Congested Network Models

Introduction

This lecture provides a general mathematical formulation of transportation supply models, based on congested network flow models. The bases for these models are graph models. Next, network models, including link performances and costs, and network flow models, including link flows, are introduced. Finally, congested network (flow) models are developed, modeling relationships among performances, costs, and flows.

Network Structure

The network structure is represented by a graph. The latter is defined by a set N of elements called nodes and by a set of pairs of nodes belonging to N, $L \subseteq N \times N$, called links. The graphs used to represent transportation services are generally oriented; that is, the links have a direction and the node pairs defining them are ordered pairs. A link connecting the node pair (i, j) can also be denoted by a single index, say a.

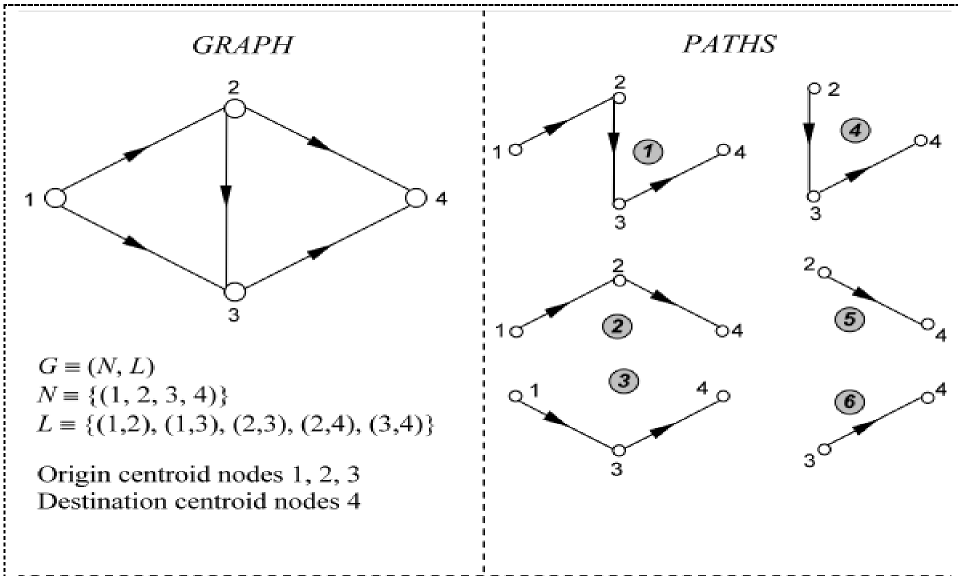
The links in a graph modeling a transportation system represent phases and/or activities of possible trips between different traffic zones. Thus, a link can represent an activity connected to a physical movement (e.g., covering a road) or an activity not connected to a physical movement (such as waiting for a train at a station). Links are chosen in such a way that physical and functional characteristics can be assumed to be homogeneous for the whole link (e.g., the same average speed). In this sense, links can be seen as the partition of trips into segments, each of which has certain characteristics; the level of detail of such a partition can clearly be very different for the same physical system according to the objectives of the analysis.

Nodes correspond to significant events delimiting the trip phases (links), that is, to the space and/or time coordinates in which events occur that they represent. In synchronic networks, nodes are not identified by a specific time coordinate, and the same node represents events occurring at different moments (instants) of time. For example, the different entry or exit times in a road segment, an intersection, or a station, may be associated with a single node, representing all the entry/exit events.

Centroid nodes represent the beginning or end of individual trips. In diachronic networks, on the other hand, nodes may have an explicit time coordinate and therefore represent an event occurring at a given instant. The graphs considered in this lecture are synchronic because diachronic networks assume a within-period system representation; diachronic graphs for scheduled services.

A trip is a sequence of several phases and, in a graph that represents transportation supply, it consists of a path k, defined as a succession of consecutive links connecting an initial node (path origin) to a final node (path destination). Usually, only paths connecting centroid nodes are considered in transportation graphs. On this basis, each path is unambiguously associated with one, and only one, O-D pair, whereas several paths can connect the same O-D pair. An example of a graph with different paths connecting the centroid nodes is depicted in Fig. 1.

A binary matrix called the link–path incidence matrix Δ can represent the relationship between links and paths. This matrix has a number of rows equal to the number of links n_L and a number of columns equal to the number of paths n_P. The generic element δ_{ak} of the binary matrix Δ is equal to one if link a belongs to path k, $a \in k$, and zero, otherwise, $a \notin k$ (see Fig.1). The row of the link–path incidence matrix corresponding to the generic link identifies all the paths including that link (columns k for which $\delta_{ak} = 1$). Moreover, the elements of a column corresponding to the generic path k identify all the links that make it up (rows a for which $\delta_{ak} = 1$).



O-D Pairs	1-4			2-4		3-4
Links \ Paths	1	2	3	4	5	6
1	1	1	0	0	0	0
2	0	0	1	0	0	0
3	1	0	0	1	0	0
4	0	1	0	0	1	0
5	1	0	1	1	0	1

Fig. 1 Example of a graph and link–path incidence matrix.

Flows

A link flow f_a can be associated with each link a. Link flow is the average number of homogeneous units using link a (i.e., carrying out the trip phase represented by the link) in a time unit. In other words, the link flow is a random variable of mean f_a. Several link flows can be associated with a given link depending on the homogeneous unit considered. User flows relate to users, such as travelers or goods, possibly of different classes. Vehicle flows relate to the number of vehicles, perhaps of different types such as automobiles, buses, trains, and so on.

For individual modes, such as automobiles or trucks, user flows can be transformed quite straightforwardly into vehicle flows through average occupancy coefficients. For scheduled modes, such as trains, vehicle flows derive from the service schedule and are often treated as an input to the supply model.

The link flow of the generic user class or vehicle type i is denoted by f_a^i. In accordance with the results of traffic flow theory, link performance and cost variables are affected by the user or vehicle flow. To allow for this dependence it is often worth homogenizing the various classes of users or various types of vehicles by defining equivalent flows associated with links. In this case, the flows of different user classes or vehicle types are homogenized to a reference class or type:

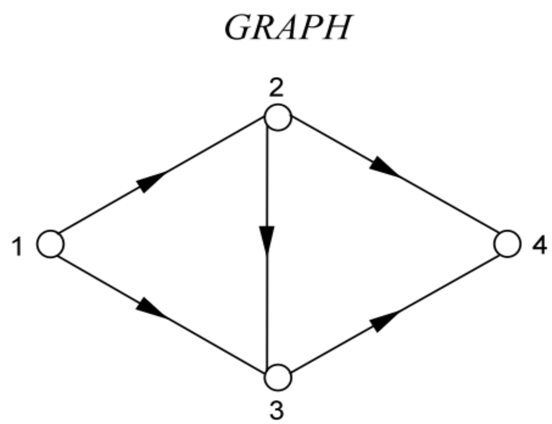
$$f_a = \sum_i w_i f_a^i$$

Where w_i is the homogenization coefficient of the users of class i with respect to their influence on link performances. For example, for road flows, automobiles are usually the reference vehicle type (w_i = 1) and the other vehicle flows are transformed into equivalent auto flows with coefficients w_i. The latter is greater than one if the contribution to congestion of these vehicles is greater than that of cars (buses, heavy vehicles, etc.), less than one in the opposite case (motorcycles, bicycles, etc.).

The vector of link flows f has, as a generic component, the flow on the link a,f_a, for each a ∈ L (see Fig. 2).

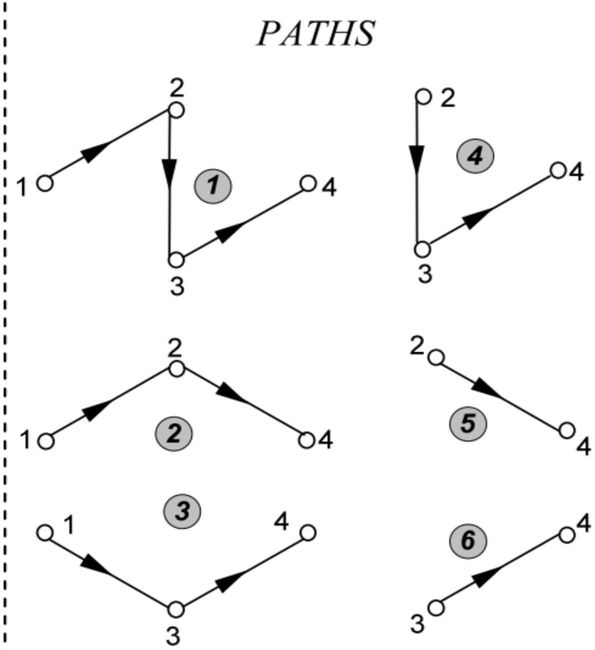
Flow variables can also be associated with paths. Under the within-day stationarity hypothesis, the average number of users, who in each subinterval travel along each path, is constant. The average number of users, who in a time unit follow path k, is called the path flow h_k . If the users have different characteristics (i.e., they belong to different classes), path flows per class i, h_k^i, can be introduced. Path flows of different user classes or vehicle types can be homogenized by means of coefficients w_i similar to those introduced for link flows; the equivalent path flow is obtained as:

$$h_k = \sum_i w_i . h_k^i$$



$G \equiv (N, L)$
$N \equiv \{(1, 2, 3, 4)\}$
$L \equiv \{(1,2), (1,3), (2,3), (2,4), (3,4)\}$

Origin centroid nodes 1, 2, 3
Destination centroid nodes 4



Link flow vector

$$\boldsymbol{f} = \begin{bmatrix} f_{12} \\ f_{13} \\ f_{23} \\ f_{24} \\ f_{34} \end{bmatrix}$$

Path flow vector

$$\boldsymbol{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix}$$

Link-path incidence matrix

$$\Delta = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1,2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1,3 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2,3 & 1 & 0 & 0 & 1 & 0 & 0 \\ 2,4 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3,4 & 1 & 0 & 1 & 1 & 0 & 1 \end{array}$$

Link flow vector

$$\boldsymbol{f} = \begin{bmatrix} f_{12} \\ f_{13} \\ f_{23} \\ f_{24} \\ f_{34} \end{bmatrix} = \Delta \boldsymbol{h} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} h_1 + h_2 \\ h_3 \\ h_1 + h_4 \\ h_2 + h_5 \\ h_1 + h_3 + h_4 + h_6 \end{bmatrix}$$

Fig. 2 Transportation network with link and path flows.

There is clearly a relationship between link and path flows. Indeed, the flow on each link a can be obtained as the sum of the flows on the various paths containing that link. This relationship can be expressed by using the elements δ_{ak} of the link–path incidence matrix as

$$f_a = \sum_i \delta_{ak}.h_k \qquad 1$$

Or in matrix terms:

$$f = \Delta h \qquad 2$$

Where h is the path flow vector.
Equation (1) or (2) expresses the way in which path flows induce flows on individual links. For this reason, it is referred to as the (static) Network Flow Propagation (NFP) model (see Fig. 1). Note that the linear algebraic structure depends crucially on the assumption of intro period stationarity (within the day static model); if this assumption is removed, the model loses its algebraic-linear nature.

Performance Variables and Transportation Costs
Some variables perceived by users can be associated with individual trip phases. Examples of such variables are travel times (transversal and/or waiting), monetary cost, and discomfort. These variables are referred to as level-of-service or performance attributes. In general, performance variables correspond to disutilities or costs for the users (i.e., users would be better off if the values of performance variables were reduced). The average value of the nth performance variable, related to link a, is denoted by r_{na} . The average generalized transportation link cost, or simply the transportation link cost, is a variable synthesizing (the average value of) the different performance variables borne and perceived by the users in travel-related choices and, more particularly, in path choices. Thus, the transportation link cost reflects the average user's disutility for carrying out the activity represented by the link. Other performance variables and costs, which cannot be associated with individual links but rather with the whole trip (path), are introduced shortly.

Performance variables making up the transportation cost are usually nonhomogeneous quantities. In order to reduce the cost to a single scalar quantity, the different components can be homogenized into a generalized cost by applying reciprocal substitution coefficients β, whose value can be estimated by calibrating the path choice model. For example, the generalized transportation cost c_a relative to the link a can be formulated as:

$$c_a = \beta_1.t_a + \beta_2.mc_a$$

Where:
t_a : is the travel time and,
mc_a : is the monetary cost (e.g., the toll) connected with the crossing of the link.
More generally, the link transportation cost can be expressed as a function of several link performance variables as:

$$c_a = \sum_n \beta_n . r_{na}$$

Different users may experience and/or perceive transportation costs, which differ for the same link. For example, the travel time of a certain road section generally differs for each vehicle that covers it, even under similar external conditions. Furthermore, two users experiencing the same travel time may have different perceptions of its disutility. If we then add the fact that the analyst cannot have perfect knowledge of such costs, we realize that the perceived link cost is well represented by a random variable distributed among users, whose average value is linked to transportation cost c_a. There may be other "costs" both for users (e.g., accident risks or tire consumption) and for society (e.g., noise and air pollution) associated with a link. It is usually assumed that these costs are not taken into account by users in their travel-related choices and are not included in the perceived transportation cost. The transportation cost is, therefore, an internal cost, used to simulate the transportation system and, in particular, travelers' choices. The other cost items are external costs, used for project design and assessment. External costs are sometimes referred to as impacts.

Different groups (or classes) of users may have different average transportation costs. This may be due to different performance variables (e.g., their speeds and travel times are different or they pay different fares) or to differences in the homogenization coefficients β_n (e.g., different time/money substitution rates corresponding to different incomes). In this case, a link cost c_l^ican be associated with each user class i. In what follows, for simplicity of notation, the class index i is taken as understood unless otherwise stated.
Link performance variables and transportation costs can be arranged in vectors. The performance vector r_a is made up of the nth performance variable for each link, its components being r_{na}. Analogously, vector c, whose generic component ca is the generalized transport cost on a link a, is known as the link cost vector.

The concepts of performance variables and generalized transportation cost can be extended from links to paths. The average performance variable of a path k, z_{nk}, is the average value of that variable associated with a whole origin-destination trip, represented by a path in the graph. Some path performance variables are linkwise additive; that is, their path value can be obtained as the sum of link values for all links making up the path.

Examples of additive path variables are travel times (the total travel time of a path is the sum of travel times over individual links) or some monetary costs, which can be associated with some or all individual links. An additive path performance variable can be expressed as the sum of link performance variables as:

$$z_{nk}^{ADD} = \sum_{a \in k} r_{na} = \sum_a \delta_{ak} r_{na}$$

or in vector notation

$$z_n^{ADD} = \Delta^T r_n$$

Other path performance variables are nonadditive; that is, they cannot be obtained as the sum of link-specific values. These variables are denoted by z_{nk}^{NA}. Examples of nonadditive performance variables are monetary cost in the case of tolls that are nonlinearly proportional to the distance covered or the waiting time at stops for high-frequency transit systems, as shown below.

The average generalized transportation cost of a path k, g_k , is defined as a scalar quantity homogenizing in disutility units the different performance variables perceived by the users (of a given category) in making trip-related choices and, in particular, path choices.

The path cost in the most general case is made up of two parts: likewise additive cost $g_k^{ADD}$and nonadditive cos, g_k^{NA}, assuming that they are homogeneous:

$$g_k = g_k^{ADD} + g_k^{NA} \qquad 3$$

The additive path cost is defined as the sum of the linkwise additive path performance variables:

$$g_k^{ADD} = \sum_n \beta_n . z_{nk}^{ADD}$$

Under the assumption that the generalized cost depends linearly on performance variables, the additive path cost can be expressed as the sum of generalized link costs. The relationship between additive path cost and link costs can be expressed by combining all the equations previously presented:

$$g_k^{ADD} = \sum_n \beta_n . z_{nk}^{ADD} = \sum_n \beta_n \sum_a \delta_{ak} r_{na} = \sum_a \delta_{lk} \sum_n \beta_n r_{na} = \sum_a \delta_{ak} c_a$$

or

$$g_k^{ADD} = \sum_a \delta_{ak} c_a \qquad 4$$

The expression (4) can also be formulated in vector format by introducing the vector of additive path costs g^{ADD} (see Fig. 3):

$$g^{ADD} = \Delta^T c \qquad 5$$

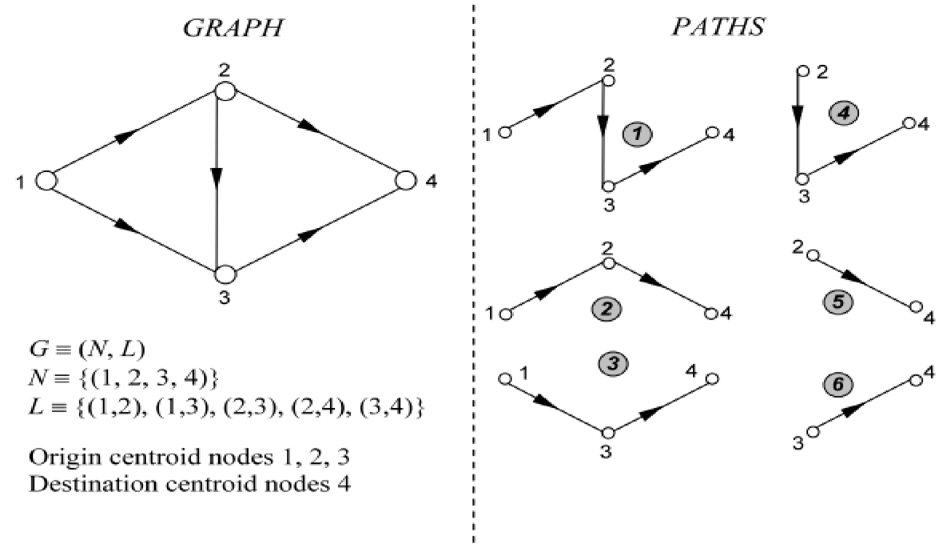
The nonadditive path cost g_k^{ADD} includes nonadditive path performance variables:

$$g_k^{NA} = \sum_n \beta_n z_{nk}^{NA}$$

Finally, the path cost vector g, of dimensions (n p × 1), can be expressed as:

$$g = \Delta^T c + \ g^{NA} \qquad 6$$

Where g^{NA} is the nonadditive path cost vector. In many applications, the nonadditive path cost vector is or is assumed to be, null.



Link cost vector

$$\boldsymbol{c} = \begin{bmatrix} c_{12} \\ c_{13} \\ c_{23} \\ c_{24} \\ c_{34} \end{bmatrix}$$

Link-path incidence matrix

$$\Delta = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1,2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1,3 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2,3 & 1 & 0 & 0 & 1 & 0 & 0 \\ 2,4 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3,4 & 1 & 0 & 1 & 1 & 0 & 1 \end{array}$$

Path cost vector

$$\boldsymbol{g}^{ADD} = \Delta^T \boldsymbol{c} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{12} \\ c_{13} \\ c_{23} \\ c_{24} \\ c_{34} \end{bmatrix} = \begin{bmatrix} c_{12} + c_{23} + c_{34} \\ c_{12} + c_{24} \\ c_{13} + c_{34} \\ c_{23} + c_{34} \\ c_{24} \\ c_{34} \end{bmatrix}$$

Fig. 3 Transportation network with link and path costs.

Link Performance and Cost Functions

Link performance attributes generally depend on the physical and functional characteristics of the facility and/or the service involved in the trip phase represented by the link itself. Typical examples are the travel time on a road section depending on its length, alignment, allowed speed, or the waiting time at a bus stop depending on the headway between successive bus arrivals. When several travelers or vehicles use the same facility, they may interact with each other, thereby influencing link performance.

Typically, the effects of congestion on link performance increase as the flow increases. For instance, the larger the flow of vehicles traveling along a road section, the more likely faster vehicles will be slowed by slower ones, thus increasing the average travel time. Moreover, the larger the flow arriving at an intersection, the longer the average waiting time; the larger the number of users on the same train, and the lower the riding comfort.

In general, congestion effects are such that the performance attributes of a given link may be influenced by the flow on the link itself and by flows on other links. Link performance functions relate the generic link performance attribute r_{na} to the physical and functional characteristics of the link, arranged in a vector b_{na}, and to the equivalent flow on the same link and, possibly, on other links, arranged in the vector f :

$$r_{na} = r_{na}(f; b_{na}, y_{na})$$

Where y_{na} is a vector of parameters used in the function.

Because the generalized transportation cost of a link c_a is a linear combination of link performance attributes, link cost functions can be expressed as functions of the same parameters:

$$c_{na} = c_{na}(f; b_{na}, y_{na}) \qquad 7$$

Where vectorsb_{na} and y_{na} have the same meaning as above.

Link performance and cost functions may have some mathematical properties, which are used to study the properties of supply-demand interaction models and to analyze the convergence of their solution algorithms. Performance and cost functions can be classified as <u>separable</u> and <u>nonseparable</u> across a link. In the former case, the performances and cost variables of a link depend exclusively on the (equivalent) flow of the link itself:

$$c_a(f) = c_a(f_a)$$

In the latter case, they also depend on the flow of other links. Examples of both types of functions are given in the following sections.

The cost function vector c (f) is obtained by ordering the n_L functions of the individual network links:

$$c = c(f) \qquad 8$$

Under the assumption that the first partial derivative of c (f) exists and is finite, the Jacobian matrix, Jac [c (f)], may be defined:

$$Jac[c(f)] = \begin{bmatrix} \frac{\partial c_1}{\partial f_1} & \cdots & \frac{\partial c_1}{\partial f_{nL}} \\ \vdots & \frac{\partial c_i}{\partial f_i} & \vdots \\ \frac{\partial c_{nL}}{\partial f_1} & \cdots & \frac{\partial c_{nL}}{\partial f_{nL}} \end{bmatrix}$$

The cost functions generally have an asymmetric Jacobian. In some cases, they may have a symmetric Jacobian: $\frac{\partial c_i}{\partial f_j}=\frac{\partial c_j}{\partial f_i}$ ∀i, j; that is, the cost varies on link a, due to a flow variation on link j, is equal to the cost variation on link j, due to a flow variation on the link i. Separable cost functions are clearly a special case, the Jacobian being a diagonal matrix: $\partial c_i/\partial f_j = 0$, $\forall\, i \neq j$.

In the case of uncongested networks, the cost functions are independent of the flows, so the partial derivatives are all equal to zero and the Jacobian is null.

Impacts and Impact Functions

Design and evaluation of transportation systems, in addition to performance variables perceived by the users, require the modeling of impacts borne by the users, but not perceived in their mobility choices, and of impacts on nonusers.

Examples of the first type include indirect vehicle costs (e.g., tire or lubricant, vehicle depreciation, etc.) and accident risks with their consequences (death, injury, material damage). The impacts for nonusers include those for other subjects directly involved in the transportation system, such as costs and revenues for the producers of transportation services, and impacts "external" to the transportation system (or market).

Examples of externalities are the impacts on the real estate market, urban structure, or the environment such as noise and air pollution. The mathematical functions relating these impacts to physical and functional parameters of the specific transportation systems and, in some cases, to link flows are called impact functions. Often these functions are named with respect to the specific impact they simulate (e.g., fuel consumption functions or pollutant emission functions). Some impacts can be associated with individual network links and depend on the flows, e_l (f). Link-based impact functions are usually included in transportation supply models. Some impact functions may be quite elementary whereas others may require complex systems of mathematical models. Examples of link-based impact functions are those related to air and noise pollution due to vehicular traffic.

General Formulation

To summarize the above points, a transportation network consists of the set of nodes N, the set of links L, the vector of link costs c, which depend on the vector r of link performances, the vector g^{NA} of nonadditive path costs, and the vector e of relevant impact variables: (N, L, c, g^{NA}, e). For congested networks, the link cost vector is substituted by the flow-dependent cost functions c(f); the same holds for flow-dependent internal and external impacts e (f), whereas the nonadditive costs vector g^{NA} is usually assumed to be independent of the flows. In this case, the abstract transportation network model can be expressed as (N, L, c (f), g NA, e(f)). Performance variables and functions are not explicitly mentioned, as they are included in the generalized transportation cost functions. The set of relationships connecting path costs to path flows is known as the supply model. The supply model can therefore be formally expressed by combining (2), (6), and (8) into a relationship connecting path flows to path costs:

$$g(h) = \Delta^T c(\Delta h) + g^N$$

Where it is assumed that nonadditive path costs, if any, are not affected by congestion. Link characteristics can be obtained through performance, cost, and impact functions for the link flows corresponding to the path flow vector. Clearly, model (9) expresses the abstract congested network model described in the previous sections. The same type of model can be used to describe other systems such as electrical or hydraulic networks.

The general structure of a supply model has depicted in Fig. 4. The graph defines the topology of the connections allowed by the transportation study and the flow propagation model defines the relationship between path and link flows. The link performance model expresses for each element (link) the relationships among performances, physical and functional characteristics, and flow of users. The impact model simulates the main external impacts of the supply system. Finally, the path performance model defines the relationship between the performances of single elements (links) and those of a whole trip (path) between any origin-destination pair.

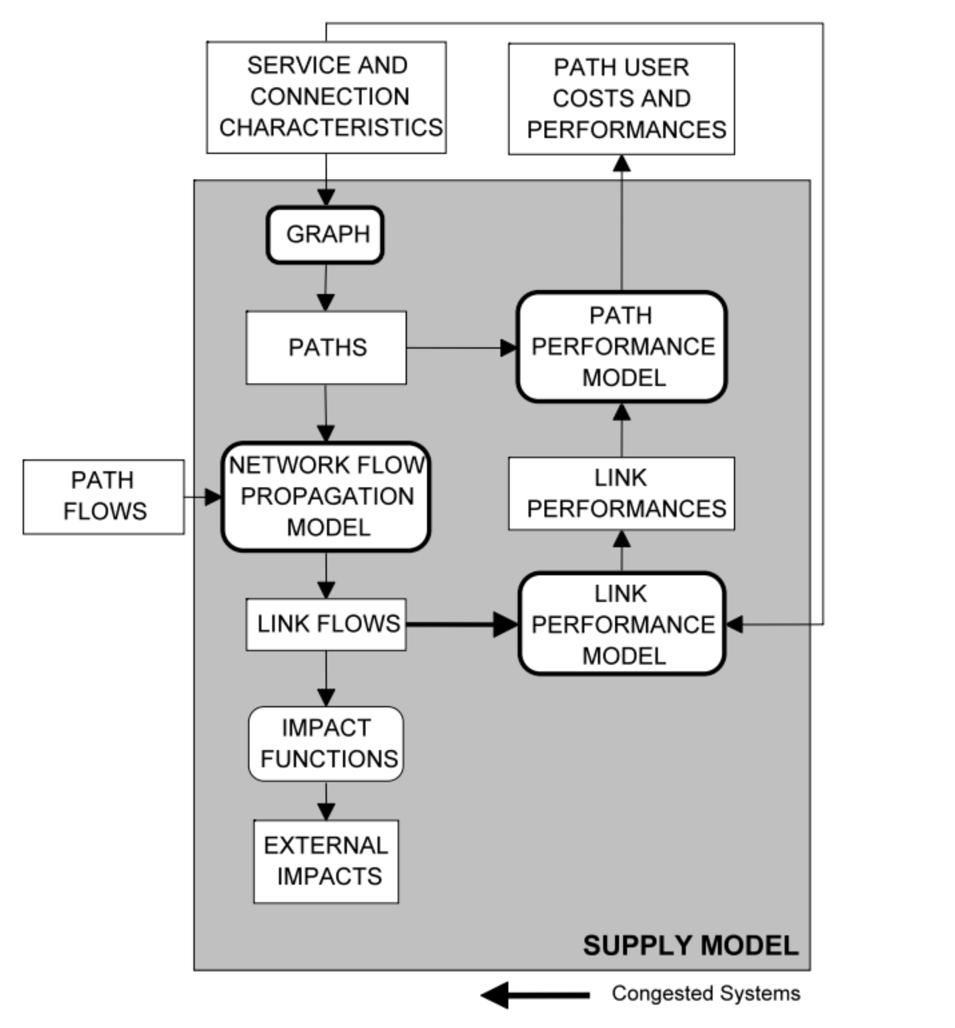


Fig. 4 Schematic representation of supply models.