STOCHASTIC MODELING OF TRAIN DELAYS FORMATION

VLADIMIR CHEBOTAREV[1], BORIS DAVYDOV[2] & KSENIYA KABLUKOVA[1]
[1]CC FEB RAS, Russia
[2]FESTU, Russia

ABSTRACT

Two stochastic models of train interaction are considered in the paper. These models are designed to study the process of calculation the arrival deviation probability density function. The stochastic models allow you to receive an adequate forecast the development of traffic situation taking random fluctuations of the train run trajectory into account. The paper proves the equivalence of two train traffic models that reflect the mechanism of formation the arrival time distribution. Both models take the scattering of train departure deviations into account. The first type model describes the result of the run time random nature, while the second model reflects the impact of short-term unplanned train stops. The study also makes an attempt to outline the regularity of formation the standard and the abnormal arrival deviation distributions. The proposed models are verified by using the results of the historical data analysis obtained at the main railway line.

Keywords: arrival deviation distribution, bimodal distribution, departure time distribution, random speed variations, running time distribution, stochastic model, train traffic.

1 INTRODUCTION

Stochastic models of train traffic are used for probabilistic forecasting of the traffic situation development in the presence of random disturbances (see, for example, [1] and [2]). They make it possible to identify inter-transport conflicts and predict their further propagation over the railway network. Train delay propagation is one of the significant negative consequences of these conflicts. The models proposed by the researchers are aimed at eliminating delays or reducing of their duration.

For each conflict situation, a stochastic model of train interaction can be created (before, at the time and after the conflict point). Distributions of the following random variables can be considered as input data of such model: departure deviation from the schedule, time headway, speed, travel time, primary delay, etc. The simulation results are the dependences of the arrival delay probability distribution on the input distributions. In turn, knowledge of the delay distribution is the basis for predicting the delay duration, as well as a set of measures to eliminate it.

In many papers (see, for example, [2]–[5]), when constructing the delay distribution function it is assumed that the types of input distributions are fixed (as a rule, this is an exponential distribution or its modifications). Such models describe only particular cases and do not cover other possible variants of input distributions observed in reality. Our study shows that a more general approach is to represent the input random variables (run time, in particular) using the gamma distribution.

Two models of a conflict situation development are proposed in the paper. Both models have two-train type and can be used when considering the pair interaction of trains on the railway network. The first model is used to describe a situation when both trains have different speeds, and the second one is for the situation when a primary delay of a random duration occurs. At the same time there are no restrictions on the form and type of input distributions.

Under the relevant conditions, the distribution of the random arrival deviation of the second train and the distribution of the total arrival deviation of both trains were found for each model. It was shown that under definite natural conditions the proposed models are equivalent in a certain sense.

In the considered models, input random variables can have arbitrary suitable distributions. This fact expands the possibilities of applying the models to various situations (when input random variables behave differently, and the essence of the conflict situation remains unchanged).

The equivalence of both models makes it possible to check only one of them for compliance with real statistics obtained from Russian Railways. The travel time is an input random variable of the model. When reviewing the statistical data and their subsequent processing, it was noticed that the histograms for the travel time often visually resemble the densities of bimodal distributions (see Table 3, column 1). Mixtures of distributions are applied for modeling the abovementioned bimodality. The paper considers both unimodal and bimodal travel time distribution. It is shown that the formula found for the arrival deviation distribution of the second train is consistent with statistical data.

2 LITERATURE REVIEW

Modeling the process of random delays occurrence and their propagation along the train chain makes it possible to assign rational traffic adjustments and predict their likely consequences. The analytical method which uses convolutions of distributions of the initial (at the site entry) and newly arising deviations to obtain the distribution of delay is proposed in [6]. But applications of this work perhaps are limited since only the uniform distribution the secondary delays is used. In reality, this assumption is rarely confirmed.

One effective approach which allows us to predict statistical characteristics of the train traffic is Situational-Heuristic Method (SEMN) [7]. Expected values of operating times are obtained from statistics of previous periods taking the current on-site situation into account. A similar approach is used in the paper [8] which models the local traffic using average values of train processing time. It should be noted that there are attempts to derive some average train schedule based on the emerging situation. Such attempts are fraught with conservation of technological problems that exist on the site.

One of the fundamental papers is devoted to the problem of traffic stochastic modeling under consideration [1]. The authors describe a model of delay formation in a train set traffic using a probabilistic approach. Total run time of the train is considered as a sum of random intervals of movement along the section elements when analyzing the process of delay propagation. The paper [3] shows that calculation of the arrival time distribution at the terminal station is based on repeated use of convoluting the distributions of random variables characterizing each of the operations.

Below mentioned papers use this approach to analyze the process of delay propagation along the train chain. The authors of the study [4] solve the problem of interacting the two trains in a flow that have different speeds and are influenced by random impacts. The result of analysis is used to model the train traffic in a dense heterogeneous train flow. This idea is developed in the papers [2], [3], [5]. The paper [2] represents trains traffic in a form occupying an intermediate position between the macro-models and models with very detailed description of the process, i.e. micro-models. Approximation method is proposed for exact representation of delay distributions. Cumulative distribution is calculated from the sequence of activities which is determined by a stochastic event-graph.

Most fully the problem of predicting the random delay appearance and propagation across the large railway network is considered in [5]. Scattering of arrival moments is treated as a random variable generated with joint accounting the departure times and the running intervals. The main subject of research is the process of adjusting the schedule at transfer stations in the presence of multiple delays. The model operates with discrete distributions of travel, dwell times and deviations from the schedule. The authors use a simplified description of input random variables so that they receive a small computing time.

The proposed approach was further developed in the study [3] which attempts to take the realistic distribution of operation time into account and choose the appropriate approximation. The approach proposed uses a probabilistic operational graph which considers transfer operation and conflict situations. The authors argue that mesoscopic modelling of traffic is the purposeful approach to compute the delay propagation. The purpose of the study is to reduce computing time. The authors consider only station and crossing as the operating points where conflicts can occur.

The general characteristic of the schedule which is realized under conditions of random disturbances is robustness. Timetable robustness is understood as its ability to absorb small current deviations [9], [10]. The compensatory mechanism begins to fail with increasing the intensity of traffic which leads to the multiplication of delays. A characteristic that shows the ability to eliminate the disturbances that occur is defined as timetable stability [11] and it is connected with the time interval that is needed to restore the normative traffic. The results obtained in these papers do not allow analytically combining the results of calculating the probability distributions after each operation. This is due to the fact that each operation corresponds to a separate stochastic model. The general model for all kinds of operations is not considered in this research.

Paper [12] is devoted to analysis of train delays and their evolution in real time. Bayesian networks are the framework for design the stochastic model in [12]. Bayesian networks are an appropriate method to represent the complex interdependencies between train events. Proposed stochastic model adequately reflects train delays due to interactions with other trains. However, this approach does not allow to identify the cause of the delay and to develop a mechanism to eliminate this cause.

3 TWO STOCHASTIC MODELS OF THE ARRIVAL DISTRIBUTION FORMATION

In this section, the reader is presented with two models of the arrival deviation formation at the destination station. Each of them reflects a conflict situation arising between a pair of trains on a track.

The conflict, which is described by the model 1, stems from the difference in train speeds. The follower, which has a higher speed, catches up with the feeder train at some point of the track. The indicated situation is schematically shown in Fig. 1(a). The lines show the possible trajectories of trains here. The lines emanating from one point on the time axis show probable trains' paths. The conflict arises at the time when the time headway between trains is reduced to the minimum allowable unit (s_0).

Model 2 serves to describe the inter-train conflict caused by the initial delay of the leader train. Figure 1(b) provides a diagram of such conflict situation occurrence. The "bundles" of parallel lines originating on the horizontal axis mean that the departure times are random variables. The horizontal lines represent the primary (upper line) and secondary (lower line) delays, which are also random. The conflict occurs at a time when further train rapprochement is impossible. The follower stops and continues moving immediately when the leader train departs.
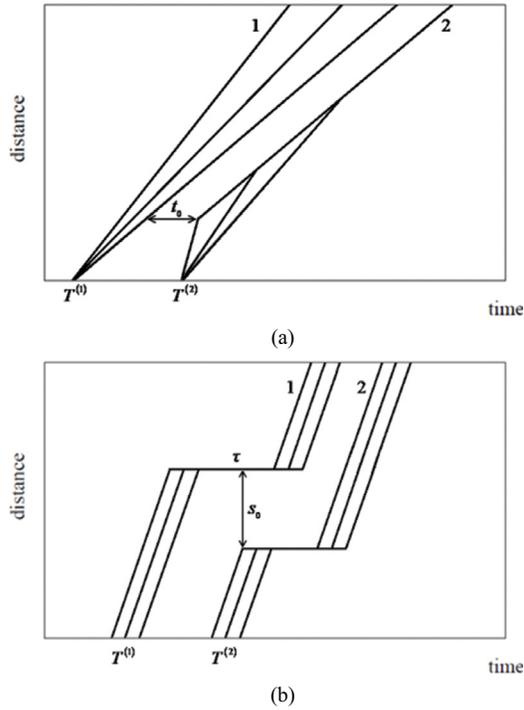
Figure 1: Conflict situations arising from different train speeds (a) and random primary delay (b).

Two models of the arrival deviation formation are described in Subsections 3.1 and 3.2. Each of models corresponds to conflict situation from Fig. 1.

3.1. Model 1 (the conflict between vehicles due to the difference in their speeds)

Consider the following two-train stochastic model. Trains with numbers 1 and 2 depart one by one from a certain starting point of the route and move in the same direction. Their scheduled time-spatial trajectories are the same. It is assumed that the condition of compliance a minimum safe time distance t_0 between trains is fulfilled. Trains are influenced by random factors that lead to scattering of departure and travel times.

The most common situation of arrival deviation occurrence is considered below. A conflict situation between two trains occurs on open tracks. The difference in train speeds and the departure deviations are the reasons of delay (i.e. of arrival deviation).

Let us introduce the following notations to describe the movement of a train with the number i ($i = 1,2$):

δ_i is the departure deviation from the schedule, $F(t) := P(\delta_i \leq t)$ is the cumulative distribution function of the random variable δ_i (it is assumed that the variables δ_1 and δ_2 are equally distributed); $f(t) = \dfrac{dF(t)}{dt}$ is the probability density function of δ_1,

ρ_i is the train travel time between starting and ending points of the route,

$L(t) = P(\rho_1 \leq t), l(t)$ is the probability density function of ρ_1,

d_i is the departure time of the ith train according to schedule,

a_i is the arrival time of the ith train according to schedule,

ξ_i is the random deviation of the ith train from the arrival schedule (note that the arrival deviation ξ_i may take different signs); then $a_i + \xi_i$ is actual arrival time of the ith train,

$U_i(t) = P(\xi_i \leq t)$ $u_i(t)$ is the probability density function of ξ_i,

$U(t) = P(\xi_1 + \xi_2 < t)$, $u(t)$ is the probability density function of the total delay $\xi_1 + \xi_2$,

t_0 is the minimum permissible time headway between trains,

$s_0 = s_0(v) = v \cdot t_0$ is the minimum safe distance between trains, corresponding to an arbitrary speed v.

The following assumptions are the basis for building the model 1.

Assumption 1. For the actual departure times the inequality $d_1 + \delta_1 < d_2 + \delta_2$ holds.

Assumption 2. Random variables ρ_1 and δ_1 are independent.

The inequality from Assumption 1 means that the departure order of trains is fixed.

Consider the movement of a pair of trains along the track. Train speeds are different. The situation when the second train is faster than the first one is described. In this case, when the time interval between trains becomes equal to t_0, a conflict situation arises, as shown in Fig. 2. The distance s_0 corresponds to the time interval t_0. Note that in Fig. 2 s_0 is depicted as a constant just for the sake of simplicity.

Let us derive equality

$$U_2(t) = I(t > 0) \int_0^\infty F(t - d_1 + a_2 - t_0 - y) l(y) dy, \tag{1}$$
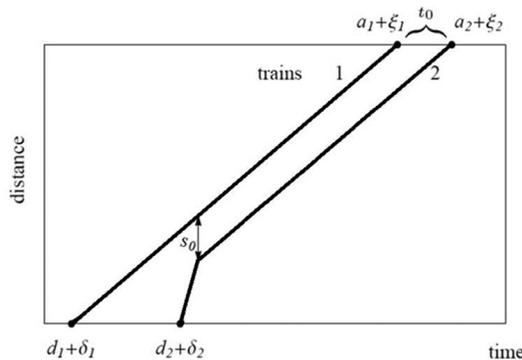
Figure 2: The layout of arrival deviation ξ_2 formation (model 1).

based on the arrival time formation layout, shown in Fig. 2 (see also [13]).

It can be seen from Fig. 2 that $a_1 + \xi_1 - (d_1 + \delta_1) = \rho_1$, hence

$$a_1 + \xi_1 = \rho_1 + \delta_1 + d_1. \tag{2}$$

Moreover, $a_2 + \xi_2 = a_1 + \xi_1 + t_0$. Thus,

$$\xi_2 = \rho_1 + \delta_1 + d_1 - a_2 + t_0, \tag{3}$$

$$U_2(t) = I(t > 0)\mathrm{P}(\rho_1 + \delta_1 < t - d_1 + a_2 - t_0). \tag{4}$$

Under the assumption of ρ_1 and δ_1 independence equality (1) holds (using the well-known fact that the distribution of the sum of independent random variables is the convolution of the initial distributions). In turn, equality (1) implies the formula for the probability density function of ξ_2,

$$u_2(t) = I(t > 0)\int_0^\infty f(t - d_1 + a_2 - t_0 - y)l(y)dy \tag{5}$$

Next, the distribution of the random sum $\xi_1 + \xi_2$ will be received.

It follows from (2) and (3) that

$$\xi_1 + \xi_2 = 2(\rho_1 + \delta_1 + d_1) + t_0 - (a_1 + a_2). \tag{6}$$

In turn, equation (6) entails next formula

$$U(t) = \mathrm{P}(\xi_1 + \xi_2 < t) = \mathrm{P}(\rho_1 + \delta_1 < 0.5(t - t_0 + a_1 + a_2) - d_1).$$

Hence, by analogy with (1), the following equality is obtained:

$$U(t) = \int_0^\infty F(0.5(t - t_0 + a_1 + a_2) - d_1 - y)l(y)dy. \tag{7}$$

Consequently,

$$u(t) = 0.5\int_0^\infty f(0.5(t - t_0 + a_1 + a_2) - d_1 - y)l(y)dy. \tag{8}$$

Thus, a model that allows us to find both the distribution of the second train delay and the distribution of the total arrival delay was obtained. Arbitrary suitable distributions can be taken as distributions of input random variables (departure deviation and travel time).

3.2. Model 2 (the train traffic disturbed by a random primary delay)

Two trains follow one track one after another in one direction from station A to station B with the same average speed v_0. The distance from the train 2 to the train 1 is denoted by $X_2 + s_0$, where $s_0 = s_0(v_0) = v_0 \cdot t_0$ is the minimal safe distance between trains, X_2 is random variable (without any assumptions about its distributions). Both trains have the same destination station. Let us also introduce the notation $\mu_2 = X_2/v_0$. Time points $T^{(j)}$, $j = 1, 2$, denote actual departure times from station A. Time headway is equal to $T^{(2)} - T^{(1)} = \mu_2 + t_0$.

Suppose that train 1 made an unplanned stop at some point in time. The random duration of this stop is called the primary delay and is denoted by τ. Depending on the value of τ, the subsequent train may receive a delay, which is called by the secondary delay. It is assumed that the following train stops when the distance to the front train is reduced to s_0. Moreover, it is assumed that as soon as the front train restore running then the next one immediately follows it.

In what follows, $\tau_1 = \tau$, and τ_2 is the secondary delay duration. It should be noted that within the framework of model 2 the deviation of the real arrival time from the planned one for every train coincides with τ_k, $k = 1, 2$. Figure 3 illustrates the process of output delay formation which is described by the model 2.

The dotted lines (lines $\bar{1}$ and $\bar{2}$) represent the scheduled trajectories of trains 1 and 2, solid lines (1 and 2) depict the real trajectories taking into account the delays. It can be seen that the arrival time of the train 1 differ from the schedule on τ and the train 2 on τ_2.

Assume that departure times from the original station are disturbed by small random influences. Let μ_2 have a density function $\psi(t)$ ($\mu_2 + t_0 = T^{(2)} - T^{(1)}$). In addition assume, the duration τ of the unscheduled primary stop has a density function $g(t)$. The condition of our model is that random variables μ_2 and τ are independent.

The problem is to find the distribution function $G_2(t) = P(\tau_2 < t)$. By the previous study [14], the resulting cumulative distribution function of delay τ_2 has the following form:

$$G_2(t) = I(t > 0) \int_{-t}^{\infty} \left[\int_0^{t+y} g(z) dz \right] \psi(y) dy. \tag{9}$$

The probability density function corresponding to eqn (9) is following:

$$g_2(t) = I(t > 0) \int_0^{\infty} g(t+y) \psi(y) dy. \tag{10}$$
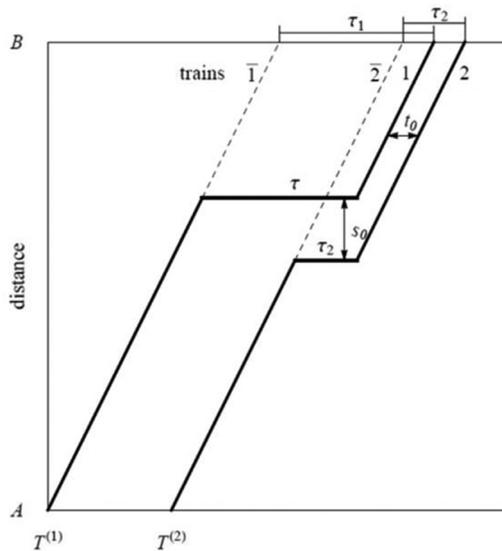
Figure 3: The layout of arrival deviation τ_2 formation (model 2).

4 EQUIVALENCE OF THE TWO CONSIDERED MODELS

The results of Section 3 are collected in Table 1. Table 1 contains the key formulas of models 1 and 2. This section will show the equivalence of the formulas from the first row of Table 1.

Let us show the equivalence of models 1 and 2. In other words, show that $u_2(t) = g_2(t)$, where the function $u_2(t)$ is defined by the equality (5), and $g_2(t)$ by the equality (10) (see Table 1 row 3).

The relationship of random variables included in both models can be described as follows. Model 1 is declared in terms of random variables δ_1 (with density $f(t)$) and ρ_1 (with density $l(t)$. Model 2 is formulated in terms of random variables τ (with density $g(t)$) and μ_2 (with density $\psi(t)$).

Comparing the values that determine the time headways in both models, derive $T^{(2)} - T^{(1)} = \mu_2 + t_0 = d_2 + \delta_2 - d_1 - \delta_1$. Therefore,

$$\mu_2 = \delta_2 - \delta_1 + d_2 - d_1 - t_0 \tag{11}$$

By (2), $a_1 + \xi_1 - (d_1 + \delta_1) = \rho_1$. Hence, since $\xi_1 = \tau$, then

$$\tau = \rho_1 - (a_1 - d_1 - \delta_1). \tag{12}$$

Figure 3 depicts that $\tau_2 + \mu_2 + t_0 = \tau + t_0$. Here of, $\tau_2 = \tau - \mu_2$. Using now the equalities (11) and (12), obtain

$$\tau_2 = \tau - \mu_2 = \rho_1 - (a_1 - d_1 - \delta_1) - (\delta_2 - \delta_1 + d_2 - d_1 - t_0).$$

Next, using the equality $a_1 - d_1 - \delta_1 = a_2 - d_2 - \delta_2$ (the planned travel times of both trains are identical according to the model) leads to

$$\tau_2 = \rho_1 - (a_2 - d_2 - \delta_2) - (\delta_2 - \delta_1 + d_2 - d_1 - t_0) = \rho_1 + \delta_1 + d_1 - a_2 + t_0 \equiv \xi_2.$$

The last equality follows from (3). This implies the equivalence of densities $u_2(t) = g_2(t)$. Figure 4 illustrates the equality $\xi_2 = \tau_2$.

The equivalence of models 1 and 2 was proved. These models describe the occurrence of the arrival delay under different input assumptions. In model 1, arrival delay ξ_2 occurs as a

Table 1: Key formulas of models 1 and 2.

	Model 1	Model 2
Investigated random variable	$\mu_2 = \delta_2 - \delta_1 + d_2 - d_1 - t_0$	$\tau_2 = \tau - \mu_2$
Cumulative distribution function	$U_2(t) = \int_0^\infty F(t - d_1 + a_2 - t_0 - y) l(y)\, dy\; t > 0$	$G_2(t) = \int_{-t}^\infty \left[\int_0^{t+y} g(z)\, dz \right] \psi(y)\, dy$ $t > 0$
Probability density function	$u_2(t) = \int_0^\infty f(t - d_1 + a_2 - t_0 - y) l(y)\, dy\; t > 0$	$g_2(t) = \int_0^\infty g(t + y) \psi(y)\, dy$ $t > 0$
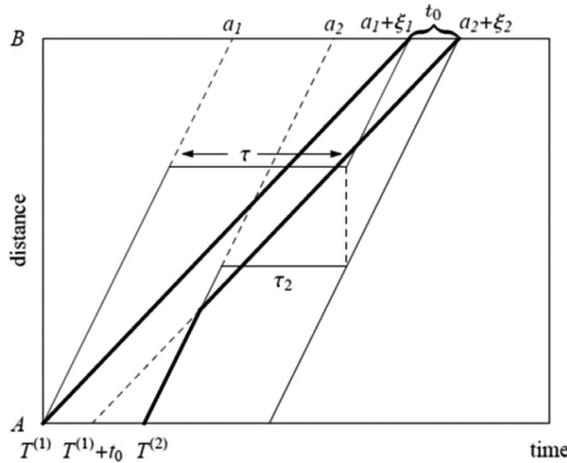
Figure 4: Graphical mapping of equivalence the models 1 and 2.

result of randomness of the departure deviations δ_1, δ_2 and travel time ρ_1. In turn, in model 2, arrival delay τ_2 occurs as a result of randomness of the time headway μ_2 and the primary delay τ. Random values in both models are interrelated: deviations δ_1, δ_2 (from model 1) with the time headway μ_2 (from model 2), travel time ρ_1 (from model 1) with the primary delay τ (from model 2). Thus, the models of the arrival delay formation are equivalent if the input random variables of these models are related to each other.

5 STATISTICAL VERIFICATION OF THE MODEL 1

In this section comparing the formula (5) (model 1) with real statistics is conducted. Statistical analysis aims to find out which of input distributions of the density function (5) is consistent with real data.

Note that in the formula (5) there are two input distributions. These are the distribution of the departure deviation $f(t)$ and the distribution of the travel time $l(t)$. Both of them must be set in accordance with the statistics of real train traffic.

The statistics collected for the trains of the suburban radial railway line Moscow-Tver are explored. Moscow-Tver line is characterized by double track style, length 167 km. The line is a part of the October Railway, linking the two largest cities in Russia: Moscow and St. Petersburg. The line is designed for both passenger and freight traffic. Passenger traffic is about 250 thousand people per month. The time headway during rush hours is from 2 to 15 min, depending on the type of train (high-speed, standard).

Let us set the actual form and parameters of the input distributions $f(t)$ and $l(t)$ in (5). Suppose that the departure deviation δ_1 obeys the exponential distribution law with a density function:

$$f(t) = I(t > 0)\lambda e^{-\lambda t}, \tag{13}$$

where $\lambda > 0$.

Consider two examples in which the function $f(t)$ will have the form (13), and the function $l(t)$ will have a unimodal (Example 1) or bimodal (Example 2) form. The unimodal density

function of the travel time is characteristic of hours when train traffic is not intense (e.g., daytime hours). Bimodality, on the contrary, is manifested in hours with massive train traffic (e.g., morning and evening rush hours) and is associated with increasing the number of random factors affecting the train traffic.

Example 1. Studies show that in the case of relatively free traffic (e.g., daytime hours), the graph of the probability density function of the travel time has a longer right branch. It is shown in [14] that such distribution is well approximated by the gamma distribution, i.e.

$$l(t) = I(t > b_1) \frac{e^{-(t-b_1)/\beta} (t-b_1)^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha}, \tag{14}$$

where $\alpha > 0$, $\beta > 0$, $b_1 > 0$.

The histogram of the travel time given in Table 2 (column 1) is fairly well approximated by shifted gamma distribution (14). Note that for the convenience of displaying in travel time histogram, the vertical axis intersects the horizontal one not at the origin, but at the point (15, 0). Table 2 contains also histogram of arrival deviations (column 2) with approximating density curve $u_2(t)$ from (5). The following calculated parameters of the resulting arrival deviation probability density function $u_2(t)$ are taken:

$$\alpha = 50, \ \beta = 0.2 \ \text{min}, \ b_1 = 10 \ \text{min}, \ \lambda = 2 \ \text{min}^{-1}. \tag{15}$$

It should be noted that the parameter α from (14) is dimensionless. The normative parameters for a given track are the following:

$$d_1 = 0 \ \text{min}, \ a_2 = 23 \ \text{min}, \ t_0 = 3 \ \text{min}. \tag{16}$$

In Table 2, travel times and arrival deviations are measured in minutes (horizontal axis).

As can be seen from Table 2, when the traffic is not intense, the travel time density is unimodal in nature. This is due to the fact that with such traffic the influence of random factors

Table 2: Histograms of the travel times and arrival deviations with approximating density functions (the case of unimodal $l(t)$).

Travel time	Arrival deviations
Daytime hours	
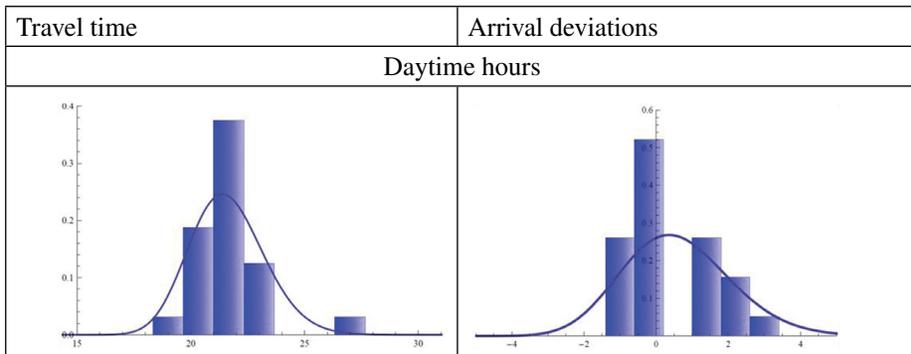
is minimal. These factors do not depend on the train flow itself, but depend on external conditions (weather, visibility, etc.).

Example 2. It turned out that bimodal travel time densities are not uncommon, but appear quite often. According to the collected statistics, at rush hours the train travel time frequently obeys some bimodal distribution law. Take the travel time density function $l(t)$ as a mixture of two gamma densities $l_1(t)$ and $l_2(t)$:

$$l(t) = l(t, p) = pl_1(t) + (1-p)l_2(t), \quad 0 \le P \le 1, \tag{17}$$

where $l_1(t)$ is determined the same as $l(t)$ in (14), $l_2(t) = l_1(t-b_2)$ with $b_2 > 0$.

Bimodality of the density $l(t)$ is determined by the influence of two random factors (e.g., weather conditions and speed mode). To simulate the degree of influence of each of them, the parameter P is used. The more P, the stronger the first factor influences and the weaker the influence of the second factor, and conversely.

Let us set the actual form and parameters of the input distribution $l(t)$ defined by (17). Table 3 contains the examples of histograms plotted from samples of travel time and arrival deviations distributions in the morning and evening rush hours. In Table 3, travel times and arrival deviations are measured in minutes (horizontal axis).

Note that for the convenience of displaying in travel time histograms, the vertical axis intersects the horizontal one not at the origin, but at the point (18, 0).

The given examples show that the appearance of a bimodal travel time distribution leads to a rise in the left branch of the arrival deviation density $u_2(t)$. If the influence of random factors on the train traffic is observed, then it leads to the occurrence of inflection points of the function $u_2(t)$. Thus, the nature of the histograms constructed on the basis of experimental data agrees well with the theoretical results, which are given in Section 3 (model 1) of this article.

Table 3: Histograms of the travel times and arrival deviations with approximating density functions (the case of bimodal $l(t)$).
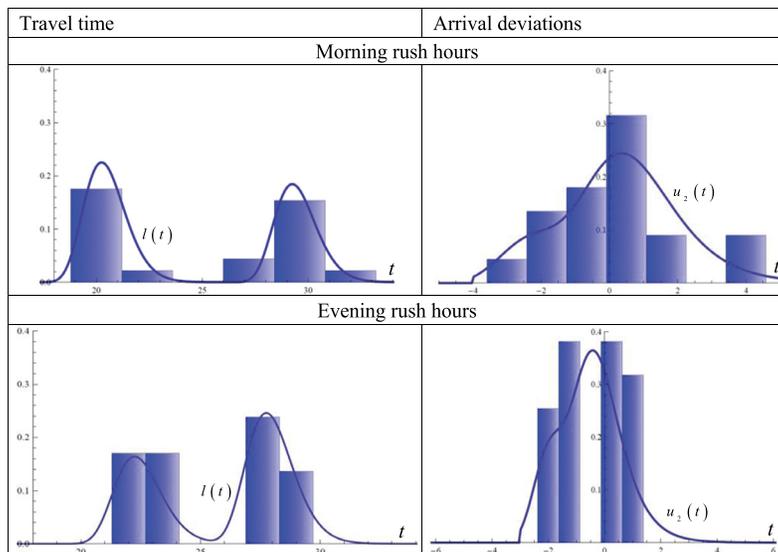
The histograms of the travel time given in Table 3 (column 1) are fairly well approximated by a mixture of gamma distributions (17) with $P = 0.3$. Table 3 contains also histograms of arrival deviations (column 2) for morning and evening rush hours with approximating density curves $u_2(t)$ from (5). For example, the following calculated parameters of the resulting arrival deviation density $u_2(t)$ characterize morning rush hours:

$$P = 0.3, \alpha = 16, \beta = 0.25 \text{ min}, b_1 = 17.5 \text{ min}, b_2 = 2.5 \text{ min}, \lambda = 1 \text{ min}^{-1}. \tag{18}$$

Moreover, the normative parameters for a given track are the following:

$$d_1 = 0 \text{ min}, a_2 = 27 \text{ min}, t_0 = 3 \text{ min}. \tag{19}$$

Parameters correspond to the considered statistical data. From Table 3, it can be seen that the theoretical density $u_2(t)$ from (5) is consistent with the experimental data.

Results similar to passenger train traffic were obtained when considering statistical data on the freight train traffic through Trans-Siberian main line. The analysis shows that during periods of increase the train flow intensity, the number of sections in which a bimodal distribution of travel time is observed reaches 20%. Accordingly, the shape of the arrival deviation density curve becomes symmetrical or has left-sided skewness.

6 CONCLUSIONS

The article considers two models of train interaction on a track. The first model describes a conflict situation arising due to the difference in train speeds (the second train catches up with the first one at the time of movement). The second model describes the situation with a primary delay of random duration.

For both conflict situations, the distribution function of arrival deviation (delay) at the destination station was found. The distribution function allows you to predict arrival delays based on the distribution of input random variables.

The paper shows that models 1 and 2 are completely equivalent.

The input random variables of model 1 (departure deviations and the travel time) are associated with the input random variables of model 2 (time headway and primary delay) by exact equalities. The fact of equivalence allows checking the consistency the one of models with statistical data (another one is considered consistent, because the first is agreed). Therefore, a verification of model 1, for which there were suitable statistics, was carried out.

To simulate the input densities of travel times, bimodal distributions are used. The paper discusses a method for specifying a bimodal density function of the travel time using a mixture of gamma distributions. Both the input (departure deviation and travel time) and the resulting (arrival deviation) densities are in good agreement with the real-world data.

The flexibility of both models is that the resulting analytical formulas allow us using arbitrary theoretical distributions that adequately reflect the influence of input factors in various real situations.

REFERENCES

[1] Carey, M. & Kwiecinski, A., Stochastic approximation to the effects of headways on knock-on delays of trains. *Transportation Research Part B*, **28(4)**, pp. 251–267, 1994.

[2] Meester, L.E. & Muns, S., Stochastic delay propagation in railway networks and phase-type distributions. *Transportation Research Part B*, **41**, pp. 218–230, 2007.

[3] Buker, T. & Seybold, B., Stochastic modelling of delay propagation in large networks. *Journal of Rail Transport Planning and Management*, **2(12)**, pp. 34–50, 2012.

[4] Boucherie, R.J. & Huisman, T., Running times on railway sections with heterogeneous train traffic. *Transportation Research Part B: Methodological*, **35(3)**, pp. 271–292, 2001.

[5] Berger, A., Gebhardt, A., Muller-Hannemann, M. & Ostrowski, M., Stochastic Delay Prediction in Large Train Networks. *Proceedings of the 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'11), Saarbrucken, Germany*, pp. 100–111, 2011.

[6] Muhlhans, E., Berechnung der Verspatungsentwicklung bei Zugfahrten/Eisenbahntechn. *Rundschau ETR*, **39(7/8)**, pp. 465–468, 1990.

[7] Shapkin, I.N., Yusipov, R.A. & Kozhanov, E.M., Simulation of train functioning on the basis of multivariate regulation of technological operations. *Bulletin VNIIZhT*, **4**, pp. 30–36, 2006 (in Russian).

[8] Karetnikov, A.D. & Vorobyev, N.A., *Schedule of Train Traffic*. Transport, Moscow, 301 p., 1979 (in Russian).

[9] Salido, M.A., Robustness in Railway Transportation Scheduling. Salido, M.A., Barber, F. & Ingolotti, L. (eds.), *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA 2008), Chongqing, China*, pp. 2833–2837, 2008.

[10] Al-Ibrahim, A., *Dynamic Delay Management at Railways. A Semi-Markovian Decision Approach*, PhD thesis, Universiteit van Amsterdam, 335 p., 2010.

[11] Goverde, R.M.P., *Punctuality of Railway Operations and Timetable Stability Analysis*, PhD thesis, Technical University of Delft, 165 p., 2005.

[12] Corman, F. & Kecman, P., Stochastic prediction of train delays in real-time using Bayesian networks. *Transportation Research Part C: Emerging Technologies*, **95**, pp. 599–615, 2018. DOI:10.1016/j.trc.2018.08.003.

[13] Chebotarev, V., Davydov, B. & Kablukova, K., Random delays forming in the dense train flow. *Proceedings of the 16th International Conference on Railway Engineering Design & Operation (COMPRAIL 2018), Lisbon, Portugal*, pp. 435–445, 2018.

[14] Chebotarev, V., Davydov, B. & Kablukova, K., Probabilistic Model of Delay Propagation along the Train Flow. *Probabilistic Modeling in System Engineering* (ed. Kostogryzov, A.), IntechOpen, pp. 171–193, 2018. DOI: 10.5772/intechopen.75494.