Transportation (2006) 33: 347–370
DOI 10.1007/s11116-005-3507-7

A topological method to choose optimal solutions after solving the multi-criteria urban road network design problem

FRANCESCO RUSSO & ANTONINO VITETTA*
Department of Computer Science, Mathematics, Electronics and Transportation,
University of Reggio Calabria, Feo di Vito – 89100, Reggio Calabria, Italy
*(*Author for correspondence, E-mail: vitetta@unirc.it)*

Key words: assignment, cluster analysis, multi-criteria, network design

Abstract. The paper proposes and applies a method for systematically sorting and reducing the number of different possible solutions to a network design problem (NDP). This is achieved first by defining a topological similarity measurement and then by applying cluster analysis. The NDP can be derived from the scientific literature. In general, the method consists of some models and subsequent algorithms that generate different solutions (enumerative, branch and bound, genetic, expert panel, ...) and evaluate for each solution an objective function (with deterministic or stochastic network assignment and with elastic or inelastic demand). The NDP, mainly in urban areas, needs multi-criteria evaluation and in each case a large set of non-dominated solutions is generated. In this paper, in order to select solutions and identify latent optimal network layouts, cluster analysis is carried out. The methodology utilises a "cluster" formation in relation to the solution topology and a "best" (representative) solutions extraction in relation to the criteria values. It can be utilised after solving the existing multi-criteria NDP and in other network problems, where the best solutions (for global or local network layouts) are extracted (with respect to the network topology) from a large set. The method is applied in a test system and on different real networks in two Italian towns, in order to analyse the goodness of the solution algorithm and assess its possible application to different networks.

1. Introduction

In recent years, various studies have been carried out to define models and algorithms which allow optimal transport network configurations to be obtained. In most cases such models generate network configurations, in terms of topology and capacity, which aim to achieve specific objectives such as a reduction in overall travel time, congestion control, limitation of pollution, optimisation of investment resources, accident limitation and increasing safety and security. The most useful models and algorithms used for the design network can be aggregated in the classical field of network design problem (NDP) (Magnanti & Wong 1984).

The urban road network design (URND) problem is an idealised version of a problem often faced by agencies responsible for road management in urban areas. The problem is how best to design such systems in response to transportation demand. The URND generally consists in determining the optimal configuration of the elements of the urban network defined by topology and capacity. The considered elements are the direction of lanes (topology) and capacities of links in terms of number of lanes allocated for each direction and the regulation strategy at junctions (capacity).

Various models and algorithms have been proposed for network design. The NDP treated in the scientific literature can be classified by different characteristics. As regards the decision variable, the problem can be classified into discrete or continuous. In the former case, an optimal topological configuration can be obtained (network topology, lines route for transit system), while in the latter case an optimal configuration in terms of capacity (lane allocation, signal setting at junctions, frequency for transit system) is performed. As regards the solution approach the NDP can be carried out using: an optimisation approach or a simulation approach. In the former case, the complete representation of user behaviour is generally simplified, employing an optimal search for the solution to the formulated problem. In the latter, user behavioural hypotheses are consistently simulated, albeit generating approximately the optimal configurations to be analysed.

In general, the NDP and the solution algorithms deal distinctly with the problem of network design for road transport and for transit services. In the former case, an optimal configuration of links and intersection capacities is generated, while in the latter an optimal configuration in terms of topology and optimal frequency of transit lines is carried out. For transit (Cantarella & Sforza 1988; Florian & Costantin 1993) it is possible to distinguish the design of lines (Ceder & Wilson 1986) from that of optimal frequencies (Russo 1998). With regard to roads, design of the link directions and link capacities has been treated separately (Billheimer & Gray 1973; Foulds 1981; Boyce et al. 1988; Chen & Alfa 1991; Poorzahedy & Turnquist 1982; Meng & Yang 2002; Drezner & Wesolowsky, 2003) from signal setting at intersections (Cantarella & Sforza 1991).

A special issue of the *Journal Transportation Research* (vol. 35B, no. 1) has been dedicated to the NDP. Friesz and Shah (2001) proposed two classes of non-traditional models for the (dis)equilibrium network design problem and used them to describe research needed to advance the design of both static and dynamic networks. They introduced non-separable elastic transportation demands for the first time. Meng et al. (2001) proposed a locally convergent augmented Lagrangian method to solve the problem. The bilevel programming model is transferred into a single level optimisation problem by virtue of a marginal function tool.

In network design models different strategies are generally considered (system manager, user, non-user). For each strategy various criteria may be identified. The effects are not in general additive and it is thus useful to obtain solutions to the problem which respect the imposed constraints and which are not, with regard to the various effects and objectives, dominated by each other. One solution dominates another solution if all criteria values of the first solution are better than the corresponding criteria values of the second solution. These solutions may be obtained by multi-criteria analysis, which tends to optimise the single criterion. A multi-criteria planning problem is thus obtained. A system manager, for example, might have similar objectives to minimise (pollution, congestion) to those of road users and at the same time different objectives to minimise (pedestrian area) to road users' objectives. System managers and road users could (but usually do not) have the same objective criteria but the system manager aims to optimise total system indicators and not individual performance, unlike most users: the manager and the users are operating at different levels of information. Multi-criteria analysis generates a large number of non-dominated solutions which also contain the best solution for each criterion. The non-dominated yet non-best solutions for a single criterion are called mixed solutions. In relation to a weight associated to each criterion, a non-dominated solutions can be the optimal mono-criterion solution.

To select the best solutions in terms of global or local road layout, a method for nesting solutions has to be identified from some researchers and professionals. The method generally proposed for URND considers solution selection in relation to criteria expressed just with numerical values. This method, even if applicable in URND, gives arbitrary solutions because similar configurations in terms of numerical values could arise from different road layouts.

In this paper a method for systematically sorting and reducing the number of different possible solutions to a URND is proposed and applied. This is achieved first by defining a topological similarity measurement and then by applying cluster analysis. Finally, the best solutions are selected in relation to topological and criteria values, which appears to provide a sound and novel way of tackling the problem of multi-criteria evaluation. The proposed topological indicator is not the only one possible but it is particularly simple. It does not take account of link capacity or flow but could easily be extended to them.

Cluster analysis is a well-established way of sorting a number of items into groups according to an appropriate similarity measure. This is the first time, to our knowledge, that it has been applied to the network design problem and in general to a network problem in which the emphasis is on the topological cluster. The quoted papers propose different models and algorithms but none of them give methods to select good solutions in a topological way from many solutions. With three criteria in a medium-sized town in one of the applications proposed in this paper, the URND generates about 102 solutions. Thus each cluster defines a latent layout of the network, where the main roads are fixed (in direction and number of lanes) and the secondary links differ in topology and/or capacity. The configurations are called latent because they are not defined in only one non-dominated configuration but they have to be extracted from a large set of topologically similar non-dominated configurations.

In this paper we propose a method to select the optimal solutions from a large set of non-dominated solutions. One of the most recent techniques used for automatic solutions generation is the URND reported in Section 2. Section 3 describes the method proposed by the authors. The solutions are identified with topological cluster analysis, and are nested and classified in relation to topological similarity. The method uses "cluster" formation in relation to the solution topology and a "best" solution extraction in each cluster, in relation to the criteria values. The main steps in the procedure are the following: definition of similarity level (Section 3.1), cluster formation (Section 3.2) and selection of best solutions (Section 3.3). The methodology was tested on two different Italian towns (Acireale and Crotone, with populations of 50,000 and 100,000 respectively) to analyse its goodness and the possibility of extending the results to different networks (Section 4). Conclusions and further developments are reported in Section 5.

2. Existing models for the road urban network design problem

2.1. *General methodology*

There is no systematic methodology using URND, though various common elements can be highlighted in the method proposed in the literature, relating to the following:

(a) control variables;
(b) objective for each group of stakeholders (decision makers);
(c) traffic assignment; and
(d) solution generations.

A new point relative to solution selection is important for URND and is proposed in this paper in Section 3.

(a) In the URND the variables can be divided into state variables and control (or decisional) variables. Control variables can be divided into continuous variables relative to the junction setting and discrete variables relative to the link layout. The continuous variables at the network level for optimal signal setting are obtained within the equilibrium model (Cantarella & Sforza 1991). The discrete variables for optimal link layout can be generated with several methods that are of a heuristic type in real-size systems.

(b) In the URND three categories of objectives can be considered:

- Users seek to reduce their travel time, stress, congestion and pedestrian paths, and seek to maximise their utility and safety;
- System managers seek to reduce system management costs, maximise system utility (safety, security) and seek to reduce traffic incidents;
- Non-users seek to maximise pedestrian areas and, in general, quality of life and sustainable development.

(c) In the URND three main constraints are considered in representing user behaviour in path choice, system capacity and demand configuration. Such constraints are simulated by means of traffic assignment model. Traffic assignment can be carried out with deterministic or stochastic user behaviour; the optimal signal setting can be obtained by considering isolated or interacting junctions. The traffic assignment component is widely treated in the scientific literature under different hypotheses and has a definitive formulation in terms of fixed point theory (Sheffy 1985; Cascetta 2001). Within traffic assignment the problem of optimal signal setting at junctions has to be solved (Cantarella & Sforza 1991).

(d) Solution generation can be carried out using the following mutually exclusive tools: a set generated synthetically by experts, exact optimisation algorithm (only for small systems), or a heuristic optimisation algorithm (also for real networks). Heuristic optimisation algorithms have received considerable attention, with the use of genetic algorithms (Goldberg 1989; Cantarella & Vitetta 1994) or with enumerative algorithms (Billheimer & Gray 1973; Foulds 1981; Poorzahedy & Turnquist 1982; Boyce et al. 1988; Chen & Alfa 1991). Solution generation is not directly treated in this paper and in the next section it is only specified in order to introduce the notation and terminology necessary for solution selection proposed in Section 3. In monocriterion URND only one optimal solution is generated if the model is formulated in order to guarantee solution existence and uniqueness. The monocriterion approach is not acceptable because several categories are present. With the developments of the multi-criteria approach, it has become necessary to find new tools that are able to assess sets of non-dominated solutions, as classical multi-criteria analysis does not give unambivalent results but confirms many solutions at the same level or range of goodness.

2.2. *Problem definition and solution: model and algorithm*

In this work, for URND we make reference to the formulation proposed by Magnanti and Wong (1984). If we consider a supply model represented by a network with

- a set $\mathbf{Q}_1$ of nodes to design in terms of junction regulation with cardinality H,
- a set $\mathbf{Q}_2$ of sequences of contiguous links (the sequence cannot overlap with other sequences and should be made up by homogeneous links in terms of capacity that have the same topologic configuration) to design in terms of topology (lane allocation in each direction) with cardinality K,
- a list $\mathbf{y}$ of supply configuration variables (state vectors) of set $\mathbf{Q}$ ($\mathbf{Q} = \mathbf{Q}_1 \cup \mathbf{Q}_2$) with element y_j and cardinality $H + K$,
- a vector $\mathbf{f}$ of flows with element f_i,
- an objective function $\Gamma(\mathbf{f}, \mathbf{y})$, the network design model may be expressed as follows:

$$\left\{\begin{array}{ll} \text{(b)} & \mathbf{f},\mathbf{y} = \arg\min_{\mathbf{f},\mathbf{y}} \Gamma(\mathbf{f},\mathbf{y}) \\ \text{(a)} & \\ \text{(c)} & \text{subject to constraints relative to:} \\ \text{(c.i)} & \quad \text{supply} \\ \text{(c.ii)} & \quad \text{demand} \\ \text{(c.iii)} & \quad \text{supply-demand interaction} \\ \text{(c.iv)} & \quad \text{budget} \end{array}\right.$$

(a) In the URND the variables can be divided into state variables and control (or decisional) variables.State variables are the link flows and the generalised cost for the run on the links and the wait at the junction. The state variables are obtained by considering a defined supply and demand configuration and applying equilibrium models.The variables obtained in the problem are the vector $\mathbf{f}$ (state variable) of flow and the list $\mathbf{y}$ (control or design variable) of supply configuration. In urban areas, the flows involve all the user classes that travel on the transportation system. For each link i, the term $\mathbf{y}_i$ defines the topological configuration and the link capacity of link i. The term $\mathbf{y}_i$ is a state vector with the following information for link i: lanes allocated in each direction and junction regulation in the final node in terms of cycle length, phase sequence and duration. To reduce the number of variables, y_i can be relative to a "sequence". This constraint also allows practical design rules (links that belong to the same main road have the same direction in the layout). A sequence is a loopless path composed by links constrained to have the same configuration, and therefore, in automatic solution generation, the links change the configuration contemporaneously.

(b) The objective function $\Gamma(\mathbf{f}, \mathbf{y})$ contains elements relative to the three groups of categories of stakeholders defined: users, system managers, non-users. Each group has a strategy to perform defined with criterion CR_s to minimize or maximize. Such components may conflict with one another and hence the additive effects are not possible. The criteria relative to each group could be: total travel time for users; total management costs for system managers; pollutant emissions for non-users. The solutions may be obtained by multi-criteria analysis, which tends to optimise all criteria together: $\Gamma(\mathbf{f}, \mathbf{y}) = \{\mathrm{CR}_1, \mathrm{CR}_2, \mathrm{CR}_3, \ldots\}$. In classical bi-level optimisation only one criterion was considered: system manager criterion in the objective function and user criteria in the constraints. Multi-criteria optimisation allows different criteria to be used in the objective function, maintaining the constraints relative to user behaviour.

(c.i) Supply constraints concern the supply system characteristics (cost functions on links, regulation on links and at intersections, etc.). The network topology is defined by a link-path incidence matrix $\mathbf{A}$ with element A_{ig} (that is 1 if link i belongs to the path g and 0 otherwise) and a list $\mathbf{y}$ of link configuration vectors of state variables. The link–path incidence matrix $\mathbf{A}$ is a function of network configuration:

$$\mathbf{A} = \mathbf{A}(\mathbf{y}) \tag{1}$$

The path flows are reported in a vector of path flows $\mathbf{F}$ with element F_g ($f_i = \Sigma_g A_{ig} F_g$) related to the vector $\mathbf{f}$ with the following:

$$\mathbf{f} = A(\mathbf{y})\mathbf{F} \tag{2}$$

The link cost vector, $\mathbf{c}$, is defined as a vector whose generic component c_i consists of the transport cost (generalised) on link i. The scalar function $c_i = c_i(\mathbf{f})$, which allows us to calculate the average transport cost c_i of each link corresponding to a link flow vector, is called a cost function and may be either separable or non-separable. The cost functions are defined in a vector of cost function $\mathbf{c} = c(\mathbf{f})$ with element c_i. Assuming a configuration of network layout **y** in congested conditions the vector of link cost **c** depends on the overall network configuration **y** and on the vector of link flow **f** (generally with a strictly monotone increasing function with respect to **f**):

$$\mathbf{c} = \boldsymbol{c}(\mathbf{f}, \mathbf{y}) \tag{3}$$

The path costs are defined in a vector of path costs **C** with element C_g (assuming that it is additive to link costs), the generic C_g can be expressed as $C_g = \Sigma_i A_{ig} c_i$ and in vector notation considering (2) and (3) is:

$$\mathbf{C} = \mathbf{A}\ \mathbf{c} = \boldsymbol{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y}) \tag{4}$$

(c.ii) Demand constraints concern the conservation of demand and path choice behaviour. Generally, demand conservation is defined by:

$$\mathbf{F} = \mathbf{P}\ \mathbf{d} \tag{5}$$

where **d** is a vector of non-zero elements of the O/D matrix and **P** is the path-O/D pairs choice matrix with element p_{gr} that represents the probability of choosing path g in the O/D pair r. The demand vector **d** and the matrix **P** depends on the vector of path cost **C** and considering (4) can be reported as:

$$\mathbf{d} = \boldsymbol{d}(\mathbf{C}) = \boldsymbol{d}[\boldsymbol{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y})] \tag{6}$$

$$\mathbf{P} = \boldsymbol{P}(\mathbf{C}) = \boldsymbol{P}[\boldsymbol{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y})] \tag{7}$$

and the demand conservation considering (5), (6) and (7) can be written in the form:

$$\mathbf{F} = \boldsymbol{P}[\boldsymbol{A}(\mathbf{y})^T (\mathbf{f}, \mathbf{y})] \boldsymbol{d}[\mathrm{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y})] \tag{8}$$

(c.iii) The supply–demand interaction constraint concerns the circular dependence between costs, demand and flows on the links. The flow vector may be determined by the use of equilibrium assignment models and relative algorithms found in the literature (Cantarella 1998). The supply–demand interaction constraints, in order to obtain the flow vector f^*, considering (2) and (8), can be expressed as the solution to the problem:

$$\mathbf{f} = \boldsymbol{A}(\mathbf{y}) \boldsymbol{P}[\mathrm{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y})][\boldsymbol{A}(\mathbf{y})^T \boldsymbol{c}(\mathbf{f}, \mathbf{y})]$$

In relation to the preliminary hypothesis on user behaviour on route choice, two models can be considered: stochastic user equilibrium (SUE) and deterministic user equilibrium (DUE). SUE concerns probabilistic choice on the route and DUE relates to deterministic choice on the route (Cascetta 2001):

- in SUE $\mathbf{P} = \boldsymbol{P}(\mathbf{C})$ is a function;
- in DUE $\mathbf{P} = \boldsymbol{P}(\mathbf{C})$ is a map.

(c.iv) Various budget constraints exist, such as financial budgets, pollutant budgets and technical budgets: financial budgets concern the operator who manages the network, pollutant budgets concern non-users, while technical budgets refer to regulations and the maximum size of the infrastructural links.

(d) The general exact solution of the URND model in terms of uniqueness cannot be guaranteed in real systems. The heuristic approach for the solution has to be considered.The URND methodology used in this paper for generating solutions, is introduced into the optimisation design approach with multi-criteria analysis. The general scheme is the following (Figure 1):

(d.i) through heuristic algorithms (Cantarella and Vitetta 1994) the solutions for analysis are generated; if the algorithm is in the first iteration in the solutions for analysis, the current network configuration is introduced;
(d.ii) the newly obtained solutions are analysed in an algorithm where: the vector **f** is obtained and the signals at the junctions are designed (Webster 1958) in an iterative algorithm within the Frank and Wolfe algorithm for the DUE model or MSA algorithm for the SUE model; various indicators are calculated; analysis of dominance is undertaken with regard to the generated solutions and those dominated are eliminated from the final solution;
(d.iii) if the test at the end of the algorithm is not respected, other solutions are generated for analysis and we return to step (d.i).

At the outer level new network configurations are generated through heuristic algorithms considering the solutions already evaluated; at the inner level the traffic signal setting and link flow assignment are carried out by an iterative method within traffic assignment.

In the following section the non-dominated solutions are analysed with the proposed cluster methodology.

3. Methodology for solution selection

In resolving URND with multi-criteria analysis, a large number of non-dominated solutions are generated: the optimal solution in relation to each

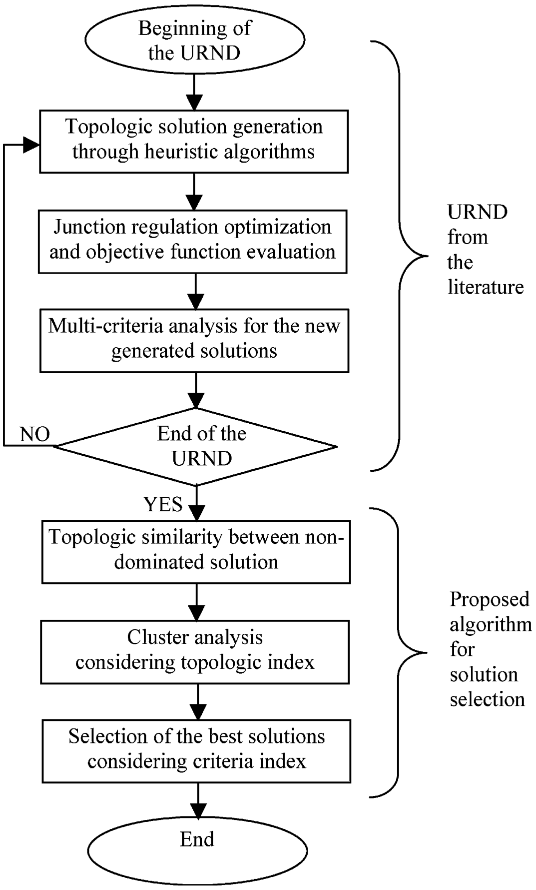


Figure 1. Algorithm for network design.

criterion and the mixed solutions. Today, according to the available literature, three dominating multi-criteria analysis methods are emphasised: ELECTRA, Analytic Hierarchy Processing and PROMETHEE. The common characteristic is numerical analysis for the comparison of the different solutions. This is the strength of this method (because it could be applied to all the problems) but also the weakness (because it considers only numerical aspects). The proposed method for multi-criteria analysis considers a topological method to compare solutions expressed in term of a network, and they could be applied in other fields of engineering systems (chemical, electrical, electronic, mechanical, military, telecommunications).

A three-step method is proposed: topological similarity evaluation between solutions, cluster formation (Kendall et al. 1983), selection of the best solution in the clusters (Figure 1). The similarity and cluster steps use topological indexes. In the similarity and cluster levels, criteria indicators are not considered as a very different network layout could generate very similar indicator values. This is, in general, the limit of consolidated methods. For example, considering a single origin/destination pair and two different network layout configurations, each with only one path (A and D of Figure 2), the two solutions A and D have different layout configurations but they have the same travel time and the same total travel cost (length) on the network. This is an example of two totally different configurations with the same value of one criterion (total travel time on the network). One of the advantages of the proposed method is that it solves multi-criteria analysis where network topology and criteria values are considered together.

3.1. *Topologic similarity evaluation between solutions*

Similarity among the non-dominated solutions generated is evaluated by comparing the solutions from a topological point of view, in order to filter out small topological differences between non-dominated solutions.

Let:

(a) $\boldsymbol{S}$ be the set of non-dominated solutions (with respect to the criteria) generated with a URND model;
(b) W and V be two URND solutions belonging to $\boldsymbol{S}$ (W, $V \in \boldsymbol{S}$);
(c) k be a generic sequence of links in the infrastructural network with $k = 1 \ldots K$;
(d) l_k be the "length" (distance, time or generalised cost) of the sequence k.

Considering 2 URND solutions W and V and a sequence k, the following variable can be defined (Figure 3):

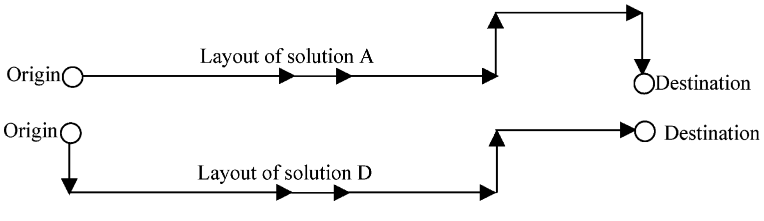


Figure 2. Example of two different layout configurations with the same value of criteria.

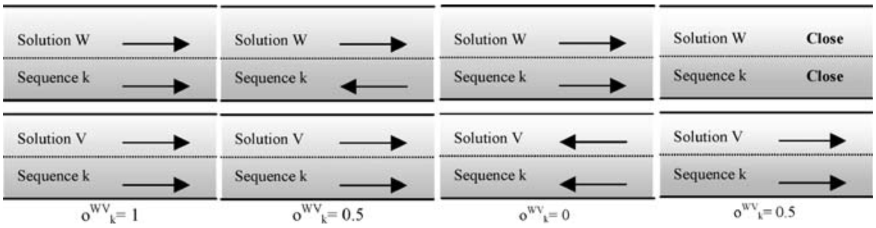


Figure 3. Example of similarity in a pair of solutions (W and V) for a sequence with two lanes.

$$O_k^{WV} = \begin{cases} 0.0 \text{ if the topologic configuration for } k \text{ in } W \text{ and } V \text{ is different} \\ \text{in the two possible directions} \\ 0.5 \text{ if the topologic configuration for } k \text{ in } W \text{ and } V \text{ is the same} \\ \text{in one of the two possible directions} \\ 1.0 \text{ if the topologic configuration for } k \text{ in } W \text{ and } V \text{ is the same} \\ \text{in the two possible directions} \end{cases}$$

We can define the similarity level between solutions W and V (variable between 0% and 100%) as:

$$\lambda_{WV} = \left[\sum_{k=1}^{K} O_k^{WV} l_k\right] / \left[\sum_{k=1}^{K} l_k\right]$$

in which K is the cardinality of the set designed sequences.

Thus we can calculate the similarity level and generate a matrix λ of cardinality $K \bullet K$, where, in the generic element λ_{WV}, the similarity level among solutions W and V is reported.

The proposed topological indicator is not the only one possible. Other topologic indicators can be used and tested. The similarity indicator proposed is particularly simple, in that it is based on the degree of commonality between solutions in terms of the sequence of links, but it takes account only of the topological features and not of the similarity of the solutions in terms of link capacity or flows, for example. The proposed indicator could be easily extended.

In Figure 4 a numerical example of three non-dominated network configurations is reported. The solutions are called A, B and C. Configurations A and B have overlapping similarity equal to 80%; configurations B and C have 90% similarity; configurations A and C 70%. The symmetric similarity matrix λ is reported in the same figure.

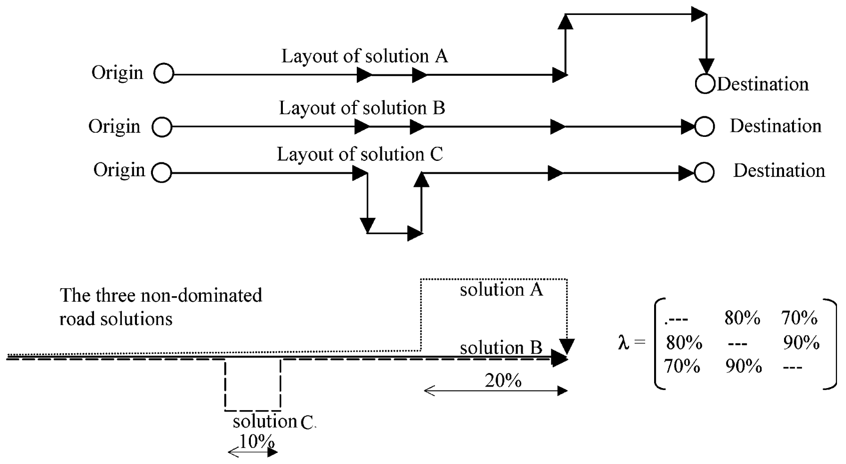


Figure 4. Example of topologic configuration of three non-dominated solutions.

3.2. *Cluster formation*

In general, cluster formation can be performed with different techniques. Having defined a similarity index λ between solutions, a solution *W* is aggregated to a cluster **R** if:

(a) *W* has a similarity level λ at least with a solution (Nearest-neighbour or single linkage) belonging to cluster **R** (hereafter referred to as N_cluster);
(b) *W* has a similarity level λ with all solutions (Furthest-neighbour or complete linkage) belonging to cluster **R** (hereafter referred to as F_cluster).

In the former case, there is a tendency to immediately group the solutions within the clusters, forming long chains of solutions. In the latter, the solutions tend to be grouped in the clusters only if they are similar to all the solutions already present in the cluster. In the same case, having defined a similarity level, all the solutions present in the cluster are certain to be included within the same similarity level.

Considering

- a set of non-dominated solutions with similarity index λ_{WV} for each pair of solutions *V* and *W*; only indexes with *V* and *W* different are considered; the index is symmetric with respect to the two solutions ($\lambda_{WV} = \lambda_{VW}$),
- a dummy variable δ_{VW} equal to 1 for all pairs of non-dominated solutions *V* and *W*; this variable is used to define the pair of solutions that has to be analysed,

The algorithm to generate the cluster can be developed with the following steps:

(1) the pair of solutions V and W is selected inside the set of solutions with δ_{VW} equal to 1 with the maximum value of the similarity index λ_{WV};

(2) one of the following sub-steps is applied considering the solutions V and W

  (2.1) if V and W do not belong to any cluster

  - in nearest and further neighbour, two solutions form a new cluster $\mathbf{R}$ with similarity index λ_{VW};

  (2.2) if one of them (solution V) belongs to the cluster $\mathbf{R}$ and one of them (solution W) does not belong to any cluster (the opposite case that V does not belong to a cluster and W belongs to a cluster is an identical case in terms of the algorithm) one of the two following steps is applied

  - in nearest neighbour, solution W is inserted in cluster $\mathbf{R}$ with similarity index equal to the maximum similarity index value between solution W and all the solutions belonging to $\mathbf{R}$;
  - in furthest neighbour, solution W is inserted in cluster $\mathbf{R}$ with similarity index equal to the minimum similarity index value between solution W and all the solutions belonging to $\mathbf{R}$;

  (2.3) if V belongs to the cluster $\mathbf{R}_1$ and W belongs to the cluster $\mathbf{R}_2$ one of the two following steps is applied

  - in nearest neighbour, clusters $\mathbf{R}_1$ and $\mathbf{R}_2$ are aggregated in the cluster $\mathbf{R}$ with similarity index equal to the maximum similarity index value between solutions belonging to $\mathbf{R}_1$ and all the solutions belonging to $\mathbf{R}_2$;
  - in furthest neighbour, clusters $\mathbf{R}_1$ and $\mathbf{R}_2$ are aggregated in the cluster $\mathbf{R}$ with similarity index equal to the minimum similarity index value between solutions belonging to $\mathbf{R}_1$ and all the solutions belonging to $\mathbf{R}_2$.

(3) one of the following sub-steps is applied considering the solutions V and W;

  (3.1) if V and W do not belong to any cluster

  - the dummy variables δ_{VW} and δ_{WV} are assumed equal to 0;

  (3.2) if one of them (solution V) belongs to the cluster $\mathbf{R}$ and one of them (solution W) does not belong to any cluster

  - all the dummy variables between the solution W and the solutions that belong to cluster $\mathbf{R}$ are assumed equal to 0;

(3.3) if V belongs to the cluster $\mathbf{R}_1$ and W belongs to the cluster $\mathbf{R}_2$

- all the dummy variables between the solutions that belong to cluster $\mathbf{R}_1$ and the solutions that belong to cluster $\mathbf{R}_2$ are assumed equal to 0.

(4) if at least one dummy variable δ_{VW} is equal to 1, return to step 1, otherwise end the algorithm.

Overlapping clusters can be identified with a dendrogram (tree diagram) showing the level of similarity at which each pair of clusters join together. In Figure 5 two dendrograms are reported in a test case with the similarity matrix λ defined *a priori* and reported in Figure 4. Solutions B and C generate a cluster with $\lambda_{AB} = 90\%$ similarity and they are the first clusters generated. With a similar algorithm in the F_cluster case, solutions B and C generate a cluster with 90% similarity and they are the first cluster generated. Solution C is aggregated to cluster containing solutions A and B with 70% similarity.

In problems concerning the topological aggregation of transportation solutions, each of the two proposed aggregation algorithms should be verified. In this paper, in the numerical application, a comparison between the two algorithms is reported in a real transportation system. An application on larger systems is reported in Section 4.

3.3. *Selection of the best solution in the clusters*

In each cluster the best solution, in the sense that it represents the latent characteristics, should be identified. Different ordering methods can be applied to select the best solution and for each method different indicators can be used.

The most widely used ordering methods refer to dominance and discordance distance indexes. All methods try to select the "best" solution, cluster representative, with respect to external decisional criteria. To select the

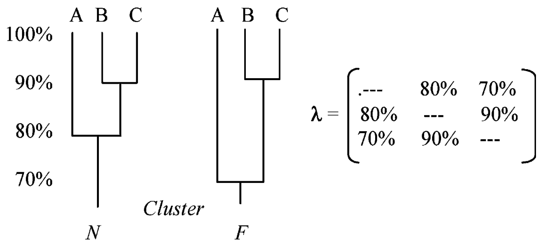


Figure 5. Example of cluster formation.

"best" solution in each cluster, network layout or criteria values indicators can be used. With network layout distance indexes, network configuration indicators have to be used (the solution with the greatest topological similarity, the solution with the least topological deviation). With criteria value indicators the "best" solution depends on the criteria values and position indicators (distance, concentration, ...) are considered.

In the proposed methodology, in order to define the "best" solution a distance index is considered. The index represents the distance between solutions with criteria values reported in the range [0,1]. Note that the criteria used can be some of those used in the objective functions. The solution that has the lowest index value in each cluster is considered the most representative for the same cluster, and is defined as the optimal cluster solution.

Let:

(a) n be the number of solutions in cluster $\mathbf{R}_i$ and w the generic solution in cluster $\mathbf{R}_i$;
(b) h be the generic indicator;
(c) $\mathbf{cr}_w$ with generic h element cr_{wh}, be the vector of criteria of solution w in the range [0,1] inside cluster $\mathbf{R}_i$ and nc the numerousness of $\mathbf{cr}_w$;
(d) p be the norm.

The index is defined as:

$$m_w = \sqrt[p]{\sum_{h=1}^{nc} (\mathrm{cr}_{wh})^p}$$

If the norm is equal to 2, the Euclidean distance is considered and the index is called the Minkowski index (Kendall et al. 1983):

$$m_w = \sqrt{\sum_{h=1}^{nc} (\mathrm{cr}_{wh})^2} = \sqrt{\mathbf{cr_w}^T \mathbf{cr_w}}$$

The best solution of cluster $\mathbf{R}$ is that with the lowest m_w.

4. Application in real urban networks

The methods for URND and for solution selection with a topological index are applied in a real system. The main aims of the application are to verify the possibility of applying the method for a real system and to analyse the goodness of the methodology and the transferability of results to different networks. A secondary aim is to obtain some key practical rules to apply when a road transportation system has to be designed.

The study areas considered are the towns of Acireale and Crotone (both in southern Italy, with populations of 50,000 and 100,000, respectively). The main road network characteristics (number of centroids, nodes and links, number of designed link sequences, number of non-dominated solutions generated with the network design model) are reported in Table 1. The link cost functions are divided into the running term and the junction waiting term depending on the link flow and on the geometric links and junctions characteristics. Transport demand is obtained through a four-step model (Cascetta 2001). The car modal O/D matrix used is relative to two peak hours: the 7:30–8:30 time slice and the 13:30–14:30 time slice.

It is hypothesised that a perturbation is introduced and a connected set of links in the central area of the town is closed to road traffic and thus is set in the design condition with a zero running speed (for example, for the creation of a pedestrian area).

The two study areas have comparable characteristics in terms of number of inhabitants, total area, level of mobility and socio-economic distribution. Although the town of Crotone is slightly larger than Acireale in terms of surface area, in our graph, the number of links in the Acireale graph is greater than the number of links in Crotone (613 and 416 links, respectively). It is hypothesised that a percentage (about 25%) of roads, in the historical town centre (in both towns), are closed to vehicular traffic. In order to compare the results, in the two towns the same number of designed links (200) were used. By applying the methodology described in Section 2, 74 non-dominated solutions were generated in the town of Acireale and 128 in Crotone.

In the objective function the criteria reported in Table 2 are considered. The criteria are relative to the users (total travel time) and to the system managers (total length of modified links as an indication of the cost of applying the configuration, and total user flow and length per user flow on the modified links as an indication of the inertia of the system to evolve towards equilibrium conditions).

In order to perform the topological steps (similarity and cluster) the methodology proposed in Section 3 can be used directly without other exog-

Table 1. The network characteristics and solutions generated.

Town	Acireale	Crotone
Number of centroids	33	22
Number of nodes	265	154
Number of links	613	416
Number of link sequences	40 (200 links)	30 (200 links)
Non-dominated generated solutions	74	128

Table 2. The indicators utilised in multi-criteria optimisation.

Criterion	Indicator	Formula
Users	Total travel time on the links in each time slice h	$\mathrm{TTh} = \sum_i t_i^h * f_i^h$
System Manager	Total length of modified links	$\mathrm{LT} = \sum_i l_i * m_i$
Inertia	Total flow of users on the modified links in each time slice h	$\mathrm{FTh} = \sum_i f_i^h * m_i$
	Total length per user flow (vehic*km) on the modified links in each time slice h	$\mathrm{KTh} = \sum_i l_i * f_i^h * m_i$

* t_i^h travel time on link i in time slice h; f_i^h flow on link i in time slice h; l_i the "length" of the sequence i; $m_i = 1$ if the configuration of link i has been modified; 0 otherwise.

enous data. To develop the "best" step the same criteria considered in the objective function are considered.

The steps followed in Crotone are reported below. The same procedure is followed for Acireale.

The current configuration was first optimised with URND and total travel time was calculated for users on the optimised network obtaining:

- 724.3 vehicle-hours/hour using a SUE assignment model;
- 663.3 vehicle-hours/hour using a DUE assignment model.

The perturbation was then introduced in terms of closing a connected set of links in the central area of the town and the indicators recalculated, obtaining a total travel time:

- 729.9 vehicle-hours/hour using a SUE assignment model;
- 666.9 vehicle-hours/hour using a DUE assignment model.

Optimal configurations were generated by using genetic algorithms, DUE and SUE-probit assignment and the four criteria reported in Table 2. By applying the design model, solutions may be obtained with overall travel times of:

- 724.6 vehicle-hours/hour with SUE assignment and a genetic algorithm;
- 663.9 vehicle-hours/hour with DUE assignment and a genetic algorithm.

Overall user travel time on the network for the optimal configurations obtained are very similar to current times without perturbation, which proves the importance of the design model. Such solutions are obtained at the cost of significant increases in the indicators relative to the other criteria used.

By applying the automatic designed model with DUE assignment, in the town of Crotone 128 non-dominated solutions are generated and in Acireale 74 non-dominated solutions are generated. These solutions are considered for solution selection.

The similarity step was carried out obtaining the λ matrix, and then the two F_Cluster and N_Cluster algorithms were applied.

In Figure 6 the dendrograms for the two towns with F_Cluster algorithm are reported. Each non-dominated solution is indicated with a number. The similarity level is reported on the left; the best solution of each cluster is marked with *. The different clusters are joined by a line at the similarity level; all solutions with similarity over 70% were grouped in the same starting clusters.

By applying the automatic designed model, in Crotone 128 non-dominated solutions are generated and the minimum level of solution aggregation is 50%; in Acireale, although 74 non-dominated solutions are generated, the minimum level of aggregation is 25%. Despite its smaller size in the number

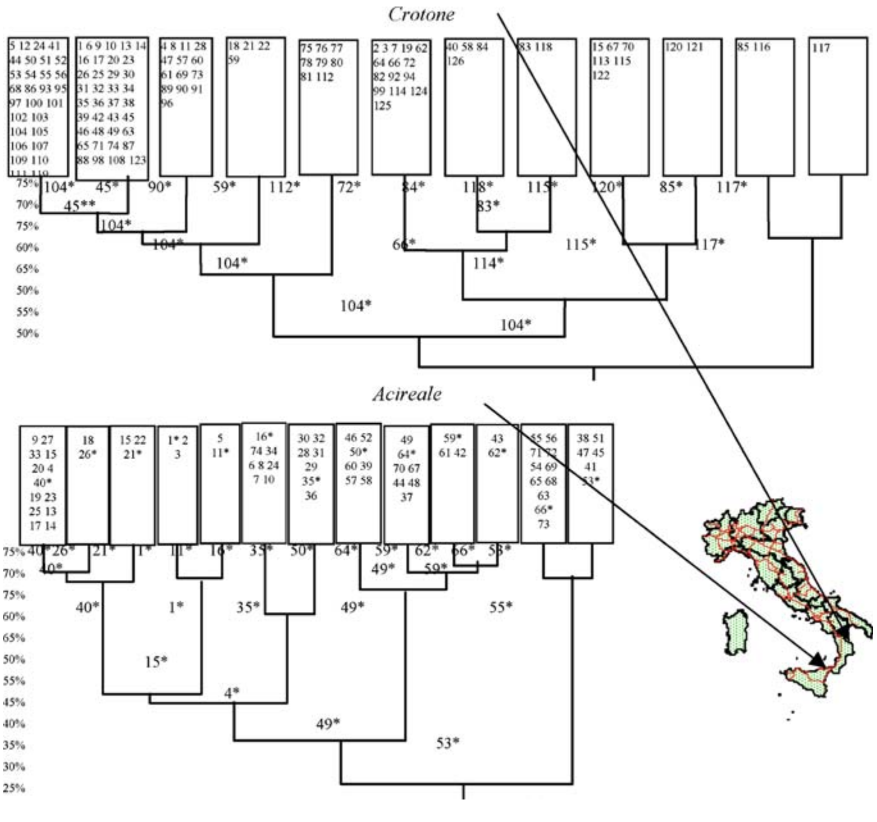


Figure 6. The cluster for the two cities.

of links, there are more non-dominated solutions generated in Crotone than in Acireale. The local topology of the town affects the total number of non-dominated solutions generated. However, from an engineering viewpoint the number of significant solutions is similar in the two cases examined. Indeed, considering the similarity level of 75%, there are 12 clusters in Crotone and 13 in Acireale. The local density of the graph and its structure affect the number of starting non-dominated solutions generated, while the number of significant solutions is comparable.

In both towns, there is a subset of sequences of "frozen" (non-topologically changing) roads in all configurations (25% of the sequences in Acireale, 50% in Crotone), i.e. the closure of some sequences (always closed or always open in all generated solutions), providing useful indications for the project. The global methodology was performed with proprietary compiled software developed in C++ and using a PC Pentium III 600 MHz and the elapsed time was 10 hours for Crotone and 8 hours for Acireale.

The number of solutions which allows the maximum compromise in the two cities is 104 (Crotone) and 53 (Acireale). In Crotone there are more non-dominated solutions than in Acireale and in Crotone all the solutions are clustered with a higher similarity level than in Acireale. This is probably due to the fact that in Crotone there are more possibilities of organising the network than in Acireale, with non-dominated criteria values, but the solutions are more similar to one another. In Crotone there are 50% of sequences frozen and in Acireale 25% of sequences frozen.

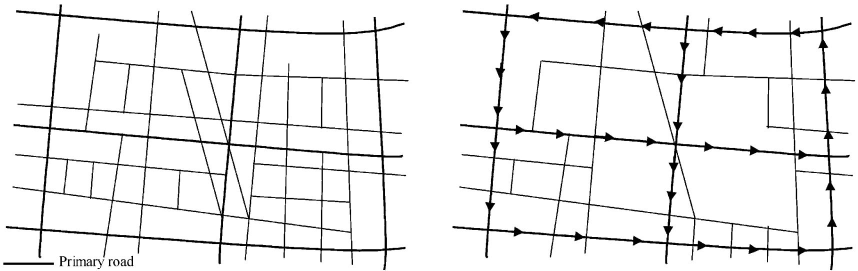
It is worth noting that a feature common to all the solutions in Crotone is that of always maintaining bidirectionality in the outermost ring. By contrast, in Acireale there is the tendency to create one-way or two-way mono-directional rings. This shows that often the urban network of a town and the lack of high-capacity roads make it impossible to create one-way outer circuits, permitting two-way vehicle flow throughout the day. This observation confirms that Crotone has fewer "real degrees of freedom" in network organisation than has Acireale.

The optimal solutions generated, in term of network configurations, are characterised by some circuits around the town centre (primary roads), and secondary roads are closed to traffic. The running capacity increases insofar as the roads are generally used in only one direction. The crossing capacity increases because the conflict points between the various flows on the intersection are reduced. The possibility of introducing dynamic network management during the day reduces the total travel time on the network since the solution generated with the minimum total travel time in the morning peak hour is different from that generated with the minimum total travel time in the afternoon peak hour. It requires a high monetary cost to control and manage the network and to inform the users, who should know different network layouts in the same area.

From the application of the methodology and analysing the layout relative to each cluster, we may extract some ideas to apply when a road transportation system has to be designed.

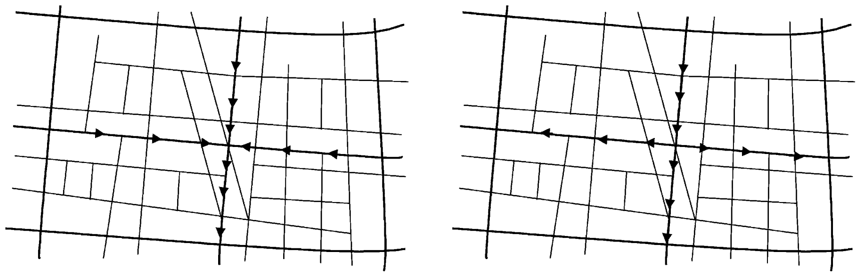
The main ideas identified are (Figure 7):

I. Pedestrian zone (links with low flow) in some areas in the secondary network;
II. Dynamic network management during the day (solution II.1 morning peak hour and solution II.2 afternoon peak hour);



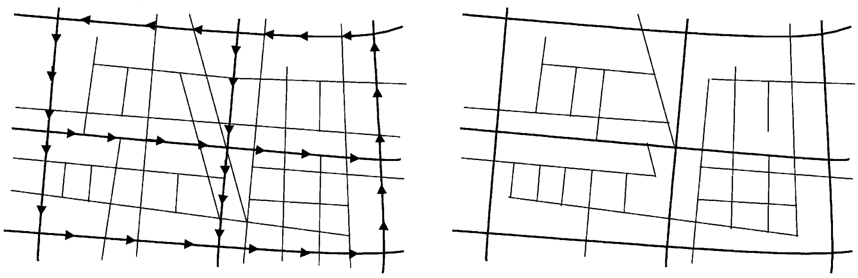
Basic graph of an urban road network

(I) Pedestrian area (links with low flow) in some areas in the secondary network



(II.1) Dynamic network management during the day (morning peak hour)

(II.2) Dynamic network management during the day (afternoon peak hour)



(III) Mono-directional link around the centre

(IV). Reducing the junction between primary and secondary network

Figure 7. Some layouts for topological network design.

III. Mono-directional links around the centre or a perturbed area in order to increase junction service level;
IV. Reducing the number of junctions between the primary and secondary network.

5. Conclusions

This study analysed non-dominated configurations of an urban road network. A method for topological cluster formation and best solution extraction is reported. Numerical application was carried out and are reported on two real road urban networks of medium size in order to verify the validity of the method.

Two different kinds of conclusion can be proposed: the first concerns the goodness of the proposed models while the second group of conclusions concerns the results of the proposed algorithm in the experimental areas. Due to the general methodology, interesting results may be obtained since the algorithm solves the problem in a topological context. Interestingly, after the similarity step, commercial software could also be used to perform cluster and best steps. It has been demonstrated that the obtained rules increase the global level of service of networks in terms of reductions in generalised cost for users (travel time, monetary cost and safety), for public network managers (management costs) and non-users (pollution).

The application was carried out on a test system and on two real networks; the algorithm in real large networks and/or with real time control has to be tested. The proposed methodology is heuristic in some respects (solution optimisation) but it can be applied to large networks. It has various restrictive hypotheses (heuristic single junction traffic signal setting optimisation, separable link cost function, inelastic demand) that can be removed in subsequent studies. Furthermore, calculation times are not low and a global code of optimisation must be designed.

In the two urban areas where the application was developed some common layouts of optimal solutions are carried out. The main characteristics of such common layouts can be summarised observing the layout of the optimal solution generated (and not reported in this paper) by a synthetic rule (or image or concept) like the maximum number of pedestrian areas, maximum number of mono-directional links and rings around the perturbed area, and minimum number of junctions between the secondary and primary network. These rules cannot be considered fixed. They should be confirmed,extended or modified in other applications on real transportation systems.

The main result of this paper concerns the possibility of applying a topological method to select non-dominated solutions. The results confirm that traditional multi-criteria analysis which considers only the criteria for solution selection could consider solutions in the same group with similar criteria values but different topological configurations. With the proposed topological approach, the group (cluster) is generated with a topological index and selection of the best solution is extracted with the criteria values.

Acknowledgements

The authors wish to thank the anonymous referees for their suggestions which were most useful in revising the paper.

Web site references An application of the proposed algorithm on larger systems is reported in http://last.unirc.it/2005/app_a.pdf.The general algorithm is reported in http://last.unirc.it/2005/app_b.pdf.

References

Billheimer JW & Gray P (1973) Network design with fixed and variable cost elements. *Transportation Science* 7: 49–74.

Boyce DE, Ben Ayed O & Blair CE III (1988) A general bilevel programming formulation of the network design problem. *Transportation Research* 22: 311–318.

Cantarella GE (1998) A general fixed point approach to multi-mode multi-user equilibrium with elastic demand *Transportation Science* 31: 107–128.

Cantarella GE & Sforza A (1988) Determinazione di tariffe e frequenze per il servizio passeggeri di una rete ferroviaria. Un approccio in programmazione non lineare. Atti delle giornate AIRO, Pisa.

Cantarella GE & Sforza A (1991) Road network signal setting: Equilibrium conditions. In: Papageorgiou M (ed) *Concise Encyclopedia of Traffic and Transportation Systems* (pp. 366–371). Pergamon Press.

Cantarella GE & Vitetta A (1994) A multicriteria analysis for urban network design and parking location. *Preprints of Tristan II Conference* (pp. 839–852). Capri, Italy.

Cascetta E (2001) *Transportation System Engineering*. Kluwer.

Ceder A & Wilson NHM (1986) Bus network design. *Transportation Research* 20B: 331–344.

Chen M & Alfa AS (1991) A network design algorithm using a stochastic incremental traffic assignment approach. *Transportation Science* 25: 215–224.

Drezner Z & Wesolowsky GO (2003) Network design: Selection and design of links and facility location. *Transportation Research* 37A: 241–256.

Florian M & Costantin I (1993) *Optimizing Frequencies in a Transit Network: a Nonlinear Bi-level Programming Approach*. Publication 914 of Crt, University of Montreal.

Foulds LR (1981) A multi-commodity flow network design problem *Transportation Research* 15B: 273–283.

Friesz TL & Shah S (2001) An overview of nontraditional formulation of static and dynamic equilibrium network design. *Transportation Research* 35B: 5–21.

Goldberg DE (1989) *Genetic Algorithms in Search. Optimization and Machine Learning*. Addison Wesley.

Kendall SR, Stuart A & Ord J (1983) *The Advanced Theory of Statistics*. Vol 3 Cap 44 London & High Wycombe: Charles Griffin & Company Limited.

Magnanti TL & Wong RT (1984) Network design and transportation planning: Models and algorithms. *Transportation Science* 18: 1–55.

Meng Q, Yang H & Bell MGH (2001) An equivalent continuously differentiable model and a local convergent algorithm for the continuous network design problem. *Transportation Research* 35B: 83–105.

Meng Q & Yang H (2002) Benefit distribution and equity in road network design. *Transportation Research* 36B: 19–35.

Poorzahedy H & Turnquist MA (1982) Approximate algorithms for the discrete network design problem. *Transportation Research* 16B: 45–55.

Russo F (1998) Transit frequencies design for enhancing the efficiency of public urban transportation systems: An optimization model and an algorithm. *Proceedings of the 31st ISATA Stuttgart*. Croydon, England: Automotive Automation Limited.

Sheffy Y (1985) *Urban Transportation Network*. Englewood Cliff: Prentice Hall.

Webster FW (1958) Traffic signal settings. Road Research Technical Paper no. 39.

About the authors

Francesco Russo obtained his first degree in Transportation Engineering in 1982. From 1987 to 1998 he was an Assistant Professor before he became an Associate Professor. He was appointed a full Professor in 1999 at the University of Reggio Calabria, Department of Computer Science, Mathematics, Electronics and Transportation. He is the director of the Ph.D. course in Transportation and Logistic in the University of Reggio Calabria. He has published more than 150 papers in transportation area and is responsible for research projects with the EU and the Italian Government.

Antonino Vitetta obtained his first degree in Transportation Engineering in 1990, and completed his Ph.D. in Transportation Engineering in 1994. From 1994 to 2001 he was Assistant Professor and since 2001 is Associate Professor at the University of Reggio Calabria, Department of Computer Science, Mathematics, Electronics and Transportation. He has published more than 100 papers in transportation area.