Automation and Remote Control, Vol. 64, No. 11, 2003, pp. 1651–1689. Translated from Avtomatika i Telemekhanika, No. 11, 2003, pp. 3–46.

REVIEWS

Mathematical Modeling of Traffic Flows

V. I. Shvetsov

Institute for Systems Analysis, Russian Academy of Sciences, Moscow, Russia
Received June 5, 2003

Abstract—Main methods and concepts of mathematical modeling of traffic flows were reviewed. Two important lines of research—modeling of loading of urban transportation network and modeling of traffic flow dynamics—were discussed. For modeling of loading, consideration was given to the models for calculation of correspondences and distribution of flows over the network including different variants of equilibrium distribution and the optimal strategy algorithm. The main classes of dynamic models—macroscopic (hydrodynamic), kinetic, and microscopic—were examined as well.

1. INTRODUCTION

The transportation infrastructure is one of the major pillars supporting life in cities and regions. In many large cities the potentialities of extensive development of the transportation network were exhausted over the last decades or are approaching completion. That is why optimal planning of networks, improvement of traffic organization, and optimization of the routes of public conveyances take on special importance. Solution of these problems cannot do without mathematical modeling of the transportation networks. The main task of the mathematical models lies in determining and forecasting all parameters of the transportation network such as traffic intensity over all network elements, volumes of traffic in the public conveyances network, mean traffic speeds, delays and time losses, and so on.

Exhaustive classification of the mathematical models of transportation networks seems to be impossible owing to diversity of their tasks, mathematical apparatus, source data, and degree of traffic detailing. On the basis of the functional role of models, that is, the problems which are solved with their use, the following three basic classes can be distinguished:

(1) prediction models;

(2) simulation models;

(3) optimization models.

The prediction models are intended to solve the following problem. Let the geometry and characteristics of a transportation network, as well as allocation of the flow-generating urban objects be known. It is required to determine the traffic flows over this network. In more detail, the forecast of transportation network loading includes calculation of the averaged traffic characteristic such as the volumes of interdistrict traffic, flow intensity, distribution of cars and passengers among the routes, and so forth. These models enable one to predict the results of changes in the transportation network or object allocation.

By contrast, simulation modeling is aimed at reproducing traffic in every detail including its time profile. At that, the averaged values of flows and the distributions among the roads are regarded as known and are used as the source data. This distinction can be epitomized as follows: the transportation network prediction models answer the question of *how many* and *where to* and their simulation counterparts detail the traffic, provided that the *how many*s and *where to*s are known. Therefore, the flow prediction and simulation are mutually complementing directions of research.

It follows from the aforementioned that a wide spectrum of the so-called *models of traffic flow dynamics* can be classified by their functional role as the simulation models. Various techniques ranging from simulation of individual cars to describing dynamics of the car density function can be used in the models of this class.

The dynamic models are characterized by a much greater degree of detailing and, correspondingly, much higher need for computer resources. These models enable one to estimate the dynamics of traffic speed, delays at the crossroads, lengths and dynamics of formation of "queues" or "congestions," and other traffic characteristics. The dynamic simulation models are used mostly to improve traffic organization, optimize light signals, and so on. Design of the real-time computer-aided traffic control systems is now topical. These systems must combine sensor information with dynamic simulation. Besides their practical use, the dynamic models are of significant scientific interest in connection with studies of the traffic flow as a physical phenomenon with complicated nontrivial properties such as spontaneous loss of stability, phenomena of self-organization and collective behavior, and so on.

The flow prediction and simulation models are aimed at adequate *reproduction* of the traffic flows. There exist, however, many models for *optimization* of the transportation networks. They optimize the routes of passenger and freight traffic, establish the optimal network configurations, and so forth. The methods of transportation network optimization represent a wide area of research lying out of the scope of this review. Its fundamentals can be found, for example, in [118, 122].

Section 2 of the present review considers the main aspects and principles of modeling the load of urban transportation networks. Methods of calculation of correspondences, distribution of the correspondences over the network, including the networks of street and extrastreet transport, are discussed. Section 3 reviews the main directions in modeling dynamics of the traffic flows, including the macroscopic (or hydrodynamic) models, kinetic models, microscopic simulation models, and those of the kind of cellular automata. Greater emphasis is laid on the main principles and basic variants of the models, whereas the existing modifications and variants of the models are briefly characterized and supplied with the necessary references.

2. MODELS PREDICTING THE LOAD OF TRANSPORTATION NETWORKS

2.1. Main Principles of Load Modeling

The traffic flows are formed by individual movements of the traffic participants or users of the transportation network. In the general case, by the movements are meant not only the trips by various kinds of transport, but pedestrian movements as well. The main factors defining the quantity of movements and their distribution over the urban transportation network are as follows:

(1) Flow-generating factors, that is, distribution of the traffic-generating objects such as place of residence, places of work, cultural and welfare facilities, and so on;

(2) Characteristics of the transportation network such as the quantity and quality of streets and roads, parameters of traffic organization, routes and transportation capacities of the public conveyances, and so on;

(3) Behavioral factors such as population mobility, preferences in choosing the transportation means and routes, and so on.

To construct mathematical models, one must describe the above factors in formal terms. Description relies on the traffic graph whose nodes correspond to the crossroads and stations of the extrastreet transport, whereas the arcs correspond to the segments of streets and lines of the extrastreet transport, as well as to the transfers from street to extrastreet transport. The route graph of public conveyances is an individual component of the transport graph whose nodes are the stops and the arcs are the segments of the routes between the stops. The stop nodes are connected to the ordinary nodes by the boarding and exit arcs.

To describe the flow-generating objects, one needs to divide the town into several conventional arrival—departure (AD) districts. Each AD district is included into the graph as a node connected to the ordinary nodes by special arcs. The total volume of traffic from one AD district to another—independently of particular ways of transportation—is called the *interdistrict correspondence*.

User behavior is modeled on the basis of the mathematical formulation of a criterion used by the user to estimate alternative ways and means of transport. This criterion is called the *generalized path cost*. Higher cost makes a path less attractive. The generalized cost is composed of the generalized costs of the component arcs. The path cost may also include the cost of transitions from arc to arc such as the cost of making a turn upon moving over the street-road network or that of boarding upon passing from the transfer arc to that corresponding to a trip.

The generalized cost of a path is defined as the weighted sum of terms representing the influence of various factors on its estimate. In the general case, it may include the following components:

(1) time of movement calculated from the given function of movement speed vs. arc load. For the roads with different physical characteristics and conditions of traffic control, different speed functions are used;

(2) additional delays on various elements of the transportation network (time of parking, waiting time, and so on);

(3) explicit costs (turnpikes, payments for entering certain urban zones, and so on);

(4) conditional penal additions of time used for modeling various features of the transportation network and measures of transport control.

Studies show that the path cost depends mostly on time. Other factors are of corrective nature; they are expressed numerically in conventional minutes added to the movement time. The path between two points of a network that has the minimum generalized cost among all possible paths is called, therefore, for simplicity the *shortest* path.

The fact that the choice of means and paths of movement made by the network users influences the same choice of other users represents the most important and fundamental feature of forming the load of a transportation network. This mutual influence is described in mathematical terms by the functions of arc cost vs. the total flow along this arcs. This fact establishes a feedback in the course of load formation: the choice of paths forming the load is based on comparing the costs of various paths, whereas the costs themselves are defined by the existing load. The transport flows that are observed in physical networks represent an *equilibrium state* of this process. This equilibrium state is determined using the iterative algorithms described below.

The following four stages are specified in the problem of modeling the traffic flows in the network of a large city:

(1) trip generation, that is, estimation of the total volumes of arrivals to and departures from each urban district;

(2) modal split, that is, split by the modes of transportation such as pedestrian, public conveyances, private cars, and so on;

(3) trip distribution, that is, determination of the correspondence matrices defining the volume of traffic between each pair of the calculated urban districts;

(4) trip assignment, that is, distribution of the correspondences over the transportation network or determination of all paths chosen by the movement participants and the amount of traffic over each path.

This four-stage division of modeling is conventional because all stages are interrelated and, generally speaking, cannot be tackled—due to the aforementioned feedbacks—as independent problems.

For example, the majority of correspondence models make use of the cost of interdistrict traffic as an important factor. Similarly, the split by kinds (for example, between the private and public conveyances) depends on the ratio of costs when using these kinds of transport. Consequently, the correspondences can be calculated and split correctly if the resulting network load is already known. All this leads to the need for iterative solution of the problem.

2.2. Models of Correspondence Calculation

The *correspondence matrix* whose elements are the volumes of traffic (cars or passengers per hour) between each pair of the conventional AD districts characterizes numerically the structure of movements over the network. The entire variety of movements over the network can be decomposed into different groups according to the following criteria:

(1) differences in the aims of movements;

(2) differences in choosing the means of transportation;

(3) differences in the preferences when choosing the paths of movement.

Among the groups of movements with different aims, the following are most important and popular:

(1) movements from the places of residence to the places of work and back (the so-called work correspondences);

(2) movements from the places of residence to the cultural and welfare facilities and back;

(3) movements between the places of work (business trips);

(4) movements between the cultural and welfare facilities.

The matrix of interdistrict correspondences is calculated for each group of movements. The total volumes of arrivals and departures in each AD district are used as the input information for the model of correspondence calculation. Estimation of the volume of arrivals and departures for different groups is concerned with the spatial distribution of the flow-generating objects and *population mobility*, that is, mean number of trips done with one or another purpose. This estimate is based on the available demographic and socio-economic data and the results of studies and for the most part precedes the mathematical modeling proper.

By various means of movement are meant, for example, pedestrian movement, movement by public conveyances, or by private car. From the standpoint of the calculation procedure, the sense of specifying the means of movement is as follows: the chosen means of movement does not vary at the stage of distributing the correspondences over the network. Therefore, the procedure of choosing the way of movement by the user is decomposed into two stages: choice of the means of movement (*modal choice*) and choice of particular way(s) on the basis of some criterion for way estimation (*criterial choice*). Modal choice is realized at the stage of correspondence distribution; criterial choice is realized at the stage of distributing the correspondences over the network.

Grouping of the movement participants by preferences gives rise to the notion of the class of transportation network users. In the general case, the transportation network users are placed among different classes if they employ different path estimation criteria. Here, the notion of "different criteria" includes also the fact that different classes of users can move along different elements of the transportation network. Below, examples of different user classes are presented:

(1) In a network with turnpikes, or paid enters to certain areas, or paid and free parkings in different areas, persons of different means and social standing prefer different ways.

(2) The users of a public-conveyances network may differ in preferences. For example, some may prefer comfortable movement with the minimum of transfers and pedestrian passages or minimized time of reaching the destination.

In order to model the complex network loading with regard for all such factors, users are divided into classes. For each class, individual correspondence matrix is calculated, and the correspondences are distributed over the network. At that, an individual criterion for path optimality is used for each class. The gravitational and entropy models, models of competing possibilities, and some others are the most commonly used models for calculation of correspondences.

2.2.1. Gravitational model [11, 102, 111, 113, 115, 120] historically was one of the first mathematical models proposed to estimate the interdistrict correspondences. Let us consider a system consisting of a set R of the AD districts connected by paths over the transportation network. The source data for calculation of the correspondence matrix are as follows:

O_i, the volume of departures from the district $i \in R$ and

D_j, the volume of arrivals to the district $j \in R$.

Volumes can be measured in cars, passengers, or other appropriate units depending on the type of correspondence. The condition for balance of the general arrival and departure

$$\sum_{i \in R} O_i = \sum_{j \in R} D_j \tag{1}$$

is assumed to be satisfied. If the source data do not meet this condition, they must be corrected by multiplying them by a constant coefficient.

The gravitational model relies on the following simple thesis: the correspondence from the ith district to the jth district is proportional to the total volume of departure from the ith center, the total volume of arrivals to the jth center, and some function $C(t_{ij})$ of the *transport distance* t_{ij} between the ith and jth centers. Intuitively, the transport distance reflects the degree of closeness of districts with regard for the speed and convenience of movements offered by the transportation network. In different variants of the model, this value may be determined by different means.

For calculation of the uniform matrix of correspondences, that is, the correspondences constructed of movements of one type and users of one class, the transport distance is represented numerically as the generalized cost (time of trip in a special case) of the optimal (shortest) way connecting two districts. If mixed correspondences—for example, including trips both by the public conveyances and passenger cars—are estimated, then it is required to calculate the optimal cost t_{ij}^k of movements by different kinds of transport, where k is the type of movement. Then, the weighted averaged value of these costs with regard for the coefficients of correspondence split by the types of movement can be used as the transport distance:

$$t_{ij} = \sum_k c^k(t_{ij}^1, \ldots, t_{ij}^K) t_{ij}^k, \tag{2}$$

where $c^k(t^1, \ldots, t^k)$ are the coefficients of correspondence split into movement types vs. the set of optimal movement times for different types; these coefficients satisfy the condition $\sum_k c^k = 1$.

We denote by F_{ij} the correspondence from the ith district to the jth one. Then, the gravitational model can be formulated as

$$F_{ij} = A_i O_i B_j D_j C(t_{ij}), \quad i, j \in R, \tag{3}$$

where the coefficients A_i, B_j obey the conditions

$$\begin{gathered} \sum_{j \in R} F_{ij} = O_i, \quad i \in R, \\ \sum_{i \in R} F_{ij} = D_j, \quad j \in R, \\ F_{ij} \geq 0, \quad i, j \in R. \end{gathered} \tag{4}$$

The function $C(t)$ is called the *gravitational function*. It is the main factor defining the distribution of movements by distance. Therefore, the term *settling curve* is also used. Some publications treat this function as the "*a priori* probability of generation of correspondence" depending on the distance, although in the general case it needs not satisfy any normalization conditions. This function is chosen in the course of model calibration by comparing the modeled output distribution of distances with the survey data. Many calibration studies were carried out for different cities [14, 90, 93, 116, 117]. The following approximation is often used in practical calculations:

$$C(t) = \exp\left(-\alpha t^{\beta}\right), \quad \alpha \geq 0, \quad \beta \geq 0. \tag{5}$$

For labor correspondences in larger cities, the suitable values of parameters are $\alpha \approx 0.065$ and $\beta \approx 1$.

In the gravitational model, calculation of correspondences comes to determining the coefficients A_i and B_j from the system (3), (4) of nonlinear equations. This problem is solved using the so-called *balancing algorithm*.

Correspondences can be split by types of movements using two methods. First, one can initially estimate the volumes O_i^k and D_j^k separately for each movement type k relying on the socio-demographic indices such as the "motorization coefficient" and others. Then, an individual correspondence matrix is calculated for each type by the balancing method. An alternative approach lies in balancing the matrix of total correspondences according to the general AD volumes and subsequent elementwise splitting by the movement types. At that, the splitting coefficients are established individually for each pair of the AD districts depending on the ratio of costs of different types of interdistrict movements. This approach takes into account the impact of different factors such as the socio-economic characteristics of population, features of district location, and organization of the transportation network on the choice of the way of movement.

2.2.2. Entropy model was proposed by A.G. Wilson [111] to apply the concept of entropy to the transport problems. This approach was developed in numerous publications [36, 83, 112, 113]. The entropy model relies on the probabilistic description of the users of transportation network who are randomly distributed over a set of possible states. For calculation of the correspondences, user membership in the correspondence from i to j can be regarded as a state. Independent random choice by all users of their states leads to one or another macroscopic states of the system. According to the basic concept of the entropy model, the actual system state is that with the highest *statistical weight*. Use of the statistical weight instead of the distribution of probability of one or another state is due to the fact that finite and normalized probability distribution needs not to exist in the entropy models. The statistical weights of states reflect the comparative probabilities of different states in the system. With due regard for this statement, the highest-weight states are often called the most probable states.

In mathematical terms, the state with the highest statistical weight is defined as that maximizing the so-called *system entropy* function in the state space. As applied to the problem of determining correspondences in the transportation network, the entropy obeys the following expression:

$$H(f) = \sum_{i,j} f_{ij} \ln\left(\frac{f_{ij}}{\nu_{ij}}\right), \quad f = \{f_{ij} | i, j \in R\}, \tag{6}$$

where f_{ij} are the numbers of *state filling*, that is, the numbers of system elements that are in the states (i, j), and the values ν_{ij} have the sense of the "*a priori* most probable" values of f_{ij}. In fact, the most probable values of F_{ij} are established by maximizing the entropy under some system of constraints on f_{ij}. In the absence of constraints, maximization provides the *a priori* values $F_{ij} = \nu_{ij}$. The constraints imposed on distributions may be of divers nature. They generally reflect the available information about the macroscopic characteristics of system state. In the system of constraints used in the entropy models of transportation networks, one can specify a group of standard linear constraints expressing the arrival-departure balance. This group is also called the *transport constraints*. In view of the aforementioned, the entropy model of calculations of correspondences is representable as

$$F = \operatorname{argmax}(H(f)), \quad f = \{f_{ij} | i, j \in R\}, \tag{7}$$

$$\sum_{j\in R} f_{ij} = O_i, \quad i \in R, \quad \sum_{i\in R} f_{ij} = D_j, \quad j \in R,$$

$$f_{ij} \geq 0, \quad i, j \in R, \quad g_n(f) = 0, \quad n = \overline{1, N}, \quad h_m(f) \leq 0, \quad m = \overline{1, M}, \tag{8}$$

where the transport constraints are explicit and also N additional constraining equalities and M constraining inequalities are included in general form. The optimization problem (7), (8) is a standard problem of mathematical programming with a convex objective function. In this problem, the system of constraints is usually linear. In the general case, the problem can be solved using the Lagrange multiplier technique. In a special case of only transport constraints, it is possible to obtain an analytical expression to solve problem (7), (8). This expression coincides with that provided by the gravitational model if the *a priori* probabilities are defined according to the gravitation function as $\nu_{ij} = C(t_{ij})$. Another method of relating the entropy and gravitational models was suggested in [112]. Let us assume that the networks between the districts of arrival, i, and departure, j, are related by some amount of "generalized costs." Let the total general costs of movement be known for each departure district:

$$\sum_{j\in R} c_{ij} f_{ij} = c_i, \quad i \in R. \tag{9}$$

These equalities, which are called the *cost constraints*, can be used as constraints for entropy maximization. Let us consider an entropy problem where the *a priori* probabilities of one or another value of correspondence are equal to $\nu_{ij} = \text{const}$ and among the constraints there are those on balance and costs. Formally, solution of this problem also coincides with the expression for the gravitational model if one assumes that $t_{ij} = c_{ij}$ and takes the following gravitational function:

$$C(c_{ij}) = \exp(-\lambda c_{ij}), \tag{10}$$

where the coefficient λ is the Lagrangian multiplier of the optimization problem; its value is established in the course of solving the problem. According to this argumentation, the entropy model can provide a statistical basis for the gravitational model and enable the choice of the gravitational function.

Within the framework of the entropy maximization problem, one can also calculate the correspondences and simultaneously split them by the types of movements. According to the general approach, we assume that the random state of the transportation network user is the membership in the correspondence from i to j and choice of the kth means of transportation. Then, the system state is defined by the three-index set f_{ij}^k. The expressions of the entropy function and the corresponding maximization problem are quite similar to (6)–(8). Here, the constraints must include additional constraints on the total volume of movements for different types [119, 123].

2.2.3. Other models. One of the disadvantages of the classical gravitational model lies in the fact that the volume of correspondence is related with the characteristics (including the transport distance) of a pair of districts taken in isolation from other districts. As was noted by many researchers, "attractiveness" of a district (or the volume of arrivals to it) can also depend on its location between the rest of them [37]. A district located in an agglomeration of many other visited districts can, for example, generate a greater correspondence than an isolated district. This idea was realized in the models of a family of *competing destinations* [27, 28]. The models of competing destinations can be regarded as generalizations of the gravitational model where (3) includes additional factors such as the *visit index* of the arrival district that follows the formula

$$I_{ij} = \sum_{k \in R, k \neq i,j} \frac{D_k}{t_{kj}}. \tag{11}$$

The closer the alternative departure districts to the visit district, the greater the visit index. Introduction of this factor enables one to model the agglomeration phenomena in the structure of correspondences. Further modifications of the model are concerned with attempts to take into account the structure of the district system under consideration. Let us consider, for example, some region comprising larger cities surrounded by smaller centers each of which in turn is surrounded by even smaller regions. The structural effect in this system can manifest itself in excessive attractiveness of the larger center to the surrounding hierarchically "subordinate" centers ("excessive" means here "greater than dictated by the accessibility factors"). This effect is modeled by "ranking" the arrival—departure districts according to their hierarchical status and updating the district visit indices [24, 25].

Another important class of models is represented by various modifications of the Stouffer *intervening opportunities* model [98] which is based on the assumption that the volume of correspondences between two centers is defined more by the number and capacity of the alternative centers of arrival along the path connecting these centers, that is, by the number of alternative *possibilities* of visits, than by the distance between them. Let us first consider a simple system with one center of departure and several centers of arrival lying along one line. Let O be the departure volume, x_n, the correspondence, and λ_n, the probability that the movement participant stops at the nth center if it is reached in the course of the trip. Then,

$$x_n = O\lambda_n \prod_{j=1}^{n-1} (1 - \lambda_j), \tag{12}$$

that is, the volume of correspondence to the nth center is proportional to the product of the probability of stopping at this center by the probability that the movement participant did not stop earlier. The continuous counterpart of the model where the destinations are continuously distributed along a ray is of interest for generalizations. In the continuous model, we consider instead of the correspondences their *density* $x(r)$, where r is the distance from the center of departure. We also denote by $y(r)$ the number of movement participants who reach the point r and by $\lambda(r)$ the value of the density of distribution of the probability of stopping at r provided that this point is reached. Then, obviously,

$$y(r) = O - \int_0^r x(\rho)\, d\rho, \qquad x(r) = y(r)\lambda(r). \tag{13}$$

The following expression of the correspondence density stems from equation (13):

$$x(r) = O\lambda(r) \exp\left(- \int_0^r \lambda(\rho)\, d\rho \right). \tag{14}$$

Different variants of the model of competing opportunities can be obtained from Eq. (14) by assuming different hypotheses about the form of the function of conditional probability density $\lambda(r)$ [94, 110]. As applied to calculation of the correspondences in a transportation network, the conditional probability of stopping at a center is usually related with the center arrival capacity, that is, the number of places of employment, servicing, and so on.

For a generalized model of multiple centers of departures and arrivals, one encounters difficulties with formal determination of the number of stop opportunities "on the way" to the given center. One of the possible approaches lies in ranking the arrival centers by the distance from each departure center. All centers that are (independently of direction) nearer to the departure center than the given one are regarded as alternative opportunities "preceding" the opportunity of stopping at the given center. We make use of (14) and return again to the discrete description of the AD centers to obtain the following expression for correspondence:

$$F_{ij} = O_i \left(\exp(-\lambda U_j) - \exp(-\lambda U_{j+1})\right), \tag{15}$$

where λ is a constant, U_j is the cumulative arrival capacity of all centers preceding—in the aforementioned sense—the jth center.

The main difference between the gravitational models and those of intervening opportunities is as follows: the gravitational models are based on calculations of the transport accessibility of the arrival centers considered mostly in isolation from the alternative centers, whereas the models of intervening opportunities take into account the mutual positions of the alternative arrival opportunities, but explicitly disregard the factor of transport accessibility (distance). In this connection, different variants of model aggregation that take into consideration both factors were proposed. In particular, a combined gravity-opportunity entropy model seeking the distribution of correspondences with the maximum statistical weight was proposed [35]. It makes use of (15) as the "*a priori* probability" of correspondence generation, the accessibility factor being taken into account by introducing the "cost" constraint (9) into the general system of constraints of the entropy problem.

2.3. Models of Flow Distribution

The load of a transportation network is defined by the number of the means of transportation using each network element (arc, turn, or stage on the route of public conveyances). Load modeling lies in distributing the interdistrict correspondences to particular paths connecting pairs of districts. The correspondence matrix or, in the general case, a set of matrices associated with movements of different kinds or different user classes is the input into the load model. Modeling aims at determining for each pair of AD districts the set of paths used for movements between these districts and the correspondence split coefficients (fractions) between these paths.

Upon calculation of the entire system of paths, the load of any network element can be obtained by summing the contributions of all correspondences using the given element. Therefore, load modeling requires a more detailed description of traffic than just determination of the loads of all elements. In the international literature, the models of correspondence distribution over the transportation network are united under the common term of *traffic assignment* .

The existing models of transportation network can be classified by the following main attributes:

(1) models based on the *normative* and *descriptive* approaches;

(2) *static* and *dynamic* models.

In the normative models, the correspondences are distributed by optimizing a global criterion for transportation network efficiency such as the total time consumed by all traffic participants, total run in car×km or passenger×km, and so forth. The descriptive approach is based on the assumption that the structure of traffic flows is formed as a result of individual decisions of the participants based on optimizing their personal criteria. It is traditionally believed that the descriptive approach should be used to model the load of physical transportation networks. The normative models may be used to plan movements in the cases where the planning authority can influence the choice of routes (for example, in planning the centralized freight traffic). The recent interest to the normative models is due to starting projects of centralized control of private car traffic by the on-board computers and satellite communications.

A model is classified with the *static* models if the load is modeled in terms of the averaged traffic characteristic for the chosen period of modeling (morning rush hours, for example). In particular, if some part αF_{ij} of correspondence makes use of the chosen route, then it is assumed that over the entire period of modeling it contributes αF_{ij} to the load of each route element. This assumption is justified if the mean time of all routes does not exceed the characteristic time over which the correspondence itself can change appreciably. If the departure dynamics varies rather rapidly and the routes are sufficiently long, then one must take into account the fact that the representatives of one or another correspondence are loading each route segment at different times. At that, both the correspondence itself and the AD volumes of each district must be defined as time functions. The models with explicit time factor and explicit description of the dynamics of calculated variables over the time of modeling are called the *dynamic* models.[1]

Superposition of each correspondence on the single optimal route connecting two districts is the simplest method of distributing them over the network ("all or nothing" method). Since this method disregards the natural scatter and is oversensitive to the characteristics of individual graph arcs, various techniques were proposed for calculating the alternative paths and scattering correspondences to them. The main difficulty of these models lies in the procedure of constructing logical alternative ways. A procedure for constructing the tree of *possible paths* from all network nodes to a fixed arrival district was proposed in [124]. The path "branches" and the correspondences scatter to the possible paths according to this tree. However, these procedures disregard the following important factors influencing the choice of movement paths.

(1) Choice of a path by some users increases the load of network elements included in this path. As the result, the generalized cost of these elements increases, which in turn affects the estimate and choice of paths by other users. Therefore, the choice of one movement participant indirectly affects the choice of another. It is important to take this factor into account when calculating the load of the street-road network because the time of movement along each its element is strongly dependent on the load of the element.

(2) The fact that the users beginning movement can make decision about their strategies, rather than a particular path, is a distinguishing feature of the system of public conveyances routes. In this case, a particular continuation of the path can depend on boarding on one or another route (simply speaking, on "what bus comes first" at the interchange node).

The model based on determining the *user-equilibrium assignment* [91, 106] is the most efficient one and in full measure takes into account the factor of mutual influence of the users. The model determining the transportation network load by calculating the behavior strategies is called the *optimal strategy* model [97].

2.3.1. Model of Equilibrium Flow Distribution. Let us consider the static variant of the model of user-equilibrium assignment. It assumes that all movement participants choose their paths in terms of the minimum of individual generalized tip cost. The values of flow over all network elements are the result of individual choices. In their turn, the values of intensity represent the main factor defining the generalized cost of elements and affecting the criteria for individual choice.

[1] Traffic modeling uses the term *dynamic* in different senses. Here, we have in view the dynamic assignment models which should not be confused with the dynamic simulation models that are discussed in detail in Section 3. Sometimes, models describing the long-term evolution of a transportation network are also referred to as the dynamic models.

It is assumed that the user-equilibrium assignment of flows establishes as the result of "trials and errors"; it has properties known as the Wardrop requirements [106]:

(1) all paths connecting the districts i and j and used by the representatives of the correspondence F_{ij} have identical costs;

(2) the cost of any path between the districts i and j that is nor used for movement exceeds the costs of the used paths.

The gist of these properties can be summarized as follows: for the user-equilibrium assignment, none of the participants can change its path so as to reduce its individual cost of trip.

The mathematical difficulty involved in determination of the user-equilibrium assignment is due to the fact that this problem is not that of optimization, that is, a definition of the global criterion minimized or maximized on the user-equilibrium assignment lacks in its formulation. However, it is possible to demonstrate that under some additional simplifying assumptions this problem can be reduced to optimization of an especially constructed global criterion.

Let us first consider the problem of distribution of the correspondence matrix of one class of users. We introduce the following notation:

I—the set of network nodes;

A—the set of network arcs;

A_i^-—the set of arcs entering the node $i \in I$;

A_i^+—the set of arcs leaving the node $i \in I$;

u_a–the total flow over the arc $a \in A$;

$u_{a_1a_2}$—the total flow at the turn from the arc $a_1 \in A$ to the arc $a_2 \in A$;

u_a^{pq}—the flow over the arc $a \in A$ of the representatives of the correspondence F_{pq};

$u_{a_1a_2}^{pq}$—the flow at the turn from the $a_1 \in A$ to the arc $a_2 \in A$ of the representatives of the correspondence F_{pq}.

The total flows over the arcs and turns are related to the flows of the representatives of individual correspondences as follows:

$$u_a = \sum_{p,q \in R} u_a^{pq}, \quad u_{a_1a_2} = \sum_{p,q \in R} u_{a_1a_2}^{pq}, \quad a, a_1, a_2 \in A. \tag{16}$$

There exist also linear relations expressing the "law of conservation of network users":

$$u_{a_1}^{pq} = \sum_{a_2 \in A_i^+} u_{a_1a_2}^{pq},\ a_1 \in A_i^-, \quad u_{a_2}^{pq} = \sum_{a_1 \in A_i^-} u_{a_1a_2}^{pq},\ a_2 \in A_i^+, \quad i \in I, p, q \in R. \tag{17}$$

This equality means that the flow over each arc entering a node is equal to the sum of flows of the turns from this arc ("movement ahead" also is regarded as a turn) and the flow over the outgoing arc is equal to the sum of flows of the turns to this arc, and the balance must be observed individually for the representatives of each correspondence. The AD districts are the "sources" and "drains" for the movement participants:

$$F_{pq} = \sum_{a \in A_p^+} u_a^{pq} = \sum_{a \in A_q^-} u_a^{pq}, \quad p, q \in R. \tag{18}$$

We introduce also the *cost functions* expressing the generalized cost vs. the total flow:

$c_a(u)$—the cost function of the arc $a \in A$;

$c_{a_1a_2}(u)$—the cost function of the turn from the arc $a_1 \in A$ to the arc $a_2 \in A$.

A simplifying assumption that the generalized cost is a function of the flow over this arc (turn) and is independent of the flows over other arcs (for example, intersecting the given arc) is essential for constructing the criterion. It is assumed that the cost functions are positive nondecreasing functions of flows. They can be used to construct the *integral* cost functions using the following formulas:

$$C_a(u) = \int_0^u c_a(\nu)\, d\nu, \quad a \in A, \tag{19}$$

$$C_{a_1a_2}(u) = \int_0^u c_{a_1a_2}(\nu)\, d\nu, \quad a_1 \in A_i^-, a_2 \in A_i^+, i \in I. \tag{20}$$

The desired global criterion is constructed as a sum of integral cost functions of all arcs and turns. With the accepted notation, therefore, the model of user-equilibrium assignment can be formulated as the problem of optimization [6]

$$\min_u F(u) = \sum_{a\in A} C_a(u_a) + \sum_{i\in I} \sum_{a_1\in A_i^-} \sum_{a_2\in A_i^+} C_{a_1a_2}(u_{a_1a_2}) \tag{21}$$

under the system (16)–(18) of liner constraints.

The integral criterion (21) used to seek the user-equilibrium assignment is a purely mathematical construct having no descriptive interpretation in the transport or decision-making terms. The reader is referred, for example, to [65] for a strict proof of the fact that minimization of this criterion really provides a user-equilibrium assignment. Here, we present a nonstrict reasoning providing an insight into this criterion.

Let some function

$$F(u) = \sum C_a(u_a) \tag{22}$$

of flow distribution be given which is the sum of functions of flows over each arc. Let us consider the flow distribution minimizing it. The minimum point of any function has the following feature: for any variation of the argument at this point, the function varies nonnegatively in linear approximation. Let us consider a simple variation of distribution where a small fraction δu of flow is moved from one path to an alternative path. Then, variations of the function on each arc in linear approximation are equal to $C_a'(u_a)(\pm\delta u)$. At that, at the arcs of the original path where the flow decreased, the variation is negative, whereas at the arcs of the alternative path it is positive. Since the total variation must be nonnegative, we obtain that the sum of the derivatives $C_a'(u_a)$ along the arcs of the alternative path is not smaller than that along the arcs of the original path. Let us select the functions $C_a(u_a)$ with the idea that their derivatives provide the costs of arcs as in (19), (20). Then, at the point of minimum the total cost of the alternative path is not smaller than that of the original path as is required by the conditions of the user-equilibrium assignment.

The problem of minimization (16)–(21) can be solved by the Frank—Wolfe convergence method [29] which is as follows: let some flow distribution $u^k = \{u_a^{pq}, u_{a_1a_2}^{pq}\}^k$ be reached at the current kth iteration. We linearize the objective function at the point u^k and denote by u_L^k the solution of the linearized problem which is used to determine the "direction of shift" in search of the current approximation, that is, the $(k+1)$st approximation is sought as

$$u^{k+1} = (1-\lambda)u^k + \lambda u_L^k, \quad 0 < \lambda < 1. \tag{23}$$

The variable λ defines the fraction of flows that at the current iteration is transferred from the previous system of paths to a new one. Determination of the fraction to be redistributed at each step is an individual problem. It deserves noting that theoretically the precise value of the fraction is of no importance for determining the user-equilibrium assignment. Namely, one can assume in advance the fraction λ_k to be redistributed at each kth step. Then, if the simple conditions

$$\lambda_k \longrightarrow 0, k \longrightarrow \infty, \quad \sum_k \lambda_k = \infty \tag{24}$$

are satisfied, the sequence of iterations converges to the user-equilibrium assignment. These conditions mean that the fractions decrease (otherwise, the distribution will infinitely oscillate about the equilibrium) yet "not too fast" (otherwise, the redistributed fractions would become too small before reaching the equilibrium). For unsuccessful choice of the sequence of λ_k, however, the algorithm can converge too slowly. Therefore, individual choice of the redistributed fraction at each step is preferable. The best λ_k at each step can be established by solving the auxiliary optimization problem

$$\lambda_k = \arg\min_\lambda F(u(\lambda)), \quad u(\lambda) = (1-\lambda)u^k + \lambda u_L^k. \tag{25}$$

Problem (25) is the general problem (21) "contracted" to a one-dimensional direction obtained at the given iteration. It is a one-dimensional problem without constraints. Moreover, as follows from monotonicity of the cost functions, the objective function is convex and can be easily solved numerically.

Let us consider the linearized problem (16)–(21). At the point u^k, the differential of the objective function is as follows:

$$dF(u^k) = \sum_{a \in A} c_a(u_a^k)\, du_a + \sum_{i \in I} \sum_{a_1 \in A_i^-} \sum_{a_2 \in A_i^+} c_{a_1 a_2}(u_{a_1 a_2}^k)\, du_{a_1 a_2}. \tag{26}$$

The costs involved in this expression correspond to the distribution u^k and in this problem are constant. We get that the linearized problem lies in seeking a distribution minimizing the sum of constant (flow-independent) costs of arcs and turns. Its solution is self-evident: the sum will be minimum if each correspondence is imposed on the single shortest path calculated for these costs. Therefore, no numerical methods of linear optimization are required to solve the linearized problem; it suffices to calculate the shortest-path system for which purpose efficient special algorithms were developed [75, 114].

The primary flows of representatives of individual correspondences $u_a^{pq}, u_{a_1 a_2}^{pq}$ are the primary variables in the problem because the constraints were formulated namely for them. However, the objective function depends only on the total flows over the arcs and turns $u_a, u_{a_1 a_2}$. Next, if the problem constraints are satisfied for the distributions u^k, u_L^k, then they are mechanically satisfied on the solution of the one-dimensional problem. Indeed, if a fixed fraction λ of the representatives of each correspondence is redistributed, then the precisely same fraction of the total flow is redistributed over each arc. The flows of representatives of individual correspondences are considered in the algorithm only when calculating u_L^k by imposing the correspondences on the shortest paths. The method of this calculation itself guarantees satisfaction of the problem constraints. Therefore, we can calculate the user-equilibrium assignment by handling only the values of total flows over the arcs and turns, which saves the main memory space. If nevertheless the complete allotment of the correspondences to the paths is required to analyze the load characteristics, then the calculated paths can be stored in RAM "as they are used," that is, without occupying the main memory.

We obtain the final algorithm for calculating the user-equilibrium assignment of flows:

(1) The initial distribution u^0 is finally generated. The simplest way to do so is to distribute all correspondences among the shortest paths calculated for the unloaded network.

The subsequent iterations are performed as follows:

(2) The costs of all network elements are recalculated according to the values of flows u^k obtained at the given iteration.

(3) The system of shortest paths between the AD centers is determined according to the new costs.

(4) The flows u_L^k resulting from superposition of the correspondences on the shortest paths are calculated (the linearized problem is solved).

(5) A new flow distribution is calculated from $u^{k+1} = (1-\lambda)u^k + \lambda u_L^k$, where λ is defined by the solution of the one-dimensional problem (25).

Let us discuss now the question of a criterion for algorithm stopping. The algorithm must stop if the distribution obtained at the current step is "sufficiently close" to the user-equilibrium assignment. If the distribution is already in equilibrium, then the sum of all costs on the shortest-path system exactly coincides with the sum of all costs on the distribution itself. Indeed, the costs of the shortest paths in equilibrium are exactly equal to the costs of all other paths used. Therefore, the degree of approaching the equilibrium can be estimated at each step by the difference between the sum of all costs for the distribution obtained and the sum of all costs over the shortest paths

$$\frac{\sum c(u^k)u^k - c(u^k)u_L^k}{\sum c(u^k)u^k} < \varepsilon, \tag{27}$$

where summation over all arcs and turns is implied, the indices being omitted for brevity. To obtain a dimensionless characteristic of convergence, the difference of the sums of costs is divided by the total sum of costs.

The following modification of the initial step of the algorithm is useful for practical calculations: the correspondences are split into a fixed number of fractions and, prior to distributing the current fraction, the shortest paths are recalculated taking into account the already obtained load. The initial distribution obtained in this manner is much closer to equilibrium than the distribution of the entire volume of correspondences over a single system of shortest paths, which appreciably reduces the number of subsequent iterations.

2.3.2. Extended models of user-equilibrium assignment. The most important extensions and lines of development of the model of user-equilibrium assignment are as follows:

(1) models of user-equilibrium assignment for multiple user classes;

(2) models of user-equilibrium assignment with variable demand for a flow;

(3) stochastic models of user-equilibrium assignment;

(4) dynamic models of user-equilibrium assignment.

The *multiuser equilibrium* models enable one to determine the user-equilibrium assignment of flows in a system of multiple user classes. We recall that different user classes comprise the movement participants for which the generalized cost differs on the same network elements (the words "cost differs" refer also to the situation where the representatives of one user class are forbidden to move along an arc; in this case it can be said that the "arc cost is equal to infinity"). As a consequence, the representatives of various user classes are distributed among different paths. It is assumed that an individual correspondence matrix is calculated for each user class. At that, since the costs of arcs and turns are functions of the total flow over the arcs (turn), the distributions of the users of different classes are not independent from each other, which gives rise to the problem of system equilibrium.

Since the carriers related with different user classes can make different contributions to the total load (for example, trucks make greater contribution to the load than passenger cars), the total flows should be measured in conventional units (conventional carriers). The correspondence matrices of all user classes must be recalculated in the same conventional units. Theoretically, the very form of the delay functions can differ for different user classes. The following constraint, however, must be observed in the multiuser equilibrium model: for different user classes, the delay functions on each arc can differ only by a constant that does not depend on the total flow. This constraint is of purely mathematical nature and enables one to construct an integral criterion similar to (21) that is minimized on the user-equilibrium assignment. The following "justification" of this constraint is accepted: different user classes have different rates of free movement, but congestion "affects all of them identically."

Let M be the set of all user classes. We denote by u_a^m the flow of the users of class $m \in M$ over a and assume that the cost function of each user class has the form

$$c_a^m(u_a) = c_a(u_a) + d_a^m, \quad a \in A, m \in M, \tag{28}$$

where d_a^m are constants defining the difference in delay between different classes in free movement. Then, in the problem of multiuser equilibrium the criterion is set down as

$$\min_u F(u) = \sum_{a \in A} C_a(u_a) + \sum_{m \in M} \sum_{a \in A} d_a^m u_a^m, \tag{29}$$

where the terms related with the turn delays are omitted for brevity. The algorithm to minimize this criterion is quite similar to the above algorithm for the case of one user class.

The models of user-equilibrium assignment with *variable demand for flow* allow one to determine both the distribution and the correspondences within the framework of a unique algorithm [6, 26, 31, 91]. Let us assume that the volume of correspondence between each pair of the arrival and departure districts is given as a function of the transport distance expressed in terms of the generalized cost of the interdistrict movement:

$$F_{pq} = D_{pq}(c_{pq}), \tag{30}$$

where c_{pq} is the generalized cost of movement from the qth district to the pth district as calculated for the shortest path. If the network flows are in equilibrium, then this cost is equal to that of all paths used to connect these two districts. This function is called the *movement demand function*. It can include parameters characterizing district capacity in arrivals and departures, as well as the parameters characterizing availability of alternative movement opportunities such as public conveyances. The main requirement on this function is that it must be monotonic decreasing (or at least nonincreasing) with the cost of interdistrict movement.

In the problem of user-equilibrium assignment with fixed demand, the correspondences F_{pq} are given as constants. In the problem with variable demand, they are variables related by (18) with flows over arcs u_a^{pq}. The problem of determining the user-equilibrium assignment with variable demand is formulated as follows: needed is to determine a flow distribution such that for each pair of the AD districts the costs of all paths used are equal and do not exceed the costs of all other paths connecting the pair of districts, that is, the conditions for ordinary user-equilibrium assignment are satisfied, and additionally the correspondences F_{pq} satisfy (30). This problem can also be reduced to minimization of the following global criterion:

$$\min_u F(u) = \sum_{a \in A} C_a(u_a) + \sum_{p,q \in R} \int_0^{F_{pq}} D_{pq}^{-1}(f)\, df \tag{31}$$

which coincides with criterion (21) where a term was added comprising an integral of the *inverse* demand function $D_{pq}^{-1}(f)$ and the term referring to turns was omitted for brevity. This function decreases monotonically with increase in correspondence, which guarantees uniqueness of the solution.

This model is mathematically elegant, but in my opinion is not very useful for practical purposes. The correspondence volumes are represented in the model by uniform values that are rigidly related to the unique variant of the network load, whereas in real life the load varies during the day (and also over the year) and the averaged load characteristics obviously affect generation of correspondences. Additionally, the composition of the correspondences themselves is nonuniform. They are the sums of matrices of movements done with different aims, the factor of generalized costs exerting different influence on these movements. For example, the labor movements are less sensitive to the time factor than those "for foodstuff shopping." On the whole, calculation of the correspondence matrices has many aspects that do not fit the procedure of determining the equilibrium.

At the same time, the problem of correspondences cannot be completely segregated from that of correspondence distribution over the network. Indeed, the existing models for calculation of correspondences among other factors take into account the interdistrict transport distances expressed in terms of the generalized movement costs which in turn depend on the load. For calculations to be correct, it is required that the generalized costs used to calculate the matrices correspond to the costs obtained by distributing these matrices over the network. This can be done by iterative calculation of the matrices and network load. At each iteration, the matrices are determined from the costs obtained at the preceding step from load calculations. At the first step, one can use the costs determined for an "empty" network. Iterations are repeated until the variations of general indices over an iteration step become sufficiently small. In applications, the moving average method stabilizes the iterative process (variations of the general indices are less than one percent) in 3 to 5 iterations.

The above equilibrium models proceed from the assumption of *deterministic* behavior of the users. They assume that the users have exact notion of the flows and delays over all arcs, identically estimate the paths, and make exact decisions when choosing a path. In actual fact, however, the users behave stochastically, that is, with an appreciable element of chance. The model of *stochastic user-equilibrium assignment* [22, 67] was proposed to take this stochasticity into account. It relies on the distinction between the *actual* path cost and that *anticipated* by one or another user. The anticipated cost may be regarded as a random variable. Stated differently, for the same arc, various users conjecture at random different costs. The condition for stochastic equilibrium is then formulated as follows: the distribution is called the user-equilibrium assignment if none of the movement participant *assumes* that it can improve its individual cost of trip by changing its route.

The problem of seeking the stochastic equilibrium can be reduced to the following optimization problem [92]:

$$\min_u F(u) = \sum_{a\in A} u_a c_a(u_a) - \sum_{a\in A} C_a(u_a) - \sum_{p,q\in R} F_{pq} S_{pq}(u), \tag{32}$$

where the functions $S_{pq}(u)$ are the expectations of the anticipated costs of movement between the pth and qth districts calculated for some flow distribution u. We denote by K_{pq} the set of alternative paths connecting the pth and qth districts and by c_k^{pq} the actual cost of the path $k \in K_{pq}$. Then, the anticipated cost is representable as $\widehat{c}_k^{pq} = c_k^{pq} + \xi_k^{pq}$, where ξ_k^{pq} is a random variable with zero expectation. Correspondingly, the expectation $M\widehat{c}_k^{pq} = c_k^{pq}$. Then,

$$S_{pq}(u) = M\left[\min_{k\in K_{pq}} \widehat{c}_k^{pq}\right]. \tag{33}$$

This function depends on the distribution u because the actual costs for which the expectation is taken depend on u.

Numerical solution of (32) is much more complicated than that of the corresponding nonstochastic problem. The main difficulty here lies in the fact that in order to calculate a current approximation one needs to know not only the total load obtained at the preceding steps, but also all interdistrict paths and costs obtained in the course of iterations. Solution is based on the method of moving average for the fraction λ of flows redistributed at each iteration [84].

Recently, emphasis in the load models is made on developing *dynamic* models [10, 30, 50]. Time used as an additional variable extremely complicates the problem. The matter concerns not only the increased problem dimensionality, but also the difficulties in theoretical definition of the basic notions such as the cost function and equilibrium itself [51]. Further, in order to keep track of the exact time of movement along one or another arc of a path, one needs a more detailed description of the conditions of movement *inside* the arc, that is, some simulation models [18]. Even for simple simulation models, the need for computer resources increases dramatically owing to the iterative nature of seeking the equilibrium. Detailed discussion of this direction of research lies outside the scope of the present review. We only note that the existing works on *practical* use of the dynamic equilibrium models rely on supercomputers and parallel computer systems [19].

2.3.3. Model of Optimal Strategies. The process of formation of the public conveyances network has some features that are not characteristic of the load of the motor transport network. Whereas the user of the street-road network in equilibrium follows the optimal path to the destination, the user of the passenger transport network can determine its *optimal strategy of behavior* in the course of moving to the destination. By the strategy of passenger behavior in the public conveyances network is meant the set of rules which enable the user to reach the destination. Movement along an *a priori* chosen path, which corresponds to the behavior of movement participants in the user-equilibrium assignment models, is a simplest example of strategy. More complicated strategies arise if in the course of movement the passenger makes one or another decision about continuing the path depending on the information obtained in the course of movement. For example, a decision made at a current interchange node may depend on the carrier that leaves first. More complicated strategies provide, for example, the possibility of changing the carrier if the passenger sees from the bus window that it is possible to take a limited-stop bus, and so on.

The standard model of optimal strategies relies on a simplified description of user behavior. According to this model, the choice of strategy is as follows:

(1) for each node where the passenger can find himself in the course of moving to the destination, some set chosen from all possible continuations is fixed. The chosen continuations will be said to be "included in the strategy." Importantly, the set must be fixed "in advance," that is, not in the course of movement, but when choosing the strategy;

(2) the users finding themselves in one or another node choose from the strategy the continuation which first offers the opportunity of service (departure of the carrier).

Since carrier arrivals and departures can be regarded as random events with some distribution law, the particular paths are realized with one or another probability. Therefore, the model of optimal strategies is intrinsically stochastic.

For mathematical description of the model, it is necessary to extend the transport graph to the *route* graph. The routes of public conveyances are described by additional nodes and arcs. We call the nodes and arcs of the ordinary graph the basic nodes and arcs of the route graph. Let us consider a basic arc along which k routes of public conveyances pass. Over each such arc, there are k route arcs corresponding each to a trip along this arcs on one of the routes. The basic arc itself is also included in the graph as a pedestrian arc.[2] In order to simplify the description, it is assumed that the stops of public conveyances are always located at the nodes of the basic graph. All the route nodes where stops are made are connected to the corresponding basic node by the conventional arcs of boarding and exit.

The above definition of strategy can be formulated in terms of the route graph as follows: at each node of the route graph, a set of some outgoing arcs included in the strategy is fixed. Upon getting to a node, the user can continue along one of the fixed arcs. Therefore, definition of strategy

[2] With the exception of the arcs corresponding to the spans of extrastreet transport along which there is no pedestrian movement.

amounts to definition of a *subgraph* of the entire route graph. We call this subgraph the *strategy graph*.

A strategy is referred to as admissible if (1) it has no "cycles," that is, one cannot return to the already passed node by moving along the arcs included in the strategy, and (2) enables attainment of the destination, that is, the user always reaches the destination node in a finite number of steps by choosing arbitrarily at each node an arc from those included in the strategy.

Two characteristics are given for each arc of the route graph:

generalized arc cost c_a including as usual the mean time of moving along the arc and other additions in minutes (in what follows, we call it simply time) and

servicing frequency f_a which makes sense only for the boarding arcs and numerically is equal to the mean number of carrier departures from the given node along the route declared for boarding. The mean departure interval is $1/f_a$.

For uniformity of describing the transport graph, one can assume that all arcs have these two characteristics, the frequency of servicing all arcs that are not those of boarding being infinity. Correspondingly, the mean waiting time at these arcs is zero.

If we assume that the time of passenger arrival to a node is uniformly distributed and the departure interval $1/f_a$ is constant, we obtain that the mean time of moving along an arc, including the time waiting, is equal to $1/2f_a+c_a$. For this strategy, the mean time of reaching the destination from any node can be calculated with the knowledge of the generalized cost and frequency of all arcs. This can be done recurrently. Let k be the number of the destination node and t_i, the mean time of reaching the node k from the node i if using the given strategy. Obviously, $t_k = 0$. Let us consider an arbitrary node i. The arcs $a = (i, j) \in A_i^+$ are possible continuations according to the chosen strategy. Let t_j be already calculated for all finite nodes of these arcs. Then, the mean time t_i obeys the recurrent formula

$$t_i = \left(1/2 + \sum f_a(c_a + t_j)\right) / \sum f_a \tag{34}$$

which enables calculation of the mean time of any node.

By the *optimal* strategy for reaching the kth destination from the ith node is meant the strategy with the least mean time among all admissible strategies. In the optimal strategy, the mean time can also be called the *potential* of the ith node relative to the kth node.

The optimal strategy has the following important property: the optimal strategy of any departure node i also is the optimal strategy for all intermediate nodes j that may occur upon moving according to this strategy. This property is quite similar to the Bellman property for the shortest paths according to which an interval of the shortest path from some intermediate node to the destination node is itself the shortest path from this intermediate node. Indeed, if addition to or elimination from the jth node of some outgoing arc reduces its potential, then obviously the potentials of all nodes from which one can get to the jth node must decrease as well. This property allows one to determine a strategy of reaching the kth destination that is optimal simultaneously for all the departure nodes. Indeed, in each network node let us label the outgoing arcs included into the optimal strategy of this node. Their totality, obviously, defines the optimal strategy of all nodes. To calculate the optimal strategy for an arbitrary destination node, that is, to specify a subgraph of the route graph, an algorithm is used that is similar to that of layerwise extension upon determining the shortest paths.

Let us consider now the problem of distributing the correspondences F_{pq} over the route graph. We assume that all movement participants (passengers) follow the optimal strategies. Since in the model of optimal strategies the arc costs and frequencies are regarded as constant, it suffices to select all the arrival districts q, calculate the optimal strategy for each of them, distribute according to this strategy the correspondences from all other districts to the given district, and sum the resulting flows. We denote by u_i the incoming flow of the node $i \in I$:

$$u_i = \sum_{a \in A_i^-} u_a. \tag{35}$$

For the departure districts, the volume of departures $u_p = F_{pq}$ is regarded as the "input" flow. The flow u_i is distributed to the outgoing arcs included in the optimal strategy. The fraction of flow distributed to each outgoing arc is proportional to the probability that it is namely this arc will be the first to offer its services. Then,

$$u_a = u_i\, f_a / \sum f_b. \tag{36}$$

To save calculations, it is required when calculating a strategy to store the list of all network nodes (including the departure districts) ordered in the potential relative to the district of arrival. The superposition of the flows on the outgoing arcs begins then from the highest-potential departure district and is continued for all nodes in the descending order of potentials. This scheme guarantees that upon passing to the current ith node the incoming flow u_i will be already calculated. By repeating the distribution procedure for all arrival centers, we accumulate the total flows on all arcs of the route graph.

3. DYNAMIC MODELS OF TRAFFIC FLOW

The majority of the dynamic existing models of traffic flow can be roughly sorted into the following three classes:

(1) macroscopic (hydrodynamic) models;

(2) kinetic (gas-dynamic) models;

(3) microscopic models.

By the macroscopic models are meant those describing car movement in averaged terms such as density, mean speed, flow, and so on. With this approach, the traffic flow becomes similar to motion of a specific fluid. Therefore, the models of this class are also called the *hydrodynamic* models.

By the microscopic models are meant those explicitly describing the movement of each car. As compared with its macroscopic counterpart, this approach theoretically enables more precise description of the movement of each car. However, its application requires more computer resources.

The intermediate place is occupied by the kinetic approach which describes the flow by the density of distribution of cars in the phase space, that is, in the space of car coordinates and speeds. Dynamics of the phase density follows the *kinetic* equation based on the averaged description of interactions of individual cars. In this sense, the equation is closer to the microlevel than the hydrodynamic equation. Theoretical importance of the kinetic models lies in the fact that they enable methodical derivation of the macroscopic models.

A special place in the class of micromodels belongs to the models of the "cellular automata" type that recently have received wide recognition. They describe car movement in time and space in very simplistic discrete terms, which makes them very effective for computations.

3.1. Macroscopic Models

The first macroscopic model based on the hydrodynamic analogy was proposed by Lighthill and Whitham in 1955 [66]. It is known in the literature as the LW-model after the names of its authors. We denote by $\rho(x, t)$ the density and $V(x, t)$ the mean speed of cars at the road point with the coordinate x at time t. These values are related by the following continuity equation expressing the law of "conservation of the number of cars" on the road:

$$\partial_t \rho + \partial_x(\rho V) = 0. \tag{37}$$

It is assumed that the mean speed is a deterministic (decreasing) function of density:

$$V(x,t) = V_e(\rho(x,t)). \tag{38}$$

By substituting (38) into (37), we obtain

$$\partial_t \rho + \left(V_e + \rho \frac{dV_e}{d\rho}\right) \partial_x \rho \tag{39}$$

which describes propagation of the nonlinear *kinematic waves* with the speed of transfer $c(\rho) = V_e(\rho) + \rho dV_e/d\rho$. The wave profile becomes with time more and more steep right up to the vertical line where it becomes discontinuous (*shock waves*) [108, 109]. Different variants of the shock wave model were used by many authors [20, 32, 73] to describe the dynamics of congestions. Although the original LW-model now seems to be oversimplified, it is namely its simplicity that allows one to employ the model as an auxiliary tool for modeling the load dynamics of larger traffic networks [18, 19].

In real life, the car density usually does not vary discretely, but is rather a continuous function of the coordinates and time. A term of the second order describing the diffusion density was added to (39) with the aim of eliminating the shock waves by smoothing the wave profile:

$$\partial_t \rho + V_e \partial_x \rho = -\rho \frac{dV_e}{d\rho} \partial_x \rho + D\partial^2_{xx}\rho. \tag{40}$$

Relation (38) which proceeds from the assumption that at each time instant the mean flow speed $V(x,t)$ corresponds to the equilibrium value V_e for the given car density is the most serious disadvantage of the model. Therefore, the model is not adequate to reality in describing the nonequilibrium situations arising near road nonuniformities such as descents and exits, as well as in the conditions of the so-called "stop-and-go" movement.

Instead of the deterministic relation (38), the differential equation modeling the mean speed *dynamics* was suggested for describing the nonequilibrium situations. This equation was proposed first in 1971 [77] as

$$\partial_t V + V \partial_x V = -\frac{C(\rho)}{\rho} \partial_x \rho + \frac{1}{\tau}\left(V_e(\rho) - V\right), \tag{41}$$

where

$$C(\rho) = -\frac{1}{2\tau}\frac{dV_e}{d\rho}. \tag{42}$$

It relies on the microscopic description of movements of individual cars according to the *follow-the-leader* model (see Section 3.4.1). The term $V\partial_x V$, which is called the *convection* term, describes the variation of speed at the given place of the road owing to kinematic transfer of the cars from the preceding road segment to the given point with the mean flow speed. The first term in the right-hand side, which is called the *anticipation* term, describes the drivers' response (braking or speedup) to the situation ahead of them. The second term in the right-hand side, which is called the *relaxation* term, describes the tendency of approaching the mean speed V to the equilibrium value $V_e(\rho)$ for the given density, τ being the characteristic relaxation time.

Numerous modifications of the speed equation and the numerical method to solve it were proposed later [16, 17, 74, 78, 79, 88]. One of the most substantial disadvantages of equation (41) lies in stability in linear approximation of the stationary uniform solution $\rho(x,t) \equiv \rho_0, V(x,t) \equiv V_e(\rho_0)$ to smaller perturbations for all values of density. Analysis of the empirical data shows, however, that for higher values of density the laminar motion of the traffic flow becomes unstable, and small perturbations lead to *fantom congestions* or wave of the stop-and-go movement. This problem can be tackled by the following modification of the anticipation term:

$$C(\rho) = \frac{d}{d\rho}\mathcal{P}_e(\rho), \qquad \mathcal{P}_e(\rho) = \rho\Theta_e(\rho), \tag{43}$$

where $\mathcal{P}$ is the *"flow" (traffic) pressure* in terms of the flow speed variation Θ. Under this replacement, the speed equation becomes as follows [80, 81]:

$$\partial_t V + V\partial_x V = -\frac{1}{\rho}\partial_x \mathcal{P}_e + \frac{1}{\tau}\left(V_e(\rho) - V\right). \tag{44}$$

According to it, the cars slow down if the flow pressure ahead of them increases, and speed up, otherwise. It is assumed that the speed variation Θ depends on the flow density. Various empirical approximations to this relation are used. In particular, the models of [52, 60, 61] use the positive constant $\Theta_e(\rho) = \Theta_0$ as the first approximation.

Like (37), (38), Eq. (44) predicts the shock waves. To smooth and avoid discontinuities in solutions, a term of the form $\nu\partial^2_{xx}V$, which is similar to the term describing viscosity in the equations of the classical hydrodynamics, is introduced into the right-hand side of the equation. Finally, the speed equation becomes as follows:

$$\partial_t V + V\partial_x V = -\frac{\Theta_0}{\rho}\partial_x \rho + \nu\partial^2_{xx}V + \frac{1}{\tau}\left(V_e(\rho) - V\right). \tag{45}$$

Analysis of stability of the stationary uniform solution of this equation shows that for values of density exceeding some critical value ρ_{cr} solution becomes unstable to small perturbations (the instability domain and the increment also depend on the perturbation wavelength). This fact allows one to model theoretically phenomena due to the fantom congestions arising in the uniform flow as the result of random perturbations.

In the class under at hand, the Kerner-Konhäuser model [52] (KK-model) is the most popular. Its formulation lies in Eqs. (37) and (45) where $\nu(\rho) = \nu_0/\rho$ (with ν_0 for the viscosity coefficient) is used instead of the constant ν in order to improve compatibility with the Navier—Stokes equation of the classical hydrodynamics.

Analysis of stability of the stationary uniform solution of the KK-model shows that it is stable to small and very large densities and has an instability domain for mean densities. Computer-aided calculations of this model enabled one to study the process of formation and development of *clusters*, that is, isolated moving domains featuring high density and low speed of flow [53–56].

All the aforementioned models, however, are subject to certain qualitative disadvantages. For example, for some values of the parameters they may predict densities that exceed the maximum admissible values ("bumper-to-bumper"). Additionally, for strong spacial nonuniformities in the initial conditions, negative speeds can arise (congestion is "dissolved backward" as the result of viscosity) [21, 39, 76].

The above macroscopic models were formulated mostly by analogy with the classical thermodynamic equations. There exists another method of constructing the macroscopic models from the description of car interaction at the microlevel using the kinetic equation. This approach was proposed by Prigogine by analogy with the derivation in statistical physics of the gas dynamics equations from the kinetic equation for the phase density. In this way, one can obtain some of the aforementioned models, as well as much more detailed ones such as models using differential equations to describe the dynamics of speed variations, models of multilane movement, and so on (see Section 3.3).

3.2. Kinetic Models

In distinction to the hydrodynamic models that are formulated in terms of density and mean flow speed, the kinetic models are based on describing dynamics of the flow *phase density*, that is, the density of car distribution both in the coordinate and individual speed. With the knowledge of the time profile of phase density, one can also calculate the macroscopic characteristics of the flow such as density, mean speed, speed variations, and others that are defined by the *moments* of phase density in speeds of different order.

We denote the phase density by $f(x, v, t)$. The ordinary (hydrodynamic) density $\rho(x, t)$, mean speed $V(x, t)$, and speed variation $\Theta(x, t)$ are related with the moments of phase density as follows:

$$\rho(x,t) = \int_0^\infty dv\ f(x,v,t), \tag{46}$$

$$V(x,t) = \rho(x,t)^{-1} \int_0^\infty dv\ vf(x,v,t), \tag{47}$$

$$\Theta(x,t) = \rho(x,t)^{-1} \int_0^\infty dv\ (v-V)^2 f(x,v,t). \tag{48}$$

The differential equation describing the time profile of phase density is called the kinetic equation. For the traffic flow, the kinetic equation was first formulated by Prigogine and his colleagues in 1961 as follows [85–87]:

$$\partial_t f + \partial_x (fv) = \left(\frac{\partial f}{\partial t}\right)_{\text{int}} + \left(\frac{\partial f}{\partial t}\right)_{\text{rel}}. \tag{49}$$

Like (37), it is a continuity equation expressing now the car conservation law in the phase space. The terms in the left-hand side describe variations of the phase density caused by the kinematic transfer, whereas the terms in the right-hand side describe "instantaneous" changes in the car speeds caused by *interaction* and *relaxation*.

According to Prigogine, by interaction of two cars is meant an event where a faster car catches up with the slower one. In this case, the driver of the faster car either overtakes the car in front or slows down to its speed. Consequently, the term $(\partial f/\partial t)_{\text{int}}$ in the right-hand side of the kinetic equation can be expressed in the number of interactions resulting in braking at the given place in time unit. The following simplifying assumptions are introduced:

(1) the possibility for overtaking is determined with some probability p; as the result of overtaking, the speed of the faster car is not changed;

(2) the speed of the car going in front does not change in any case as the result of interaction;

(3) the interaction occurs at a point (the sizes of cars and the distance between them are negligible);

(4) as the result of interaction, the speeds change in a moment;

(5) consideration is given only to the pair interactions, simultaneous interactions of three or more cars being ruled out.

Under these assumptions, the term describing interactions can be formulated as

$$\left(\frac{\partial f}{\partial t}\right)_{\text{int}} = \int_v^\infty dv'(1-p)(v'-v)f_2(x,v',x,v,t) - \int_0^v dv'(1-p)(v-v')f_2(x,v,x,v',t), \tag{50}$$

where $f_2(x, v, x', v', t)$ is the joint density of distribution of car pairs with the respective coordinates and speeds (x, v) and (x', v'). The first integral in (50) describes the increase in the phase density owing to deceleration of the faster cars to the speed v. The second integral describes the process of reduction in density owing to deceleration of the cars cruising at the speed v to lower speeds as the result of interaction. If we assume now additionally that

(6) by analogy with the hypothesis of *"molecular chaos"* in the classical gas-dynamics, the hypothesis of *"car chaos"*

$$f_2(x, v, x', v', t) = f(x, v, t)f(x', v', t),$$

is valid, that is, the speeds of cars in the flow are not correlated prior and after the interaction, then the integrals in (50) can be formally united into one integral from 0 to ∞ and the interactive term takes the form

$$\left(\frac{\partial f}{\partial t}\right)_{\text{int}} = f(x, v, t)\int_0^\infty dv'(1-p)(v'-v)f(x, v', t). \tag{51}$$

Besides interactions leading to braking, spontaneous increase in the speeds of cars that are not currently interacting with a car ahead of them occurs in the traffic flow. We denote by $F(v; x, t)$ the distribution of speeds in the flow at the point x at the time t. The distribution of speeds satisfies the following relations:

$$f(x, v, t) = \rho(x, t)F(v; x, t), \qquad \int_0^\infty dvF(v; x, t) \equiv 1. \tag{52}$$

Empirical data show that a natural (equilibrium) speed distribution $F_0(v; x, t)$ establishes in the free flow which can be interpreted as a distribution of drivers between the *"desired"* speeds. By the "desired" speed is meant that with which each driver would move in the absence of hindrances and interactions. According to Prigogine, the tendency of all drivers to move at the desired speeds leads to the effect of collective relaxation of the actual speed distribution to the "desired" distribution. If the characteristic time of this process is denoted by τ, then the term $(\partial f/\partial t)_{\text{rel}}$ in the right-hand side of the kinetic equation is representable as

$$\left(\frac{\partial f}{\partial t}\right)_{\text{rel}} = -\frac{f(x, v, t) - \rho(x, t)F_0(v; x, t)}{\tau}. \tag{53}$$

Two parameters—the probability p of overtaking and the relaxation time τ—remain undefined in the kinetic equation. It is usually assumed that they characterize collective behavior and, therefore, are functions of the *macroscopic* flow characteristics such as flow density and mean speed and are independent of the *individual* values of speed. With this assumption, $(1 - p)$ can be factored out of the integral in (51), as well as the parameters p and τ can be factored out of the sign of integration when integrating the kinetic equation with respect to speeds, which substantially simplifies derivation of the macroscopic equations (see Section 3.3).

Summation of (49), (50), and (53) provides the following kinetic equation:

$$\partial_t f + v\partial_x f = -\frac{f - \rho F_0}{\tau} + (1-p)\rho(V - v)f. \tag{54}$$

Its examination reveales nontrivial linear properties admitting its descriptive interpretation in terms of traffic flow. For different values of density, this equation admits two stationary spatially uniform solutions with basically different properties of which one solution has a speed distribution corresponding to the desired distribution and the other solution provides identical speed to all cars. They are interpreted as the *modes of individual* and *collective movements*, respectively. The general stationary uniform solution is a superposition of the above solutions. With increase in density, the fraction of cars participating in the collective movement increases and that of the freely moving cars decreases. Division of flow into two fractions is similar to condensation in some models of statistical physics. Stability analysis shows that the stationary uniform solution is stable for smaller densities (in the individual movement mode). As density exceeds some critical value, instability of long-wave perturbations occurs. Finally, in the purely collective movement the flow is unstable to perturbations of any wavelength.

The works of Prigogine and his collaborators were of great importance for developing the kinetic approach to modeling the traffic flow, but Eq. (54) itself time and again was subjected to criticism. One of the weakest points of this theory is the hypothesis of "car chaos" which assumes that the car speeds between interactions are not correlated. It is, however, obvious that if overtaking is impossible, then the overtaking car is forced to follow the car in front of it. During the following time interval their speeds cannot be regarded as statistically independent. Various modifications of the kinetic equation were proposed to take this effect in account. For example, it was proposed to divide the common phase density into the flow densities of freely moving and "blocked" cars [1]. Methodical development of this approach led to the need for modeling the dynamics of density of distribution of small clusters of cars. At that, the distribution of clusters of a fixed length n obeys an individual distribution function for each n. The processes of cluster increase owing to new cars joining from behind and decrease owing to overtaking are modeled by corresponding modifications of the interactive terms in the right-hand sides of equations [2, 7, 49, 64].

Another challenged aspect is the principle of collective relaxation of speed distribution. The kinetic equation with the relaxation term of the form (53) was shown [76] to describe inadequately behavior of the traffic flow if the flow is spatially nonuniform. Let us consider, for example, a free road segment to which a flow arrives. Obviously, the faster cars will be the first to reach the final road segment, whereas the slower ones need some time. According to (54), however, the distribution of speeds along the entire road will approach the desired distribution from the first instants, that is, slower cars will immediately appear all along the road. One might say that the desired speed distribution stands in the model as the property of the road, rather than the drivers. To remove this drawback, Paveri-Fontana [76] proposed an important improvement in the kinetic equation based on replacing the collective relaxation of speed distribution by the *individual relaxation* of the speed of each driver to the desirable speed.

We assume that each driver has the desirable individual speed w. Then, we can consider the density of car distribution in the *extended phase space* $g(x, v, w, t)$, where x is the coordinate, v is the actual speed, and w is the desirable speed. The phase density is related with the extended phase by the following equation:

$$f(x, v, t) = \int_0^\infty dw \; g(x, v, w, t). \tag{55}$$

Time variations of the extended phase density obey the following equation:

$$\partial_t g + v\partial_x g + \partial_v (a(v, w)g) = \left(\frac{\partial g}{\partial t}\right)_{\text{int}}, \tag{56}$$

where $a(v, w)$ is the acceleration of the car moving with the speed v and having the desirable speed w. The main distinction of (56) from (49) lies in the fact that a term related with individual acceleration of cars arises in the left-hand side instead of the term describing the collective relaxation. We can assume as a first approximation that

$$a(v, w) = \frac{w - v}{\tau}, \tag{57}$$

that is, each driver tries to approach the desirable speed; here, τ is the characteristic relaxation time.

The interactive term is obtained similar to (50) by calculating the interactions per time unit:

$$\left(\frac{\partial g}{\partial t}\right)_{\text{int}} = \int\limits_{v}^{\infty} dv'(1-p)(v'-v)g(x,v',w,t)f(x,v,t)$$
$$-\int\limits_{0}^{v} dv'(1-p)(v-v')g(x,v,w,t)f(x,v',t). \tag{58}$$

By integrating (56)–(58) with respect to w, one can obtain the equation for the ordinary phase plane that is called the *reduced* Paveri-Fontana equation:

$$\partial_t f + v\partial_x f + \partial_v\left(\frac{W-v}{\tau}f\right) = (1-p)\rho(V-v)f, \tag{59}$$

where $W = W(x,v,t)$ is the mean value of the desired speed at the given point of the phase space:

$$f(x,v,t)W(x,v,t) = \int dw\; wg(x,v,w,t). \tag{60}$$

The reduced equation is not, strictly speaking, closed in f because calculation of W requires the extended density g. However, it is namely the reduced form that is used in many works dealing with derivation of the macroscopic equations from the Paveri-Fontana equation; at that, simplifying assumptions are made to calculate W. In the simplest case, one may assume that all cars have identical desirable speed or that the cars are divided into classes of passenger and freight cars and that the cars of one class have the same desirable speed.

Both the original Prigogine equation and that of Paveri-Fontana consider cars as point objects, their interaction being regarded as localized at a point. In the real flow, however, the sizes of cars are not negligible. Additionally, the car responds to an obstacle in front of it at some distance that also cannot be regarded as negligible, especially for high flow density. Helbing [40, 42] modified the kinetic equation with the aim of taking this factor into consideration. To describe traffic, he used the methods employed by statistical physics to describe motion of free-flowing materials such as sand or grains. Modification concerns the interactive term where the intensity of interaction of fast cars ($v' > v$) with the car moving at the speed v obeys the formula

$$\chi(x,t)\int\limits_{v'>v} dv'(v'-v)f_2(x,v',x+s,v,t), \tag{61}$$

where f_2 is the pair density of distribution of the interacting cars. In contrast to (50), the coordinate of the second (moving ahead) car is shifted by some "interaction distance" s which can also be interpreted as the "safe distance"

$$s = \frac{1}{\rho_{\max}} + TV, \tag{62}$$

where $\rho_{\max}$ is the maximum ("bumper-to-bumper") density and T is the safe time interval observed by the drivers.[3] The multiplier $\chi(x,t)$ plays the part of *effective section of interactions* and reflects the increase in the virtual number of interactions with increase in the flow density. The fundamental work [40] proposed the expression

$$\chi(x,t) = \frac{1}{1-\rho(x+TV,t)s} \tag{63}$$

[3] Empirical studies show that the distance observed by the drivers depends on the flow speed, whereas the time interval is invariant.

according to which the interaction intensity becomes infinite if the mean distance between the cars in flow becomes equal to the interaction distance. Subsequent works [43, 47] suggested another expression providing correct asymptotics of solutions for $\rho \to 0, \rho \to \rho_{\max}$.

The kinetic models also were generalized to the case of multilane movement and movement of a mixed flow consisting of users of different classes such as different types of carriers or drivers with different types of behavior [42, 49, 95]. With that end in view, it is necessary to introduce the phase densities of the cars of different user classes on each lane and add to the right-hand sides terms describing the process of lane changing. We describe below the multilane model of [95].

Let $f_i^a(x, v, t)$ be the phase density of the cars of the class $a \in \{1, \dots, A\}$ on the lane $i \in \{1, \dots, N\}$. Dynamics of the phase density obeys the following kinetic equation:

$$\partial_t f_i^a + \partial_x (f_i^a v) + \partial_v \left(f_i^a \frac{V_{0i}^a - v}{\tau_i^a} \right) = \left(\frac{\partial f_i^a}{\partial t} \right)_{\text{int}} + \left(\frac{\partial f_i^a}{\partial t} \right)_{\text{lc}}, \tag{64}$$

where V_{0i}^a is the desired speed of the users of the class a on the ith lane and τ_i^a is the corresponding time of relaxation. The first term in the right-hand side describes the process of braking as the result of interaction, and the second term, the process of lane changing. Three types of changing the lane are discussed:

(1) change of lane as the result of interaction with a slower car ahead,

(2) "spontaneous," reflecting driver's desire to move along a certain lane,

(3) "mandatory," near lane junctions, at the descents and exits from a highway, and so on.

Correspondingly, we obtain

$$\left(\frac{\partial f_i^a}{\partial t} \right)_{\text{lc}} = \left(\frac{\partial f_i^a}{\partial t} \right)_{\text{lc}}^{\text{int}} + \left(\frac{\partial f_i^a}{\partial t} \right)_{\text{lc}}^{\text{spont}} + \left(\frac{\partial f_i^a}{\partial t} \right)_{\text{lc}}^{\text{mand}}. \tag{65}$$

The terms $(\partial f_i^a/\partial t)_{\text{int}}$ and $(\partial f_i^a/\partial t)_{\text{lc}}^{\text{int}}$ are defined in terms of interaction intensity. We denote by $I_i^{ab}(x, v, t)$ the intensity of interactions of the cars of the class a moving at the speed v at the point x on the ith lane with all cars of the class b moving before them with the speeds $w < v$, and by $J_i^{ab}(x, v, t)$, the intensity of interactions on the ith lane of all cars of the class a moving at the speeds $w > v$ and the cars of the class b moving before them with the speed v. The intensities $I_i^{ab}(x, v, t)$ result in a lower phase density $f_i^a(x, v, t)$ owing to lane changing and deceleration to a lower speed, whereas $J_i^{ab}(x, v, t)$ result in increased phase density owing to decelerating the faster cars of the class a to the speed v. The intensities obey companion equations of (61):

$$I_i^{ab} = \chi_i \int\limits_{v > v'} dv' \, (v - v') f_i^{ab}(x, v, x + s_i^a, v', t), \tag{66}$$

$$J_i^{ab} = \chi_i \int\limits_{v' > v} dv' \, (v' - v) f_i^{ab}(x, v', x + s_i^a, v, t). \tag{67}$$

We denote by p_i^a the probability that a car of the class a on the ith lane can change it without delay. Obviously, this probability is equal to the sum $p_i^a = p_{i,i-1}^a + p_{i,i+1}^a$ of the probabilities of changing to the right, $(i-1)$, or left, $(i+1)$, lane. These probabilities are assumed to be functions of macroscopic variables such as density and mean lane speeds.

In terms of the interaction intensities, the interactive terms in the right-hand side of the kinetic equation take the form

$$\left(\frac{\partial f_i^a}{\partial t}\right)_{\text{int}} = (1 - p_i^a) \sum_b \Big[\mathcal{J}_i^{ab}(x, v, t) - \mathcal{I}_i^{ab}(x, v, t)\Big], \tag{68}$$

$$\left(\frac{\partial f_i^a}{\partial t}\right)_{\text{lc}}^{\text{int}} = \sum_b \Big[p_{i-1,i}^a \mathcal{I}_{i-1}^{ab}(x, v) - p_{i,i-1}^a \mathcal{I}_i^{ab}(x, v) + p_{i+1,i}^a \mathcal{I}_{i+1}^{ab}(x, v) - p_{i,i+1}^a \mathcal{I}_i^{ab}(x, v)\Big]. \tag{69}$$

The rates of spontaneous lane change are proportional to the phase densities. The proportionality coefficients defining the frequency of lane changes can be conveniently represented as $1/T_{i-1,i}^a$. The values $T_{i-1,i}^a$ have dimension of time and are interpreted as the characteristic mean times between changes of lanes. The mean times also are assumed to be the functions of macroscopic variables. With regard for this notation, the spontaneous terms in the right-hand side of the kinetic equation are as follows:

$$\left(\frac{\partial f_i^a}{\partial t}\right)_{\text{lc}}^{\text{spont}} = \frac{f_{i-1}^a(x, v, t)}{T_{i-1,i}^a} - \frac{f_i^a(x, v, t)}{T_{i,i-1}^a} + \frac{f_{i+1}^a(x, v, t)}{T_{i+1,i}^a} - \frac{f_i^a(x, v, t)}{T_{i,i+1}^a}. \tag{70}$$

The expression

$$\frac{1}{T_{i,j}^a} = g_{i,j}^a \left(\frac{\rho_i}{\rho_i^{\max}}\right)^{\beta_1} \left(1 - \frac{\rho_j}{\rho_j^{\max}}\right)^{\beta_2}, \tag{71}$$

which is in good agreement with the empirical data [96], is suitable for the mean times.

The form of terms concerned with forced changes of lanes depends on the particular characteristics of the road as is exemplified by road narrowing and merging of two lanes into one [95]. Among other modifications of the kinetic equation we note [57, 71, 107] where it was attempted to describe the processes of acceleration similar to those of deceleration, that is, as a result of the stochastic process of car interaction.

3.3. Derivation of the Macroscopic Models from the Kinetic Equation

Direct use of the kinetic equations for modeling the traffic flow is not always convenient owing to high dimensionality of the problem. Great theoretical importance of these equations is due the fact that they enable one to derive systematically equations for the macroscopic variables such as density and mean speed.

According to (47), (48), the macroscopic variables can be expressed in terms of the moments of phase density in the variable speed. An arbitrary moment of the order k is as follows:

$$m_k(x, t) = \int_0^\infty dv v^k f(x, v, t). \tag{72}$$

By multiplying both sides of the kinetic Eq. (54) by v^k and integrating with respect to dv, we obtain a differential equation for the kth moment:

$$\partial_t m_k + \partial_x m_{k+1} = -\frac{m_k - m_{0k}}{\tau} + (1 - p)(m_1 m_k - \rho m_{k+1}), \tag{73}$$

where m_{0k} is the kth moment of distribution of the desirable speeds which is regarded as a given function in the Prigogine model. For $k = 0$, Eq. (73) provides the continuity Eq. (37). For $k = 1$, the equation can be reduced by some transformations to the speed Eq. (44) where the “internal flow pressure" is related with the speed variation in the flow by the expression $\mathcal{P}(\rho) = \rho\Theta(\rho)$ and the equilibrium speed V_e follows the expression

$$V_e = V_0 - (1-p)\tau\rho\Theta. \tag{74}$$

As follows from this expression, the equilibrium speed is equal to the mean desirable speed from which the term describing the movement conditions was subtracted. The higher the density or scatter of speed, the greater the decrease in speed. Additionally, an increase in the relaxation time or the overtake probability leads to a longer time of relaxation. Therefore, the kinetic approach allows one to justify the form of the macroscopic equations and also obtain analytical expressions and meaningful interpretations of the functions and model parameters.

The moment Eqs. (73) make up an infinite chain where dynamics of the moment of the order $k+1$ is required to calculate the dynamics of the moment of the order k. To break this chain and close the equation system for some k, an approximate expression is introduced which relates the $(k+1)$st moment with lower-order moments. The chain most frequently is interrupted for $k = 1$ using some approximate expression for the speed variation vs. the density $\Theta = \Theta(\rho)$, although some works consider the dynamic equation of variation.

The macroscopic equations were first obtained from the Prigogine kinetic equation [81], and the equations of moments, by integrating the Paveri-Fontana equation [76]. However, these equations were not calibrated by the actual data, and this study was not continued. The interest in this line of research was aroused in the mid-1950's owing to the studies of Helbing and his colleagues [38, 39, 40, 42, 43, 95]. He relied on the kinetic Paveri-Fontana equation to propose some of its modifications and generalizations providing eventually well-calibrated macromodels of the traffic flow.

The Paveri-Fontana equation (56) was formulated for the density of distribution in an *extended* phase space, that is, for the distribution in the actual, v, and desirable, w, speed. The moments of kth order in the actual speed and of the lth order in the desirable speed are described by the expression

$$m_{k,l}(x,t) = \int\limits_0^\infty dv v^k \int\limits_0^\infty dw g(x,v,w,t). \tag{75}$$

In particular, the car density is the moment of zero order: $\rho = m_{0,0}$; the mean speed and the mean desirable speed are defined by the moments of the first order: $\rho V = m_{1,0}$, $\rho W = m_{0,1}$; the variations and correlations of the actual and desirable speeds are defined in a similar manner, and so on.

In the general case, the derivation of equations for the moments from the kinetic equation is not as simple, as in the above case of the Prigogine equation. An important distinction between the Paveri-Fontana and Prigogine equations is that the integrals (58) making up the interactive term cannot be formally united into a single integral from 0 to ∞ taken with respect to the variable speed. As the result, multiplication of the kinetic equation by $v^k w^l$ and integration with respect to v and w does not provide a closed equation, except for the moments of the form $m_{k,0}$, that is, moments associated with distribution of the actual speeds for which this derivation is possible [76]. A similar difficulty arises with other modifications of the kinetic equation, for example, for the "nonpoint" interactions and so forth. However, approximate derivation of the macroscopic equations is possible by using the iterative method known in statistical physics as the Chapman–Enskog method. The so-called "zero approximation" of this method relies on the assumption that the distribution of the actual and desirable speeds in the flow is a locally normal distribution of the form

$$g(x,v,w,t) = \frac{1}{2\pi\sqrt{\Delta}} \exp\left\{-\frac{1}{2\Delta}(\Theta_{vv}(\delta V)^2 - \Theta_{vw}\delta V\delta W + \Theta_{ww}(\delta W)^2)\right\}, \tag{76}$$

where $\Delta = \Theta_{vv}\Theta_{ww} - \Theta_{vw}^2$. Substitution of this expression into the integral of interactions of the kinetic equation and integration with respect to speeds provide the equations of moments.

An equation for correction of the first order to the zero approximation is derived at the next step, and so on. As was shown in the kinetic gas theory, the zero approximation to the solution of the kinetic equation leads to a hydrodynamic Euler equation, and correction of the first order, to the Navier–Stokes equation. As was demonstrated by calculations, the transport equations for the zero and first approximations are in many respects similar to the corresponding hydrodynamic equations. That is why they are called the Euler-like and Navier–Stokes-like equations of transport dynamics. Derivations of equations for different variants of the kinetic equation can be found in [38, 41, 46, 95, 103, 104].

Among the macroscopic models obtained by integrating the kinetic equation, the gaz-kinetic traffic model [47, 95] is best studied and calibrated to the actual data. We present here a formulation of its one-lane version. Allowance for the "nonlocality" of interactions is an essential distinction of the model. The cars located at the point x interact with the cars located ahead at the interaction distance $x + s(\rho, V)$. In the formulas below the variables whose values are taken at the shifted point $x + s$ are primed for brevity. The desirable speed is assumed to be constant for all drivers, and a bivariant normal distribution similar to (76) is taken in the Euler approximation for the joint distribution of the speeds V and V'. Integration of the kinetic equation in the speed space provides the desired macroscopic model of the traffic flow. Its mathematical formulation represents an ordinary continuity Eq. (37) for density and equation of the form (44) for speed where the "internal pressure" of the flow is related by $\mathcal{P}(\rho) = \rho\Theta(\rho)$ with the speed variation in the flow and the equilibrium speed V_e is as follows:

$$V_e = V_0 - \tau(1-p)B(\delta V). \tag{77}$$

The term $B(\delta V)$, which is pivotal to the model and called the *Boltzmann factor*, stems from the interaction integral in the kinetic equation and describes the interaction of cars. It depends on the effective dimensionless difference δV of speeds at the points x and $x + s$, as well as the variation of flow speed Θ. The Boltzmann factor is as follows:

$$B(\delta V) = \chi(\rho)\rho' S\left[\delta V N(\delta V) + (1 + \delta V^2)E(\delta V)\right], \tag{78}$$

$$\delta V = (V - V')/\sqrt{S}, \tag{79}$$

$$S = \Theta - 2k\sqrt{\Theta\Theta'} + \Theta', \tag{80}$$

where $N(x)$ is the standard normal distribution and $E(x)$ is an inexact integral of this distribution:

$$N(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \tag{81}$$

$$E(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} dy\, e^{-\frac{y^2}{2}}. \tag{82}$$

The Boltzmann factor is negligible for negative difference of speeds (in the accelerating flow) and increases dramatically with increase in the positive difference of speeds (in the decelerating flow). A that, the effective difference of speeds (79) decreases with increase in the variation Θ or correlation k of the flow speeds. The great scatter of individual speeds can be said to "blur" the effect of regular reduction of the mean speed in the direction of movement, which manifests itself in a reduction in the effective difference of speeds, although on the whole the Boltzmann factor increases with increase in variation owing to the factor S in (78).

Empirical data-based approximating expressions are used for the variation and the correlation coefficient. In particular, $\Theta(\rho) = \alpha(\rho)V^2$, where $\alpha(\rho)$ is a function that is close to the step function and has a small value for low densities and a jump in the domain of mean flow densities. "Effective interaction section" $\chi(\rho)$ is another important function that needs approximation. The following expression was obtained in [47] by calibrating the model and taking into account the desired asymptotics for low and high densities:

$$[1 - p(\rho)]\chi(\rho) = \frac{V_0 T^2}{\tau\alpha(\rho_{\max})} \frac{\rho}{(1 - \rho/\rho_{\max})^2}. \tag{83}$$

Studies demonstrated that the model adequately reproduces the observed properties of the traffic flow such as origination of the stop-and-go movement, hysteresis upon passing from free to congested movement, and so forth [43, 47, 48]. In particular, it was used to analyze the flow dynamics near the source of nonuniformity (descents and exits from highway) and provided a diagram of various qualitative states of the flow for different values of the incoming and exit flows [44]. The reader is referred to [95] for derivation and calibration of a more general variant of the model for multiple lanes and multiple classes of users.

3.4. Microscopic Models

First microscopic models were proposed in the 1950's [82, 89]. Let the cars be numerated by the subscript n according to their order on the road. It is assumed in the micromodels that acceleration of the nth car is defined by the states of the neighbor cars, the greatest effect being produced by the immediately preceding car $n - 1$. This car is often called the *leader*, and the entire class of micromodels, the *"follow-the-leader"* models.

We denote the coordinates and speeds of cars, respectively, by x_n and v_n. The following factors mostly define acceleration of the nth car:

(1) speed of the given car with respect to the leader $\Delta v_n(t) = v_n(t) - v_{n-1}(t)$;

(2) intrinsic car speed $v_n(t)$ defining the safe interval of movement;

(3) distance to the leader $d_n(t) = x_{n-1}(t) - x_n(t)$ or, as a variant, the "pure" distance $s_n(t) = d_n(t) - l_{n-1}$ taking into account the car length l_n.

Correspondingly, in the general form the movement of cars obeys a system of ordinary differential equations:

$$\dot{v}_n(t) = f(v_n, \Delta v_n, d_n). \tag{84}$$

The differences between the models are defined by the form of the function f.

3.4.1. Follow-the-Leader Models. In the first variants of the follow-the-leader model, it was assumed that each driver adapts its speed to that of the leader:

$$\dot{v}_n(t) = \frac{1}{\tau}\left[v_{n-1}(t) - v_n(t)\right], \tag{85}$$

where τ is the characteristic time of adaptation.

Equation (85) was obtained in [82] by differentiating with respect to time the relation

$$\Delta x_n(t) = x_{n-1}(t) - x_n(t) = (\Delta x)_{\text{safe}} + \tau v_n(t) \tag{86}$$

which expresses the desire of drivers to maintain the speed-dependent safe distance $(\Delta x)_{\text{safe}}$ to the leader. However, this simple model does not describe properties of the real flow such as instability and origination of congestion waves. Several modifications were proposed. For example, it was suggested in [12] to introduce in the left-hand side of the equation a delay $\Delta t \approx 1.3$ s of the argument to reflect the driver's response to variations in the leader's speed. Now, the multiplier $1/\tau$ in (85) can be interpreted as a *sensitivity coefficient* $\mathcal{S}$ characterizing the rate of driver's response to variations in the leader's speed. In the general case, this coefficient is a dynamic variable depending on the speed and the current distance to the leader. In view of the aforementioned, the model can be set down as the following differential equation with biased argument:

$$\dot{v}_n(t + \Delta t) = \mathcal{S}[v_{n-1}(t) - v_n(t)]. \tag{87}$$

Under sufficiently great values of Δt, equations of this kind are usually unstable. For $S = 1/T =$ const, in particular, condition for instability of Eq. (87) has the form $\Delta t/T > 1/2$. In principle, instability allows one to model the congestion waves. However, an equation where the sensitivity coefficient S is assumed to be constant does not reproduce many properties of the actual flow such as the fundamental diagram, that is, dependence of the flow on density. A more adequate model can be obtained by assuming that sensitivity increases with smaller distance to the leader. An expression for this coefficient that is in agreement with the experimental data was proposed in [33] as

$$\frac{1}{T} = \frac{1}{T_0}\frac{[v_n(t+\Delta t)]^{m_1}}{[x_{n-1}(t) - x_n(t)]^{m_2}}, \tag{88}$$

where the constants m_1 and m_2 are selected empirically.

Let us consider a uniform stationary flow with identical speeds v_n and intervals d_n. Then, the flow density is $1/d_n$, and the equilibrium speed $V_e = v_n$. Hence, we get an equilibrium relation for speed and density:

$$V_e(\rho) = V_0[1 - (\rho/\rho_{\max})^{m_2-1}]^{1/(1-m_1)}, \tag{89}$$

where V_0 is the speed of free movement (desirable speed) and $\rho_{\max}$ is the maximum admissible car density. Fundamental diagrams that are close to the empirical data were obtained for the following parameters: $m_1 \approx 0.8$, $m_2 \approx 2.8$ [68] or $m_1 = 0.953$, $m_2 = 3.05$ [62, 63].

The follow-the-leader model is used now in the software package MITSIM. For a more detailed review of the model's variants and history the reader is referred to [9].

3.4.2. Optimal-speed Models. Incorrect description of the dynamics of a single car is one of the disadvantages of the follow-the-leader model where in the absence of a leader the car acceleration is zero, whereas it is reasonable to assume that the drivers want to approach their speed to the desired value v_n^0. Models of another type are based on the assumption that each driver has a "safe" speed $v'_e(d_n)$ which depends on the distance to the leader. It is also called the *optimal speed*. These models are oriented to adapting the speed to the optimal value, rather than to that of the leader. The leader's influence is expressed indirectly via the dependence of the optimal speed on the distance. This model was first proposed in [72] where it was planned to perform speed adaptation with a time delay:

$$v_n(t + \Delta t) = v'_e(d_n(t)) = v_e(s_n(t)). \tag{90}$$

The differential equation

$$\dot{v}_n(t) = \frac{1}{\tau}(v'_e(d_n(t)) - v_n(t)), \tag{91}$$

where the optimal speed obeys the expression

$$v'_e(d) = \frac{v_0}{2}(\tanh(d - d_c) + \tanh d_c) \tag{92}$$

with appropriate constants v_0 and d_c, was proposed in [4, 5]. This equation can be obtained by expanding (90) into the Taylor series to the terms of the first order in $\tau = \Delta t$: $v_n(t + \Delta t) \approx v_n(t) + \Delta t\, dv_n(t)/dt$.

One can demonstrate that if the instability condition

$$\frac{dv'_e(d_n)}{d\,d_n} = \frac{dv_e(s_n)}{ds_n} > \frac{1}{2\tau} \tag{93}$$

is satisfied, that is, the relaxation time τ is large or the dependence of the optimal speed on the distance $v_e(s_n)$ is steep, then a small perturbation of the optimal-speed model leads to congestion. Similar results were obtained in [3, 105] for similar models with an additional response time delay Δt.

3.4.3. Treiber Model. The standard optimal-speed model has some disadvantages. Specifically, it is very sensitive to a particular choice of the functional dependence of the optimal speed on the $v'_e(d_n)$, as well as to the choice of τ. For greater τ, cars begin to collide, whereas for too small values, unrealistic accelerations arise. As a matter of fact, the characteristic acceleration times exceed the characteristic deceleration times approximately by the factor of five. Additionally, in reality the drivers keep a greater distance and, for a high speed $\Delta v_n(t)$ relative to the leader, pull up earlier. Numerous variants of the model were developed to take into account this and other features of the real driver behavior [8, 34, 45, 58, 59, 99].

The Treiber "intelligent driver model" (IDM) [100, 101] may be regarded as one of the most successful micromodels. Its calibration and experiments demonstrated that it is stable to parameter variations, manifests realistic behavior upon acceleration and deceleration, and reproduces the main observed characteristics of the traffic flow. It is assumed in the IDM that the car acceleration is a continuous function of the speed v_n, "pure" distance to the leader $s_n = d_n - l_{n-1}$, and the speed relative to the leader Δv_n:

$$\dot{v}_n = a_n \left[1 - \left(\frac{v_n}{v_n^0}\right)^{\delta} - \left(\frac{s_n^*(v_n, \Delta v_n)}{s_n}\right)^2\right]. \tag{94}$$

The term $a_n[1 - (v_n/v_n^0)^{\delta}]$ in the right-hand side of this equation describes dynamics of car acceleration on a free road, whereas $f_{n,n-1} = -a_n[s_n^*(v_n, \Delta v_n)/s_n]^2$ describes deceleration caused by interaction with the leader. Choice of the parameter δ enables one to calibrate acceleration behavior. The magnitude $\delta = 1$ corresponds to the exponential time of acceleration which is characteristic of the majority of models. When this parameter increases, acceleration does not decrease exponentially in the course of acceleration (at the limit, for $\delta \to \infty$ acceleration a_n is constant until reaching the desired speed v_n^0), which is in better agreement with the real behavior of drivers. The braking term depends on the ratio of the "desirable," s_n^*, and actual, s_n, distances, the former obeying the following expression:

$$s_n^*(v_n, \Delta v_n) = s'_n + s''_n\sqrt{\frac{v_n}{v_n^0}} + T_n v_n + \frac{v_n \Delta v_n}{2\sqrt{a_n b_n}}. \tag{95}$$

The model parameters can be chosen individually for each nth car, which allows one to take into account the individual characteristics of drivers and carriers. Many common flow characteristics, however, can be obtained by considering identical drivers. The parameters have a descriptive interpretation such as the desirable speed v_0, safe time interval T, maximum acceleration a, "comfortable" (not emergency) braking b, "sensitivity" index for acceleration δ, "congestion" distance s' and s'', and car length l. In order to reduce the number of parameters, one can simplify the model by assuming that $\delta = 1$, $s'' = 0$, and $l = 0$. At that, the model mostly remains adequate.

In the equilibrium flow where $\dot{v}_n = 0$ and $\Delta v_n = 0$, the drivers aim at retaining the speed-dependent equilibrium distance $s_e(v) = s^*(v, 0)[1 - (v/v_0)^{\delta}]^{-1/2}$ to the leader. From this relation one can establish the equilibrium speed and construct the fundamental diagram. In the special case of $\delta = 1$ and $s' = s'' = 0$, in particular, one can establish an analytical expression for the equilibrium speed:

$$v_e(s) = \frac{s^2}{2v_0T^2}\left[-1 + \sqrt{1 + \frac{4T^2v_0^2}{s^2}}\right]. \tag{96}$$

The equilibrium flow $Q_e(\rho) = \rho V_e(\rho)$ as a function of density, that is, the fundamental diagram, is obtained from this expression and an obvious expression relating the distance with the car density $s = (d - l) = (1/\rho - l) = (1/\rho - 1/\rho_{\max})$. The coefficient δ affects the form of the diagram so that the passage from the free mode to the loaded mode becomes sharper with greater δ. For $\delta \to \infty$ and $s'' = 0$, the fundamental diagram at the limit becomes triangular with the angle at the point of maximum flow: $Q_e(\rho) = \min(\rho v_0, [1 - \rho(l + s')]/T)$. For smaller δ, the form of the curve smooths and approaches the observed form.

3.4.4. Cellular Automata. By the cellular automaton (CA) is meant an idealized representation of a physical system with discrete time and space and the system elements having a discrete set of feasible states. The CA concept was introduced by von Neumann in the 1950's in connection with the abstract theory of self-reproducing computers. Later on the CA's were used to model many dynamic systems in diverse areas. Modeling of traffic flow by CA was first suggested in [15]. Active development and investigation of the CA in the transport area was started by Nagel and Schreckenberg [70]. There exists a vast literature on CA's that is discussed, for example, in detail in [13].

In the CA models, the car coordinate and speed, as well as time, are discrete variables. The road (lane) is decomposed into conventional "cells" of identical length Δx, each cell being at each time instant either empty or occupied by a single "car." The states of all cells are simultaneously (concurrently) updated according to a certain set of rules at each time step $t \longrightarrow t+1$. The choice of one or another set defines the diversity of CA variants. This description is, of course, very simplistic, but enables a very high computer performance. Owing to their computational efficiency, the CA models attract since recently attention of researchers and designers of the software for transport calculations.

We present a formulation of the original Nagel–Schreckenberg model. Let us denote by x_n and v_n the coordinate and speed of the nth car and by $d_n = x_{n+1} - x_n$, the distance to the leader. Speed can assume one of $v_{\max} + 1$ admissible *integer* values $v_n = 0, 1, \ldots, v_{\max}$. The state of all cars in the system is updated at each step $t \longrightarrow t+1$ according to the following rules.

Step 1: *Acceleration.* If $v_n < v_{\max}$, then the speed of the nth car is increased by one, if $v_n = v_{\max}$, it remains unchanged:

$$v_n \longrightarrow \min(v_n + 1, v_{\max}).$$

Step 2: *Deceleration.* If $d_n \le v_n$, then the speed of the nth car is reduced to $d_n - 1$:

$$v_n \longrightarrow \min(v_n, d_n - 1).$$

Step 3: *Random perturbations.* If $v_n > 0$, then the speed of the nth car can be reduced by one with the probability p; speed remains unchanged if $v_n = 0$:

$$v_n \xrightarrow{p} \max(v_n - 1, 0).$$

Step 4: *Movement.* After Steps 1–3. each car moves ahead by the number of cells corresponding to its new speed

$$x_n \longrightarrow x_n + v_n.$$

Step 1 reflects the general tendency of the drivers to move as fast as possible without exceeding the maximum admissible speed. Step 2 guarantees lack of collisions with the cars ahead. Step 3 introduces an element of stochasticity, thus taking into consideration the differences in drivers' behavior—excessive braking, in particular. For realistic description of the traffic flow, the typical cell length is taken 7.5 m, which approximately corresponds to the distance occupied by one car in congestion. At that, the time step 1 s corresponds to $v_{\max} = 5$. It is usually assumed that $p = 0.5$ for highways and $p = 0.2$ and $v_{\max} = 2$ for street traffic [23].

The given model is "minimal" in the sense that the above four steps are required to reproduce the basic characteristics of the flow. For adequate modeling of more complex aspects of flow dynamics, however, additional rules are required (or the existing rules must be modified). Numerical examples show that the flow is stable for small densities and collapses for their higher values. Process stochasticity plays the key role in the development of instabilities, that is, p must be other than zero for congestions to evolve. For $p = 0$, the flow remains stable for all values of density [69, 70]. This fact can be regarded as a serious theoretical disadvantage of the CA models as compared with the macromodels or the follow-the-leader models where fluctuations play the role of the initial push, the further evolution of the congestion being explained by instability (of a quite deterministic) equilibrium solution.

REFERENCES

1. Alberti, E. and Belli, G., Contributions to the Boltzmann-like Approach For Traffic Flow—A Model For Concentration Dependent Driving Programs, *Transpn. Res.*, 1978, vol. 12, pp. 33–42.

2. Andrews, F.C., A Statistical Theory of Traffic Flow on Highways—II. Three-Car Interactions and the Onset of Queuing, *Transpn. Res.*, 1970, vol. 4, pp. 367–377.

3. Bando, M., Hasebe, K., Nakanishi, K., and Nakayama, A., Analysis of Optimal Velocity Model With Explicit Delay, *Phys. Rev. E*, 1998, vol. 58, pp. 5429–5435.

4. Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y., Structure Stability of Congestion in Traffic Dynamics, *Jpn. J. Indust. Appl. Math.*, 1994, vol. 11, pp. 203–223.

5. Bando, M., Hasebe, K., Nakayama, A., Shibata, A., and Sugiyama, Y., Dynamical Model of Traffic Congestion and Numerical Simulation, *Phys. Rev. E*, 1995, vol. 51, pp. 1035–1042.

6. Beckmann, M.J., McGuire, C.B., and Winsten, C.B., *Studies in the Economics of Transportation*, New Haven: Yale Univ. Press, 1956.

7. Beylich, A.E., Elements of a Kinetic Theory of Traffic Flow, in *Rarefied Gas Dynamics*, Campargue, E., Ed., Paris: Commissariat à l'Energie Atomique, 1979, vol. 1, pp. 129–138.

8. Bleile, T., Traffic Simulation Supporting Urban Control System Development, *Proc. 4th World Congr. Intelligent Transp. Syst.*, Brussels: ITS Congr. Association, 1997.

9. Brackstone, M. and McDonald, M., Car Following: A Historical Review, *Transpn. Res.*, 2000, vol. 2, pp. 181–196.

10. Carey, M., Nonconvexity of the Dynamic Traffic Assignment Problem, *Transpn. Res.*, 1992, vol. 26, pp. 127–133.

11. Carrothers, G.A.P., An Historical Review of the Gravity and Potential Concepts of Human Interaction, *J. Am. Inst. Planners*, 1956, vol. 22, pp. 94–102.

12. Chandler, R.E., Herman, R., and Montroll, E.W., Traffic Dynamics: Studies in Car Following, *Oper. Res.*, 1958, vol. 6, pp. 165–184.

13. Chowdhury, D., Santen, L., and Schadschneider, A., Statistical Physics of Vehicular Traffic and Some Related Systems, *Phys. Rep.*, 2000, vol. 329, pp. 199–329.

14. Cliff, A.D., Martin, R.L., and Ord, J.K., Evaluating the Friction of Distance Parameter in Gravity Models, *Region. Studies*, 1974, vol. 8, pp. 281–286.

15. Cremer, M. and Ludwig, J., A Fast Simulation Model for Traffic Flow on the Basis of Boolean Operations, *Math. Comput. Simul.*, 1986, vol. 28, pp. 297–303.

16. Cremer, M. and Meissner, F., Traffic Prediction and Optimization Using an Efficient Macroscopic Simulation Tool, in *Modeling and Simulation*, Pave, A., Ed., Ghent: Soc. Comput. Simul. Int., 1993, pp. 515–519.

17. Cremer, M. and Papageorgiou, M., Parameter Identification for a Traffic Flow Model, *Automatica*, 1981, vol. 17, pp. 837–843.

18. Daganzo, C.F., The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamical Theory, *Transpn. Res. B*, 1994, vol. 28, pp. 269–287.

19. Daganzo, C.F., The Cell Transmission Model, Part II: Network Traffic, *Transpn. Res. B*, 1995, vol. 29, pp. 79–93.

20. Daganzo, C.F., A Finite Difference Approximation of the Kinematic Wave Model of Traffic Flow, *Transpn. Res. B*, 1995, vol. 29, pp. 261–276.

21. Daganzo, C.F., Requiem for Second-order Fluid Approximations of Traffic Flow, *Transpn. Res. B*, 1995, vol. 29, pp. 277–286.

22. Daganzo, C.F. and Sheffi, Y., On Stochastic Models of Traffic Assignment, *Transpn. Sci.*, 1977, vol. 11, pp. 253–274.

23. Esser, J. and Schreckenberg, M., Microscopic Simulation of Urban Traffic Based on Cellular Automata, *Int. J. Mod. Phys. C*, 1997, vol. 8, pp. 1025–1036.

24. Fik, T.J., Hierarchical Interaction: The Modeling of Competing Central Place System, *The Ann. Reg. Sci.*, 1988, vol. 22, pp. 48–69.

25. Fik, T.J. and Mulligan, G.F., Spatial Flows and Competing Central Places: Towards a General Theory of Hierarchical Interaction, *Envir. Plan. A*, 1990, vol. 22, pp. 527–549.

26. Florian, M., A Traffic Equilibrium Model of Travel by Car and Public Transit Modes, *Transpn. Sci.*, 1977, vol. 11, pp. 166–179.

27. Fotheringham, A.S., A New Set of Spacial-Interaction Models: The Theory of Competing Destinations, *Envir. Plan. A*, 1983, vol. 15, pp. 15–36.

28. Fotheringham, A.S., Modeling Hierarchical Destination Choice, *Envir. Plan. A*, 1986, vol. 18, pp. 401–418.

29. Frank, M. and Wolfe, P., An Algorithm for Quadratic Programming, *Nav. Res. Log. Q.*, 1956, vol. 3, pp. 95–110.

30. Friesz, T.L., Luque, F.J., Tobin, R.L., and Wie, B.W., Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem, *Oper. Res.*, 1989, vol. 37, pp. 893–901.

31. Gartner, N.H., Optimal Traffic Assignment with Elastic Demands: A Review; Part II: Algorithmic Approaches, *Transpn. Sci.*, 1980, vol. 14, pp. 192–208.

32. Gazis, D.C., *Traffic Science*, New York: Wiley, 1974.

33. Gazis, D.C., Herman, R., and Rothery, R.W., Nonlinear Follow the Leader Models of Traffic Flow, *Oper. Res.*, 1961, vol. 9, pp. 545–567.

34. Gipps, P.G., A Behavioural Car Following Model for Computer Simulation, *Transpn. Res. B*, 1981, vol. 15, pp. 105–111.

35. Gonçalves, M.B. and Ulysséa-Neto, I., The Development of a New Gravity-Opportunity Model for Trip Distribution, *Envir. Plan. A*, 1993, vol. 25, pp. 817–826.

36. Harris, B. and Wilson, A. G., Equilibrium Values and Dynamics of Attractiveness Terms in Production-constrained Spatial-interaction Models, *Envir. Plan. A*, 1978, vol. 10, pp. 371–388.

37. Haynes, K.E. and Fotheringham, A.S., Gravity and Spacial-interaction Models, *Sci. Geography Series 2*, Beverly Hills: Sage, 1984.

38. Helbing, D., High-fidelity Macroscopic Traffic Equations, *Physica A*, 1995, vol. 219, pp. 391–407.

39. Helbing, D., Improved Fluid-Dynamic Model For Vehicular Traffic, *Phys. Rev. E*, 1995, vol. 51, pp. 3164–3169.

40. Helbing, D., Derivation and Empirical Validation of a Refined Traffic Flow Model, *Physica A*, 1996, vol. 233, pp. 253–282. pp. 253–282.

41. Helbing, D., Gas-kinetic Derivation of Navier-STokes-Like Traffic Equations, *Phys. Rev. E*, 1996, vol. 53, pp. 2366–2381.

42. Helbing, D., *Verkehrsdynamik*, Berlin: Springer, 1997.

43. Helbing, D., Hennecke, A., Shvetsov, V., and Treiber, M., MASTER: Macroscopic Traffic Simulation Based on a Gas-Kinetic, Non-Local Traffic Model, *Transpn. Res. B*, 2001, vol. 35, pp. 183–211.

44. Helbing, D., Hennecke, A., and Treiber, M., Phase Diagram of Traffic States in the Presence of Inhomogeneities, *Phys. Rev. Lett.*, 1999, vol. 82, pp. 4360–4363.

45. Helbing, D. and Tilch, B., Generalized Force Model of Traffic Dynamics, *Phys. Rev. E*, 1998, vol. 58, pp. 133–138.

46. Helbing, D. and Treiber, M., Enskog Equations for Traffic Flow Evaluated up to Navier-Stokes Order, *Gran. Matt.*, 1998, vol. 1, pp. 21–31.

47. Helbing, D. and Treiber, M., Gas-kinetic-based Traffic Model Explaining Observed Hysteretic Phase Transition, *Phys. Rev. Lett.*, 1998, vol. 81, pp. 3042–3045.

48. Helbing, D. and Treiber, M., Jams, Waves, and Clusters, *Science*, 1998, vol. 282, pp. 2001–2003.

49. Hoogendoorn, S. and Bovy, P.H.L., Gas-kinetic Model for Multi-Lane Heterogeneous Traffic Flow, *Transpn. Res. Rec.*, 1999, vol. 1678, pp. 150–159.

50. Janson, B.N., Dynamic Traffic Assignment for Urban Road Networks, *Transpn. Res. B*, 1991, vol. 25, pp. 143–161.

51. Jayakrishnan, R., Tsai Wei K., and Chen, A., A Dynamic Traffic Assignment Model with Traffic-Flow Relationships, *Transpn. Res. C*, 1995, vol. 3, pp. 51–72.

52. Kerner, B.S. and Konhäuser, P., Cluster Effect in Initially Homogeneous Traffic Flow, *Phys. Rev. E*, 1993, vol. 48, pp. R2335–R2338.

53. Kerner, B.S. and Konhäuser, P., Structure and Parameters of Clusters in Traffic Flow, *Phys. Rev. E*, 1994, vol. 50, pp. 54–83.

54. Kerner, B.S., Konhäuser, P., and Schilke, M., Deterministic Spontaneous Appearance of Traffic Congestions in Slightly Inhomogeneous Traffic Flow, *Phys. Rev. E*, 1995, vol. 51, pp. R6243–R6246.

55. Kerner, B.S., Konhäuser, P., and Schilke, M., Formation of Traffic Jams Caused by Fluctuations and Random Processes in Traffic Flow, in *Proc. 7th World Conf. Transp. Res.*, Hensher, D., Ed., Oxford: Pergamon, 1995, vol. 2, pp. 167–182.

56. Kerner, B.S., Konhäuser, P., and Schilke, M., 'Dipole-layer' Effect in Dense Traffic Flow, *Phys. Lett. A*, 1996, vol. 215, pp. 45–56.

57. Klar, A. and Wegener, R., Enskog-like Kinetic Models for Vehicular Traffic, *J. Stat. Phys.*, 1997, vol. 87, pp. 91–114.

58. Krauss, S., Wagner, P., and Gawron, C., Continuous Limit of the Nagel-Schreckenberg Model, *Phys. Rev. E*, 1996, vol. 54, pp. 3707–3712.

59. Krauss, S., Wagner, P., and Gawron, C., Metastable States in a Microscopic Model of Traffic Flow, *Phys. Rev. E*, 1997, vol. 55, pp. 5597–5602.

60. Kühne, R.D., Macroscopic Freeway Model for Dense Traffic—STop-Start Waves and Incident Detection, in *Proc. 9th Int. Sympos. Transp. and Traffic Theory*, Hamerslag, R., Ed., Utrecht: VNU Science, 1984, pp. 21–42.

61. Kühne, R.D., Freeway Speed Distribution and Acceleration Noise—Calculations from a Stochastic Continuum Theory and Comparison with Measurements, in *Proc. 10th Int. Sympos. on Trans. and Traffic Theory*, Gartner, N.H., Ed., New York: Elsevier, 1987, pp. 119–137.

62. Kühne, R.D. and Kroen, A., Knowledge-based Optimization of Line Control Systems for Freeways, *Int. Conf. Artificial Intelligence Appl. Transp. Engineering*, San Buenaventura, California, June 20–24, 1992, pp. 173–192.

63. Kühne, R.D. and Rödiger, M.B., Macroscopic Simulation Model for Freeway Traffic with Congestions and Stop-Start Waves, in *Proc. 1991 Winter Simulation Conf.*, Nelson, B.L., Ed., Phoenix: Society for Comput. Simul. Int., 1991, pp. 762–770.

64. Lampis, M., On the Kinetic Theory of Traffic Flow in the Case of a Nonnegligible Number of Queueing Vehicles, *Transpn. Sci.*, 1978, vol. 12, pp. 16–28.

65. LeBlanc, L.J., Morlok, E.K., and Pierskalla, W., An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem, *Transpn. Res. B*, 1975, vol. 9, pp. 309–318.

66. Lighthill, M.J. and Whitham, G.B., On Kinematic Waves: II. A Theory of Traffic on Long Crowded Roads, *Proc. Roy. Soc. London, Ser. A*, 1955, vol. 229, pp. 317–345.

67. Maher, M.J. and Hughes, P.C., A Probit-based Stochastic User Equilibrium Assignment Model, *Transpn. Res. B*, 1997, vol. 31, pp. 341–355.

68. May, A.D., Jr., and Keller, H.E.M., Non-integer Car-following Models, *Highway Res. Rec.*, 1967, vol. 199, pp. 19–32.

69. Nagel, K. and Herrmann, H.J., Deterministic Models for Traffic Jams, *Physica A*, 1993, vol. 199, pp. 254–269.

70. Nagel, K. and Schreckenberg, M., A Cellular Automaton Model for Freeway Traffic, *J. Phys. I France*, 1992, vol. 2, pp. 2221–2229.

71. Nelson, P., A Kinetic Model of Vehicular Traffic and Its Associated Bimodal Equilibrium Solutions, *Transp. Theor. Stat. Phys.*, 1995, vol. 24, pp. 383–409.

72. Newell, G.F., Nonlinear Effects in the Dynamics of Car Following, *Oper. Res.*, 1961, vol. 9, pp. 209–229.

73. Newell, G.F., A Simplified Theory of Kinematic Waves in Highway Traffic, I General Theory, II Queueing at Freeway Bottlenecks, III Multi-destination Flows, *Transpn. Res. B*, 1993, vol. 27, pp. 281–313.

74. Papageorgiou, M., *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control*, Berlin: Springer, 1983.

75. Pape, U., Implementation and Efficiency of MOor-Algorithms for the Shortest Route Problem, *Math. Program.*, 1974, vol. 7, pp. 212–222.

76. Paveri-Fontana, S.L., On Boltzmann-like Treatments for Traffic Flow. A Critical Review of the Basic Model and an Alternative Proposal for Dilute Traffic Analysis, *Transpn. Res.*, 1975, vol. 9, pp. 225–235.

77. Payne, H.J., Models of Freeway Traffic and Control, in *Mathematical Models of Public Systems*, Bekey, G.A., Ed., La Jolla: Simulation Council, 1971, vol. 1, pp. 51–61.

78. Payne, H.J., A Critical Review of a Macroscopic Freeway Model, in *Research Directions in Computer Control of Urban Traffic Systems*, Levine, W.S., Ed., New York: Am. Soc. of Civil Engineers, 1979, pp. 251–265.

79. Payne, H.J., FREFLO: A Macroscopic Simulation Model of Freeway Traffic, *Transpn. Res. Rec.*, 1979, vol. 722, pp. 68–77.

80. Phillips, W.F., A Kinetic Model for Traffic Flow with Continuum Implications, *Transpn. Plan. Technol.*, 1979, vol. 5, pp. 131–138.

81. Phillips, W.F., A New Continuum Traffic Model Obtained from Kinetic Theory, *Proc. 1978 IEEE Conf. Decision and Control*, New York: IEEE, 1979, pp. 1032–1036.

82. Pipes L.A., An Operational Analysis of Traffic Dynamics, *J. Appl. Phys.*, 1953, vol. 24, pp. 274–281.

83. Popkov, Yu.S., *Macrosystems Theory and Its Applications*, Berlin: Springer, 1995.

84. Powell, W.B. and Sheffy, Y., The Convergence of Equilibrium Algorithms with Predetermined Step Sizes, *Transpn. Sci.*, 1982, vol. 16, pp. 45–55.

85. Prigogine, I., A Boltzmann-like Approach to the Statistical Theory of Traffic Flow, in *Theory of Traffic Flow*, Herman, R., Ed., Amsterdam: Elsevier, 1961.

86. Prigogine, I. and Andrews, F.C., A Boltzmann-like Approach for Traffic Flow, *Oper. Res.*, 1960, vol. 8, pp. 789–797.

87. Prigogine, I. and Herman, R., *Kinetic Theory of Vehicular Traffic*, New York: Elsevier, 1971.

88. Rathi, A.K., Lieberman, E.B., and Yedlin, M., Enhanced FREFLO: Modeling of Congested Environments, *Transpn. Res. Rec.*, 1987, vol. 1112, pp. 61–71.

89. Reuschel, A., Fahrzeugbewegungen in der Kolonne, *Österr. Ingen.-Archiv.*, 1950, vol. 4, pp. 193–215.

90. Sen, A., Maximum Likelihood Estimation of Gravity Model Parameters, *J. Regional Sci.*, 1986, vol. 26, pp. 461–474.

91. Sheffy, Y., *Urban Transportation Networks*, Englewood Cliffs: Prentice Hall, 1985.

92. Sheffy, Y. and Powell, W.B., An Algorithm for the Equilibrium Assignment Problem with Random Link Times, *Network*, 1982, vol. 12, pp. 191–207.

93. Sheppard, E.S., Gravity Parameter Estimation, *Geographical Anal.*, 1979, vol. 11, pp. 120–132.

94. Shvetsov, V.I. and Dubov, Yu.A., Expected Distributions in the Intervening Opportunities Model, *Envir. Plan. A*, 1997, vol. 29, pp. 1229–1241.

95. Shvetsov, V.I. and Helbing, D., Macroscopic Dynamics of Multilane Traffic, *Phys. Rev. E*, 1999, vol. 59, pp. 6328–6339.

96. *Spurwechselvorgänge auf zweispurigen BAB-Richtungsfahrbahnen*, Sparmann, U., Ed., Bonn-Bad Godesberg: Bundesministerium für Verkehr, 1978.

97. Spiess, H. and Florian, M., Optimal Strategies: A New Assignment Model for Transit Networks, *Transpn. Res. B*, 1989, vol. 23, pp. 83–102.

98. Stouffer, S.A., Intervening Opportunities: A Theory Relating Mobility and Distance, *Am. Sociolog. Rev.*, 1940, vol. 5, pp. 845–867.

99. Tomer, E., Safonov, L., and Havlin, S., Presence of Many Stable Nonhomogeneous States in an Inertial Car-Following Model, *Phys. Rev. Lett.*, 2000, vol. 84, pp. 382–385.

100. Treiber, M. and Helbing, D., Explanation of observed Features of Self-Organization in Traffic Flow, *e-print cond-mat/9901239*, 1999.

101. Treiber, M., Hennecke, A., and Helbing, D., Congested Traffic States in Empirical Observations and Microscopic Simulations, *Phys. Rev. E*, 2000, vol. 62, pp. 1805–1824.

102. Voorhees, A.M., *A General Theory of Traffic Movement*, ITE, 1955.

103. Wagner, C., A Navier-Stokes-like Traffic Model, *Physica A*, 1997, vol. 245, pp. 124–138.

104. Wagner, C., Hoffmann, C., Sollacher, R., Wagenhuber, J., and Schürmann, B., Second Order Continuum Traffic Flow Model, *Phys. Rev. E*, 1996, vol. 54, pp. 5073–5085.

105. Wang, B.-H., Wang, L., Hui, P. M., and Hu, B., Analytical Results for the Steady State of Traffic Flow Models with Stochastic Delay, *Phys. Rev. E*, 1998, vol. 58, pp. 2876–2882.

106. Wardrop, J.G., Some Theoretical Aspects of Road Traffic Research, *Proc. Inst. Civil Engineers II*, 1952, pp. 325–378.

107. Wegener, R. and Klar, A., A Kinetic Model for Vehicular Traffic Derived from a Stochastic Microscopic Model, *Transp. Theor. Stat. Phys.*, 1996, vol. 25, pp. 785–798.

108. Whitham, G. B., *Linear and Nonlinear Waves*, New York: Wiley, 1974.

109. Whitham, G. B., *Lectures on Wave Propagation*, Berlin: Springer, 1979.

110. Wills, M.J., A Flexible Gravity-opportunities Model for Trip Distribution, *Transpn. Res. B*, 1986, vol. 20, pp. 89–111.

111. Wilson, A.G., A Statistical Theory of Spatial Distribution Models, *Transpn. Res.*, 1967, vol. 1, pp. 253–270.

112. Wilson, A.G., *Entropy in Urban and Regional Modelling*, London: Pion, 1970.

113. Wilson, A.G., A Family of Spatial Interaction Models and Associated Developments, *Envir. Plan. A*, 1971, vol. 3, pp. 255–282.

114. Dinits, E.A., Economic Algorithm to Determine the Shortest Paths in Network, in *Transportnye sistemy* (Transport Systems), Moscow: VNIISI, 1978.

115. Livshits, V.V., A Mathematical Model for Random-deterministic Choice and Its Use for Calculating the Labor Correspondences, in *Avtomatizatsiya protsessov gradostroitel'nogo proektirovaniya* (Computer-aided Urban Planning), Moscow: TSNIIP 1973.

116. Livshits, V.V., Urban System Concentration and Mathematical Modeling of Adaptive Behavior of the Urban Population, in *Ispol'zovanie prikladnogo sistemnogo analiza v proektirovanii i upravlenii razvitiem gorodov* (Using Applied System Analysis in Design and Management of Urban Development), Moscow: Stroiizdat, 1974.

117. Livshits, V.V. and Strel'nikov, A.I., *Kalibrovka i proverka gravitatsionnoi statisticheskoi modeli trudovykh korrespondentsii* (Calibration and Testing of the Gravitational Statistical Model of Labor Correspondences), Moscow: TSNIIP Gradostroitel'stva, 1983.

118. *Avtomatizatsiya planirovaniya i upravleniya transportnymi sistemami* (Computer-aided Planning and Control of Transport Systems), Livshits, V.N., Ed., Moscow: Transport, 1987.

119. *Matematicheskoe obespechenie gradostroitel'nogo proektirovaniya* (Mathematical Support of Urban Design), Leningrad: Nauka, 1989.

120. Popkov, Yu.S., Posokhin, M.V., Gutnov, A.E., and Shmul'yan, B.L., *Sistemnyi analiz i problemy razvitiya gorodov* (System Analysis and Problems of Urban Development), Moscow: Nauka, 1983.

121. Romm, A.P., *Avtomatizatsiya protsessov gradostroitel'nogo proektirovaniya* (Computer-aided Urban Design), Moscow: TSNIIP Gradostroitel'stva, 1983.

122. Steenbrink, P.A., *Optimization of Transport Networks*, New York: Wiley, 1974. Translated under the title *Optimizatsiya transportnykh setei*, Moscow: Transport, 1981.

123. Fedorov, V.P., Mathematical Model of Formation of Passenger Flows, *Izv. Akad. Nauk SSSR, Tekh. Kibern.*, 1974, no. 4, pp. 17–26.

124. Yakovlev, L.A., *Paket prikladnykh programm dlya proektirovaniya sistem gorodskikh putei soobshchenii* (An Applied Software Package for Designing Urban Communications), Moscow: TSNIIP gradostroitel'stva, 1983.

This paper was recommended for publication by V.M. Vishnevskii, a member of the Editorial Board