The Passengers' Turnout Simulation for the Urban Transport System Control Decision-Making Process

Ilya V. Stepanchenko(✉), Elena G. Krushel, and Alexander E. Panfilov

Kamyshin Technological Institute, Volgograd State Technical University, Kamyshin 403874, Russian Federation
stilvi@mail.ru

Abstract. The report deals with the problems of the small city passengers' municipal transport system simulation. The particular purpose of the report concerns to the mathematical model of the passengers' turnout elaboration. The proposed stochastic model of passengers' arrivals at the stops of the transport net as well as the model of the passengers' transition between the stops is applicable to the passengers' turnout simulation for the decision-making on the municipal transport system efficiency raise. The results of the long-term observations of the real transport system operation and the urban transport passengers flow big data cognitive processing have applied for the model parameters identification and adequacy checkup. The model has the form of tabular probability distribution function depending on the passengers' start and destination stops names, the type of the weekday, the number of time interval within the day, and the passengers' social group belonging. The model testing shows the high correlation of the simulated passengers' flow with the real observations results. The confidence level of the accordance between the simulated data probability distribution and the experimental data empiric probability distribution is moderate. In spite of it, the presented model is applicable for the passengers' turnout simulation.

Keywords: Simulation · Passengers' transport system · Passengers flow · Stochastic tabular model · Turnout

1 Introduction

The appropriation of the municipal passengers' transport consists in the passengers' transition within the city bounds. The transport service as well as citizens supplying gas, electricity, and water is the essential social facility, which should be the subject of the municipal checkup and control.

However, the passengers' transition fulfillment sometimes is not profitable for the following reasons. First, the upper level of the trip price may be strictly limited by the citizens' financial possibilities restriction. Second, the fuel and spare parts costs may be over the top acceptable for the profit receiving. Third, the faults of the transport system

A. Kravets et al. (Eds.): CIT&DS 2017, CCIS 754, pp. 389–398, 2017.
DOI: 10.1007/978-3-319-65551-2_28

control may take place, e.g. the high level of the exploitation expenses, the deterioration of the transport facilities, the unsuccessful transport schedule etc.

Nevertheless, the municipal passengers' transport profitability is extremely desirable. The reasonable economic efficiency of the private transport exploitation comparable with the municipal one can offer the indirect proof of the municipal transport profitableness potential attainment.

One of the ways of municipal transport system's efficiency raise consists of the carrying out and further maintenance of the rational transport facilities schedule. Such approach is quite perspective from the control possibilities point of view since the corresponding problems are the part of transport enterprise administration's scope of functions.

The effective schedule of the passengers' transport facilities traffic calls for the passengers' flow simulation instead and in addition to the real experiments since the durable experiments with the operating transport system should disturb the transport service regularity leading to the citizens' dissatisfaction [1–3]. The appropriate model carrying out and further computational experiments execution are based on the general simulation methods of the queue systems study as well as on the random series generation technique. The corresponding model acquisition needs for the real passengers flow reception (e.g. by means of the known passengers' flow observation table method [1]) and further data mining of the observation results.

The passengers' transition model consists of the following submodules. The first submodule presents the model of passengers flow at the stops allocations on municipal transport system net (briefly named further as "model of passengers' arrivals"). The second submodule presents the model of the single transport facility trip. The presented report deals with the first model justified by the real passengers flow observations carried out in the Kamyshin city of Volgograd region, Russia [1].

The problems of the urban transport systems' operation effectiveness raise as well as the urban transport facilities flow control are widely presented in the scientific publications concerning to the motor transport flow simulation and study [4–6]. Particularly, several academic journals deal purely with motor flow's dynamics [7–9] (e.g. "Transportation Research", "Transportation Science"). The analysis is addressed both to the transport nets as a whole and to the separate elements of the net. The discussion touches the following problems:

- Transport nodal points location and transport road crossing control [4–6];
- Effective control of the transport cargo traffic [10];
- Optimization of passengers' routes [11, 12];
- Passengers' behavior strategy estimation and the ways of purposeful on it [13, 14].

Two main approaches have formed historically by now: firstly deterministic and secondly probabilistic (stochastic) ones. The deterministic approach supposes the existence of the dependence between the particular factors (e.g. the correspondence of the rate and distances amongst the motors in the transport flow). The stochastic models treat the transport flow as the probability process.

The small towns possess the set of specific transport problems, which the scientific community almost ignores. The main of them are follows:

- The insufficient approach to the generalized consideration of urban transport control system as a whole, particularly: the lack of the systematic analysis of the motors’ utilized capacity, ill-founded motors distribution between the routes, unjustified trip price value, the groundless choice of the motors types etc.
- The low economic efficiency (oftentimes the absence of profitableness) of the municipal transport systems caused by the transport facilities incomplete filling as well as the competition between the social, private, and personal transport facilities.
- The existence of the extremely g-loaded highways’ sections between the points of the intensive passengers flows (particularly in the small towns such sections are usually located near the central markets). The exhausts from the motors pipes cause the dangerous pollutant concentrations in the living zones nearby the highways.

2 Problem Statement

The passengers’ arrivals model ought to estimate the number of passengers forthcoming to the particular transport’s stop during the given time interval and to settle the name of destination stop for each of forthcoming passengers. The number of the forthcoming passengers should depend on the following factors: the start as well the destination stops, the time interval within a day, the weekday name, and the passenger’s social group. The factors roster was determined in accordance with the results presented in [1]).

The results of the complete passengers flow observations gave the following model factors details [1]:

- Each start and terminal stop of transport net may treat as the subject of simulation.
- The number of the time interval (its duration is 1 h) lays within the bounds from 5:00 until 20:00 accordingly to the real municipal transport operation mode.
- The weekday name varies from Sunday to Saturday.
- The social groups of passengers are following: pupil, student, worker, unemployed, and pensioner. Each group consists of the two subgroups depending on the availability vs. the absence of the concessionary ticket.

The following efforts are necessary for the model of passengers’ arrivals at the stops of the transport system’s net elaboration:

- develop the model form and structure for the formal arrival’s model reception;
- generalize the data of the passengers’ destination stops choose for each stop treated as the start of the passenger’s arrivals on the base of the primary observation results [1];
- estimate the statistical characteristics (sufficient for the probability density estimation) of the passengers’ arrivals process as well as the destination stop choice depending on the stop allocation, passenger’s social group, and daytime;
- analyze the possibility of the parameters’ averaged values application instead of the particular parameters values checked up during the real observations;
- justify the model adequacy by means of the simulation results analysis.

The purpose of the parameters' averaged values application consists in the decrease of the parameters number by means of the particular parameters values replacement by the set of the aggregated parameters.

The model justification should be carried out for the several stops with the essential distinction of the passengers' turnout.

The criteria for the model adequacy justification are following:

- The chi-square test and the standard deviation value application as the fitting criteria of the probability laws choice for the simulation of the passengers flow intensity and passengers' social groups belonging.
- The relative error's and the correlation coefficient values application for the validation of mean values of the passengers flows intensity and the number of passengers in the different social groups.

3 The Passengers' Arrivals Model Formalization

3.1 The Formal Model of the Passengers' Arrival

In accordance to [1] the general model form may be presented as follows:

$$x_{i,j,s,k,m} = F(V_i, V_j, T_s, D_k, K_m), \tag{1}$$

where V is the set of all stops at the urban transport net; i, j are the indices of the start and terminal stops correspondingly, $i,j \in 1\ldots|V|$; T is the set of the non-overlapping time intervals the join of which cover the whole workday of the municipal passengers transport operation, $s \in 1\ldots|T|$; D is the weekdays set (from Sunday to Saturday), $k \in 1\ldots|D|$; K is the set of passengers types classified by the social groups belonging and by the concessionary ticket availability, $m \in 1\ldots|K|$; $x_{i,j,s,k,m}$ is the number of the m^{th} type passengers arrived at the i^{th} stop during the s^{th} time interval of k^{th} weekday and selected the j^{th} stop as the terminal point.

The abstract function $F(\ldots)$ in (1) belongs to the stochastic functions group. The determination of $F(\ldots)$ in the general mathematical form of the certain random values probability distribution law is possible but the lack of data for its reliable identification should lead to the essential inaccuracy. The acceptable accuracy is obtainable only in the case of the values of four arguments fixation from five ones in (1). Hence, in this case, the number of analyzing functions should raise. Therefore, the choice of the function $F(\ldots)$ should be compromising.

Since the function $F(\ldots)$ describes the certain distribution of the random passengers' arriving at the transport stop, it is reasonable to use the tabular form of $F(\ldots)$ representation.

Although the tabular form of $F(\ldots)$ in (1) ensures the correct approximation of the $x_{i,\ j,\ s,\ k,\ m}$ probability density function, several disadvantages are attended:

1. The model is able to forecast the passengers' number, the arrivals of which is expected during the s^{th} time interval. However, the exact moment of the individual passenger's arrival within the s^{th} time interval is out of the model's possibilities.

At the same time, just this parameter is necessary for the transport system simulation based on the queue system theory.
2. The tabular form of $F(\ldots)$ requires considerable memory volume to store the corresponding table. Consequently, the speed of simulation process decreases.
3. The adequacy of the model (1) depends on the experimental data representativeness. However, the large-scale observations of the whole set of routes at the whole week and for time intervals within the whole workday is mostly unacceptable because of unreasonable time and resources expenses.

First, of the failings can be eliminated by the appropriate choice of the time scale. Thus, assume the time intervals s in (1) as equal to the simulation time discrete (e.g. one minute). Therefore, the moment of the passengers' $x_{i,\ j,\ s,\ k,\ m}$ arrivals at the s^{th} time discrete becomes fixed and hence the simulation of the passengers' transportation system based on the queue systems theory results would be admissible.

3.2 The Stochastic Parameters Choice

The parameters of the tabular function $F(\ldots)$ in (1) can be determined using the statistical processing of the observations results concerning the data of the passengers' arrivals at the stops and their transition between the stops.

The estimation of the following statistical parameters is in need:

- The expectation of the arrivals number per time unit (a minute in our case) of the passengers from the definite social group at the definite workday for each destination stop:

$$Mx_{i,k,m} = \frac{1}{60 \cdot |T|} \sum_{j=1}^{|V|} \sum_{s=1}^{|T|} x_{i,j,s,k,m}; \tag{2}$$

- The expectation of the arrivals number per time unit (a minute in our case) of the passengers from the definite social group for all days cumulatively (i.e. for observations period as a whole):

$$Mx_{i,m} = \frac{1}{|D|} \sum_{k=1}^{|D|} Mx_{i,k,m}; \tag{3}$$

- The pair correlation coefficient between the expectations' estimates of the arrivals number per time unit (a minute in our case) of the passengers from the definite social group at the differing workdays:

$$rmx_{i,k1,k2} = \frac{E(Mx_{i,k1,m} \cdot Mx_{i,k2,m}) - E(Mx_{i,k1,m}) \cdot E(Mx_{i,k2,m})}{\sqrt{Var(Mx_{i,k1,m}) \cdot Var(Mx_{i,k2,m})}}, \tag{4}$$

where $E(\ldots)$ is the designation of the mathematical expectation function, $Var(\ldots)$ is the designation of the variance, $k1$, $k2$ are the workdays indices, $k1, k2 \in D$.

- The pair correlation coefficient between expectations' estimates of the arrivals number per time unit (a minute in our case) of the passengers from the definite social group for all days cumulatively:

$$rmx_{i,k} = \frac{E(Mx_{i,k,m} \cdot Mx_{i,m}) - E(Mx_{i,k,m}) \cdot E(Mx_{i,m})}{\sqrt{Var(Mx_{i,k,m}) \cdot Var(Mx_{i,m})}}; \tag{5}$$

- The passengers' arrivals at each hour of the definite workday as the part of the arrivals within the whole of the workday (regardless the passengers' social group):

$$dx_{i,s,k} = \sum_{j=1}^{|V|} \sum_{m=1}^{|K|} x_{i,j,s,k,m} \Bigg/ \sum_{j=1}^{|V|} \sum_{m=1}^{|K|} \sum_{s=1}^{|T|} x_{i,j,s,k,m}; \tag{6}$$

- The passengers' arrivals expectation at each hour as the part of the arrivals within the whole of the workday (regardless the passengers' social group):

$$dx_{i,s} = \frac{1}{|D|} \cdot \sum_{k=1}^{|D|} dx_{i,s,k}; \tag{7}$$

- The pair correlation coefficient between the parts of passengers arrivals for each hour of the differing workdays (regardless the passengers' social group):

$$rdx_{i,k1,k2} = \frac{E(dx_{i,s,k1} \cdot dx_{i,s,k2}) - E(dx_{i,s,k1}) \cdot E(dx_{i,s,k2})}{\sqrt{Var(dx_{i,s,k1}) \cdot Var(dx_{i,s,k2})}}; \tag{8}$$

- The pair correlation coefficient between the parts of passengers' arrivals for each hour of the definite day and the parts of passengers' arrivals for each hour of all days cumulatively (regardless the passengers' social group):

$$rdx_{i,k} = \frac{E(dx_{i,s,k} \cdot dx_{i,s}) - E(dx_{i,s,k}) \cdot E(dx_{i,s})}{\sqrt{Var(dx_{i,s,k}) \cdot Var(dx_{i,s})}}; \tag{9}$$

- The arrivals amount of the passengers transiting to the definite destination stop during the definite day:

$$Kx_{i,j,k} = \sum_{S=1}^{|T|} \sum_{m=1}^{|K|} x_{i,j,s,k,m}; \tag{10}$$

- The expectation's estimate of a number of the passengers transiting to the definite destination stop during all days cumulatively:

$$Kx_{i,j} = \frac{1}{|D|} \cdot \sum_{k=1}^{|D|} Kx_{i,j,k}; \tag{11}$$

- The pair correlation coefficient between the expectations' estimates of the amount of the passengers transiting to the definite destination stop during all days cumulatively and arrivals amount of the passengers transiting to the definite destination stop during the definite day:

$$rkx_{i,k} = \frac{E(Kx_{i,j,k} \cdot Kx_{i,j}) - E(Kx_{i,j,k}) \cdot E(Kx_{i,j})}{\sqrt{Var(Kx_{i,j,k}) \cdot Var(Kx_{i,j})}}. \quad (12)$$

The enumerated above statistical characteristics can be useful also for the detection of the model (1) several arguments independence (e.g. time intervals within the workday, social group of the passenger, weekday, and destination stop). If the justification of the independence should be proofed the decrease of the tabular function $F(\ldots)$ dimension would be decreased.

4 The Passengers' Arrivals Model Simplification

The computation and analysis of the descriptive statistics (2)–(12) based on the actual big data of the urban transport passengers flow to justify the following conclusions:

- Although the number of the passengers depends on the arrivals at the definite starting stop, the arrivals at the differing stops are mutually independent. Therefore, the averaging of the arrivals at starting stops is unacceptable;
- The argument "weekday" almost linearly depends on the other arguments of the model (1) (the correlation coefficient between K and D equals 0.96…0.99, between V and D equals 0.93…0.97, between T and D equals 0.82…0.97). Hence, this argument can be replaced by the constant k_day_k, i.e. the arrivals number at k^{th} day can be estimated as the part of averaged daily arrivals number:

$$x_{i,j,s,k,m} = F(V_i, V_j, T_s, K_m) \cdot k\_day_k; \quad (13)$$

- The social group of the passenger as well as the destination stop name are his individual characteristics. The influence of latter on the passengers' forthcomings is a week and may be neglected:

$$x_{i,s,k} = F_1(V_i, T_s) \cdot k\_day_k; \quad (14)$$

$$v_n = F_2(V_i, T_s, K_m); \quad (15)$$

$$k_n = F_3(V_i, V_j, T_s); \quad (16)$$

where $n = 1 \ldots x_{i,s,k}$ is the index of the passenger arriving at the i^{th} stop of k^{th} day's s^{th} minute;
- The destination stop v_n and the passenger's type k_n in (15) and (16) are mutually dependent. As the result, the difficulties of the practical applications may occur

because of the uncertainty of the order of these parameters fixing. Substituting (15) in (16) instead V_j, leads to:

$$k_n = F_3(V_i, F_2(V_i, T_s, K_m), T_s) = F_4(V_i, T_s); \tag{17}$$

i.e. passenger's type k_n is independent of the destination stop V_j. Similarly, the independence of v_n from K_m leads to

$$v_n = F_2(V_i, T_s, F_3(V_i, V_j, T_s)) = F_5(V_i, T_s); \tag{18}$$

- The correlation coefficients between the definite hour values of passengers' types and their daily averaged values are significant (equal 0.5…0.95). Hence, T_s in $F_4(\ldots)$ can be replaced by the time averaged argument *K_sr_t:*

$$k_n = F_4(V_i, K\_sr\_t); \tag{19}$$

- Similarly, the correlation coefficients between the passengers' destination stops per hour and their daily averaged values are significant (equal 0.45…0.95). Hence, T_s in $F_5(\ldots)$ can be replaced by the time averaged argument *Vj_sr_t:*

$$v_n = F_5(V_i, Vj\_sr\_t). \tag{20}$$

Thus, the model of the passengers' arrivals at the starting stops and their transit to the destination stops turns into the following form:

$$\begin{cases} x_{i,s,k} = F_1(V_i, T_s) \cdot k\_day_k, \\ k_n = F_4(V_i, K\_sr\_t), \quad n = 1..x_{i,s,k}, \\ v_n = F_5(V_i, V\_sr\_t), \quad n = 1..x_{i,s,k}. \end{cases} \tag{21}$$

The model (21) contains the reduced arguments number comparable with (1). Hence, the program implementation of (21) and its further application would be easier.

5 The Simulation Results Analysis

In the course of the model adequacy testing, we have passed through ten simulation sessions for each stop of the city's transport net. The results lead to the following conclusions:

1. The model (21) adequacy is better for the stops with significant passengers' turnover. For example, all adequacy criteria for the stop with 560 average forthcoming passengers for workday are better than for the stop with 130 average forthcomings. In turn, all adequacy criteria of latter are better than for the stop with 50 average forthcomings.

2. The accuracy of the model (21) is best for the case of passengers arrivals distribution via daytime (confidence level value is 0.87). The confidence level values of passengers arrivals distribution at the destination stops and at the passengers' types are comparably less (0.69 and 0.51 correspondingly).
3. The correlation between the real observation data and the simulation data for each of testing distributions is very significant (the corresponding correlation coefficients values are more than 0.83). Hence, the simulation results application is acceptable for the practical applications instead and in addition to the real observations results.
4. The accordance between the empiric probability distribution laws of the real and simulated data (estimated by the chi-square test) is not enough to the confident conclusion on the model (21) adequacy. The best confidence level demonstrates the probability distribution laws of passengers' flow distribution via daytime (0.61… 0.83), and the worst one is for the distribution via the passengers' types (0.48… 0.51).
5. The model generates random time and corresponding parameters of the passengers' turnover with minor standard deviation. The latter shows the model tolerable accuracy and stability.

Because of the minor statistical accordance, the model (21) is more suited for the passengers' arrivals simulation for the conditions of the abstract workday with the averaged features. The way of model statistical adequacy raise lays in the model (21) complication by means of the refusal from the passengers' forthcoming number averaging accordingly to the types and workday time. However, in this case, the model's dimension should raise significantly while the adequacy would improve for 15-20%.

In spite of slight coinciding of the passengers' arrivals probability distribution laws, the presented model is applicable for the passengers' transition simulation.

6 Conclusion

The report presents the adequate model of the passengers' arrivals at the stops and their transition between the stops of the small city municipal transport net. The model configuration appears as the tabular probability distribution function with the parameters identified by the data acquisition of the passengers' turnout prolonged real observations.

The model shows the high correlation with the real observations results. However, the coinciding of the empiric probability distribution laws of the simulating and real results is not so high. The way of the model adequacy raise lies in the model's complication by means of the refusal from the passengers' forthcomings number averaging accordingly to the passengers' types and workday time.

In spite of slight coinciding of the passengers' arrivals probability distribution laws, the presented model is applicable for the passengers' turnout simulation.

Acknowledgment. The article was based on works supported by grants of Russian Foundation for Basic Research Povolzhie No. 14-07-97011, No. 15-47-02321.

References

1. Krushel, E.G., Stepanchenko, I.V., Panfilov, A.E., Berisheva, E.D.: An experience of optimization approach application to improve the urban passenger transport structure. In: Kravets, A., Shcherbakov, M., Kultsova, M., Iijima, T. (eds.) JCKBSE 2014. CCIS, vol. 466, pp. 27–39. Springer, Cham (2014). doi:10.1007/978-3-319-11854-3_3
2. Parygin, D., Sadovnikova, N., Kravets, A., Gnedkova, E.: Cognitive and ontological modeling for decision support in the tasks of the urban transportation system development management. In: 6th International Conference on Information, Intelligence, Systems and Applications (IISA), Greece, pp. 1–5. IEEE Xplore, Corfu (2015)
3. Sadovnikova, N., Parygin, D., Kalinkina, M., Sanzhapov, B., Ni, T.N.: Models and methods for the urban transit system research. In: Kravets, A., Shcherbakov, M., Kultsova, M., Shabalina, O. (eds.) Creativity in Intelligent, Technologies and Data Science. CCIS, vol. 535, pp. 488–499. Springer, Cham (2015). doi:10.1007/978-3-319-23766-4_39
4. Ding, H., Wang, W., Luo, T., Yang, Z., Li, Y., Li, Z.: Cellular automata based modeling for evaluating different bus stop designs in China. Discrete Dyn. Nat. Soc. **2015**(10) (2015)
5. Di, D., Dongyuan, Y.: Dynamic traffic analysis model of multiple passengers for urban public transport corridor. Adv. Mech. Eng. **7**(11) (2015)
6. Maximey, I.V., Sukach, E.I., Galushko, V.M.: Design modeling of service technology for passenger flow by city transport network. J. Autom. Inf. Sci. **41**(4), 67–77 (2009)
7. Lewe, J.H., Hivin, L.F., Mavris, D.N.: A multi-paradigm approach to system dynamics modeling of intercity transportation. Transp. Res. Part E: Logist. Transp. Rev. **71**, 188–202 (2014)
8. Horn, M.: An extended model and procedural framework for planning multi-modal passenger journeys. Transp. Res. Part B: Methodol. **37**, 641–660 (2003)
9. Herbon, A., Hadas, Y.: Determining optimal frequency and vehicle capacity for public transit routes: a generalized newsvendor model. Transp. Res. Part B: Methodol. **71**, 85–99 (2015)
10. Zheng, Y., Xi, X., Zhuang, Y., Zhang, Y.: Dynamic parameters cellular automaton model for passengers in subway. Tsinghua Sci. Technol. **20**(6), Article no. 7349931, 594–601 (2015)
11. Feng, Y., Li, X., Ma, N.: Research of traffic share in different models of transport based on the Agent theory. In: 2012 Joint Rail Conference, JRC 2012, pp. 813–822. ASME, Philadelphia (2012)
12. Li, D., Yuan, J., Yan, K., Chen, L.: Monte Carlo simulation on effectiveness of forecast system for passengers up and down buses. In: 3rd International Symposium on Intelligent Information Technology Application, IITA 2009, vol. 1, Article no. 5369635, pp. 359–361. IEEE, NanChang (2009)
13. Schelenz, T., Suescun, A., Wikström, L., Karlsson, M.: Application of agent-based simulation for evaluating a bus layout design from passengers' perspective. Transp. Res. Part C: Emerg. Technol. **43**, 222–229 (2014)
14. Bure, V.M., Mazalov, V.V., Plaksina, N.V.: Estimating passenger traffic characteristics in transport systems. Autom. Remote Control **76**(9), 1673–1680 (2015)