



HAL
open science

Solving the Transit Network Design problem with Constraint Programming

Alexandre Barra, Luis Carvalho, Nicolas Teypez, Van-Dat Cung, Ronaldo
Balassiano

► **To cite this version:**

Alexandre Barra, Luis Carvalho, Nicolas Teypez, Van-Dat Cung, Ronaldo Balassiano. Solving the Transit Network Design problem with Constraint Programming. 11th World Conference in Transport Research - WCTR 2007, Jun 2007, Berkeley, United States. hal-00189690

HAL Id: hal-00189690

<https://hal.science/hal-00189690v1>

Submitted on 21 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Solving the Transit Network Design problem with Constraint Programming

A. Barra^{1,2*}, L. Carvalho³, N. Teypaz⁴, V.-D. Cung⁴, R. Balassiano¹

¹*PET/COPPE, Universidade Federal do Rio de Janeiro*

²*Laboratoire PRISM-CNRS, Université de Versailles*

³*Division of Applied Mathematics, Brown University*

⁴*Laboratoire GILCO, Institut National Polytechnique de Grenoble*

Abstract

Transit network design (TND) is a complex combinatorial optimization problem. The goal is to obtain an optimized public transport network, taking into consideration main characteristics of the system, passenger's desires, budget constraints and several possible levels of service. The author proposes a comprehensive novel model, based on Constraint Satisfaction, with essential and complementary constraints. Using Constraint Programming (CP), some tests are made using the proposed model in an attempt to open new perspectives for solution alternatives. The tests show that the chosen CP package is apparently unable to process large instances, although it works well with small ones. Other CP packages and search strategies should be tried by CP specialists in order to comprehensively validate the conclusions of the present work.

Keywords: Public transport; Transit Network Design; Constraint Programming.

Acknowledgments

The first author thanks (i) the Brazilian Foundation CAPES of the Brazilian Ministry of Education for the sponsorship (*bolsa de doutorado sanduíche PDEE*) in France; (ii) the AOC team of the PRISM-CNRS Laboratory of the *Université de Versailles*, for the *séjour* of 19 months and a partial sponsorship after CAPES' funding of this research; (iii) the Transportation Department of the *Universidade Federal do Ceará* for Fortaleza's instances cession; (iv) the National Confederation of Industry – CNI (Brazil) for the funding to make it possible the paper presentation and the WCTR'07 organization staff for the partial scholarship for the conference attendance.

* Corresponding author. He is currently a policies Analyst at the National Confederation of Industry - CNI. *Mail address:* SBN, quadra 1, bloco B, 12th floor, 70041-902 Brasília, DF, BRAZIL. *Tel.:* +55 61 81511246. *Fax:* +55 61 33179456. *E-mail address:* alexandrebarra@gmail.com

Solving the TND problem with Constraint Programming

1. INTRODUCTION

From the pioneer work of Lampkin and Saalmans (1967) till nowadays, the problem of optimizing public transportation (transit) networks has been exhaustively studied. It continues to attract researchers all over the world, probably because of the impossibility to find “the” optimal solution for a real-scale network, and the difficulty of finding “near-to-optimum” solutions. The Transit Network Design (TND) is a combinatorial problem and it is classified as NP-hard.

The Transit Network Design (TND) is more complicated than the traditional Network Design Problem. Besides finding which arcs to be included in the network, TND involves arcs grouping in fixed routes, also determining the frequency of service of these routes (transit lines).

The result of the public transportation network design, then, should include a set of routes and their frequencies. Most commonly, the problem is formulated on a graph with nodes, links, and (subsequently) routes (Desaulniers and Hickman, 2003). Let $G = (N,A)$ be a graph with N the set of nodes and A the set of links, and R represents the set of routes. Nodes represent intersections (e.g., road intersections), but can also represent zone centroids where a geographic zone is represented by a single point (the centroid) and the bus stops, train stations etc. A link between nodes represents a particular mode of transport between nodes, and a route represents a sequence of nodes and links of a single mode.

'As input to this problem, the formulation typically assumes that there is an existing origin-destination (O-D) matrix, covering the demand between a set of nodes or zones, either on a daily basis or for a specific period within the day. Alternately, it is reasonable (although complicated) to assume that demand is endogenous and determined as an equilibrium problem, in which the flows are a function of the network design. With the origin-destination flows and an assignment of these flows to routes, the set of routes and their frequencies must be determined' (Desaulniers and Hickman, 2003).

As stated by other authors, the transit demand assignment plays an important role in the solution quality. More than that, we are convinced that it plays an essential role in the optimization process and directly affects the solution search, i.e. the design of the transit network.

Naturally transit planning professionals base upon their own experiences, with or without the assistance of a simulation software, to conceive new transit networks or to point out minor improvements in a network. For them, normally it is easier to adopt empirical strategies than to formulate the TND like a complex optimization problem. This is justified because this modeling process is not yet spread out since TND is deep mathematical problem, and not so appealing to the technical corps of transit agencies (Barra and Kawamoto, 1999).

If one takes the urban case, it is even more difficult to the transit specialist of a large-sized city to incorporate into his/her work all the complexity of the transit system without a rational method based on a combinatorial optimization model.

By the means of a simulation package, the transit planner can use his/her own

experience and knowledge to test possible configurations, to predict user's behaviors, and to assess tested alternatives, with a good interaction in the whole process. Nevertheless, some optimization methods can also ensure a good interactivity, with the additional advantage of not to be necessarily dependent on the specialist's subjectivity.

Using optimization approaches, not simulation ones, the planner can take, besides his/her own experience, the alternatives automatically found by the algorithm that can be accepted or not by him/her, on its whole or in some extent. These alternatives could be brought from solution spaces that human mind have not imagined, since our minds can be guided many times by our habits and existent conformations.

TND is a complex problem since a huge number of discrete variables and constraints must be considered for real large-sized cities. The actual demand and existent infrastructure characteristics should be considered. Both level of service parameters and cost/operational restrictions must be taken into consideration. In a more sophisticated model, one could also consider fare policy, e.g. lower fare for poverty reasons in poor areas. If one considers that infrastructure could be changed, making it possible new subway lines for instance, the problem becomes even more complicated.

The use of approximate methods is then justified by the impossibility of applying exact methods in real-scale transit network optimization. Purely exact methods can only deal with small networks and simplified demand data, as stated by Chakroborty (2003) and other researchers.

The authors propose constraint satisfaction model for the TND problem, which is tested using a constraint programming (CP) system, which seems to have not been applied before to TND, although it has been used on the traveling salesman problem, another combinatorial optimization problem, with success.

CP is at the same time a declarative programming language and an optimization technique, based normally on branch-and-bound search and, but not necessarily, heuristics. CP can assign passengers into routes simultaneously with the itinerary design process, which is impossible in techniques like Genetic Algorithm or Tabu Search.

In the proposed model, several levels of service parameters are considered like: minimum frequency, maximum number of transfers, etc., as well as operational restrictions like maximum routes length or minimum ridership per line.

The method allows expert's interaction, using his/her experience to enhance the results by changing variables and constraints parameters, but do not depend on his/her expertise.

The authors' initial intent was to apply this method in real data from Fortaleza, a Brazilian city with 2.5 million inhabitants. Four road networks were generated to be used as an input, and represent different levels of accuracy the road network of the city. Actual O-D matrix, bus and train lines could also be considered as input data to test the method. Unfortunately, the chosen CP system has demanded a huge computational time to process these original instances, so authors had had to reduce the instances complexity in order to obtain some results and to assess them.

Nevertheless, the results show that Constraint Programming is a promising technique since it considers both route and scheduling design together. Improvements must be made to, in one way, reduce computational time cost and in the other hand to include other constraints. Finally, CP's flexibility enables the use of the proposed method in other instances, with few modifications.

2. TND MODELING

In the Transit Network Design problem, the involved variables and constraints are, in great portion, formed by integer values. So models are generally of Mixed-Integer Programming – MIP, since there are also some linear variables.

If there was a linear problem, using Linear Programming it would be easier to build mathematical models. But for discrete problems, like TND, it isn't so evident to find a mathematical model that represents the reality with fidelity.

According to Wagner (1986), with sufficient ingeniousness, one can always imagine a non trivial integer programming representation of a combinatorial problem. Frequently, those integer problems are too big (i.e., there are a huge number of constraints and variables) and, than, formulation is of a negligible computational interest. But several of the algorithm techniques for integer programming problems can be directly applied to combinatorial models without been *a priori* converted in integer programming models.

In deed, some TND authors prefer to implement the solution technique without previously formulating a mathematical model. Teypez (2004) presents a MIP formulation, but also a solution method based on Tabu Search which has no math modeling *a priori*.

Several models have been presented in the literature for optimization combinatorial problems, but not everytime with a complete mathematical representation of all relations between discrete variables. That is not a problem itself, since it is possible to find 'good' solutions without mathematical modeling. Anyway, to represent the reality, some decisions must be made by the TND researcher, like the choice of the demand assignment model and data modeling.

2.1 Route Choice and the Demand Assignment Model

Given a Origin-Destination (O-D) matrix, i.e., a matrix with the origin node, the destination node, and the quantity of passengers desiring to travel between these two nodes (normally traffic zones' centroids) in that specific time period (24 h, morning peak-period, etc), planner must choose the demand assignment model to the resulting routes from the optimization process.

Generally, the assignment of O-D pairs to the resulting routes network is done in sequence, i.e., the demand assignment starts after the route set is found. In the present work, we show that Constraint Programming makes it possible the simultaneous search of the best routes and of the demand assignment to the best routes. Actually, best routes are found in accordance with the demand assignment, which probably is a novel way of solution search.

According to the chosen approach, based mainly on the users' behavior model, all O-D pairs are assigned to the transit route set. If the different forms of demand behavior modeling are combined with the different forms of network's arcs costs modeling, one can use the following classification (adapted from Mauttone, 2005): All or Nothing - AON), Stochastic Network Loading - SNL, Deterministic User Equilibrium - DUE, Stochastic User Equilibrium - SUE, Deterministic Dynamic Process - DDP and Stochastic Dynamic Process - SDP.

road network	assignment criteria	route choice	
		deterministic	stochastic
without congestion	transit loading	AON	SNL
with congestion	equilibrium	DUE	SUE
	dynamic	DDP or SDP	DDP or SDP

Figure 1: Assignment models classification

AON model type assigns all O-D pair demand to one route only, typically the one with the best travel time. SNL model type distributes the demand among the route set, generally in function of each route's frequency. When network costs, i.e., travel time on board depends on lines' ridership (passenger flows), assignment models are DUE or SUE. DUE type is solved applying variables inequalities and fixed-point. SUE type uses a fixed-point approach to be solved, and allows congestion effects modeling in onboard travel time as well as in effective service frequencies. DDP and SDP types include attribute learning process, which allow to update passengers' option mechanisms.

For further considerations on route choice, traffic assignment or transit demand assignment, one can see Bovy and Stern (1990); Bell (1995); Aragón and Leal (1999; 2003).

2.2 Network Modeling

The road/rail system representation, formed by roads, railroads, intersections, natural barriers, etc must be constructed in the most possible reliable way in accordance with an existent road/rail system. Flow directions, as well as links capacity and speed limits should be also preferably taken into consideration.

Other infrastructure requirements are the bus stops, subway stations, rail terminals. Traffic zones centroids must also be present and linked to transit nodes.

Besides that, there is congestion caused by transit routes operation and loading, which is different from congestion caused by general traffic that transit is subjected in shared roads (low level of right-of-way). The transit congestion itself is caused, for instances, by the overtime spent on boarding operation or by over ridership, bringing it lower the vehicle speed. Or also, by not observing the minimum headway between vehicles, causing a vehicle convoy of the same transit line, with a loss to the normal operation (see Lam et al, 1999).

In most large-sized cities in the world, bus transit is one of the main transit modes, and its operations mostly occur in shared road ways. When there is some segregation level, like in bus corridors, congestion effect is mitigated. Anyway, even bus corridors or subway lines (the highest level of right-of-way) have capacity restrictions, i.e., there is a limit to the maximum quantity of vehicles that can ride on that link in a chosen time period. So network flow models (see Ahuja et al., 1993) can represent congestion effect, neither on shared roads or in segregated ones.

It is not possible, obviously, to use simultaneously all these elements in a TND model. Each researcher picks up some of the most convenient constraints and variables, in accordance with his/her beliefs. If the proposed method and model succeeds, than the author can

2.3 Instances for TND and Present Challenges

Mandl's data, based on a Swiss road network, were presented in 1979 and have been used as a reference instance for TND community since then (Mandl, 1979; Marwah et al., 1984; Baaj and Mahmassani, 1991; Shih and Mahmassani, 1994; Pattnaik et al., 1998; Kidwai, 1998; Chakroborty, 2003; Tey paz, 2004; Mauttone, 2005). Unfortunately, this only instance is not sufficient, because it has 15 nodes and 21 links, while real-scale instances may have hundreds or even thousands of nodes and links.

In the other hand, TND researchers have been using more complex models and larger instances, in an attempt to increase fidelity (like Zhao and Gan, 2003). To understand and to evaluate methods and models developed in the near past by these researchers, one must be able not only to reproduce their experiences, but to be sure too that the studied instances fulfill some minimum requirements that each model may require. Besides, the same evaluation algorithm must be used to guarantee direct comparison.

In the present days, there is a lack of standardization in both steps (instance definition and solution evaluation). This problem produces several inconveniences: the difficulty of results comparison, the impossibility of cooperation between different research teams that use different both instance and evaluation standards, and the impossibility of cooperation on instance evolution itself that could be produced if there was a common basis for instance definition.

The scientific community may need, in these days of great possibilities opened by increasing capacity of computer processing, to work together in cooperation to transgress actual models' limits.

Metropolitan cities around the planet are facing, with an increasing progress, traffic and environmental problems. More than ever, governments, transportation experts and citizens must work together to reduce congestions, accidents and pollution. Regional transportation is not much better than that either.

TND community must accept the actual challenge of producing more robust models, instances with higher fidelity to real-scale urban networks, and better solutions that could be really applied. If one succeeds, which means that proposed transit network increases its attractiveness, the benefits for suburban and urban populations of large-sized cities could be significant.

Mauttone (2005) says that there is no pattern instances¹ for the TNDP: '*most papers about the TNDP in which numerical results are presented, these are generated from instances based on fictitious data*'. The researcher claims that the difficulty of obtaining realistic instances lay on high costs of reliable data obtainment, specially the Origin-Destination matrix obtained from residences' poll. Anyway, since realistic instances are available to one team, TND community must effort to make it available to everyone.

In this sense, we are constructing a website (www.det.ufc.br/tndlib) which may be a meeting point to the TND community, the *TND Lib*. In the *TND Lib* website, one will find all instances proposed by all authors who want to have their instances published, for free. The idea is to follow good experiences, sharing data as well as research results.

¹ Given the literature, probably Barra et al. (2005) were the pioneer to propose standardization of instances as well as evaluation functions to the Transit Network Design problem.

With the present work regarding on CP opened perspectives, and the *TND Lib*, we expect to improve the scientific cooperation between TND researchers, since it will allow a common basis for the community of operations research, transportation engineering, maths and economics to plan regional and urban transit systems.

2.4 Solutions in TND

A solution of the TND problem is ideally formed by at least two main components: (i) the routes set; and (ii) the paths set, as explained below.

The routes set is composed by all routes found in the feasible solution, and each transit route/line is composed by its itinerary (sequence of nodes or sequence of arcs from the initial to the final ones) and by its frequency (number of departures in the chosen period of time).

The paths set is composed by the path that the demand of each O-D pair takes in the transit network, i.e. the sequence of transit lines the demand of the O-D pair takes to have its travel desires satisfied. Obviously, there may be paths formed by a segment of a single route (direct travel). The path can also be formed by walking paths, e.g. from residence to the initial bus stop, in a large transfer station, and from the final bus stop to office/school.

3. THE CONSTRAINT SATISFACTION MODEL

The model proposed was inspired in several models presented in the literature, and it takes into consideration also the authors' practical experience in their professional fields (consulting, teaching, research) to the variables and constraints choice. It differs from others since there is no objective function, it is based on the concept of constraint satisfaction².

The way of presenting the model formulation, from a systematic view, was inspired by Fan and Machemehl (2006a,b).

Sets/Indices

$i, j \in N = \text{nodes}$; $r_k \in R = \text{transit routes set}$; $p_o \in P = \text{paths set}$

Input Data

$nVertices$ = the number of vertices in the instance; $nArcs$ = the number of arcs in the instance; A [orig, dest, DistCost, Speed] = the description of each arc, with origin node, destination node, distance between nodes and speed at the arc; $nRoutes$ = number of transit routes in the solution ($k = 1 \dots nRoutes$); $minRouteCost$ = minimum route length; $maxRouteCost$ = maximum route length; $minRouteRidership$ = minimum route ridership; $maxRouteCap$ = maximum route capacity; $minH$ = minimum headway; $maxH$ = maximum headway; $maxTransfers$ = maximum number of transfers; $fcoef$ = coefficient to enable some variation in the minimum fleet demanded by the solution; $existFl$ = existent operational fleet in the actual transit network; $maxPathCost$ =

² Constraint satisfaction problems are typically identified with problems based on constraints on a finite domain, and to be coded it is recommended a Constraint Programming language. Such problems are usually solved via search, in particular a form of backtracking or local search.

maximum travel time; $maxTotalPathCost$ = maximum travel time in the whole transit network; $odPairs$ {orig, dest, flow} = O-D matrix, with origin and destination nodes, and the demand; $minSpeed$ = minimum speed of each O-D pair in the transit network; CTU = capacity of the transit unit.

Given a connected directed graph $G = (N, A)$ consisted by a finite set of N nodes and A arcs which connect pairs of nodes, find a solution formed by a transit network (routes set) and a paths set, subject to:

(a₁ and a₂) $\forall r_k \in \mathbb{R}, minRouteCost \leq RouteCost_k \leq maxRouteCost$ (route length is limited on its minimum and maximum values);

(b) $\forall r_k \in \mathbb{R}, RouteRidership_k \leq maxRouteCap$ (the daily ridership in each route must be equal or less than the capacity of the line $maxRouteCap = maxF_k \times CTU_{kv}$), while $maxF_k$ is the maximum frequency of line k (see below) and CTU_{kv} is the capacity of the transit unit v (each vehicle or rail composition of route k used in the considered time period). Discrete;

(c) $\forall r_k \in \mathbb{R}, RouteRidership_k \geq minRouteRidership$ (the daily ridership in each route must value at least $minRouteRidership$, in order to have financial feasibility). Discrete;

(d) $\forall r_k \in \mathbb{R}, H_k \geq minH$ (headway H of each line k is at least $minH$ of the transport mode). Discrete;

(e) $\forall r_k \in \mathbb{R}, H_k \leq maxH$ (headway H of each line k is limited by $maxH$). Discrete;

(f) $\forall p_o \in P, \#Transfers_o \leq maxTransfers$ (number of transfers $\#Transfers$ of each O-D pair o is limited). Discrete;

(g) $\forall p_o \in P, PathCost_o \leq maxPathCost$ (total travel time of each O-D pair o is limited);

(h) $\forall p_o \in P, Speed_o \geq minSpeed$ (speed on transit network of each O-D pair o obeys a minimum speed), $Speed_o = \frac{DistCost_o}{TimeCost_o}$, $DistCost$ is the distance cost of pair o , and

$TimeCost$ is the total travel time of pair o ;

(i) $TotalPathCost \leq maxTotalPathCost$ (the daily traveled through kilometers in the transit network $TotalPathCost$ is limited),

where $TotalPathCost = \sum_{k \in K} RouteCost_k \times DayTrips_k$; $DayTrips_k = \sum_p Dep_{kp}$ (the daily

number of trips $DayTrips$ of line k is the sum of departures in one day (24 h).

(j) $TotalNeedFl \leq fcoef \times existFl$ (this constraint enables some variation in the minimum fleet demanded by the solution, and $fcoef > 0$ and its value is near 1; where $TotalNeedFl = \sum_{k \in K} needFl_k$ and $needFl$ is the necessary fleet to operate the feasible transit

route found). Discrete.

Some input data can be fixed to all $nRoutes$, like $maxRouteCap$, CTU or F .

The headway H is the time (generally in minutes) between two departures (the time gap). $minH$ is related to technical restrictions of each transport mode, and it depends on the used controlling technology, segregation level from general traffic and congestion influence. If $H < minH$, it is likely to happen accidents (e.g. a two trains collision) or at least operational problems (like two buses of the same route forming a not planned convoy). In the contrary, $maxH$ is related to the desired level of service.

The frequency F is the number of vehicles per hour in one route, $F_k = 60 / H_k$ and $maxF = 60 / minH$.

To calculate $needFl$, it is necessary to know the total travel time in a round trip of route k in the chosen time period, i.e. the cycle time T_c . One must take the peak period, and how many minutes it takes, $Period$:

$$needFl = \frac{T_c}{H}, H = \frac{Period}{Dep},$$

$$Dep = \frac{Pass_{ps}}{RF \times CTU}, RF = \frac{Pass_{ps}}{Pass_{cr}}, \text{ then } Dep = \frac{Pass_{cr}}{CTU}$$

and finally $needFl = \frac{T_c \times Pass_{cr}}{Period \times CTU}$

where:

$Period$ = considered period [minutes]; Dep = necessary number of departures in the period per direction to satisfy the demand; $Pass_{ps}$ = number of passengers per period per direction; CTU = capacity of the transit unit/vehicle (seats + maximum number of on foot passengers); RF = renewal factor; $Pass_{cr}$ = number of passengers per period per direction in the critical session (critical ridership).

$PathCost_o$ is the total time cost of each pair p_o , i.e., it depends on the path, and it values the sum of $WaitTime_o$, $AboardTime_o$ and eventually $TransferPenal_o$.

$TotalPathCost$ is calculated summing the total travel time (or total distance) of every trip made by each route, along the day, including transfer penalties $TransferPenal_o$, if any.

$Speed_o$ is a way to weight the time cost according to the traveled distance, i.e. it is the average speed of each O-D pair o .

$RouteRidership$ must attain at least $minRouteRidership$ in the whole day in order to have financial feasibility. In some cities or regions, this feasibility can be reach by delimited areas (that can be called transit *basins*, like a river basin), instead of by single routes. In those locations, $minRouteRidership$ must be replaced by a minimum ridership of each basin, i.e. all routes in the delimited area.

The upper bound ($maxRouteCost$) and the lower one ($minRouteCost$) of the route length ($RouteCost_k$) should be defined by the specialist according to technical restrictions and particularities of the city or region.

The proposed constraints were classified in accordance with its importance (essential or complementary) and the main reason of its choice (for technical/operational reasons, for a better level of service or for a lower operational cost), like the following:

importance	constraint	level of service	technical	cost	
essential	a1 and a2		X		
	b		X		
	d		X		
	f	X			
	g	X			
	i				X
	e	X			
complementary	h	X			
	c			X	
	j			X	

Figure 2: Classification of the proposed constraints

So the authors have selected 11 constraints (since constraint ‘a’ is actually two grouped constraints), being 7 essentials, 4 complementaries; being 3 due to technical reasons, 4 regarding the level of service and 3 the operational (fixed and variable) costs.

To know which of the feasible solutions found is the ‘best’ one, it can be used one or more evaluation functions, like the functions proposed by Barra et al. (2005).

4. USING CONSTRAINT PROGRAMMING TO THE TND

Transit Network Design methods generally work in a bi-level process. Generally, there is an optimization level, in which is made the search for the optimal (or sub-optimal) routes set; and a simulation level, in which is made the transit demand assignment. Simulation, also known as analytic method, is also used in models involving network flow (like in Loureiro, 1994).

Only after the demand assignment one can assess the quality of the solution found. The explanation for this assertion is based on the fact that, if in the first level all itineraries of the routes are already known, only after the complete transit demand assignment one knows the routes’ frequencies, or the necessary fleet. So, only after the demand assignment process it is achieved the complete solution of the problem, assessed by one or more evaluation functions.

Using optimization methods, also known as synthetic methods, the design of transit networks has been solved using heuristics as in Andréasson (1977), Mandl (1979), Baaj and Mahmassani (1991), Shih and Mahmassani (1994), Carmo et al. (2002) and Wan and Lo (2006); or meta-heuristics like in Pattnaik et al. (1998), Kidwai (1998), Dwivedi (2000), Chakroborty and Dwivedi (2002), Teypaz (2004), Fan and Machemehl (2004, 2006a, 2006b) and Mauttone (2005; 2006).

4.1 Constraint Programming: a good alternative?

Besides heuristics and meta-heuristics, which other techniques could be used in for the Transit Network Design? This question has appeared naturally in our research, that involves other techniques like Tabu Search. After all, if by one side the scientific community must improve recognized efficient methods, the researchers must also try other alternatives that potentially could bring even better results.

Several researchers from India have been developing implementations based on Genetic Algorithm (GA), and have been reaching better results along the years. The second way was adopted by researchers that have tried other meta-heuristics than GA,

like Zhao and Gan (2003), Zhao and Ubaka (2004) and Teypaz (2004), with Tabu Search, or Mauttone (2005), with GRASP.

The Constraint Logic Programming on finite domains – CLP (FD) has already showed a good efficiency in many practical implementations of Combinatorial Optimization problems, specially due to its characteristics of dealing well with discrete problems. As examples, there have been applications to some logistic problems, like the choice of optimum stock level or the fleet scheduling.

The Constraint Programming has appeared in the beginning of 1960's decade, with the technique of "iterative graphs", but it was its joint with the Logic Programming back in 1987 that it has happened a great development in programming and in applications. Today the main available solvers for the Logic Constraint Programming (or simply Constraint Programming) are: GNU Prolog, CHIP, SICStus, ECLiPSe, Oz, Ilog Solver and Koalog Constraint Solver (Georget, 2002). The first three are based on declarative programming language Prolog. SICStus includes a Constraint Handling Rules (CHR), a constraint handler, which is a set of rules for simplification, propagation and "simpagation" (simultaneous simplification and propagation) of constraints. In all these languages, the programmer must define relation rules between variables, as well as possible domain for each one.

The declarative programming is a paradigm of the Logic Constraint Programming, or simply Constraint Programming (see Abdennadher, 2001; Abdennadher e Fruhwirth, 2004), where relations between variables can be stated in the form of constraints. It differs to procedural-based programming (like C/C++ or Pascal/Delphi) because it is not subjected to the need of defining all steps the machine must follow to execute the algorithm. Declarative programming is based on relations of objects defined by user. These objects are the variables and constraints of the model.

While constraints are pieces of syntax used to model real world behaviour, a constraint solver determines if a constraint has a solution. Basically, the Constraint Programming can deal with three kinds of constraints: arithmetic constraints, tree constraints and constraints in finite domains. Constraints in finite domains are largely used on constraint programming which involves choices, letting possible values to the variable to correspond to different choices (Marriot and Stuckey, 1998).

Constraint Programming was chosen to the implementation of the proposed Constraint Satisfaction Model by its semantic simplicity and a theoretical low computational cost, besides (as far as we know) it hasn't been used to solve TND problems. The intention is to open new perspectives, in an exploratory study.

4.2 The Choice of the Implementation Toolkit

To implement the method, it was chosen the toolkit ILOG Solver (Ilog) and the modeling tool OPL Studio (Ilog). Teypay and Cung (2005) have made some tests using the Mandl's data with OPL Studio but with another optimization engine, ILOG CPLEX, used also by Wan and Lo (2003).

The well known ILOG CPLEX (used to exact methods) is unable to handle large networks for the TND problem (see Teypaz and Cung, 2005).

The choice of ILOG OPL Studio was based on its simplicity to code the model, so any transport specialist could easily model the problem, even those who are not likely with mathematical modeling. So it is easier to access ILOG Solver (which is a C++

library) via ILOG OPL Studio, than directly in this engine or another one like GNU Prolog, since OPL Studio does not require good expertise in computers programming.

ILOG's systems devoted to optimization problems has been showed efficient in some cases (see for instance Halatsis et al., 1996), and since a full licence is available at GILCO Lab (www.gilco.inpg.fr), we have chosen to test them.

4.3 Computational Performance and Code Improvement

Testing the coded model, with several small instances derived from Mandl's network, OPL Studio has produced good results. Using a PC with Sempron 2800 processor, 515 MHz RAM, the system gives all possible solutions for these instances in up to 0.08 seconds in some cases.

Nevertheless, for bigger networks, of even for the modified Mandl's network (oriented graph with only 15 nodes and 42 arcs), it takes more than one whole week executing without finding any solution.

In an attempt to accelerate the processing execution, other constraints were added in the code, to observe its behavior, expecting that solutions could be found, and in a reasonable time. The new constraints, that could be redraw or reintroduced by the specialist before each run, are described as following.

The first new constraint is the determination, by the specialist, of the route cardinality, i.e. the number of arcs than some of the routes must take. It is similar to constraint 'a', but it is different since it doesn't matter the cost of the taken arc, just the number of arcs is fixed for some chosen routes.

The second added constraint is the mandatory initial arcs for some routes, i.e. the initial node and the direction to the second node are chosen by the specialist. The third and last extra constraint is the mandatory arc for a specific route or for any route, i.e. a determined arc (or more than one) is chosen to be a segment of a route, no matter if it is the initial arc, the final one or in other segment of the route.

Another try was the simplification of the model, using only the essential constraints of the Constraint Satisfaction model, but there was any practical advantage in the computing cost.

The results, however, have continued to be unsatisfactory, since this solution "guidance" takes away the meaning of the whole optimization process. Then the surplus in limiting the search space will reflect the specialist's desire to achieve one or another solution.

Nevertheless, similar constraints can be used by the specialist to fix, among existent transit routes, which of them may be maintained in the optimization process. This is positive considering punctual needs to improve transit offer quality, but not in a complete design of a transit network.

In the model coding, we had decided not conditioning the search starting from the existent transit network, if any. Several other approaches (GA, Tabu Search, etc) require an initial transit network, so the try was to test an optimization strategy without this restriction that could also "guide" the search, and/or to eventually maintain performances' vices of the current transit system. This could have played a role in the huge computational cost showed by our code to OPL Studio / ILOG Solver, since it must search and backtrack in a very large search space.

Anyway, it is positive to prove that it is possible to implement a TND model

without necessarily having an existent transit network to be optimized. Also, it is good to have the option, to the specialist, to choose or not some mandatory nodes or arcs for some routes. It is relevant to remember that some methods require from the specialist the initial and/or the final node of each route, like in Jerby and Ceder (2006) for instance.

Another positive statement is that the routes set and the paths set can be found simultaneously. Many implementations, however, treat them in sequence, as discussed before about the traditional work realized in the practice of transit planning agencies. The ideal, in our opinion, is to choose implementation tools that can treat both questions in a concurrent and simultaneous way, which is possible using a Constraint Programming language. That is an important step toward new challenges for the scientific community since they are mutually dependents, but it mustn't make the implementation impracticable.

4.4 Open Perspectives

Even with the modeling simplification and the added constraints it wasn't possible to reduce the search space in a way to obtain a 'good' solution in an appropriate computing time. Nevertheless, Constraint Programming allows that search heuristics to be introduced to 'accelerate' the possibility of finding solutions. This way wasn't explored yet.

A next would be then to refine the search strategy. This kind would be similar to a proof search in the solution tree, while the one that was explored regards a search in width in a tree which has been 'guided' by the simplification or by the added constraints.

To refine the search mode might reveal a better way than the adopted strategy. In this sense, the original Constraint Satisfaction model would be preserved, and Constraint Programming would allow easily the specification of several forms of search, which would amplify its utility in the TND problem's solution. Unfortunately, this direction wasn't the preferred in the chosen approach, since this approach seemed to be more promising.

Actually, both strategies can be considered as a pattern in numerical solutions of optimization problems. Given a initial point, one can execute the search in two ways: (i) the search length is fixed and the better direction is picked up (known as *trusted region search*), or (ii) the better direction is fixed and the search length is chosen (known as *line search*).

Constraint Programming shows itself as a promising tool to this kind of problem. CP is a language with smaller semantic gap, and it is in deed more adequate to optimization problems that can be solved by constraint satisfaction, specially in finite domains, without the need *a priori* of the minimization of an objective function.

To confirm or not this impressions in this exploratory study, many other tests should be made by the scientific community, including big instances after successful results with the small ones.

5. CONCLUSIONS AND RECOMMENDATIONS

Transit network design (TND) is a combinatorial optimization problem. To solve

it, the authors have presented a comprehensive novel model, based on Constraint Satisfaction, with essential and complementary constraints, eleven in total. These constraints take in to consideration main characteristics of the system, passenger's desires (represented by the Origin-Destination matrix), budget constraints (as total km daily traveled by all routes in the transit network, or necessary fleet to operate the resulting network) and several possible levels of service (varying the maximum headway or the minimum average speed and other characteristics in accordance with specialist choices).

Using a Constraint Programming toolkit, OPL Studio / ILOG Solver, some tests are made using the proposed model in an attempt to open new perspectives for solution alternatives. The tests show that the chosen CP toolkit is apparently unable to process large instances, although it works well with small ones. Other CP packages and search strategies should be tried by CP specialists in order to comprehensively validate the conclusions of the present work.

As concluding remarks, some assertions can be made:

- (i) other relevant aspects of the problem, besides the essential and complementary constraints, should be considered to make the model and the implementation closer to the reality. There is obviously a trade-off between fidelity and computational cost, as discussed before about the simultaneous search of routes set and paths set;
- (ii) other search strategies should be tried, in order to achieve non explored spaces, also reducing computational processing time;
- (iii) the interactivity of the specialist with the search should be improved, since in the present work the interactivity is limited on the choice of the input values. Ideally, it should be possible to the specialist to make some interference in the optimization process during its execution, as long as he/she identifies good possibilities (based on quantitative assessments and qualitative measures given during the process) and directions that would make it easier to the algorithm to find better solutions. Here there is another trade-off between the specialist positive aid based on objective reasons and the negative guidance to a desired solution based on subjective beliefs;
- (iv) the system must be more friendly to the specialist, in order to amplify the number of possible interested transit planners and public transport planning agencies. Specially, some effort should be done in the visualization of candidate routes and search space, using a Geographic Information System for instance.

So for future researches, it is recommended: (a) the test of other possible search strategies in Constraint Programming, using heuristics in addition to *branch-and-bound and backtracking*, to solution quick finding; (b) the test of other Constraint Programming toolkits, to be compared with ILOG Solver and OPL Studio; (c) the consideration of other constraints, regarding the fare policy for example.

It is also important a more dynamic scientific cooperation between different teams in different countries, maybe through the *TND Lib*, with a intense exchange of experiences to boost the development of this so complex and fascinating field.

REFERENCES

- Abdennadher, S. (2001) Rule-based constraint programming: theory and practice. *Habilitationsschrift*. Institut für Informatik, Universität München. July.
- Abdennadher, S.; Fruhwirth, T. (2004) Integration and optimization of rule-based constraint solvers. In Bruynooghe, M. (ed). *Logic Based Program Synthesis and Transformation - LOPSTR 2003*, Revised Selected Papers, LNCS. Springer.
- Andréasson, I. (1977) A method for the analysis of transit networks. *Advances in Operations Research*. pp. 1-8.
- Baaj, M. H. and Mahmassani, H. S. (1991) An AI-based approach for transit route system planning and design. *Journal of Advanced Transportation*, 25[2]:187-210.
- Barra, A. (2003) Rational method to the multimodal network design problem on urban public transport. *Présentation aux "Séminaires AMAP"*. PRISM-CNRS, Université de Versailles Saint-Quentin-en-Yvelines, Versailles, France. May 23th.
- Barra, A. and Kawamoto, E. (1999) Bus transit routing in Brazil: current practice and perspectives. *Proceedings of the International Conference -Modelling and Management in Transportation*. Poznan-Kraków : Euro Working Group on Transportation, v. 2. p. 225-230. Poland.
- Barra, A. and Kawamoto, E. (2000) Roteirização de ônibus urbano: escolha de um método apropriado às cidades brasileiras. *Proceedings of the XI PANAM - Panamerican Conference of Traffic & Transportation*. Gramado, 19 a 23 de novembro. Brazil.
- Barra, A.; Cung, V.-D.; Balassiano, R. and Teypaz, N. (2005) Standardization of instances and solution evaluations for the public transport network design problem. *Proceedings of the PLURIS 2005 - 1º Congresso Luso-Brasileiro para o Planejamento Urbano, Regional, Integrado e Sustentável*. 28-30/set. Universidade de São Paulo - São Carlos. Brazil.
- Barra, A. and Nassi, C. D. (2002) Public transport fare policy to reduce poverty in urban environment. *Proceedings of X CODATU: Urban Mobility For All. X COoperation pour le Développement et l'Amélioration des Transports Urbains et périurbains – CODATU*. Lome, Togo.
- Carmo, M. R. R.; Netto, P. O. B.; Portugal, L. S. (2002) Uma heurística interativa para a geração de caminhos em grafos com restrição de grau: aplicação ao projeto de sistemas metroviários. *Pesquisa Operacional*, 22[1]:09-36. São José dos Campos.
- Chakraborty, P. (2003) Genetic Algorithms for Optimal Urban Transit Network Design. *Computer-Aided Civil and Infrastructure Engineering*, 18:184-200.
- Chakraborty, P. and Dwivedi, T. (2002) Optimal Route Network Design for Transit Systems Using Genetic Algorithms. *Engineering Optimization*, 34(1):83-100.
- Desaulniers, G. and Hickman, M. D. (2003) *Public transit*. Report. Les Cahiers du GERAD, G-2003-77. Groupe d'études et de recherche en analyse des décisions. November. Available at: www.gerad.ca
- Dwivedi, T. (2000) A genetic algorithm based method for design of routes for urban transit networks. *Master's Thesis*. Department of Civil Engineering, Indian Institute of Technology, Kanpur, India.
- Fan, W. and Machemehl, R. B. (2004) *Optimal transit route network design problem: algorithms, implementations, and numerical results*. Report No. SWUTC/04/167244-1. University of Texas at Austin. May. 267 p.
- Fan, W. and Machemehl, R. B. (2006a) Optimal transit route network design problem with variable transit demand: genetic algorithm approach. *Journal of Transportation Engineering*, 132[1]:40-51. January.
- Fan, W. and Machemehl, R. B. (2006b) Using a Simulated Annealing algorithm to solve the transit route network design problem. *Journal of Transportation Engineering*, 132[2]:122-132. February.
- Georget, Y. (2002) *Une introduction à la programmation par contraintes*. Exposé. ENST, le 7 juin. 36 p.
- Halatsis, C.; Stamatopoulos, P.; K., Isambo et al. (1996) Crew scheduling based on constraint programming: the PARACHUTE experience. *Proceedings of the 3rd Hellenic-European Conference on Mathematics and Informatics HERMIS'96*, pages 424-431. Disponível em <http://citeseer.ist.psu.edu/halatsis96crew.html>
- Jerby, S. and Ceder, A. (2006) Optimal planning and design of radial bus routes. *Proceedings of the 10th International Conference on Computer-Aided Scheduling of Public Transport – CASPT 2006*. Leeds, UK, June 21-23.
- Kidwai, F. A. (1998) *Optimal design of bus transit network: a genetic algorithm based approach*. Ph.D. Dissertation. Indian Institute of Technology. Kanpur, India.
- Lampkin, W. and Saalmans, P. D. (1967) The design of routes, service frequency and schedules for a

- municipal bus undertaking: a case study. *Operational Research Quarterly*, 18(4):375-397. Dec.
- Loureiro, C. F. G. (1994) *Modeling investment options for multimodal transportation networks*. Ph.D. Dissertation. The University of Tennessee, Knoxville. Major Professor: Arun Chatterjee. August. 166 p.
- Mandl, C. E. (1979) *Applied network optimization*. Academic Press, New York.
- Marriot, K.; Stuckey, P. J. (1998) *Programming with constraints: an introduction*. The MIT Press, Cambridge, Massachusetts, USA. 467 p.
- Marwah, B. R.; Umrigar, F. S.; Pattnaik, S. B. (1984) Optimal design of bus routes and frequencies for Ahmedabad. *Transportation Research Record*, 994:41-47.
- Mauttone, A. (2005) *Optimización de recorridos y frecuencias en sistemas de transporte público urbano colectivo*. Tesis de Maestría en Informática. RT 05-09. Instituto de Computación - Facultad de Ingeniería - Universidad de la República. Montevideo, Uruguay. 177 p. Julio. / www.fing.edu.uy/inco/pedeciba/
- Mauttone, A. and Urquhart, M. E. (2006) A multi-objective metaheuristic approach for the transit network design problem. *Proceedings of the 10th International Conference on Computer-Aided Scheduling of Public Transport – CASPT 2006*. Leeds, UK, June 21-23.
- Pattnaik, 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. Jul./Aug.
- Shih, M.-C. and Mahmassani, H. S. (1994) *A design methodology for bus transit networks with coordinated operations*. Center for Transportation Research, University of Texas at Austin, EUA, 185 p. Aug.
- Teypez, N. (2004) *Étude sur la conception d'un réseau optimisé de transport public urban*. Mémoire (Master's in Operations Research, Combinatorial and Optimization). GILCO, Institut National Polytechnique de Grenoble. Under supervision of Van-Dat Cung and Alexandre Barra. June. France.
- Teypez, N. and Cung, V.-D. (2005) A linear model for the transit network design problem. *Proceedings of the EWG ECCO XVIII - 18th Conference of the European Chapter on Combinatorial Optimization*. Minsk, Belarus. p.72
- Wan, Q. K. and Lo, H. K. (2003) A mixed integer formulation for multiple-route transit network design. *Journal of Mathematical Modelling and Algorithms*, 2(4): 299-308.
- Wan, Q. K. and Lo, H. K. (2006) Congested multimodal transit network design. *Proceedings of the 10th International Conference on Computer-Aided Scheduling of Public Transport – CASPT 2006*. Leeds, UK, June 21-23.
- Zhao, F; Gan, A. (2003) *Optimization of transit network to minimize transfers*. Final Report. Contract no. BD-015-02. Miami. December.
- Zhao, F.; Ubaka, I. (2004) Transit Network Optimization – Minimizing Transfers and Optimizing Route Directness. *Journal of Public Transportation* 7(1):63-82.

Barra, A.; Carvalho, L., Teypaz, N.; Cung, V.-D.; Balassiano, R. (2007) Solving the Transit Network Design problem with Constraint Programming. *Proceedings of the 11th World Conference in Transport Research - WCTR 2007 - June 24-28 - University of California, Berkeley, USA.*

Cover Page

Alexandre Barra, D.Sc

*Alumnus, PET/COPPE, Universidade Federal do Rio de Janeiro
Former researcher, Laboratoire PRISM-CNRS, Université de Versailles Saint-Quentin-en-Yvelines
Analyst, National Confederation of Industry – CNI (Brazil)*

E-mail address: alexandrebarra@gmail.com

SBN, quadra 1, bloco B, 12th floor
70041-902 Brasília, DF, BRAZIL

Luis Carvalho, M.Sc

Doctorate candidate, Division of Applied Mathematics, Brown University

E-mail address: carvalho@dam.brown.edu

Division of Applied Mathematics, Brown University
182 George St, Box F
02912 Providence, RI, USA

Nicolas Teypaz, M.Sc

Doctorate candidate, GILCO lab., ENSGI – Institut National Polytechnique de Grenoble

E-mail address: teypaz@gilco.inpg.fr

Laboratoire GILCO, ENSGI-INPG
H building, Office H123
46, avenue Félix-Viallet
38031 Grenoble, FRANCE

Van-Dat Cung, D.Sc

Professor, GILCO lab., ENSGI – Institut National Polytechnique de Grenoble

E-mail address: cung@gilco.inpg.fr

Laboratoire GILCO, ENSGI-INPG
H building, Office H123
46, avenue Félix-Viallet
38031 Grenoble, FRANCE

Ronaldo Balassiano, PhD

Professor, PET/COPPE, Universidade Federal do Rio de Janeiro

E-mail address: ronaldo@pet.coppe.ufrj.br

Programa de Engenharia de Transportes – COPPE/UFRJ
Centro de Tecnologia Bloco H - Sala 106 Cidade Universitária
21949-900 Rio de Janeiro, RJ, BRAZIL