, as transit vehicles become more crowded, the probability of not being able to board the next arriving vehicle increases, so the effective frequency decreases and the waiting time becomes gamma rather than exponentially distributed (see Trozzi et al, 2013). Define:

(16) $$\mathbf{G} = \text{Diag}\{g_{bs}\}$$

where

(17) $$g_{bs} = \begin{cases} \infty, b \notin BL \subset A \\ \frac{f_b}{\mu_b}, b \in BL \subset A \end{cases}, \forall s \in S$$

and μ_b is the bus (first, second, etc.) on which the passenger can expect to board on boarding link b, and $BL \subset A$ is the set of boarding links. When $\mu_b = 1$ the passenger can expect to board the first bus to arrive, etc. Note that $\mu_b \geq 1$ does not have to be an integer. The waiting time is given by the gamma distribution[1]. We can now write the problem as P$_2$:

(18) P$_2$: $\min_{\mathbf{x},\mathbf{w}} \mathbf{c}^{\mathrm{T}}\mathbf{x} + \mathbf{1}^{\mathrm{T}}\mathbf{w}$ subject to

(19) $\mathbf{z} = \mathbf{D}\mathbf{x}$

(20) $\mathbf{E}\mathbf{w} \geq \mathbf{G}^{-1}\mathbf{x}$

(21) $\mathbf{x} \geq \mathbf{0}, \mathbf{w} \geq \mathbf{0}$

P$_2$ may be solved in the same was as before by replacing the diagonal frequency matrix **F** by the diagonal effective frequency matrix **G**. If $\boldsymbol{\mu}$ is given, then there are no further complications. If $\boldsymbol{\mu}$ is not given, then the following iterative scheme can be applied: Set $\mu_b = 1$ when the flow on the line is at or below the line capacity, otherwise increase μ_b above 1 where the flow on the line would otherwise exceed line capacity. The complication here is that the capacities apply to lines rather than legs. Lines comprise a sequence of links, whereby a link connects two consecutive stops on a line. The loading on each link should therefore not exceed the line capacity, which in turn is equal to the product of vehicle carrying capacity (in terms of number of passengers) and line frequency (in terms of services per unit time).

When $\boldsymbol{\mu}$ is given, we can reformulate P$_2$ as P$_3$:

(22) P$_3$: $\min_{\boldsymbol{\xi} \geq \mathbf{0}} \mathbf{h}^{\mathrm{T}}\boldsymbol{\xi}$

subject to

(23) $(\mathbf{DGE})^{-1}\mathbf{z} + (\mathbf{DGE})^{-1}\mathbf{D}\boldsymbol{\xi} \geq \mathbf{0}$

where

(24) $\mathbf{B} = \mathbf{I} - \mathbf{GE}(\mathbf{DGE})^{-1}\mathbf{D}$

(25) $\mathbf{h}^{\mathrm{T}} = \mathbf{c}^{\mathrm{T}}(\mathbf{GE}(\mathbf{DGE})^{-1}\mathbf{D} - \mathbf{I}) + \mathbf{1}^{\mathrm{T}}(\mathbf{DGE})^{-1}\mathbf{D}$

(26) $\mathbf{k} = \mathbf{GE}(\mathbf{DGE})^{-1}\mathbf{z}$

Algorithm A$_0$ may then be applied as before, replacing **F** by **G**.

For each line $r \in R$ and link $l \in L$, we can partition the legs into two sets; those in A_{rl} and those not. The sum of the flows on all $a \in A_{rl}$ should not exceed θ_{rl}, the line capacity. The flow on those legs which start earlier than the link take priority over the flow that boards at that link, so where capacity is exceeded we need to reduce the flow on the relevant boarding link b until the line capacity is met. Thus the line capacity constraint causes a metering effect, limiting the ability to board the next vehicle to arrive. Metering reduces the effective frequency.

[1] The gamma distribution has the form $f(w;\mu,\lambda) = \frac{\lambda}{\Gamma(\mu)}(\lambda w)^{\mu-1}e^{-\lambda w}$ which reduces to the exponential distribution when $\mu = 1$. If buses arrived according to a Poisson process, then the waiting time for the μ^{th} bus is described by the gamma distribution. The expected value of w is μ/λ and the variance is μ/λ^2.

Proposition 2: There exists a feasible set of effective frequencies $\boldsymbol{\mu}$.

Proof 2: Suppose not. Let $b(rl)$ be the boarding link for line $r \in R$ and link $l \in L$. By induction, if there is no effective frequency for boarding link $b(rl)$ compatible with θ_{rl}, then the flow on the upstream link l' must have exceeded the line capacity constraint. If the flow on the upstream link exceeds $\theta_{rl'}$, then the flow upstream of the upstream link must also have exceeded line capacity, if this link exists, etc. For a given $r \in R$, suppose the flow on link $l \in L$ exceeds its line capacity constraint. Then the flow on link $l' \in L$, where l' is upstream of l, must exceed the line capacity constraint, if such a link exists. However, there is always an effective frequency for first boarding link of the line that leads to a flow on the first link, which is compatible with the line capacity constraint. This implies that there must always be a feasible set of effective frequencies $\boldsymbol{\mu}$.

□

Since the effective frequencies determine the link or leg flows, which determine the effective frequencies, etc., the following iterative algorithm A_1 is proposed:

Algorithm A_1

1. *Initialisation*: Set $\mu_{b(rl)} \leftarrow 1, \forall r \in R, \forall l \in L$
2. *Set effective frequencies*: Solve (30) for **G**
3. *Run* A_0 *after replacing* **F** *by* **G**: Calculate **w** and **x**
4. *Find new effective frequencies*: $\mu_{b(rl)} \leftarrow \max\left\{1, \frac{\mu_{b(rl)} \sum_{a \in A_{rl}} x_a}{\theta_{rl}}\right\}, \forall r \in R, \forall l \in L$
5. *Termination*: Return to Step 2 until satisfactory convergence is achieved

□

To start, the given frequencies are used and a solution is sought (Steps 1 through 3). Where the demand for a link is not equal to the capacity, a new μ is set equal to the maximum of 1 and the product of old μ and the ratio of the demand over capacity (Step 4). If satisfactory convergence has not yet been achieved, the algorithm returns to Step 2 and repeats the calculations with the latest effective frequencies. Proposition 2 proves that there is at least one vector $\boldsymbol{\mu}$ which satisfies the line capacity constraints. Algorithm A_1 increases μ when demand outstrips supply and reduces μ when supply outstrips demand, subject to $\mu \geq 1$. This algorithm may be applied whether or not legs are used.

Example 4. Consider the example presented in Figs. 1 and 2 when the capacity on line 1 is reduced to 50 passengers per minute. After 100 iterations of A_1, the results presented in Tables 7 and 8 are produced for Fig. 1. The flow on link 3 is reduced to 50 passengers per minute, with the rest arriving by link 6. One third of the passengers set off on link 1 with the rest choosing link 4. Line 3 is still unused. The objective function value is now 2450 passenger-mins. Passengers are expected to be able to board the first arriving bus except in the case of passengers attempting to board line 1 at stop C, who are expecting to board the third arriving bus ($\mu_{12} = 3$). Convergence of algorithm A_1 is rapid.

Table 7: Link travel time (mins), frequency (bus/min), h-, xi- and x-vectors

A	c_a	f_a	h_a	ξ_a	x_a
1	5	1e10	0.0000	0.0000	33.3333
2	5	1e10	0.5714	0.0000	33.3333
3	5	1e10	4.8571	0.0000	50.0000

4	10	1e10	0.0000	0.0000	66.6667
5	10	1e10	0.0000	0.0000	0.0000
6	5	1e10	0.0000	0.0000	50.0000
7	0.5	0.1	6.1905	0.0000	33.3333
8	0.5	0.2	1.9048	0.0000	66.6667
9	0.5	1e10	-0.5714	33.3333	0.0000
10	0.5	0.2	5.0000	0.0000	0.0000
11	0.5	1e10	0.0000	0.0000	66.6667
12	0.5	0.1	3.8571	0.0000	16.6667
13	0.5	0.1	8.7143	0.0000	50.0000
14	0.5	1e10	-4.8571	50.0000	0.0000

Table 8: Constraints and waiting times (passenger-mins) on the boarding links

i	*Links**	z_i	w_i
1	7,8	100	333.3330
2	-7,1	0	0.0000
3	-1,2,9	0	0.0000
4	-2,12,3,14	0	0.0000
5	-8,4	0	0.0000
6	-4,11	0	0.0000
7	-9,10	0	0.0000
8	-10,5	0	0.0000
9	-11,-14,12,13	0	500.0000
10	-13,6	0	0.0000

*Links included in each constraint; negative refers to an inflow and positive to an outflow. Hence the first constraint is $x_7 + x_8 = 100$, the second is $x_1 - x_7 = 0$, etc.

When legs replace onboard links in Fig. 1 to obtain Fig. 2 essentially the same results are produced after 100 iterations (see Tables 9 and 10). The objective function is again 2450 passenger-mins and passengers attempting to board line 1 at stop C can expect to board the third bus that arrives. Convergence of algorithm A_1 is as before rapid.

□

Table 9: Link travel time (mins), frequency (bus/min), h-, xi- and x-vectors

A	c_a	f_a	h_a	ξ_a	x_a
1	5	1e10	-0.5000	33.3333	0.0000
2	15	1e10	5.5000	0.0000	33.3333
3	5	1e10	0.0000	0.0000	16.6667
4	10	1e10	0.0000	0.0000	66.6667
5	10	1e10	0.0000	0.0000	0.0000
6	5	1e10	0.0000	0.0000	50.0000
7	0.5	0.1	5.3333	0.0000	33.3333
8	0.5	0.2	2.3333	0.0000	66.6667
9	0.5	1e10	0.0000	0.0000	0.0000
10	0.5	0.2	5.0000	0.0000	0.0000
11	0.5	1e10	0.0000	0.0000	66.6667

12	0.5	0.1	7.5000	0.0000	16.6667
13	0.5	0.1	7.5000	0.0000	50.0000
14	0.5	1e10	0.0000	0.0000	0.0000
15	10	1e10	-5.0000	33.3333	0.0000

Table 10: Constraints and waiting times on the boarding links.

i	*Flows**	z_i	w_i
1	7,8	100	333.3330
2	-7,1,2,15	0	0.0000
3	-1,9	0	0.0000
4	-12,3	0	0.0000
5	-8,4	0	0.0000
6	-4,9,11	0	0.0000
7	-9,10	0	0.0000
8	-10,5	0	0.0000
9	-11,-14,12,13	0	500.0000
10	-13,6	0	0.0000
11	-15,14	0	0.0000

*Flows included in each constraint; negative refers to an inflow and positive to an outflow. Hence the first constraint is $x_7 + x_8 = 100$, the second is $x_1 + x_2 + x_{15} - x_7 = 0$, etc.

Fig. 7 shows convergence of A_1 over 90 iterations for Fig 1. Series 1 shows how μ_7, μ_8, μ_{10} and μ_{13} remain unchanged at 1. Series 2 shows how μ_{12} converges to 3. Convergence for Fig. 2 is very similar.

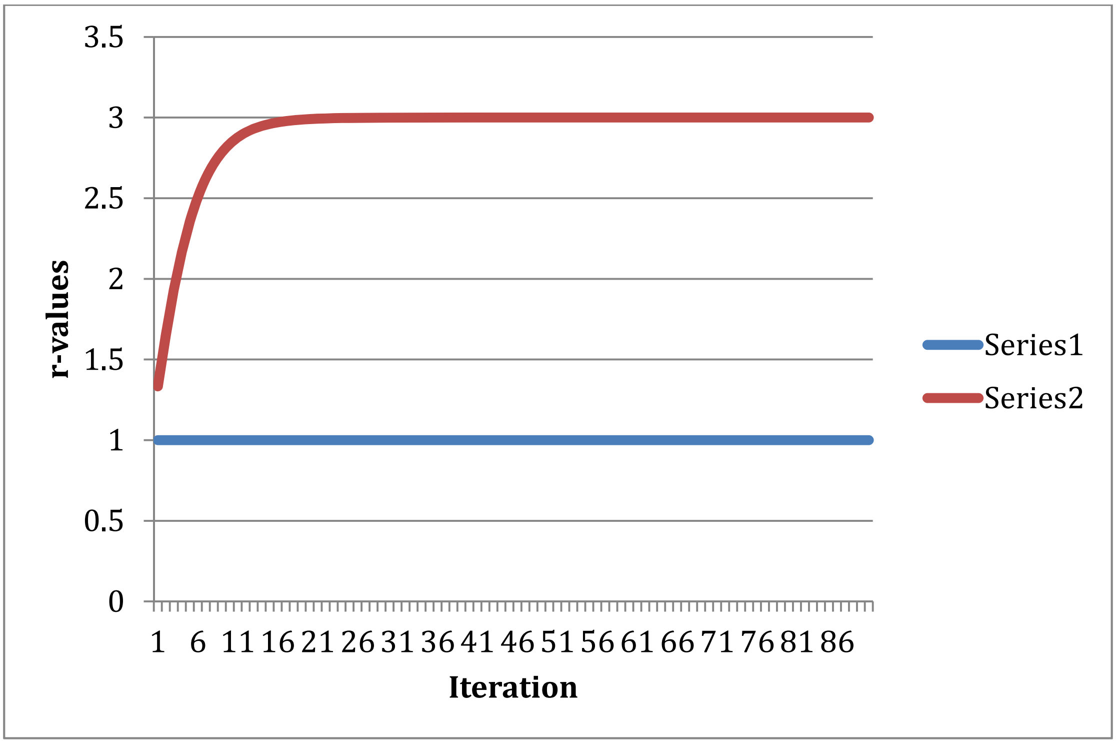

Fig. 7: Convergence of A1 for Fig. 1

Conclusions

This paper reformulates the Spiess and Florian frequency-based transit assignment method in matrix algebra. This reveals some important dimensional properties of the method, namely that the number of destination-specific passenger wait times at stops is equal to the number of flow conservation constraints (Proposition 1). A new way of solving the frequency-based transit assignment model is presented together with an iterative method for finding equilibrium effective frequencies when there are capacity constraints. Both procedures are easily implementable in a scripting language that handles matrices, like R. The existence of a feasible set of effective frequencies is proven (Proposition 2). It is shown that a wider range of fare schemes, for example flat fares, can be modeled by the use of legs.

Numerical examples are presented to illustrate important features of the method. While convergence of the iterative method for finding equilibrium effective frequencies is demonstrated by example, a proof of convergence is still being sought.

References

Bell, M. G. H., 1995 Stochastic user equilibrium assignment in networks with queues. Transportation Research Part B, 29 (2), 115-125.

Bell, M. G. H., Liu, X., Angeloudis, P., Fonzone, A., Hosseinloo, S. H., 2011. A frequency-based maritime container assignment model. Transportation Research Part B 45 (8), 1152-1161.

Bell, M. G. H., Liu, Xin, Rioult, J., Angeloudis, P., 2013. A cost-based maritime container assignment model. Transportation Research Part B, In Press.

Billi, C., Gentile, G., Nguyen, S., Pallottino, S., 2004. Rethinking the wait model at transit stops. Proc., TRISTAN-Workshop, Le Gosier, Guadeloupe.

Chriqui, C., Robillard, P., 1975. Common bus lines. Transportation Science, 9 (2), 115–121.

Cominetti, R., Correa, J., 2001. Common-lines and passenger assignment in congested transit networks. Transportation Science, 35 (3), 250–267.

de Cea, J., Fernandez, J. E., 1993. Transit assignment for congested public transport systems: an equilibrium model. Transportation Science, 27 (2), 133–147.

Dial, R. B., 1967. Transit pathfinder algorithm. Highway Research Board, 205, 67–85.

Gentile, G., Nguyen, S., Pallottino, S., 2005. Route choice on transit networks with online information at stops. Transportation Science, 39 (3), 289–297.

Jansson, K., Ridderstolpe, B., 1992. A method for the route-choice problem in public transport systems. Transportation Science, 26 (3), 246–251.

Lam, W. H. K., Gao, Z. Y., Chan, K. S., Yang, N., 1999. A stochastic user equilibrium assignment model for congested transit networks. Transportation Research Part B, 33 (5), 351–368.

Last, A., Leak, S. E., 1976. Transept: a bus model. Traffic Engineering & Control, 17 (1), 14–17.

le Clercq, F. (1972) A public transport assignment model. Traffic Engineering & Control, 14(1), pp. 91–96.

Liu, Y., Bunker, J., Ferreira, L., 2010. Transit users' route-choice modeling in transit assignment: A review. Transport Reviews, 30 (6), 753-769.

Marguier, P. H. J., Ceder, A., 1984. Passenger waiting strategies for overlapping bus routes. Transportation Science, 18 (3), 207–230.

Nguyen, S., Pallottino, S. , 1988. Equilibrium traffic assignment for large scale transit networks. European Journal of Operational Research, 37 (2), 176–186.

Nielsen, O. A., 2000. A stochastic transit assignment model considering differences in passengers utility functions. Transportation Research Part B, 34 (5), 377–402.

Nielsen, O. A., Frederiksen, R. D., 2006. Optimisation of timetable-based, stochastic transit assignment models based on MSA. Annals of Operations Research, 144, 263–285.

Nökel, K., Wekeck, S., 2009. Boarding and alighting in frequency-based transit assignment. Paper presented at the 88th Annual Transportation Research Board Meeting, Washington, DC, January.

Schmöcker, J.-D., Bell, M. G. H., Kurauchi, F., 2008. A quasi-dynamic capacity constrained frequency-based transit assignment model. Transportation Research Part B, 42 (10), 925–945.

Spiess, H., Florian, M., 1989. Optimal strategies: a new assignment model for transit networks. Transportation Research Part B, 23 (2), 83–102.

Trozzi, V., Gentile, G., Bell, M. G. H., Kaparias, I., 2013. Dynamic user equilibrium in public transport networks with passenger congestion and hyperpaths. Transportation Research Part B, 57, 266-285.

Wu, J. H., Florian, M., Marcotte, P., 1994. Transit equilibrium assignment: a model and solution algorithms. Transportation Science, 28 (3), 193–203.

Yang, L., Lam, W. H. K., 2006. Probit-type reliability-based transit network assignment. Transportation Research Record, 1977, 154–163.