

Transit Network Design And Scheduling: a Global Review

Valérie Guihaire^{a,b,*} Jin-Kao Hao^b

^a*Perinfo*

41 avenue Jean Jaurès, 67100 Strasbourg, France

^b*LERIA, Université d'Angers*

2 Boulevard Lavoisier, 49045 Angers Cedex 01, France

Abstract

This paper presents a global review of the crucial strategic and tactical steps of transit planning: the design and scheduling of the network. These steps influence directly the quality of service through coverage and directness concerns but also the economic profitability of the system since operational costs are highly dependent on the network structure. We first exhibit the context and the goals of strategic and tactical transit planning. We then establish a terminology proposal in order to name sub-problems and thereby structure the review. Then, we propose a classification of 69 approaches dealing with the design, frequencies setting, timetabling of transit lines and their combinations. We provide a descriptive analysis of each work so as to highlight their main characteristics in the frame of a two-fold classification referencing both the problem tackled and the resolution method used. Finally, we expose recent context evolutions and identify some trends for future research. This paper aims to contribute to unification of the field and constitutes a useful complement to the few existing reviews.

Key words: public transportation, network design, network scheduling.

1. Introduction

Public transportation planning covers a very wide research area. From the design of networks to the rostering of crews, from the evaluation of demand to the transit trip assignment, from mathematical methods of resolution to evolutionary ones, the process of generating a public transportation system has been approached from many sides. This thorough interest is partly due to the fact that the development of public transportation is a crucial topic in the modern society. Confronted to traffic congestion, urban parking problems and increasing pollution, car drivers might consider switching to public transit if they had an affordable and good quality system at their disposal. It is the duty and goal of transit agencies to provide such conditions, by adequately adjusting their systems, so as to maximize the quality of service to users while minimizing the costs. Tradeoffs thus need to be made and this is where various optimization techniques come into the game.

* Corresponding author. Tel.: +33 388449618; Fax: +33 388449601

Email addresses: vguihaire@perinfo.com (Valérie Guihaire), hao@info.univ-angers.fr (Jin-Kao Hao).

From the users' perspective, the system should meet the demand by providing cheap and direct service to passengers. Criteria for using public transport can also include vehicle and transfer terminal comfort, regularity, service coverage and frequency level. From the operator's perspective on the opposite, the objective is for the system to make as much profit as possible. It is the main challenge in transit planning to find an equilibrium between these conflicting objectives.

As the literature assesses, the public transit planning process is usually divided in a sequence of five steps: (1) the design of routes, (2) the setting of frequencies, (3) the timetabling, (4) the vehicle scheduling and (5) the crew scheduling and rostering (see Table 1). This review addresses the three first and thus fundamental elements of the public transit planning process, also called strategic (step 1) and tactical (steps 2 and 3) planning respectively. All the information needed by the passengers, namely the transit routes network, the frequencies and departure times, is determined during these phases. One could therefore think that these steps are essentially user-oriented. However, the problem remains multiobjective since financial objectives must also be taken into consideration. Even inside the restricted area of our problem, numerous approaches have been proposed, integrating different constraints, aiming for various objectives, and combining heterogeneous features.

In the domain of transit planning, several interesting reviews are available. Desaulniers and Hickman (2007) focus mainly on mathematical methods for each individual steps of the planning process. Fan and Machemehl (2004) and Zhao and Gan (2003) present reviews of the transit network design problem as an introduction to their applied research. Finally, Chua (1984) surveys the approaches limited to network design and frequencies setting used by British operators for urban bus services in the 80's.

There are essentially three main reasons motivating this review. First, the previous reviews all gather interesting and useful information on a particular part of the whole transit network design and scheduling problem. However, it is difficult to apprehend more global problems and the way subproblem resolution methods can be complementary or not. Second, the evolutions in public transit policies, in particular the development of intermodality, integration and deregulation, raise important new questions on transit planning. Third, the recent years have seen important advances in solution methods, leading to the development of innovative approaches whose efficiency deserves to be emphasized.

This review thus aims at dealing with the transit network design and scheduling problem in the goal of providing the reader with a broader and more complete view of the field. Given that the domain does not possess a standardized convention for terminology, we will first propose a terminology for each of its subproblems. We then establish a classification of the existing approaches according to two important criteria: the target problem and the solution method developed. Moreover, for each reviewed work, a descriptive analysis is given, so as to outline their characteristics, originality, application domain and limitations.

For a complete appreciation and broader understanding of the domain, we open the review to another level of perspective by pointing out recent policy developments impacting transit network problems, and by defining challenges and opportunities for future work.

The structure of this paper is as follows. First of all, we describe in part 2 the search methodology, and then detail the transit planning process in part 3. We explain in part 4 our two classification axes: problem tackled and resolution method used. There, we formulate terminology proposals for each problem and subproblem of strategic and tactical planning and present models and resolution method categories. In section 5, the main features of the approaches reviewed are summarized in a table before actually being exposed and analyzed individually in the frame of a problem categorization in part 6. To conclude the review, recent policy developments, challenges and opportunities for future work are presented in section 7.

2. Methodology

We followed a three-fold search methodology: international review databases search, free web search, and ancestry approach. International review databases were used to find out reference papers. It is however arguable that a part of the applied research does not appear in such sources. As an attempt to overcome this bias in the results, we additionally conducted a free web search. This way, we obtained much broader access to works from all origins.

The search terms we selected are combinations of keywords intended to cover alternate denominations of similar problems, and include: public transport; mass transportation; transit; bus; (routes) network; design; frequencies; timetabling; scheduling; transfer; synchronization; coordination; as well as terms related to optimization methods.

We excluded from the set of results the studies dealing with the following subjects: freight transportation, on-demand service, vehicle and crew scheduling, as well as pickup and delivery problems. Finally, we carried out an ancestry approach to collect references from already found papers. In order to target high quality publications in this review and overcome the language barrier, we have tried to focus on international English literature. Several papers of interest in different language are however also included. Although the free web search provides an interestingly wide coverage, it also has a number of limitations, since most studies carried out before 1995 will not be found on the web. Additionally, firms try to protect their products and commercial rights limit access to research results. This barrier can hardly be overrun whatever the search methodology. However, we have cited several commercial products in this review as well as combined academic and industrial research papers. In the end, we propose a collection of the most meaningful contributions that we could find in the domain but there is no denying that this review could not be exhaustive, considering the numerous works on the topic.

3. The Global Transit Planning Process

General transit planning is an extremely complex and difficult problem; even its sub-problems considered separately are already NP-hard from a computational complexity point of view (Magnanti and Wong, 1984; Quak, 2003). The first approach to the line planning problem was published more than 80 years ago (Patz, 1925) but more advanced and powerful methods have been developed along the years.

The global transit planning process is based on the following input: a public demand, an area with topologic characteristics, a set of buses and a set of drivers. The goal is to obtain a set of lines and associated timetables to which buses and drivers are assigned. According to Ceder and Wilson (1986), the global transit planning process can be decomposed into a sequence of five components as shown in Table 1.

Table 1
Transit Planning Process - Ceder and Wilson (Ceder and Wilson, 1986)

Independent Inputs	Planning Activity	Output
Demand data Supply data Supply data Route performance indicators	Network design	Route changes New routes Operating strategies
Subsidy available Buses available Service policies Current patronage	Frequencies setting	Service frequencies
Demand by time of day Times for first and last trips Running times	Timetable development	Trip departure times Trip arrival times
Deadhead times Recovery times Schedule constraints Cost structure	Bus scheduling	Bus schedules
Driver work rules Run cost structure	Driver scheduling	Driver schedules

Ideally, all those steps should be treated simultaneously so as to ensure interaction and feedback, thus leading to better quality results. However, due to the exceptional complexity of the process, this global approach appears intractable in practice. As a result, various subproblems have been defined over the years so as to solve the planning problem in a sequential manner, although it thereby loses any optimality guarantee. Our concern here relies on the first three steps of this framework.

3.1. Transit Network Design

3.1.1. Purpose

The goal in this component is to define a set of bus routes in a particular area, each route being determined by a sequence of bus-stops.

3.1.2. Input

Topology- The area's topology can be defined by the roads, possible areas for bus stops and transfer zones, and sometimes also the location of depots that serve as extreme terminals.

Origin-destination (OD) matrices- OD matrices are needed to define a transit routes network that satisfies as much as possible the community's demand. An OD matrix has the set of stop points as coordinates. Abscissas correspond to the origins and ordinates to the destinations of the users. The OD matrix contains the number of passengers willing to go from each origin to each destination in a given time period. The more precise the data, the more adequate the solution; therefore let us detail matrix characteristics that enhance results: *Matrix coordinates-* Coordinates of the matrices could correspond to the exact origin and destination wishes of the community. However, this is not realistic since buses cannot stop at every desired point but only at pre-defined bus stop locations. Therefore, OD matrices should have the set of possible transfer zones and bus stop locations as coordinates. *Survey conditions-* To fulfill the expectations of all potential passengers, the survey should be led through the entire community and not only through the current users of public transport.

3.1.3. Main Constraints and Objectives

Depending on the politics of the transit agency, constraints and objectives might intermingle. Since there is no assessed rule to differentiate them, we list them as a unique set of features.

Historical background- The existing network, if any, can play a role, in the sense that for some (political) reasons, it might be undesirable to disrupt service on already existing lines.

Area coverage- It measures the percentage of the estimated demand that can be served by public transit. This ratio can be computed in several ways (Spasovic et al., 1993) but usually depends on characteristics such as route length, density, bus stop and route spacing (Murray, 2003; Benn, 1995). The rules-of-thumb often used consider that people living within 400 to 500 meters from a bus stop are part of this percentage. Some plans (Murray et al., 1998) aim for a 90% ratio.

Route and trip directness- Limits are imposed on the distance that one user can cover in the transit network with consideration to one's trip demand. From the users' point of view, the bus network should indeed enable them to travel as directly as possible from their origin to their destination and to walk the shortest distance to reach the first and final bus stops. Different definitions can be used to evaluate this feature. Directness can depend on the route's deviation from a linear path (Benn, 1995) considering the additional mileage incurred by a bus trip compared to the same trip by car or another means of transportation. The number of transfers is also a recurrent criterion. Note that to compute trip directness for each user, it is necessary to go through a passenger trip assignment process. This consists in assigning routes and transfers to passengers with respect to some objective such as shortest path or smallest number of transfers (Desaulniers and Hickman, 2007).

Demand satisfaction- It is obviously a crucial issue. When users' origin or destination are too distant from bus stops, or when trip directness is insufficient, the demand can be considered unsatisfied. Note that similarly to trip directness evaluation, computing demand satisfaction requires to go through a transit trip assignment process. In a general manner, if a trip requires more than two transfers, it is assumed that the user will switch to another means of transportation.

Number of lines or Total route length- A general objective of the operator is to minimize the total route length in the perspective of reducing the number of vehicle and crew resources needed to sustain the global transit system. The number of lines can alternately be considered. Moreover, routes should neither be too short nor too long for profitability reasons.

Operator-specific objectives- For some reason, transit agencies may want to develop a network with a particular shape. Radial, rectangular, grid and triangular (Van Nes, 2002) are common shapes used.

3.2. *Transit Network Frequencies Setting*

3.2.1. *Purpose*

This step provides frequencies for every line in the network and for each time period. As a consequence, the number of line runs are roughly defined during this phase. A line run corresponds to one scheduled service of the line. The inverse of the frequency over a determined period is called the headway. It corresponds to the time elapsing between consecutive line run departures.

3.2.2. *Input*

Transit routes network- The transit routes network constitutes the main input for the current matter.

Public demand- Detailed OD matrices are needed in this step. They should provide data according to uniform demand time periods. These periods vary according to the following criteria: time of the day (Peak / Off-peak period), day of the week (Monday–Friday / Saturday–Sunday), time of the year (Seasons / Vacation Periods / Others). Since demand is time-dependent and elastic (see ??), the survey should be led on extensive periods of time and regularly updated. Such a process is necessary to achieve an efficient network with satisfying service. However, it represents a real charge for the transit agency, since collecting this data is a very complex and expensive task. Such data is therefore rarely freely available (see 6.8).

Bus fleet- In most approaches, line frequencies also depend on the available fleet size and buses capacities. In this case, a description of the vehicles used is needed, especially if the fleet is heterogeneous. The period-dependent bus running times associated to each route of the network must also be provided.

3.2.3. *Main Constraints and Objectives*

Demand satisfaction- The lines frequencies should match the demand at best so as to avoid overcrowding and excessively large headways, and thereby reduce waiting and transfer times.

Number of line runs- The number of runs for each line is an example of the multiobjective nature of the problem. While from the operator’s point of view, it is desirable to minimize this number for resource-related reasons, users for convenience wish to benefit from the widest offer of line runs.

Headway bounds- The operator can be imposed minimum and/or maximum headways on some lines or areas by regulating authorities.

Historical background- In a similar way as in 3.1, historical line runs can be imposed.

3.3. *Transit Network Timetabling*

3.3.1. *Purpose*

This step yields a timetable that includes departure times from all the stops served by each line run in the network. Each line run’s timetable consists of a departure time from the initial terminal, the expected departure times from each bus-stop on the route, and an expected arrival time at the final depot.

3.3.2. *Input*

Transit routes network- The transit routes network constitutes the main input for the current matter. Running times associated to this network are also necessary to compute the timetables.

Public demand- The lines frequencies determined in the TNFS step define the time coverage of the line. Additionally, the level of importance of each transfer is needed to secure a better quality of service through the minimization of passengers waiting times. This level of importance can be deduced from detailed period-dependent OD matrices.

3.3.3. *Main Constraints and Objectives*

Demand satisfaction- Timetable setting permits to compute passengers travel time for the first time in the process. These values should be minimized to enhance the passengers' mobility. If they are too high for some particular trip demand, this one can be considered unsatisfied.

Transfer coordination- The global network timetable should take into account each transfer zone and associated lines so as to permit smooth transfers between lines in time and space. Transfers among lines can be favored according to various criteria such as number of passengers involved.

Fleet size- The vehicle schedules are created from the line runs of the transit network timetable. Therefore, it can be useful to consider resource usage from the TNT step.

Historical background- In a similar way as in 3.1, historical line timetables can be imposed.

3.4. *Vehicle Scheduling*

The purpose of this component is to obtain a feasible sequence of line runs, also called bus service, thereby determining the number of buses required for the considered period, usually of one day long. For further information on vehicle scheduling, we refer to Bunte et al. (2006).

3.5. *Crew Scheduling and Rostering*

This component aims at assigning drivers to bus services. Each driver ends up with a timetable for a given period. The scheduling phase is usually based on a one-day-period while the rostering phase concerns longer periods so as to include other types of constraints, for instance maximum number of consecutive working days. For further information on crew scheduling and rostering, we refer to Wren and Rousseau (1993).

3.6. *The Focus in this Review*

Transit planning is a multiobjective problem, where the users' and the operator's interests conflict. In many countries, regulating authorities pay transport operators for their services. While the former require a certain level of quality of service, the latter wish to minimize their expenses. The usual practice for the regulating authorities consists in imposing lines with given route configuration and frequencies so as to secure a certain level of coverage in time and space, while the operator then adjusts the line runs departure times so as to match its resources employment in an economic manner. Consequently, transfer synchronization receives little concern. This is pretty damageable to the users: transfer synchronization is a crucial element of service quality, and it is entirely dependent on departure times setting. Therefore, timetables could be a criterion imposed to the operator the way routes and frequencies already are.

Considering this matter of fact, this review will concentrate on the first three steps of the planning process, namely the transit network design, frequencies setting and timetabling, also known as strategic and tactical planning. The two last steps (i.e. vehicle scheduling, crew scheduling and rostering), in contrast, are related to operational planning and will be left aside. The detail of the internal structure of our review area, the transit network design and scheduling, will be presented in the next section with the associated terminology.

4. **TNDSP - An Overview**

The transit network design and scheduling problem corresponds to the first steps of the transit planning process. It consists in creating (from scratch or from an existing network) a set of line routes and timetables. The need to find a balance between the users' and the operator's satisfaction makes this problem multi-objective, adding to its complexity. The purpose of this review being to classify and analyze the different studies led on the topic, we chose to categorize them according first to the problem tackled and then to the resolution method used. Before actually reviewing the articles, we will therefore briefly present the different

problems and classes of optimization methods in a general way. But first of all, let us present some models that have been used to formalize transit network problems.

4.1. *Models used to formalize Transit Network Problems*

In terms of mathematical modelling, a variety of proposals have been made depending on the problem addressed. Each one being characterized by its specific objectives and constraints, we do not give here a general MP program and refer the reader to each of these papers for additional information. We only give here an overview of the types of models that have been used in the literature. The transit network timetabling problem can be represented through a quadratic semi-assignment problem (QSAP) (Klemm and Stemme, 1988; Bookbinder and Désilets, 1992; Daduna and Voss, 1995) and through a mixed integer non linear program (MINLP) combining binary variables with real ones corresponding to departure and arrival times (Chakroborty et al., 1997, 2001, 1995; Deb and Chakroborty, 1998; Jansen et al., 2002). A MINLP is also used by Constantin and Florian (1995) for the TNFSP, and by Fan and Machemehl (2004) for the TNDFSP. In (Ceder et al., 2001; Eranki, 2004), the transit network scheduling problem is modeled as a mixed integer linear program. Borndörfer et al. (2005) propose a multicommodity flow model based on a passenger route graph for the TNDFSP. A multicommodity flow model is also applied to the TNDFSP by Yan and Chen (2002) with multiple time-space networks formulating bus movements and passenger flows. Generalized models for the transit network design problem such as node-arc, path and tree formulation are summarized by Kim and Barnhart (1999).

To be more comprehensive, let us add a few words on the treatment of demand responsiveness in mathematical models. In any economic system, changes in the offer impact the demand. In our case, modifications applied to the components of a transit network impact the users and the way they travel. The reaction is a dynamic process whose magnitude fluctuates along a timeline, depending among others on the intensity of the change as perceived by each user. Studies suggest clear differences between short-run (1-2 years), medium-run (5-7 years) and long-run (12-15 years) elasticities, implying that effects on the demand cannot be deduced from a simple before-after comparison. This asserted continuity conflicts with the common practice to assume that every user reacts immediately, objectively and identically to change. Indeed, behavioral adaptation is a more complex phenomenon, as "individual responses are influenced by habit, by the desire to experiment (or not to), by ignorance, preferences, and by domestic and economic constraints which are very powerful at any point in time, but change and evolve" (Goodwin, 1999). A shortcoming of many mathematical models is that intermediate buffer states are neglected, and the results are evaluated in an immediate rather than cumulative manner. This problem is a general modeling one and is not restricted to the domain of transportation, which explains why we will not further develop on the subject in this review, however for more extensive information, we refer the reader to Balcombe et al. (2004).

4.2. *Categorization of Optimization Methods*

In terms of optimization strategies, a variety of methods have been used in the reviewed papers. Exact search methods are often tested when the targeted problem can be formulated with a known mathematical model. These methods typically include linear programming and some forms of integer and mixed integer programming. Notice that such a practice allows little flexibility and is often restricted to situations based on a number of simplifying assumptions.

For this reason, a large number of reviewed papers, in particular the ones published in recent years, use solution methods based on heuristics and metaheuristics (Reeves, 1993). These solution methods can be classified in four big families.

- specific and ad-hoc heuristics which often follow the greedy construction principle,
- neighborhood search such as Simulated annealing and Tabu search,
- evolutionary search such as Genetic algorithms,
- hybrid search which combines two or more solution methods.

One notable advantage of the heuristics and metaheuristics is the fact that they are not designed for a particular mathematical model. Indeed, these solution methods define very general search frameworks and can adapt to almost any form of constraint and objective.

Finally, one notices that simulation is also used in hybrid search strategy.

In the rest of this review, we classify the reviewed papers according to the following categories: mathematical approaches, heuristic approaches, neighborhood search approaches, evolutionary algorithms and other approaches. When both a mathematical model and a solution method are proposed within the same paper, we categorize it following the solution method employed.

4.3. Terminology Proposals for Transit Network Problems

In the literature, various terms can be employed to describe the different steps of the transit planning process and their combinations. For instance, a problem in which transit routes and frequencies are set can be named "Bus Transit Routes Network Design" (Fan and Machemehl, 2004), "Transit Route Network Design" (Baaaj and Mahmassani, 1995) or "Line Planning in Public Transport" (Borndörfer et al., 2005). Additionally, many articles do not explicitly name the problem they are addressing in the context of the global transit planning. Therefore, we propose the terminology shown in Fig.1 to organize denominations and relations among problems and sub-problems related to strategic and tactical transit planning in the remainder of this document.

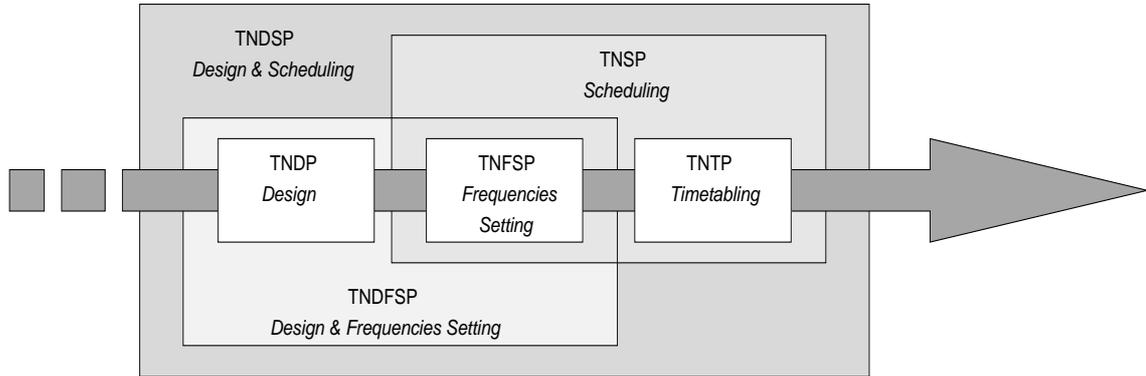


Fig. 1. Transit Network Problems (TNP) Structure

We begin with three basic transit network problems: design (TNDP), frequency setting (TNFSP) and timetabling (TNTP). Then, we introduce two combined problems: design and frequencies setting (TNDFSP=TNDP+TNF) and scheduling (TNSP=TNFSP+TNTP). Finally, the whole design and scheduling problem (TNDSP) is defined as the composition of the three basic problems.

5. TNDSP and Sub-Problems - A Classification

Many different situations and models have been considered inside the domain of transit network design and scheduling. To provide an overview of the approaches reported in the literature and a summary of their main features, we have classified these approaches in Table 5 according to six elements. The main differentiation factor clearly is the decision variables (or problem tackled). The optimization method is the second criterion that we use in the following review to classify the approaches. In Table 5, "Me." stands for "Method", "M" for "Mathematical", "H" for "Heuristic", "NS" for Neighborhood Search, "E" for "Evolutionary" and "O" for "Other". Then, objectives and constraints are listed since they differ according to the goal of the study and choice of the authors. The level of theory or application and a specificity concerning either the resolution method or the context of the study are also reported.

First Author	Year	Problem	Me.	Objectives	Constraints	Application	Specificity
Patz	1925	TNDP	H	Number of empty seats	Bus capacity, Demand	Small example	Optimal on specific example
Lampkin et al.	1967	TNDFSP	H	Number of direct passengers, Total travel time	Fleet size	Ex: G.-B.	Skeleton method
Salzborn	1972/80	TNFSF	M	Fleet size, Passengers waiting time	Passengers arrival rate	-	Feeder service
Silman et al.	1974	TNDFSP	M	Fleet size, Journey time, Overcrowding	Budget	Ex: Haifa	Skeleton method
Rapp et al.	1976	TNSP	H	Transfer optimization	-	Ex: Basel	Interactive graphics system
Sonntag	1977	TNDP	H	Average travel time, Number of transfers	Restricted set of possible lines	Ex: Dusseldorf	Railway
Mandl	1979	TNDP	H	Routes directness, Travel time	Constant frequency, Service coverage	Example	Urban
Bel et al.	1979	TNDFSP	H	Passengers total travel time	Budget	2 ex (France)	Replanning problem
Hasselstrom	1979/81	TNDFSP	M	Number of transfers, Number of passengers	Budget	Software	Variable demand
Scheele	1980	TNFSF	M	Passengers travel time	Capacity, Fleet size	-	-
Furth et al.	1982	TNFSF	M	Number of passengers, Waiting time	Fleet size, Maximum headway, Budget	Example	-
Han et al.	1982	TNFSF	H	Minimizing the maximum "occupancy level" at the maximum load point for each route	Fleet size, Capacity for each route	-	Trip assignment to overlapping routes
Koutsopoulos et al.	1985	TNSP	M	Waiting time, Operator costs, Vehicle crowding	Fleet size, Capacity per route	Simple ex.	-
Ceder et al.	1986	TNDFSP	H	Excess travel time, Transfer and Waiting time, Vehicle costs	Minimum frequency, Fleet size, Route length	Theoretical	Public transit process model
Klemm et al.	1987	TNTP	H	Transfer synchronization	-	Example	QSAP modelling
Van Nes et al.	1988	TNDFSP	M	Fulfill the demand, Number of direct trips	Fleet size	Example	-
Pape et al.	1992	TNDP	H	Number of lines, Number of direct passengers	Service coverage	Ex: Dusseldorf	Concept of corelines
Bookbinder et al.	1992	TNTP	O	Waiting time	Routes network, Fixed headways	Theoretical	Random travel times
Xiong et al.	1993	TNDP	E	Total travel time, Cost	-	Benchmark	Network improvement
Shih et al.	1994/98	TNDFSP	O	Travel time, Satisfied demand, Fleet size	-	Example	Trip assignment model for timed transfer terminals
Constantin et al.	1995	TNFSF	M	Users total expected travel and waiting time	Fleet size	Examples (3)	Urban
Daduna et al.	1995(a)	TNTP	NS	Transfer time	Lines, Frequencies	Examples	-
Baaj et al.	1995	TNDFSP	O	Number of direct trips, Waiting time, Transfer time	Headway, Fleet size, Capacity	Ex: Austin	Skeleton method
Chakraborty et al.	1995/97/98	TNTP	E	Passengers total waiting time	Fleet size, Maximum headway, Stopping time bounds, Maximum transfer time	Simple example	Several cases are studied
Bussiek	1998	TNDFSP	M	Number of direct passengers, Operator costs	Level of service, quality, Number of resources	Examples	Railway, Periodic schedules
Carrese et al.	1998	TNDFSP	H	Users waiting time and excess time compared to minimum path, Operator costs	Demand satisfaction, Routes length, Number of transfers, Total travel time, Fleet size	Ex: Rome	Urban
Patnaik et al.	1998	TNDFSP	E	Operator costs, Passengers travel time	Headway, Load factor	Example	Urban
Dhingra et al.	1999	TNDSP	E	In-vehicle travel time, Number of buses and standees, Demand satisfaction, Walking time	No transfers, Headway bounds, Maximum load factor, Fleet size, Routes length limits	Simple	Feeder service
Lee et al.	2000	TNDFSP	H	Users travel time	Network, Fixed total demand	-	Variable trip demand
Chowdhury et al.	2001	TNFSF	H	Transfer coordination	Fleet size, Capacity	Example	Intermodality
Chakraborty et al.	2001	TNTP	E	Transfer coordination (waiting time), Fleet size	Total fleet size, Stopping time, Headways, Transfer time	Theoretical	Fleet size considerations
De Palma et al.	2001	TNTP	M	Riders' total schedule delay costs	Desired boarding times	Theoretical	Single transit line
Ceder et al.	2001	TNSP	H	Number of simultaneous arrivals	Headway bounds	Examples	Simultaneous arrivals
Chakraborty et al.	2002	TNDP	E	In-vehicle time, Unsatisfied demand, Number of direct, 1-transfer, 2-transfer passengers	Road network	Benchmark	-

Bielli et al.	2002	TNDFSP	E	Average travel time, Fleet size, Network performance	Predefined possible lines only	Ex: Parma	Multicriteria analysis, Fitness function evaluation
Fusco et al.	2002	TNDFSP	O	Overall system cost	Level of service, Satisfied demand, Lines configuration, Frequency	Theoretical	Medium size towns
Yan et al.	2002	TNDSP	M	Profit	Rights of way, Fleet size, Service requirements	Ex: Taiwan	Intercity
Murray	2003	TNDP	M	Number of bus stops	Service coverage, Access	Ex: Brisbane	Bus stops location
Zhao et al.	2003/04/06a	TNDP	NS	Number of transfers, Routes directness, Service coverage	Historical lines, Number of routes, Routes directness	Software: OptNet	Urban
Guan et al.	2003	TNDP	M	Total length of transit lines, Total number of transit lines taken, Total length traveled by passengers	Transit capacity, Number of transfers for each OD pair, Line length bounds	Ex: HongKong	Railway, line assignment
Gao et al.	2003	TNFSF	H	Deterrence, Cost, System performance	Equilibrium	Example	Bilevel programming
Ceder	2003(b)	TNDFSP	H	Operator and Users costs, Fleet size	Route length, Deviation from shortest path	Example	Short turning
Ngamchai et al.	2003	TNDFSP	E	Fleet size, In-vehicle and Waiting time	Service coverage	Benchmark	Problem-specific operators
Tom et al.	2003	TNDFSP	E	Operator costs, Passengers total travel time	-	Example	-
Wan et al.	2003	TNDFSP	M	min sum operating costs	frequency bounds, capacity requirements	small example	From an existing network
Ceder	2003(a)	TNSP	H	Demand satisfaction, Number of line runs	Headway bounds	Theoretical	Includes vehicle scheduling
Quak	2003	TNDSP	H	Drive time, Number of line runs, Detour time, Overloading, Fleet size, Waiting time	Headway limits	Software	No initial network
Castelli et al.	2004	TNTP	H	Passenger transfer waiting time	Number of runs per line, Transit network	Randomly generated example	Mainly devoted to commuters, Difficult to obtain demand data
Fleurent et al.	2004	TNTP	H	Trip synchronization, Historical runs fleet size, Deadhead, Layover	Passengers waiting time bounds	Software	Hastus
Wong et al.	2004	TNTP	M	Waiting time	Known transfer time per passenger and per period	Example	Railway
Fan et al.	2004/06a-b	TNDFSP	NS-E	Waiting, Walking, In-vehicle time, Number of buses, Unsatisfied demand costs	Route length	Examples	Variety of NS algorithms
Eranki	2004	TNSP	H	Number of simultaneous arrivals, Passengers waiting time	Waiting time limits for passengers	Benchmark	Ceder follow-up study
Yu et al.	2005	TNDP	O	Number of transfers, Maximum passenger flow per unit length	Line length, Route directness	Example	Urban
Park	2005	TNFSF	O	Fleet size, Travel time, In-vehicle and Waiting time	Lines, Stochastic buses arrival times	Case study	Urban
Borndorfer et al.	2005	TNDFSP	M	Passengers traveling time, Operating costs	Demand satisfaction	Ex: Potsdam	Path-based model
Jansen et al.	2005	TNTP	NS	Total transfer time	Preset routes and Frequencies, Fixed headways	Ex: Copenhagen	Intermodal
Bachelet et al.	2005	TNDSP	O	Users waiting time, Operating costs	Total routes length	Theoretical	Urban
Cevallos et al.	2006	TNTP	E	Transfer waiting time	Fixed headways	Ex: Broward County	Modifies an existing timetable
Zhao	2006b	TNDFSP	NS	Number of transfers, Service coverage	Route directness	Ex: Miami Dade	Large-scale problems
Barra et al.	2007	TNDP	M	-	Passengers demand, budget, level of service	Benchmark	Constraint Programming, only small instances

Table 2: A Classification of the papers dealing with TNSD Problems

6. TNDSP - A Review

6.1. Transit Network Design

A transit network design problem consists in determining a set of routes, each being formed of two terminals and a sequence of intermediate bus-stops, given a trip demand distribution, an area's topologic characteristics and a set of objectives and constraints. From the users' perspective, a transit network should cover a large service area, be highly accessible, offer numerous direct-through trips, hardly deviate from shortest paths, and should globally be able to meet the demand. The number of routes or total route length should also be kept under a certain bound so as to reduce the operator's costs.

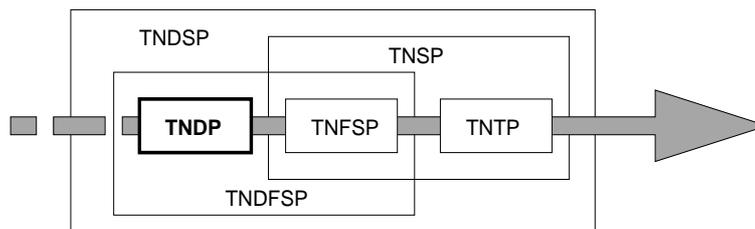


Fig. 2. TNDP positioning in the TNDSP

6.1.1. Mathematical Approaches

Determining the location and density of bus-stops is a crucial step in the network design process since it influences the transit system's accessibility, performance and level of service. The higher the frequency of bus stops on a route, the better the accessibility, but the lower the travel speed. Murray (2003) studies two variations of this problem. In the first part, the relocation of bus stops in an existing network is considered with the objective of minimizing the number of bus stops. A Location Set Covering Problem and a Maximal Coverage Location Problem are used to model the issue. The second part deals with the optimal location of bus stops to create or extend the network. Given a fixed number of additional bus stops to locate, the objective is to maximize the extra service access provided to non-covered areas. A hybrid Set Covering Problem formulation is proposed to model this second issue. The method is very interesting since it permits to expand both service access and accessibility.

The problem of simultaneous transit line configuration and passenger line assignment is dealt with by Guan et al. (2003). Focus is on large city railways and tests are carried out in Hong Kong city. Their work attempts to jointly model the transit line planning and the passenger transferring process through a linear binary integer program that can be solved by a standard branch and bound method. Given route length bounds, maximum number of transfers and capacity constraints, the objective is to minimize a function of the total length of transit lines, total number of transit lines taken and total length traveled by passengers. To keep the problem tractable under a mathematical resolution method, a precomputation of a restricted set of possible paths is achieved. However, since at this stage, frequencies and timetables are not known yet, passenger line assignments cannot be very accurate, questioning the whole process.

In the previous approaches, it is acknowledged that the size of the instances becomes a serious limitation for real-world problems. Heuristic approaches are then usefully employed.

6.1.2. Heuristic Approaches

Patz (1925) was probably the first to tackle the transit network design problem using heuristics. He proposes an iterative procedure to generate a lines network using penalties. Initially, the network contains a line for each origin-destination pair. For each one, a penalty is calculated based on the occupancy level and the number of passengers who need to transfer to complete their journey. Lines are iteratively deleted from the network based on this value and the capacity for passengers thereby freed is reassigned to other lines.

The small instance (about ten nodes linked in a specific manner) on which Patz’s tests are based permits the method to attain optimality. However, it is not extensible to larger networks.

In (Sonntag, 1977), the TNDP is approached with a heuristic procedure originally created for railway systems. Starting from a network containing a line for each origin–destination pair (like Patz (1925)), lines are iteratively deleted and passengers are reassigned to short paths in terms of travel time. This leads to a network of appropriate size with small average travel time and a small number of transfers.

In opposition to (Patz, 1925; Sonntag, 1977) and similar methods, Mandl (1979) tackles the TNDP starting with an empty routes network. He proposes a heuristic algorithm to define a transit network given a constant frequency on all bus lines. In a first part, routes on the shortest path that connects a pair of terminals and serve the greatest number of origin–destination pairs are iteratively selected. They can then be adapted to fulfill service coverage and directness objectives. In a second part, the transit network is also iteratively modified to minimize the total travel time of passengers. The method was applied on a real network in Switzerland, consisting of 15 nodes and 15570 trip demands per day.

Following the objective of maximizing the number of direct travelers, Pape et al. (1992) present a heuristic method for the TNDP. This constructive heuristic is based on the concept of corelines, which describe lines or parts of lines with a large number of passengers. Corelines are combined in a complete enumeration scheme and the best partial line plan considering the number of lines and the number of direct passengers is generated. This line plan is then extended for service coverage purpose.

6.1.3. *Neighborhood Search Approaches*

An aggregated metaheuristic approach to the transit network design problem is considered by Zhao and Gan (2003), Zhao and Ubaka (2004) and Zhao and Zeng (2006a) with the objective of minimizing the number of transfers and optimizing route directness while maximizing service coverage. The concept of keynode is defined to elaborate neighborhoods in the context of metaheuristics resolution methods. An Integrated Simulated annealing, Tabu and Greedy search algorithm is proposed in (Zhao and Gan, 2003; Zhao and Zeng, 2006a) while Basic greedy search and Fast hill climb search are implemented by (Zhao and Ubaka, 2004). These algorithms are tested on benchmark instances and on data from Miami Dade County, Florida.

6.1.4. *Evolutionary Algorithms*

Xiong and Schneider (1993) present an innovative method to select additional routes for an existing network. Their method is based on an improvement from the ordinary Genetic algorithm, called the cumulative GA. The principle is to collect all non-dominated solutions throughout the process and return this set instead of the last generation as is commonly done. The second proposal of this paper is to use neural networks instead of a passenger trip assignment algorithm to evaluate fitness functions (here waiting time and costs incurred). This proves to give accurate results and achieve a substantial time gain. The combination of these two improvements creates a very powerful tool, which can provide a set of interesting solutions to the human planner. This work is tested on a benchmark from LeBlanc et al. (1975).

Chakroborty and Dwivedi (2002) also propose a Genetic algorithm based method. In a first part, an initial set of routes is determined heuristically. Then a process consisting of an evaluation and modification procedure is iterated. A solution is a set of routes and its fitness function is evaluated through five criteria: the average or total in-vehicle time incurred (including transfer time), the percentage of passengers achieving a direct trip, a trip with one or two transfers and the percentage of unsatisfied demand. If the global quality is not good enough, routes or parts from routes are exchanged in such a way that new solutions are explored and eventually a solution of better quality will appear. Tests have been carried out on the Swiss benchmark from Mandl (1979).

6.1.5. *Another Approach*

Yu et al (2005) develop an ant colony based optimization model for bus network design that aims at achieving minimum transfers and maximum passenger flow per unit length with route length and directness constraints. This way, benefits of both passengers and transit companies are considered. A vacant network is first established, in which routes are then added according to the objectives, until all passengers are

distributed to the network or some given constraint is overrun. A coarse-grain parallel ant colony algorithm (CPACA) is used to solve the problem and tested with survey data of Dalian city.

Transit network design, as it is underlined by Yu et al (2005), is a fundamental step in the global planning process, since bus routes are rarely modified. TNDP constitutes indeed the strategic stage of this process, with a validity range of 5 to 15 years. The next step in the process is to set the lines frequencies.

6.2. Transit Network Frequencies Setting

The transit network frequencies setting problem consists in determining adequate frequencies for each line of the network and each time period. A time period is defined according to hour of the day, day of the week or time of the year. A suitable frequency assignment should provide sufficiently regular service to satisfy the users and sufficiently sparse service to reduce the required fleet size, and thereby the operator's costs.

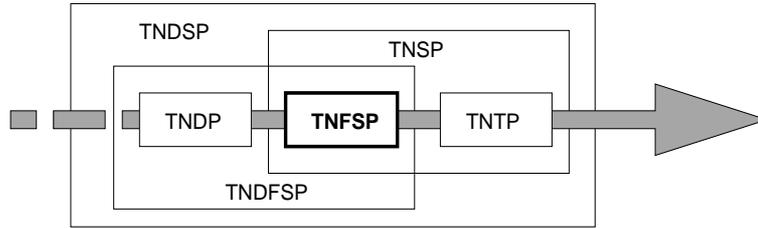


Fig. 3. TNFSP positioning in the TNDSP

6.2.1. Mathematical Approaches

In a first study, Salzborn (1972) determines frequencies given passengers' arrival rate so as to minimize the required fleet size and passengers waiting time. In a later study, Salzborn (1980) analyzes the problem of a feeder bus system along a single inter-town transfer route and determines requirements for the feasibility of such a system. The proposed model is highly simplified.

Scheele (1980) also deals with the TNFSP using mathematical resources. A nonlinear model is proposed with the objective of minimizing the total generalized passenger travel time and the passenger trip assignment is solved simultaneously with the frequencies setting problem.

Furth and Wilson (1982) present another mathematical method for the TNFSP. The objective is to maximize the net social benefit, consisting of ridership benefit and waiting time saving. Constraints are imposed on fleet size, maximum headway and total budget. The problem is solved through an algorithm using the Kuhn-Tucker conditions on a relaxation of a non linear program where the maximum headway and fleet size constraints are relaxed. The result is an optimal allocation of buses to routes.

Constantin and Florian (1995) present a model and resolution method for the TNFSP with the goal to minimize the passengers total expected travel and waiting time under fleet size constraints. A nonlinear nonconvex mixed integer programming model is formulated. A projected sub-gradient algorithm is then used to find optimal line frequencies considering the passengers route choices. Tests are performed in urban context in Stockholm, Winnipeg and Portland.

6.2.2. Heuristic Approaches

Han and Wilson (1982) deal with the TNFSP using a heuristic method. Line frequencies are set given the total fleet size and a capacity for each route. They consider the problem as one of allocating vehicles among the routes of the network. They propose a two stage heuristic to reach their objective, minimizing the maximum "occupancy level" at the maximum load point for each route. In the first phase, minimal frequencies are set so as to satisfy all the demand. In the second phase, frequencies are increased uniformly among lines so as to utilize all the available vehicles.

For the first time, Ceder and Wilson (1986) defined and presented a conceptual model for the whole bus planning process as a systematic decision sequence (Table 1), which consists of five levels: network design,

frequencies setting, timetable development, bus scheduling and driver scheduling and rostering. Minimum frequency and fleet size were given as constraints. A two level methodological approach was presented for the design of the bus routes network. The first level considers only the passengers' viewpoint and is handled by an enumeration method which generates routes whose length is within a certain factor from the length of the shortest path. The second level considers both the passengers and operator's viewpoints and relies on heuristic techniques. This algorithm aims at minimizing the difference between the number of users served directly and indirectly (i.e. using the shortest path or not).

Chowdhury and Chien (2001) consider transfer coordination for intermodal transit networks by optimizing both headways and slack times. Slack time is the additional time added into the schedule for a given trip in order to compensate potential bus delays and thus increase the probability of schedule adherence. A mathematical programming model is first developed and then a procedure is exposed that first, optimizes headways without taking coordination into account, and then optimizes slack times in the context of intermodal transit.

Recently, Gao et al. (2003) proposed a bilevel programming technique to deal with the TNFSP. In the upper-level problem the objective is to minimize the total deterrence of the transit system (consisting of in-vehicle and waiting time) and the cost caused by frequencies setting. The lower-level model is a transit equilibrium assignment model used to describe the path alternatives to transit users. A heuristic solution based on sensitivity analysis is designed to solve this model to optimize frequencies settings. The algorithm is designed to help the transit planners to adjust an existing transit network to evolutions in the demand and in various other parameters.

6.2.3. Another Approach

Park (2005) uses Genetic algorithms and simulation to optimize bus schedules in a urban transit network. Two different cases are considered. When buses arrive following a deterministic process, a simple Genetic algorithm is used in combination with problem-specific operators to determine optimized headways. On the opposite, when buses arrive stochastically, a simulation-based Genetic algorithm is used to optimize both headways and slack times. Problem-specific genetic operators include coordinated headway generator, crossover and mutation.

6.3. Transit Network Timetabling

A transit network timetabling problem consists in setting a timetable for each line of a predetermined transit network, with departure times for each line run. This process can be dependent on pre-set frequencies if those exist or just skip the frequencies setting step. For the users, a timetable is suitable if it enhances transfer conditions, through both number of possibilities and synchronization. From the operator's perspective, an appropriate departure times setting can help reduce the required fleet size. However, this latter point of view lacks treatment in the literature, since resource considerations are rarely taken into account but in a superficial manner (e.g. including a fixed cost per line run or considering that vehicles only run back and forth on lines, thus preventing interlining and imposing identical frequencies on both directions of a line).

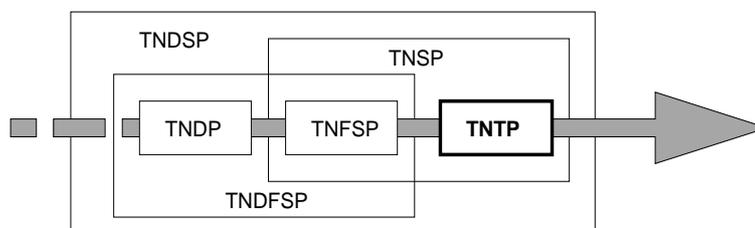


Fig. 4. TNTP positioning in the TNDSP

In order to reduce passengers' waiting times at transfer zones, transfers can be coordinated, especially when headways are large. This is often the case in rural areas. Indeed, in large cities (Chakroborty, 2003), frequencies tend to be high and missing a connection only extends one's waiting time by a few minutes. Moreover, due to congestion, coordinating transfers in order to reduce waiting time may result in more insecure transfers. On the opposite, in lower density areas (Yan and Chen, 2002), missing a connection incurs larger delays and the absence of synchronization may even discourage people from using public transport at all. Furthermore, depending on the frequency level, the behavior of passengers also varies. When headways are small, users tend to arrive randomly at bus stops and wait for the next connecting bus. This initial waiting time represents a cost for the transit user. On the opposite, when headways are longer, users tend to be aware of the schedule and reach the bus stops just before the bus itself. Therefore, no costs are usually incurred for initial waiting time.

Whatever the context, two options arise to tackle transfer coordination: timed transfers and transfer optimization (Bookbinder and Désilets, 1992). In the former case, buses from different lines are scheduled to meet at certain transfer points through the addition of layover, while in the latter case lines departure times are scheduled to minimize the total passengers transferring time. Moreover, studies have been carried out to analyze the influence of transfer coordination. Knoppers and Muller (1995) study the bus feeder system in an intercity context. They found that coordination was worthwhile only when the arrival time standard deviation on the feeder line at the transfer point is less than 40% of the headway on the major service network. Ting and Schonfeld (2005) obtain a similar conclusion.

A particular way of coordinating transfers is to normalize headways so that vehicles from different routes arrive at some transfer station simultaneously. When headways are fairly similar, a common headway could be sought. On the opposite when the network comprises lines with notably different frequencies, headway values could be fixed on integer multiples of the smallest headway (Ting and Schonfeld, 2005).

6.3.1. *Mathematical Approaches*

DePalma and Lindsey (2001) consider the transit network timetabling problem on a single transit link. Each transit rider is assumed to have an ideal boarding time, and to incur a varying schedule delay cost from traveling earlier or later. An analytical model is developed to handle the two-step process: in a first part, a demand allocation problem is solved to assign individuals to line runs. In the second part, an optimization process is led to set departure times in such a way as to minimize the riders' total schedule delay costs. Unfortunately, the needed data is extremely difficult to obtain for a real-world application.

The static timetabling problem is tackled by Wong and Leung (2004) who present a timetabling method which aims at maximizing synchronization between lines and enabling smooth transfers with minimum waiting time. Although the method especially suits railway systems, it can be applied to mass transit systems in general. A known transfer time at each interchange station for all passengers and periods is considered, as well as a few simplifying assumptions such as unlimited vehicle capacity and exact adherence to the schedule. A passenger trip assignment is achieved according to the following criteria: number of transfers and number of stops on the trips. A mixed integer programming optimization model is proposed with run times, dwell times and dispatch time of each train as decision variables. The LP relaxation of the MIP formulation is solved using CPLEX, in order to obtain values of the integer variables. The resulting linear programming formulation is finally solved to obtain timetables for all trains in the entire schedule.

6.3.2. *Heuristic Approaches*

Klemm and Stemme (1988) innovate in the domain and deal with the TNTP in the perspective of optimizing transfers. An integer programming model (quadratic semi-assignment) and a heuristic resolution method are proposed. The latter is a constructive process in which lines are scheduled one at a time with consideration to transfer synchronization.

Also focusing on minimizing transfer waiting times, Castelli et al. (2004) present a mathematical model for the TNTP. The specificity of their approach is that the number of line runs is not dependent on frequency bounds but on the quality of service mitigated by a fixed cost per line run. The quality of service is evaluated through the sum of the transfer waiting times for all the passengers, for whom all the desired pickup or de-

livery times are supposed to be available. However, despite restrictive assumptions, the authors acknowledge that their model is intractable for a network-wide optimization given the large number of variables. Thus, a Lagrangian based heuristic resolution method is proposed, in which lines are scheduled one at a time in the network.

Using the idea of weighing transfers, Fleurent et al. (2004) describe concepts that are implemented in the commercial software Hastus (<http://www.giro.ca/en/products/hastus/index.htm>) to generate synchronized transit timetables. The scheduling process includes two types of parameters: the waiting time and the relative importance of transfers (number of passengers involved, historical importance of the connections and so on). Two innovative concepts are defined: "trip meets" that describe a possible connection between two trips at a transfer place. For each transfer, a weight factor as well as minimum, maximum and ideal waiting times are provided. "Meet builders" are then created to group trip meets sharing common features. The resolution approach involves solving multiple network flow problems, Lagrangean relaxation, as well as several heuristic mechanisms, but is hardly described here.

6.3.3. *Neighborhood Search Approaches*

Daduna and Voss (1995) propose a method to determine departure times in a predetermined network with line frequencies. The goal is to synchronize arrival times at transfer zones so as to minimize the waiting time incurred. However, this objective can be refined and for example include weights on the different transfers or on the maximum waiting time at a transfer zone. A mathematical model based on quadratic semi-assignment is used to formulate the TNTP. Due to its complexity, a heuristic resolution is preferred. A variant of a regret heuristic is thus used to compute initial solutions, which is then improved by a Simulated annealing as well as different versions of Tabu search.

Jansen et al. (2002) propose a method to synchronize bus timetables so as to minimize passengers transfer time given a routes network with fixed headways. Stopping and in-vehicle time being assumed constant and deterministic, the problem is reduced to setting departure times for the first run of each line. First, a non linear mixed integer model is developed. Then, Tabu search is applied and tests are performed on the city of Copenhagen.

6.3.4. *Evolutionary Algorithms*

Chakroorty et al. (1995), Chakroorty et al. (1997) and Deb and Chakroorty (1998) put the focus on the application of Genetic algorithms to determine departure times in a transit network in a coordinated manner, so as to minimize the passengers' total waiting time. Considered constraints include a predetermined fleet size, policy headway, bounds on the possible stopping time, and a maximum transfer time. In a first part, mixed non linear programs are formulated to model the problem. However, it is underlined by the authors that a Branch-and-Bound method would be computationally intractable, even for small size instances. Therefore, Genetic algorithms (hybrid ones in (Chakroorty et al., 1997)) are chosen to tackle the problem. Reasons for this choice are especially put forward in (Chakroorty, 2003). The genetic representation of a complete schedule is composed of a series of sets of binary digits representing headways and stopping time between pairs of consecutive vehicles on all routes. It is however suggested that this representation could become a burden for very large transit systems. The author also underlines that some simplifying assumptions are made, such as unlimited vehicle capacity and fixed fleet size, and highlights the fact that a network-wide transfer optimization model would be desirable.

A slightly different problem is defined in (Chakroorty et al., 2001), where transfer coordination is combined with vehicle scheduling. This is modeled as a non linear mixed integer problem and solved using a binary coded Genetic algorithm. The problem here is to determine the fleet size and the schedules for each line such that, given the demand, the transit system provides the shortest waiting time to the passengers within the available resources. Constraints are set on total fleet size, stopping time, headways, and transfer time. Note that the considered case of the single transfer stop with multiple lines is restrictive. It is also felt that one of the shortcomings of the proposed procedure is that the chosen genetic representation will become cumbersome for large transit systems.

A network-wide approach is presented by Cevallos and Zhao (2006) to tackle the Bus Transfers Synchronization Problem. They consider the objective of minimizing transfer waiting time in an existing timetable under the constraints of strictly fixed headways between line runs. This restriction deprives the search of some flexibility but complemented with the fact that resource concerns are left aside, it makes the problem more tractable than the ones presented before, especially in the case of large networks.

6.3.5. Another Approach

A hybrid transfer optimization method is proposed by Bookbinder and Désilets (1992) to set lines departure times in a given transit network with fixed headways. It combines the flexibility of simulation methods with the power of mathematical optimization. The simulation procedure permits to take into account the buses stochastic travel times while the mathematical one (a relaxation of the QAP) permits to optimize transfers according to various waiting time related objective functions.

6.4. Transit Network Design & Frequencies Setting Problem

A transit network design and frequencies setting problem consists in determining a set of line routes associated with frequencies in a particular area, given a public demand, and relatively to a set of objectives and constraints, usually related to route directness, service coverage and operating costs. In (Yu et al, 2005), it is argued that bus routes network design and frequencies setting should not be dealt with simultaneously, since the routes network is a more stable component in a transit system. Thus, its construction should not be influenced by flexible parameters such as frequencies. However, it is also arguable that the interaction made possible by handling both problems at the same time is beneficial to the final results. Moreover, to design the transit network, a passenger trip assignment is generally needed. However, it is hardly realistic to assign paths to passengers at this stage of the process since frequencies and timetables are not established yet. This assignment can thus be more efficiently done when frequencies are involved.

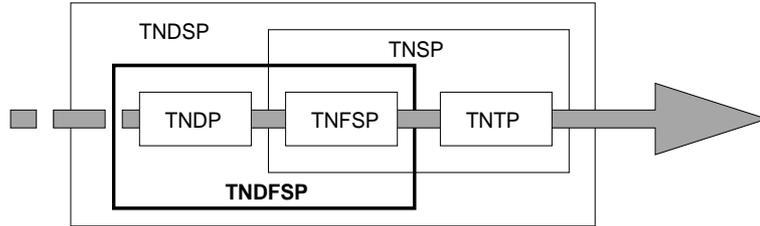


Fig. 5. TNDFSP positioning in the TNDSP

6.4.1. Mathematical Approaches

In (Hasselström, 1979, 1981), a real-world applicable three step model is proposed to solve the TNDFSP. First, a full link network is considered and then a large set of possible routes are generated. Finally, the routes of the network are selected by assigning frequencies using linear programming. Routes and frequencies are thus determined simultaneously. In the first study, the objective is to minimize the number of transfers, while in the second one the objective is the maximization of consumer surplus, i.e. the number of passengers in the context of a variable demand.

Initially focusing on rail transport, Bussieck (1998) proposes mathematical models to create line routes and frequencies that can more generally be applied to mass transit systems. A specificity of this approach is that it is based on periodic schedules, meaning fixed time intervals separate the departure times of vehicles on the same route. In the first part, the objective consists in maximizing the number of direct passengers under resource-related constraints. In the second part, Bussieck seeks to minimize the operator's costs with respect to a given level of service and quality. Mathematical programming methods such as relaxation and branch-and-bound are applied in combination with commercial solvers. In this paper, Bussieck refers to

DONS (Van den Berg and Odijk, 1994) from Rained, a software package that helps the schedulers to compute timetables satisfying sets of constraints.

Borndörfer et al. (2005) propose two multi-commodity flow models for this problem, which they call line planning problem. These models minimize a combination of total passenger travel time and operating costs. The first one uses arc variables for both line and passengers paths whereas the second one uses paths variables. In opposition to most previous approaches, these models do not use system-split (that fixes the traveling paths of the passengers before the lines are known) but allow to freely route passengers in the routes network. Routes are generated dynamically and frequencies are implicitly handled in order to compute an optimal routing. LP relaxations of the models in combination with a commercial solver are chosen to attempt to solve the problem.

In a complementary manner, (Wan and Lo, 2003) study the problem of modifying the structure and lines frequencies of an existing transit network. A Mixed Integer formulation is proposed and linearized so it can be solved by commercial solvers on small size instances.

Most recently, (Barra et al., 2007) proposed a Constraint Satisfaction model to be used in combination with a commercial Constraint Programming solver. While no objective is sought, eleven types of constraints, evolving around passengers demand, budget limits and level of service, define the problem. This paper represents a first work towards CP resolution of the TNDP and in the current state can not present any substantial result, since the CP package used was only able to process small instances. However, paths are defined that could lead to more efficient searches.

6.4.2. *Heuristic Approaches*

Lampkin and Saalmans (1967) propose a sequential approach of the TNDP. In the first phase, a heuristic algorithm based on a skeleton method is developed to design the transit network by iteratively adding routes with the objective of maximizing the number of direct passengers. In the second phase, frequencies are assigned to the routes using a random greedy-based search procedure aiming at minimizing the total travel time given the fleet size and an unlimited vehicle capacity.

Also using a skeleton method to define a bus routes network and its frequencies, Silman et al. (1974) aim at minimizing overcrowding and passengers' journey time while satisfying budget constraints. This sequential skeleton method consists in choosing endpoints for a route and then linking them by passing through different zones using shortest paths routes. The objective at this point is to obtain a network that provides short travel times to passengers and which routes are not excessively long so as to limit the required number of buses. In a second step, which concerns frequencies setting, the objective is to minimize the passengers' total journey time and inconvenience due to bus overcrowding, while respecting the available number of buses. This approach was tested on data from Haifa.

In a similar way, Bel et al. (1979) propose a sequential approach to the TNDP. Three sub-problems are identified, namely the set of streets selection, the set of lines determination and the optimal frequencies setting. To choose an interesting subset of streets, a constructive heuristic procedure is proposed, which adds links to the subset with the objective of minimizing the total travel time of the passengers under budgetary constraints. The objective function is evaluated using an all-or-nothing passenger trip assignment on the shortest path. Then, a maximal set of routes is generated from the subset of streets. Finally, line frequencies are set using a gradient-based search heuristic and considering the minimization of waiting times.

VanNes et al. (1988) present a method to design simultaneously a routes network and the associated frequencies with the objective of fulfilling the demand and maximizing the number of direct trips given a certain fleet size. Their solution approach combines mathematical and heuristic elements. An analytical model is proposed and mathematical programming methods are applied. In the beginning, a network is set that includes all possible routes with for each one a null frequency. At each step, the marginal efficiency of one frequency unit is calculated for each route, taking into account both the extra number of passengers and the extra costs incurred. Eventually, a routes exchange routine is introduced to increase the quality of the solution obtained through this purely constructive method.

Focusing on large cities, ? describe a complex urban bus network design procedure. Indeed, three levels of public transportation are differentiated: express, main and feeder. A sequential heuristic procedure is used

to compute the three routes networks and their frequencies, with respect to costs such as users waiting time and excess time compared to minimum path, and operator service costs. The model is applied to an area of Rome, resulting in substantial savings.

An AI-based heuristic approach is proposed by Lee and Vuchic (2005) to deal with the TNDFSP in the context of variable demand through an iterative approach. The objective is to minimize users total travel time. Their resolution method is inspired by the work from Rea (1972).

Dealing with the original concept of short-turning, Ceder (2003a) tackles the TNDFSP using a twofold method. The first part consists in designing routes at the network level, while the second one focuses on the so-called "short-turning" strategy, during which some line runs skip initial or terminal bus stops. This strategy can be profitable when demand varies a lot along the same line. The first part of the method aims at designing a direct and demand-oriented network. First, all feasible routes and transfers are generated, then smaller subsets maintaining connectivity are isolated, and appropriate frequencies are calculated for each line. A transit trips assignment is then achieved and eventually, the selection of the most suitable lines subset is done according to optimization criteria arising from the operator, the users and the community. The second part aims at reducing the required fleet size.

6.4.3. *Neighborhood Search Approaches*

In (Fan and Machemehl, 2004, 2006b), the TNDFSP is studied in the context of the "distribution node" level, in opposition to earlier approaches which aggregated zonal travel demand into a single node. A multiobjective nonlinear mixed integer model is used to formulate the problem. The objective considered is the minimization of the sum of user-related costs (waiting, walking and in-vehicle time), operator-related costs (number of buses) and unsatisfied demand costs (which is a rare feature in such problems). Demands are categorized as "unsatisfied" when they require more than two long walks, transfers or combination of both. For efficiency reasons, a heuristic resolution method is preferred to the mathematical one. A three component solution framework is proposed that can be applied to both fixed and variable transit demand models. First, all candidate routes satisfying length constraints are generated using algorithmic procedures. Second, the network structure is analyzed and evaluated so as to determine the line frequencies. Finally, metaheuristics are used to select an optimal set of routes from the solution space. Five different algorithms are used: Tabu search, Simulated annealing, Genetic algorithms, Random search and Local search. The Simulated annealing algorithm is particularly detailed by Fan and Machemehl (2006b).

The problem of large-scale TNDFSP is addressed by Zhao (2006b) in a follow-up study of (Zhao and Gan, 2003). Using an Integrated Simulated annealing and Genetic algorithm solution search method, they seek to obtain a transit network minimizing transfers with reasonable route directness while maximizing service coverage. Their computational tool is effectively tested both on literature benchmarks and realistic data from a transit operator.

6.4.4. *Evolutionary Algorithms*

Pattnaik et al. (1998) propose a Genetic algorithm based method to simultaneously determine the transit routes network and frequencies. The objective is to minimize both operator costs and passengers travel time, given headway constraints. The design is done in two parts. First, a set of candidate routes competing for the optimum solution is generated, and then the optimum set is selected using a Genetic algorithm in which the solution route set size and the set of solution routes are found simultaneously.

A Genetic algorithm based approach of the TNDFSP is also presented by Bielli et al. (2002), who propose a method to compute fitness function values by means of a multicriteria analysis led on performance indicators. The objective of the model is to maximize the performance of the network and minimize the required fleet size and average travel time. The genetic representation of the network is based on genes representing lines characterized by two values: frequency and belonging to the network. A numerical experience is led on the city of Parma.

Also using Genetic algorithms, Ngamchai and Lovell (2003) propose a method to create transit routes and frequencies so as to minimize total costs incurred by fleet size and in-vehicle and waiting time. The process is threefold: first, a route generation algorithm creates more or less randomly a set of routes respecting the

network connectivity constraint. Then, a route evaluation algorithm associates service frequencies to the lines and coordinates headways at transfer points. For computational reasons, headways are only coordinated at one single transfer point on each route. Finally a route improvement algorithm modifies the network using a pool of problem-specific genetic operators in order to improve its efficiency. The proposed method is tested on a benchmark from Pattnaik et al. (1998). A nice extension of this work would be to consider several transfer points at a time.

Tom and Mohan (2003) present a Genetic algorithm based method that simultaneously determines transit lines routes and frequencies, with the objective of minimizing both operating costs and passengers total travel time. The distinctiveness in this approach is the coding scheme that incorporates the line frequency as a variable, in order to enhance the performance of the model.

Finally, the application of Genetic algorithms to the TNDFSP is studied by Fan and Machemehl (2004, 2006a). Given a variable transit demand, three components (that are briefly described in 6.4.3) guide the candidate solution generation process using Genetic algorithms. The network redesign issue is also addressed in this paper.

6.4.5. *Other Approaches*

Baaj and Mahmassani (1991, 1995) propose a three step hybrid AI-based heuristic algorithm to deal with the TNDFSP. This approach combines experts' knowledge and algorithmic procedures. The procedure starts with a route generation algorithm that determines a set of routes with respect to a combination of users and operator costs. This algorithm is composed of a skeleton method that connects the origin-destination pairs with the highest demand. A passenger trip assignment is then used to determine the frequency and number of buses required on each line, followed by an analysis procedure called to evaluate the effectiveness of the network (in terms of number of direct trips, total waiting time, transfer time etc). Finally, a route improvement algorithm modifies the network structure heuristically in order to improve its global effectiveness.

A follow-up study of Baaj and Mahmassani (1991) is presented in (Shih and Mahmassani, 1994; Shih et al., 1998). The approach is still multi-objective and considers the total travel time, total satisfied demand and required fleet size. A four phase framework is used, starting with a route generation procedure. Then, a network analysis procedure evaluates the set of routes, a transit center selection procedure defines zones where passengers can transfer, and finally a network improvement procedure modifies the network. In addition to what was done by Baaj and Mahmassani (1991), advanced concepts like route coordination, variable vehicle size, and demand responsive service are considered. However, this method still relies much on the scheduler and is not applicable to real-world problems.

More recently, Fusco et al. (2002) present a three-step sequential procedure combining simple heuristic techniques and a Genetic algorithm to generate a transit lines network that approaches the minimum overall system cost. To start with, a heuristic algorithm is used to generate a set of feasible routes. Then, a Genetic algorithm selects the optimal subset of routes with associated frequencies. Finally, route modifications are tested to improve the network configuration (extension, shortening, expansion to other nodes). Their method is more detailed in Italian by Fusco et al. (2001).

6.5. *Transit Network Scheduling Problem*

Given a bus routes network and a public travel demand, the transit network scheduling problem consists in determining the network's timetable, including frequencies and departure times for each line, with respect to a set of objectives and constraints.

Frequencies and departure times settings are dealt with simultaneously especially in the case where policy headways are provided, meaning that a minimum frequency is imposed as a constraint to the schedulers. Maximum frequencies can also be provided. In these cases, frequencies are to be set into the predetermined intervals and departure times are set so as to secure transfers among bus routes.

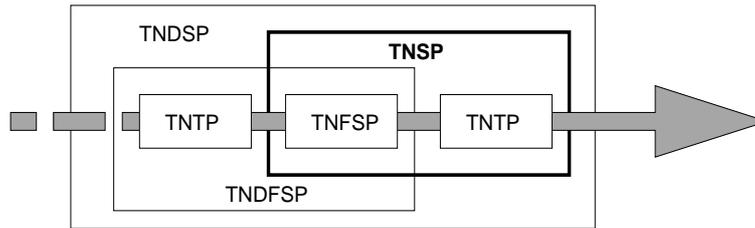


Fig. 6. TNSP positioning in the TNDSP

6.5.1. *Mathematical Approach*

Considering most NP-hard problems, mathematical resolution methods require intractable computational time. In the case of frequencies and departure times setting, the problem can most of the time be formulated by an analytical model, but can hardly be resolved by exact methods.

However, Koutsopoulos et al. (1985) present a mathematical modeling and resolution method for a simplified TNSP. Passengers waiting time costs, operating costs and vehicle crowding costs constitute the objective function to minimize.

6.5.2. *Heuristic Approaches*

Rapp and Gehner (1976) present a heuristic resolution method to optimize transfers in the context of a TNSP. At each iteration, the scheduling of one line is modified until no further improvement can be made.

Ceder et al. (2001) propose a heuristic method to set frequencies and departure times in order to maximize the number of simultaneous arrivals at transfer nodes between buses from different lines. In a first part, the problem is formulated as a mixed integer linear programming problem, but a heuristic algorithm is rather developed to solve the problem. Lines departure times are set at each iteration in a constructive manner. A limitation of this approach is that exclusively exact simultaneous arrivals will be considered, when sometimes one-way transfers might be interesting as well.

A variation of this work is presented by Eranki (2004). A model and a heuristic method are developed to create bus timetables with the objective of creating the maximum number of quasi-simultaneous arrivals and minimizing the passengers' waiting time. Indeed, in this research it is assumed that at each transfer node there is a minimum and maximum waiting time limit for passengers, thereby defining the quasi-simultaneous arrivals. The problem is modeled as a mixed integer linear programming problem and a heuristic based on (Ceder et al., 2001) is developed. The main difference with the previous approach is the inclusion of the waiting time parameter.

In a later study, Ceder (2003b) proposes three procedures based on earlier work to set frequencies and departure times of buses in a given network with frequency constraints and several objectives, among which meeting the variations in public demand. The basic criteria are to provide sufficient seating capacity and to assure a minimum service frequency. Ceder describes two point-check and two ride-check methods based on a curve representation of the passenger load for calculating frequencies. Depending on the type of headway, the chosen frequency determination method and special requests, alternative timetables can be obtained. The interesting problem of setting departure times between adjacent time periods is also addressed here.

6.6. *Transit Network Design & Scheduling Problem*

A transit network design and scheduling problem consists in providing a bus transit system readily available for operational planning. A set of lines with the associated timetables must be set up in such a way as to optimize a group of objectives and satisfy a set of constraints. In particular, the demand must be satisfied by a service of good quality and the resources requirements must be limited.

The combination of the three first steps of the global public transit planning process is very cumbersome. Thus, the models considered often have to be size-limited or simplified. Despite the demanding character of the task, we will see that several studies have tackled the problem, using various methods.

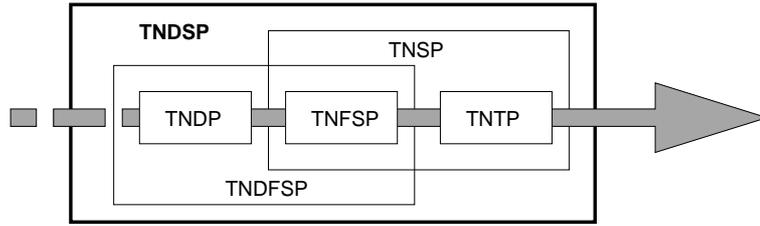


Fig. 7. TNDSP and its internal structure

6.6.1. Heuristic Approaches

Focusing on inter-city bus systems, Yan and Chen (2002) propose a mathematical scheduling model and resolution method. Given data such as rights of way, inelastic passenger trip demands and available fleet size, the objective is to determine optimal lines and timetables so as to maximize the profit. Since a rural bus system is considered, no waiting time costs are incurred. The model employs multiple time-space networks to model bus movements and passenger flows and manages the interrelationships between passenger trip demands and bus trip supplies. The TNDSP is formulated as a mixed integer multicommodity flow problem. The algorithm developed is based on a Lagrangian heuristic and a flow decomposition algorithm.

Quak (2003) presents a sequential method for solving the TNDSP. This work was realized for the largest bus company in the Netherlands. His approach is two-fold but relies on a common objective function based on total drive time, number of line runs and detour time. First, a lines network is constructed from scratch, an innovative perspective of the problem. The algorithm, the line set construction method, is based on (Pape et al., 1992) and the concept of corelines. Before tackling the scheduling problem, passengers are assigned to lines through a simulation process. In the second phase, the first step consists in setting frequencies. This is achieved using point-check and ride-check methods (Ceder, 1984). With respect to headway limits, the goal is to prevent overloading while minimizing the number of required buses. When setting the departure times in a second step, the waiting time for the passengers is minimized with respect to the defined frequencies and the weights attributed to each transfer. A line-oriented departure setting method is developed, which is inspired by the synchronization method from Ceder and Tal (1999). Note that although all the steps are considered in this approach, it is still done in a sequential manner.

6.6.2. Evolutionary Algorithm

A bus-feeder system is considered by Dhingra and Shrivastava (1999) for a small section of suburban railway corridor. The aim of the study is to define a model to design, schedule and coordinate optimal feeder routes for each railway station. In this particular context, transfers between buses are prevented. Objectives in the model are the minimization of in-vehicle travel time, of the number of buses and standing passengers. The proposed framework consists of three steps. First, routes are created to serve desired destinations in the shortest possible time while meeting maximum demand and minimizing passengers walking time. Then, the fleet size is minimized and the level of service is maintained while minimizing the number of standing passengers. Finally, transfer time between trains and buses is minimized. The use of Genetic algorithms is only suggested for this simplified problem.

6.6.3. Another Method

In (Bachelet and Yon, 2005), a hybrid method is developed to create bus routes with the associated timetables. Given the complexity of the problem, a straightforward mathematical optimization resolution appeared to be very fast on simplified problems, while simulation and optimization techniques proved to find relatively good quality solutions in a long computational time. Thus, the authors have combined both methods in a model called "model enhancement". Mathematical resolution methods gives near optimal results onto a simplified part of the problem and simulation permits a quick evaluation of the solution applied to the global problem (in terms of travel and waiting time) while the model enhancement heuristic works on the constraints. This hybridization permits to obtain a solution of good quality in a reasonable time, however convergence is not guaranteed.

6.7. *Software applications*

Many commercial products focus on the operational planning steps. Indeed, in Microbus (www.ivu.de) from IVU, Heurs (http://www.lumiplan.com/Site/heures_offre.htm) from Lumiplan or in Austrics (<http://ap.trapezgroup.com/index.php>) from Trapeze, transit network timetabling is integrated with vehicle and duty scheduling. The first steps of the transit planning thus have to be dealt with manually and are considered as data for these products.

Another category of software products that can be considered related to transit network design and scheduling problems addresses the school bus routing problem. However, constraints and objectives are quite different since no transfer can take place, each selected demand must be satisfied, and all trips occur almost simultaneously.

A different approach is based on the simulation of entire transportation systems considering individual passengers behaviour. Among the software packages that rely on this principle, we can cite MADITUC (<http://www.polymtl.ca/recherche/rc/unites/details.php?NoUnite=58>) from Ecole Polytechnique de Montréal. This model simulates urban transit networks and treats origin–destination matrices. The transportation planning software VISUM (Friedrich, 1999) also permits to analyze a transportation system through simulation. It includes some optimization algorithms that can help designing and scheduling the system.

6.8. *Benchmarks*

The most difficult information to obtain in transit routing and scheduling certainly is the detailed public demand. The usual way of representing this public demand is the origin–destination matrix. It consists in synthesizing, for each origin–destination pair, the number of passengers willing to achieve the trip for a particular period. Methods to get this data include mostly surveys and statistical forecasting. Since surveys are excessively expensive to carry out and need periodical updating to preserve adherence to the actual demand, transit agencies are reluctant to sharing their data. Additional data is also needed to create a transit network, namely topological information such as possible bus stops and transfer zones associated with the in-between distances or travel time. In this part, we provide a short list of existing benchmarks and public demand data.

The most used benchmark in the transit network design literature might be the one proposed by Mandl (1979). It was then reused for comparison purpose by Baaj and Mahmassani (1991); Kidwai (1998); Chakroborty and Dwivedi (2002); Zhao and Gan (2003). Later, Pattnaik et al. (1998) published another benchmark consisting of a 25 nodes network. Ngamchai and Lovell (2003) tested their method on this same instance, adding a randomly generated demand matrix since it lacked in the original proposition. A third benchmark instance can be cited here: LeBlanc et al. (1975) suggested an instance composed of a network and a fixed travel demand as a benchmark problem, which in turn Xiong and Schneider (1993) used as a testing basis for their work. Ceder et al. (2001) also carry out numerical experiments on an example network which is used as a benchmark instance by Eranki (2004) with only some modifications.

7. **New context, challenges and paths for future work**

In order to determine how and whether the technical advice will be applied, it is critical to expose the contextual environment in which planners must make decisions, and to highlight the gap which can exist between models and reality. To cover this additional level of perspective, we examine recent policy developments impacting transit network problems, and define challenges and opportunities for future work.

7.1. *Recent policy developments*

Over the last two decades, the world of public transport has gone through major organizational changes that have impacted all its actors from regulating authorities to users through operators and more globally

the entire population. For an overview of the evolutions in transport policies, we refer to Goodwin (1999). In order to better understand the stakes and objectives of today's public transport problem, the two most important trends will be reviewed here:

- the privatization and deregulation movement,
- the development of integration and intermodality among transit networks.

In regulated markets, public transport is usually the responsibility and monopoly of state-financed local authorities, who are in charge of organizing a service of quality at low fares. Social objectives such as the inclusion of captive users (the part of the population dependent on public transport) are predominant and services are ensured even though they are unprofitable. Nevertheless, since the mid-80's, a series of privatization and deregulation policies has come into practice in various parts of the world, with the state withdrawing from planning and control rules. Governments who have chosen to deregulate expected that introducing competition would improve efficiency (lower prices and better quality), increase user choice and also reduce subsidy requirements.

In reality, the intensity of the competition has proved to not always be beneficial to the users. Depending on the area, it sometimes leads to new situations of monopoly (often accompanied by fare increases) after one of the competitors succeeds in eliminating the others. Some other times, it leads to cooperation between operators and in the best cases to integrated service. However, most of the time, users have to undergo an indefinite period of transition during which service can be highly unstable. White and Farrington (1998) examine the consequences observed in Great-Britain, one of the first countries to have deregulated and privatized its bus industry.

In order to face competition and make profit, bus companies need to meet the challenge of providing an attractive service at low costs. These two parameters are highly influenced by three elements: the network design, the service frequencies and the timetable. On a short-time scale, operators might be tempted to engage in predatory or pre-emptive behaviors (actions taken with the intent of stealing business from a rival company or preventing a rival company from stealing their business), as it has been observed by Ellis and Silva (1998). In the context of intense competition, these additional factors are then prone to skew the original objectives of the problem.

This contributes in explaining why, in the public transit problem, the objectives and their relative weights might vary among operators: in regulated areas they can tend to favor the quality of service while in others, economic objectives might prevail.

The second important evolution of recent years is the development of the need for integrated intermodal transport systems. Parallel to the population growth, ever-rising urbanization and car ownership rates, there has been an increasing awareness on the various negative impacts of transport, mainly with respect to safety, environment, economic welfare, efficiency (congestion) and health. One of the strategies usually adopted towards more environmentally sustainable transport patterns is to convince car users to shift to public transport. Typically, the decision of the user is made on two factors: time and cost. In order to compete with private modes of transportation, it is crucial to provide a service that enhances mobility, through better and more numerous connections and through integrated tickets among modes. This higher quality of service is obviously also beneficial to current users of the system and participates in securing a better social inclusion. Finally, attracting new users can also be economically profitable to the operator on a long-term basis.

However, as discussed just before, deregulation tends to lead first to competition rather than to cooperation. Thus, it is likely to impede rather than to enhance integrated transport, a conclusion reached by Everett (2006) in the case of Australia. Even when bus companies realize that there exist mutual benefits in working together, creating an adequate intermodal system remains a challenge. This is why mechanisms are now being developed and promoted to entice local authorities and operators to create partnerships for a better coordination.

Successfully integrating several modes requires adjustments such as modifying line itineraries and shifting timetables to facilitate connections. Depending on the situation and on the operator, it can lead to different parameterizations of the objective functions in each step of the public transport problem.

For more information on the cross-implications of deregulation and integrated transport, we refer the interested reader to Potter and Enoch (1997).

To conclude on this part, we can say that beyond the theoretical objectives usually studied in Transit Network Design and Scheduling Problems, a variety of extra parameters can be added to each problem and/or be assigned a different level of importance. It implies that proposed models and solution methods should present great flexibility, for lack of which they will not be able to adapt to the polymorphic expectations of the transport planners.

7.2. *Challenges and opportunities for future work*

Transit network design and scheduling is an active domain of operations research. New policies create new needs from the operators, and new challenges for the programmers and research community. Thanks to permanent evolutions in techniques and technologies, new possibilities arise to better address transit network problems, through different definitions, models and resolution methods. We expose here a sample of these paths for research.

7.2.1. *Giving special attention to specific sub-problems*

Synchronization management - Transfers are among the most important criteria in the decision of using public transport. To reduce the inconvenience incurred, studies have been led on the synchronization of buses stopping times in transfer zones. However, due to the demanding nature of the problem, transfers are still often only synchronized in one point of the network. A network-wide optimization of the transfer possibilities would be desirable.

Intermodality - Cooperation between transit agencies, especially those working on different levels (e.g. train and bus), is an elementary need in the prospect of developing public transit. People will use mass transit more easily if they can get clear information and benefit from scheduled transfers between modes. New policies encouraging integration have led for instance to the development of bus feeder systems to service train stations (Peng and Fan, 2004). This parameter can still be emphasized in all the steps of transit planning.

Intercity transit - Studies often focus on urban transit planning rather than interurban transit planning. However, contextual differences limit the transferability of the approaches from one case to the other. Indeed, in interurban transit networks, frequencies tend to be very direction-dependent, meaning usually in the morning headways are smaller in one direction than in the other, and the opposite mechanism is respected in the evening. This impacts several steps of transit planning (especially transit network frequencies setting and vehicle scheduling) and deserves a specific treatment.

7.2.2. *Taking advantage of technological progress*

Intelligent Transportation Systems (ITS)- The development of ITS has led to equip buses with GPS and track the geographic position of the vehicles. This way, the dynamic handling of the traffic has emerged, enabling the operator to deal with unexpected variations in the schedule. This real-time management proves efficient in improving the performance of the system (Dessouky et al., 1999; Fu et al., 2003) and on-line planning is an important path for research. In addition, tracking technology provides accurate data on travel times, stopping times, and passenger counts, that can help in securing feasible timetables in terms of punctuality and occupancy.

7.2.3. *Integrating more practical guidelines*

Model adaptability - Many studies focus on theoretical problems, which greatly contribute to the performance of the solutions proposed. However, the exposed models tend to lack the flexibility required to match the operators needs. More practical guidelines could be integrated to improve the optimization system effi-

ciency and applicability.

Stochasticity - Stochasticity in the arrival time of buses and in run times is more realistic than the deterministic way and can help securing transfer possibilities and minimizing passengers transfer time. This point can be worthy of deeper studies.

Passenger assignment - The passenger trip assignment model is important in many ways. Single paths and multiple paths assignments involve different levels of realism. The model definition and moment of use in the process can greatly influence the objective function and planning results. The usage of this model is very much linked to demand responsiveness.

Demand responsiveness - Despite the limitation commented upon in 4.1, most analytical tools are based on equilibrium models. Different approaches can be tested to overcome this limitation. Some tools make use of estimation models to forecast the demand elasticities, but elaborating an accurate model is an arduous task. Forecasting methods, such as trend extrapolation and regression, are seen to be too dependent on a projection of the past into the future to be useful for anticipating changes. Furthermore, they only suggest a single view of the future. In addition, a forecasting model depends so highly on time and space, competitors, mode, sample size, date of data collection, that they can hardly be re-used from network to network. Nonetheless, in order to improve the results obtained through these methods and better forecast trends in the demand, it is desirable to better understand, model and integrate all the components that can play a role in behavioral evolution, in particular the processes related to the transmission of habits, cultural values and travel patterns among individuals (Goodwin, 1999).

7.2.4. Tailoring more efficient solution methods

Process Integration - As we mentioned earlier, the transit planning process is traditionally a sequential one. The integration of consecutive steps is a predominant path to be explored in future research. Compromises between tractability and problems integration are among the major points of interest in transit planning.

Solution Methods - Progress has been made in commercial solvers for mathematical resolution. The development of metaheuristic methods has also made it possible to tackle large size problems more efficiently. Their flexibility enables them to adapt to any type of constraints and objectives and be applied to almost any transit planning problem. Hybrid methods also hold a lot of promises in terms of tackling problems that used to be intractable. Progress in solution methods is obviously the gate to the integration of all the previously mentioned interesting paths.

8. Conclusion

In this review paper, we have presented a non-exhaustive classification and analysis of studies on public transport strategic and tactical planning. 69 reviewed references are quoted in the bibliography and provide an extensive overview of the advances in the field of transit network design and scheduling.

A terminology proposal for describing the transit network design and scheduling problems was made and has contributed in structuring the literature review. Indeed, the studies have been grouped first by problem tackled and then by resolution method. Both theoretical and practical approaches have been considered and a descriptive analysis has been provided for each of them.

This approach differs from most of the reviews we have read on the subject in that it considers and classifies all the works related to the determination of the data necessary to the transit users (namely the routes network, the frequencies and timetables) whatever the optimization methods used.

We have put forward that the global problem is computationally intractable and can hardly be tackled at once, thus preventing to guarantee overall optimality. There are basically two ways to undertake this complex problem: innovative resolution methods and pertinent problem subdivision. Features such as decision variables, objective function, constraints and assumptions must then be carefully chosen.

A selection of benchmark references was also established. However, most of them consist of basic data and can mostly be used for algorithmic comparison purpose. For real-world applications, it would be preferable to get more detailed data, for instance matrices depending on uniform demand periods.

Finally we have highlighted a sample of paths that represent challenges in public transit planning systems. Indeed, interaction between researchers and transportation companies will continue to improve the efficiency of the methods and thereby the profits of the companies as well as the quality of service to the users.

Acknowledgements

This work was partially supported by the French Ministry for Research and Education through a CIFRE contract (number 400/2005). We thank Phil Goodwin and the reviewers for their insightful comments and suggestions that contributed to improve this paper.

References

- Baaj, M. H., Mahmassani, H. S., 1991. An AI based approach for transit route system planning and design. *Journal of Advanced Transportation* 25 (2), 187–210.
- Baaj, M. H., Mahmassani, H. S., 1995. Hybrid route generation heuristic algorithm for the design of transit networks. *Transportation Research C* 3, 31–50.
- Bachelet, B., Yon, L., 2005. Enhancing theoretical optimization solutions by coupling with simulation. In *Proceedings of the 1st OICMS, Clermont-Ferrand, France*, 331–342.
- Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. *The demand for public transport: a practical guide*. TRL Report 593.
- Barra, A., Carvalho, L., Teypaz, N., Cung, V.D., Balassiano, R. 2007. Solving the Transit Network Design problem with Constraint Programming. In *e-Proceedings of the 11th World Conference in Transport Research*, June 24-28. University of California, Berkeley, USA.
- Bel, G., Dubois, D., Llibre, M., 1979. A set of methods in transportation network analysis and synthesis. In: *Int. J. of Operational Research Society* 30. 797–808.
- Benn, H. P., 1995. *Bus route evaluation standards*. Tech. rep., Transportation Research Board, Washington.
- Bielli, M., Caramia, M., Carotenuto, P., 2002. Genetic algorithms in bus network optimization. *Transportation Research C* 10 (1), 19–34.
- Bookbinder, J. H., Désilets, A., 1992. Transfer optimization in a transit network. *Transportation Science* 26, 106–118.
- Borndörfer, R., Grötschel, M., Pfetsch, M. E., 2005. *A path-based model for line planning in public transport*. Tech. Rep. Report 05-18, ZIB.
- Bunte, S., Kliewer, N., Suhl, L., 2006. An overview on vehicle scheduling models in public transport. In *Proceedings of the 10th International Conference on Computer-Aided Scheduling of Public Transport*, Leeds UK, 2006. Springer Verlag.
- Bussieck, M. R., 1998. *Optimal lines in public rail transport*. Ph.D. thesis, TU Braunschweig.
- Carrese, S., Gori, S., 2002. An urban bus network design procedure. *Applied Optimization* 64, 177–196.
- Castelli, L., Pesenti, R., Ukovich, W., 2004. Scheduling multimodal transportation systems. *European Journal of Operational Research* 155, 603–615.
- Ceder, A., 1984. Bus frequency determination using passenger count data. *Transportation Research A* 18, 469–453.
- Ceder, A., 2003a. *Designing public transport network and routes*. Pergamon Imprint, Elsevier Science Ltd, 59–91.
- Ceder, A., 2003b. *Public transport timetabling and vehicle scheduling*. Pergamon Imprint, Elsevier Science Ltd, 31–57.
- Ceder, A., Golany, B., Tal, O., 2001. Creating bus timetables with maximal synchronization. *Transportation Research A* 35, 913–928.

- Ceder, A., Tal, O., 1999. Timetable synchronization for buses. *Computer-Aided Scheduling of Public Transport*, Lecture Notes in Economics 471, Springer-Verlag Pub. N.H.M.Wilson (ed), 245–258.
- Ceder, A., Wilson, N. H. M., 1986. Bus network design. *Transportation Research B* 20, 331–344.
- Cevallos, F., Zhao, F., 2006. Minimizing Transfer Times in a Public Transit Network with a Genetic Algorithm. *Transportation Research Record 1971*, Journal of the Transportation Research Board, National Research Council, Washington, D.C., 74–79.
- Chakroborty, P., 2003. Genetic algorithms for optimal urban transit network design. *Journal of Computer Aided Civil and Infrastructure Engineering* 18, 184–200.
- Chakroborty, P., Deb, K., Porwal, H., 1997. A genetic algorithm based procedure for optimal transit systems scheduling. In *Proceedings of the 5th International Conference on Computers in Urban Planning and Urban Management*. Mumbai, India, 330–341.
- Chakroborty, P., Deb, K., Sharma, R. K., 2001. Optimal fleet size distribution and scheduling of urban transit systems using genetic algorithms. *Transportation Planning and Technology* 24 (3), 209–226.
- Chakroborty, P., Deb, K., Subrahmanyam, P. S., 1995. Optimal scheduling of urban transit systems using genetic algorithms. *Journal of Transportation Engineering* 121 (6), 544–553.
- Chakroborty, P., Dwivedi, T., 2002. Optimal route network design for transit systems using genetic algorithms. *Engineering Optimization* 34 (1), 83–100.
- Chowdhury, S., Chien, S., 2001. Optimization of transfer coordination for intermodal transit networks. In: *80th Annual Meeting*, Transportation Research Board. Washington, DC, paper No 01-0205.
- Chua, T. A., 1984. The planning of urban bus routes and frequencies: A survey. *Transportation* 12 (2), 147–172.
- Constantin, I., Florian, M., 1995. Optimizing frequencies in a transit network: a nonlinear bi-level programming approach. *International Transactions in Operational Research* 2 (2), 149–164.
- Daduna, J. R., Voss, S., 1995. Practical experiences in schedule synchronization. *Lecture Notes in Economics and Mathematical Systems* 430, 39–55.
- Deb, K., Chakroborty, P., 1998. Time scheduling of transit systems with transfer considerations using genetic algorithms. *Evolutionary Computation* 6 (1), 1–24.
- Desaulniers, G., Hickman, M., 2007. Public transit. *Handbooks in Operation Research and Management Science*, 69–120.
- Dessouky, M. M., Hall, R., Nowroozi, A., Mourikas, K., 1999. Bus dispatching at timed transfer transit stations using bus tracking technology. *Transportation Research C* 7 (4), 187–209.
- Dhingra, S. L., Shrivastava, P., 1999. Modelling for coordinated bus train network. In *Proceedings of the 6th International Conference on Computers for Urban Planning and Urban Management (CUPUM99)*. Venice, Italy.
- Ellis, C., Silva, E., 1998. British Bus Deregulation: Competition and Demand Coordination. *Journal of Urban Economics* 43, 336–361.
- Eranki, A., 2004. A model to create bus timetables to attain maximum synchronization considering waiting times at transfer stops. Master’s thesis, University of South Florida.
- Everett S., 2006. Deregulation and reform of rail in Australia: Some emerging constraints. *Transport Policy* 13, 74–84.
- Fan, W., Machemehl, R., 2004. Optimal transit route network design problem: Algorithms, implementations, and numerical results. Tech. Rep. SWUTC/04/167244-1, Center for Transportation Research, University of Texas.
- Fan, W., Machemehl, R., 2006a. Optimal transit route network design problem with variable transit demand: Genetic algorithm approach. *Journal of Transportation Engineering* 132 (1), 40–51.
- Fan, W., Machemehl, R., 2006b. Using a simulated annealing algorithm to solve the transit route network design problem. *Journal of Transportation Engineering* 132 (2), 122–132.
- Fleurent, C., Lessard, R., Seguin, L., 2004. Transit timetable synchronization: Evaluation and optimization. In *Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT)*. San Diego, California.
- Friedrich, M., Haupt, T., Nökel, K., 1999. Planning and analyzing transit networks - an Integrated Approach Regarding Requirements of Passengers and Operators. *Journal of Public Transportation* 2(4), 19–39.

- Fu, L., Liu, Q., Calamai, P., 2003. Real-time optimization model for dynamic scheduling of transit operations. 82nd Annual Meeting of the Transportation Research Board. Washington, DC. 48–55.
- Furth, P. G., Wilson, N. H. M., 1982. Setting frequencies on bus routes: Theory and practice. *Transportation Research Record* 818, 1–7.
- Fusco, G., Gori, S., Petrelli, M., 2002. An Heuristic Transit Network Design Algorithm for Medium Size Towns. In *Proceedings of the 13th Mini-EURO Conference*, Bari, Italy.
- Fusco, G., Gori, S., Petrelli, M., 2001. Un modello di progetto di rete di trasporto pubblico per le città di medie dimensioni. In *Proceedings of IV Convegno annuale "Metodi e Tecnologie dell'Ingegneria dei Trasporti"*. Reggio Calabria, Facoltà di Ingegneria, Italy.
- Gao, Z., Sun, H., Shan, L., 2003. A continuous equilibrium network design model and algorithm for transit systems. *Transportation Research B* 38, 235–250.
- Goodwin, P., 1999. Transformation of transport policy in Great Britain. *Transportation Research A* 33, 655–669.
- Guan, J. F., Yang, H., Wirasinghe, S. C., 2003. Simultaneous optimization of transit line configuration and passenger line assignment. *Transportation Research B* 40(10), 885–902.
- Han, A. F., Wilson, N., 1982. The allocation of buses in heavily utilized networks with overlapping routes. *Transportation Research B* 16, 221–232.
- Hasselström, D., 1979. A method for optimization of urban bus route networks. Tech. rep., Volvo Bus Corporation. Göteborg.
- Hasselström, D., 1981. Public transportation planning - a mathematical programming approach. Ph.D. thesis, University of Göteborg, Sweden.
- Jansen, L. N., Pedersen, M. B., Nielsen, O. A., 2002. Minimizing passenger transfer times in public transport timetables. In *Proceedings of the 7th Conference of the Hong Kong Society for Transportation Studies : Transportation in the information age*. Hong Kong, 229–239.
- Kidwai, F. A., 1998. Optimal design of bus transit network: a genetic algorithm based approach. Ph.D. thesis, Indian Institute of Technology Kanpur, India.
- Kim, D., Barnhart, C., 1999. Transportation service network design: Models and algorithms. *Lectures notes in Economics and Mathematical Systems* 471, *Proceeding of the 7th International Workshop on Computer-Aided Scheduling of public transport*. Springer Verlag, 259–283.
- Klemt, W. D., Stemme, W., 1988. Schedule synchronization for public transit networks. *Computer-Aided Transit Scheduling*, *Proceedings of the 4th International Workshop on Computer-Aided Scheduling of Public Transport*. Springer Verlag, Hamburg, Germany, 327–335.
- Knoppers, P., Muller, T., 1995. Optimized transfer opportunities in public transport. *Transportation Science* 29 (1), 101–105.
- Koutsopoulos, H. N., Odoni, A., Wilson, N. H. M., 1985. Determination of headways as function of time varying characteristics on a transit network. *North-Holland, Amsterdam*, 391–414.
- Lampkin, W., Saalmans, P. D., 1967. The design of routes, service frequencies, and schedules for a municipal bus undertaking: A case study. *Operational Research Quarterly* 18 (4), 375–397.
- LeBlanc, L., Morlok, E., Pierskalla, W., 1975. An efficient approach to solving the road network equilibrium traffic assignment problem. *Transportation research* 9 (5), 309–318.
- Lee, Y. J., Vuchic, V. R., 2005. Transit network design with variable demand. *Journal of Transportation Engineering* 131 (1), 1–10.
- Magnanti, T. L., Wong, R. T., Feb. 1984. Network design and transportation planning: models and algorithms. *Transportation Science* 18 (1), 1–55.
- Mandl, C. E., 1979. Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research* 5, 396–404.
- Mandl, C. E., 1979. *Applied Network Optimization*. London, Academic Press.
- Murray, A. T., 2003. A coverage model for improving public transit system accessibility and expanding access. *Annals of Operations Research* 123, 143–156.
- Murray, A. T., Davis, R., Stimson, R. J., Ferreira, L., 1998. Public transportation access. *Transportation Research D* 3, 319–328.
- Ngamchai, S., Lovell, D., 2003. Optimal time transfer in bus transit route network design using a genetic

- algorithm. *Journal of Transportation Engineering* 129 (5), 510–521.
- De Palma, A., Lindsey, P., 2001. Optimal timetables for public transportation. *Transportation Research B* 35 (8), 789–813.
- Pape, U., Reinecke, E., Reinecke, Y., 1992. Entwurf und implementierung eines linienplanungssystems für den busverkehr im pnv unter einer objektorientierten grafischen entwicklungs-umgebung. Gruppendifferenzarbeit.
- Park, S. J., 2005. Bus network scheduling with genetic algorithms and simulation. Master’s thesis, University of Maryland.
- Pattanaik, S. B., Mohan, S., Tom, V. M., 1998. Urban bus transit route network design using genetic algorithm. *Journal of Transportation Engineering* 124 (4), 368–375.
- Patz, A., 1925. Die richtige Auswahl von Verkehrslinien bei großen Straßenbahnnetzen. *Verkehrstechnik* 50/51.
- Peng, Fan, 2004. A decision support system for the design of urban inter-modal public transit network. In *Proceedings of Codatu XI*. Bucarest.
- Potter, S., Enoch, M., 1997. Regulating transport’s environmental impacts in a deregulating world. *Transportation Research D* 2(4), 271–282.
- Quak, C. B., 2003. Bus line planning. Master’s thesis, Delft University of technology, The Netherlands.
- Rapp, M. H., Gehner, C. D., 1976. Transfer optimization in an interactive graphic system for transit planning. *Transportation Research Record* 619, 27–33.
- Rea, J. C., 1972. Designing urban transit systems: an approach to the route-technology selection problem. *Highway Research Record* 417, 48–59.
- Reeves, C. R., 1993. *Modern Heuristic Techniques for Combinatorial Problems*. Blackwell Scientific Press, Oxford.
- Salzborn, F. J. M., 1972. Optimum bus scheduling. *Transportation Science* 6 (2), 137–148.
- Salzborn, F. J. M., 1980. Scheduling bus systems with interchanges. *Transportation Science* 14 (3), 211–220.
- Scheele, S., 1980. A supply model for public transit services. *Transportation Research B* 14, 133–146.
- Shih, M., Mahmassani, H. S., 1994. A design methodology for bus transit networks with coordinated operations. Tech. Rep. SWUTC/94/60016-1, Center for Transportation Research, University of Texas, Austin.
- Shih, M., Mahmassani, H. S., Baaj, M., 1998. A planning and design model for transit route networks with coordinated operations. *Transportation Research Record* 1623, 16–23.
- Silman, L. A., Barzily, Z., Passy, U., 1974. Planning the route system for urban buses. *Computers & OR* 1 (2), 201–211.
- Sonntag, H., 1977. Linienplanung im öffentlichen personennahverkehr. Ph.D. thesis, Technische Universität Berlin.
- Spasovic, L. N., Boile, M. P., Bladikas, A. K., 1993. A methodological framework for optimizing bus transit service coverage. 73rd Annual Meeting of the Transportation Research Board.
- Ting, C. J., Schonfeld, P., 2005. Schedule coordination in a multiple hub transit network. *Journal of Urban Planning and Development* 131 (2), 112–124.
- Tom, V. M., Mohan, S., 2003. Transit route network design using frequency coded genetic algorithm. *Journal of Transportation Engineering* 129 (2), 186–195.
- Van den Berg, J. H. A. V., Odijk, M. A., 1994. Dons: Computer aided design of regular service timetables. In *Proceedings of CompRail94*, Madrid, 109–116.
- Van Nes, R., Hamerslag, R., Immers, B. H., 1988. Design of public transport networks. *Transportation Research Record* 1202, 74–83.
- Van Nes, R., 2002. Design of multimodal transport networks, a hierarchical approach. TRAIL Thesis Series T2002/5, DUP, Delft University, The Netherlands.
- Wan, Q. K., Lo, H. K., 2003. A mixed integer formulation for multiple-route transit network design. *Journal of Mathematical Modelling and Algorithms*, 2(4), 299–308.
- White, P., Farrington, J., 1998. Bus and coach deregulation and privatization in Great Britain, with particular reference to Scotland. *Journal of Transport Geography* 6(2), 135–141.
- Wong, R. C. W., Leung, J. M. Y., 2004. Timetable synchronization for mass transit railway. In *Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT)*. San

Diego, California.

- Wren, A., Rousseau, J.M., 1993. Bus driver scheduling - an overview. In Daduna, J. R., Branco, I., and Paixao, J. M. P., editors, *Computer-Aided Transit Scheduling*, 173–187. Springer, Berlin, Germany.
- Xiong, Y., Schneider, J. B., 1993. Transportation network design using a cumulative algorithm and neural network. *Transportation Research Record* 1364, 37–44.
- Yan, S., Chen, H. L., 2002. A scheduling model and a solution algorithm for inter-city bus carriers. *Transportation Research A* 36, 805–825.
- Yu, B., Yang, Z., Cheng, C., Liu, C., 2005. Optimizing bus transit network with parallel ant colony algorithm. In *Proceedings of the Eastern Asia Society for Transportation Studies*. Vol.5, 374–389.
- Zhao, F., 2006b. Large-scale transit network optimization by minimizing user cost and transfers. *Journal of Public Transportation* 9(2), 107–129.
- Zhao, F., Gan, A., 2003. Optimization of transit network to minimize transfers. Tech. Rep. BD015-02, Florida Department of Transportation, Center for Transportation Research, Florida international University.
- Zhao, F., Ubaka, I., 2004. Transit network optimization- minimizing transfers and optimizing route directness. *Journal of Public Transportation* 7 (1), 67–82.
- Zhao, F., Zeng, X., 2006a. Simulated annealing-genetic algorithm for transit network optimization. *Journal of Computing in Civil Engineering* 20 (1), 57–68.