Computers & Operations Research 40 (2013) 3000–3009

Contents lists available at ScienceDirect

Computers & Operations Research

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

Rapid transit network design for optimal cost and origin–destination demand capture

Gabriel Gutiérrez-Jarpa [a], Carlos Obreque [b], Gilbert Laporte [c], Vladimir Marianov [d,*]

[a] School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
[b] Department of Industrial Engineering, Universidad del Bío-Bío, Concepción, Chile
[c] GERAD, CIRRELT and HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T 2A7
[d] Department of Electrical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

Available online 10 July 2013

Keywords:
Location
Coverage
Networks
Transit systems
Traffic capture

ABSTRACT

This paper proposes a tractable model for the design of a rapid transit system. Travel cost is minimized and traffic capture is maximized. The problem is modeled on an undirected graph and cast as an integer linear program. The idea is to build segments within broad corridors to connect some vertex sets. These segments can then be assembled into lines, at a later stage. The model is solved by branch-and-cut within the CPLEX framework. Tests conducted on data from Concepción, Chile, confirm the effectiveness of the proposed methodology.

1. Introduction

Many cities throughout the world have invested in recent years in the construction, expansion or modernization of rapid transit systems, such as metros, light trains, pre-metros, commuter trains and monorails. This is partly in response to increased traffic congestion and to the need to reduce carbon emissions. There is no clear definition of a rapid transit system, but these are commonly defined as networks that do not interfere with road or pedestrian traffic and are primarily designed to serve city needs. This definition therefore excludes bus lines and inter-city railway networks. There are around 400 light rail systems worldwide and 116 cities with metro systems [13], the most important of which, measured in terms of length or number of stations, are located in New York, Paris, Shanghai, Madrid and Seoul.

The design and construction of rapid transit systems are major endeavours requiring long-term planning as well as the involvement of several players such as urban planners, geologists, engineers, politicians and various interest groups. This process involves a large amount of uncertainty, particularly in the case of underground metros where unexpected delays are frequent. From a methodological point of view, the design problem is a large-scale often non-linear multi-objective problem which cannot be solved optimally unless major simplifications are made. As argued by Vuchic [35] and Laporte et al. [18], operational research tools can be put to use to suggest alternative designs among which the decision makers can choose, or to solve some well-defined subproblems. For surveys on rapid transit network design, see [20,22]. A survey of related problems can be found in [6]. Note that this paper is uniquely concerned with the design of rapid transit systems from a strategic point of view, whereas the Transit Network Design Problem (TNDP) [15], also related to these systems, is more concerned with operational issues such as frequencies on individual lines, rolling stock, and so on.

A frequently used objective in rapid transit network design is to maximize the covered population, i.e., population located close to the stations, rather than maximizing covered trips. This is operationalized by estimating the population living close to a potential station, irrespective of travel demand. Under this approach, important locations such as city centres, hospitals, universities, shopping malls, etc., must be assigned large populations, such as the number of people frequenting them during the day. Maximizing population coverage requires relatively few data and usually yields tractable models (see e.g., [28,10,26,27,9,18]). This is also the case with the literature on stop location [29,32,36]. In particular, Wu and Murray [36] propose a model for reducing the travel time in an existing system by decreasing the number of stops. The travel time is not weighted by the traffic. A second objective maximizes covered population, rather than origin–destination traffic, which is not taken into account. The population coverage objective also makes sense from a planning perspective. Since rapid transit systems are designed for the long term, they are located to cover dense population areas rather than focus on present day travel patterns. Once a system is in place, one can argue that people will tend to relocate over time in order to satisfy

* Corresponding author. Tel.: +562 23544974; fax: +562 25522563.
E-mail addresses: gabriel.gutierrez@ucv.cl (G. Gutiérrez-Jarpa), cobreque@ubiobio.cl (C. Obreque), gilbert.laporte@cirrelt.ca (G. Laporte), marianov@ing.puc.cl, vladimir.marianov@gmail.com (V. Marianov).

0305-0548/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cor.2013.06.013

their travel requirements, no matter what the shape of the system is. However, if the optimal layout of the lines is to be found, seeking a better traffic coverage is a must. Origin/destination (O/D) demand forecasts can be used, and these are typically available in the transportation sector.

In fact, whereas population coverage constitutes a sensible and easy to operationalize objective, it does not adequately reflect the aim of a rapid transit system which is to improve population's mobility. As a result, several authors have proposed models that explicitly maximize traffic coverage. One way to operationalize this concept is to first compute an O/D matrix and to design a system that will cover as much traffic as possible, subject to budgetary and operational constraints. This is what was done in Laporte et al. [24,23]. The first of these two papers proposes a heuristic for the design of a single-line in the absence of a competing mode and classifies the single line problem as NP-Hard, whereas the second paper models and solves a similar problem in the presence of competition, that is, users associated with a given O/D pair will opt for the fastest mode. Guan [11], Schöbel and Scholl [33], and Borndörfer et al. [2] have described algorithms to select a set of lines from a pool to connect several O/D pairs under a budget constraint. Gutiérrez-Jarpa et al. [12] proposed an integer programming formulation for the single-line traffic capture problem, extended here to multiple lines. The candidate lines under consideration take directness of travel into account.

Since the single-line problem is NP-Hard, so is the multiple-line problem. The scientific contribution of this paper is the introduction of a tractable model for the multiple-line problem, in which travel cost is minimized and traffic capture maximized, and a simple constraint relaxation mechanism is used, capable of solving it efficiently on a realistic data set. This is the first model that finds the exact shape of a set of lines (not one but several), minimizes cost and captures maximum origin–destination traffic (as opposed to capturing population living close to stops), optimally locates stops, and solves the problem exactly on instances of reasonable size (real data are used). Other researchers have not addressed this problem in such detail. We remark that knowing how many users will choose a transportation system is not possible without considering origin–destination traffic, as opposed to considering origins and destinations separately, as do most researchers in the field of metro location. Naturally, an exact method like the one proposed in the paper, if applicable, is better than a heuristic, as in all previous research on traffic capture.

As for most formulations for rapid transit network design, our mathematical model works on a graph in which some of the nodes are potential locations for stations. From practical and computational points of view, it makes sense to restrict the set of these locations. For example, broad corridors corresponding to heavy traffic flows can be predefined for some of the lines, leaving the determination of the specific station locations to the optimization process. These lines can be combined in several ways to yield various network topologies such as a star, a cartwheel or a triangle (Fig. 1). Laporte et al. [19] have analyzed these and other configurations from a generic standpoint and have developed a number of measures to assess them. One of their conclusions is that cartwheels and triangles are preferable to stars and grids in terms of directness and effectiveness when a uniform O/D distribution is assumed; if this is not the case, then their conclusions may not hold.

Bruno and Laporte [4] and Bruno et al. [3] have designed a heuristic and a decision support system enabling planners to design a rapid transit network possessing a given topological structure. The user first defines a configuration using a menu, and a tabu search heuristic is then applied to locate an alignment within each of the corridors specified by the user. Tests were successfully performed on data from the City of Milan. It is of course possible to replace the heuristic part of the system with an exact optimizer, provided the instance size is not too large. This is essentially the approach taken in this paper except that we simultaneously minimize cost and maximize travel demand coverage. We also include a median objective to the problem, i.e., we consider the cost for users of reaching their closest station.

The remainder of this paper is organized as follows. The problem just described is formally defined and modeled in Section 2. Computational results on data from the city of Concepción, Chile, are presented in Section 3, followed by conclusions in Section 4.

2. Formal problem definition

The rapid transit network design problem (RTNDP) is defined on an undirected graph $G(N,E)$, where $N=\{1,\dots,n\}$ is a node set and $E=\{(i,j):i,j\in N, i<j\}$ is an edge set. Let d_{ij} be the length of edge (i,j). The transit network consists of one or several *lines* with stations in N. In the latter case, the lines intersect at *intersection points* so that the transit network is connected and passengers can transfer from one line to another. Each line is made up of *segments* which are chains of edges connecting two intersection points of a line, or an intersection point with an end-point. The star configuration of Fig. 1, for example, contains six segments, one intersection point and six end-points. The cartwheel configuration is made up of twelve segments, five intersection points and four end-points. There are typically several ways of combining segments into lines [35,18]. For example, the star configuration of Fig. 1 (a) could operate as three intersecting lines or as six lines meeting at a common intersection point, among several possibilities. Similarly, the cartwheel configuration of Fig. 1(c) could have three lines (one of them circular), or more. The output of the model is a set of segments. Their combination into lines is beyond the scope of this study and is contingent upon considerations related to train scheduling, capacity, and passenger ride times, with or without transfers.

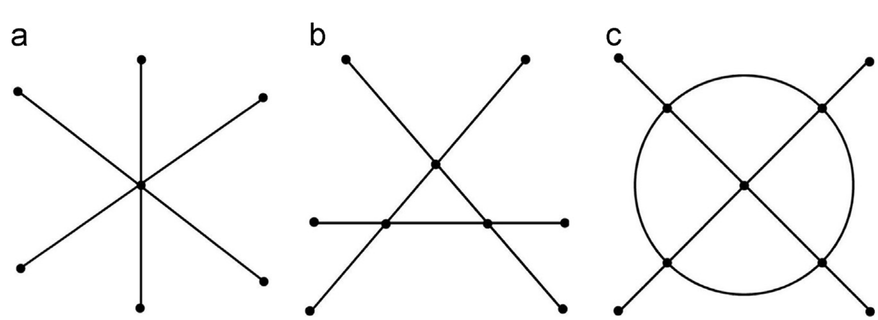

Fig. 1. Three basic configurations for a rapid transit system. (a) star, (b) triangle and (c) cartwheel.