Development of a Technology for Monitoring Passenger Traffic in the Context of Intelligent Transport Systems

Mykyta Volodarets[1(✉)], Igor Gritsuk[2], Sergii Pronin[3], and Alona Yurzhenko[2]

[1] Pryazovskyi State Technical University, Universytetska Str., 2, Mariupol 87555, Ukraine
volodarets.nikita@yandex.ru

[2] Kherson State Maritime Academy, Ushakov Ave., 20, Kherson 73000, Ukraine
{gritsuk_iv,sergiy9977}@ukr.net

[3] Kharkiv National Automobile and Highway University, St. Yaroslava Mudrogo, 25, Kharkov 61002, Ukraine
mz@ksma.ks.ua

Abstract. The article shows the results of the analysis of the development of municipal passenger transport. The regulations on the existence and the need to overcome the contradictions between the desired and existing systems for the provision of transport services to the population are formed on the basis of this analysis. The expediency of creating a decision support system for the subjects of the passenger transportation market is shown. It was based on the creation of a single information space for traffic participants in transport systems. A new method for assessing the passenger's choice of travel path based on the MNL (multinomial model) has been developed. The developed method makes it possible to obtain and predict the form of the attractiveness function of urban routes in any city for its further use in models for the formation of rational route networks. A new intelligent technology for monitoring passenger traffic has been proposed. The practical significance of the results obtained lies in the original approach to monitoring passenger traffic, proposed on the basis of the use of information technologies and distributed multidimensional databases. The result of the application of appropriate information and communication technologies for monitoring the movement of passenger transport is useful for the development of transport infrastructure in large cities. This will make it possible to profit from the rational distribution of resources for the maintenance of the route network of cities and regions.

Keywords: Intelligent transport system · Vehicle · Urban passenger transport · Passenger transport monitoring · Data warehouse technology · Automated questionnaire survey

1 Introduction

Monitoring of passenger traffic consists in obtaining information about the state of traffic, tracking and its dynamics forecasting. Monitoring is aimed at improving transport

O. Prentkovskis et al. (Eds.): TRANSBALTICA 2022, LNITI, pp. 23–32, 2023.
https://doi.org/10.1007/978-3-031-25863-3_3

services for residents of large cities and regions by improving and coordinating the work of transport organizations. The monitoring system has two components: a source of information and a system for storing, evaluating and organizing the information received. The current level of development of the transport infrastructure of cities and regions requires the intellectualization of all components of monitoring, the introduction of new information technologies for monitoring passenger traffic, taking into account the collective opinion of the transportation process participants [1–4].

Monitoring of passenger traffic in a city or region should be considered as an intelligent DSS for the decision maker's choice of alternative options for organizing transport processes, route networks and rolling stock, information development of the transport infrastructure of cities and regions. These systems combine Online Analytical Processing (OLAP) and DataMining. This combination is a new information technology that is focused primarily on users with their actual specific needs.

2 Formation of a Utility Criterion When Choosing a Route

As a result of the analysis of scientific and technical literature [5–18] devoted to the issues of modeling options for the movement of passengers in cities, two main approaches to the choice of a model can be distinguished. The first approach allows determining the probability of a passenger choosing a travel path based on the general characteristics of each route option. Within the framework of this approach, the most widespread are standard or linear models, in which the probability of a passenger choosing a travel path is directly determined by the parameters of the travel path. Another variant of the implementation of this approach is the MNL models, in which the probability of a passenger choosing a travel path is proportional to the exponent of its parameters. The second approach is based on the assumption that each passenger, when choosing a route option, uses the first vehicle from the routes that suit him. In this case, the frequency of route selection is proportional to the traffic intensity, taking into account the level of travel conditions. However, this approach is more consistent with the case of choosing a specific route on a section of the transport network common for several routes than choosing the entire path. Therefore, the models of the first approach were adopted for further development.

Analysis of modeling methods for the procedure for choosing a passenger's route of travel showed the similarity of approaches to solving this issue, but did not answer the question of how to obtain the coefficients of the models. The difference between MNL-models and linear models lies only in the general form of the function of attractiveness or cost of routes, however, the mechanism of their use to determine the probability of a passenger choosing a route of travel is the same in principle.

$$P_i = A_i / \sum_{j}^{n} A_i, \tag{1}$$

where Pi – choice probability of i route; n – number of variants of the route; Ai – calculation function of route attractiveness.

The value of the function A_i is determined depending on the model of distribution of passenger flows. For linear model.

$$A_i = \prod_i, \tag{2}$$

where Π_i – attractiveness function of route.

For MNL – model:

$$A_i = exp(A_i') \tag{3}$$

The problem is that in order to obtain the type of function, the methods of survey of intentions were mainly used, the result of which is just the definition of the attractiveness function. It remains only to obtain the value of the coefficients of the models (most often linear ones) and substitute them in the corresponding expressions.

To do this, firstly, it is necessary to determine the value of the attractiveness function through the probability of the passenger's choice of the movement path. For the general case, one can write down a system of equations with probabilities as coefficients and attractiveness functions as unknowns.

$$\begin{cases} (P_1 - 1) \cdot A_1 + P_1 \cdot A_2 + \ldots + P_1 \cdot A_n = 0; \\ P_2 \cdot A_1 + (P_2 - 1) \cdot A_2 + \ldots + P_2 \cdot A_n = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots . \\ P_n \cdot A_1 + P_n \cdot A_2 + \ldots + (P_n - 1) \cdot A_n = 0. \end{cases} \tag{4}$$

This is a homogeneous system of n linear equations with n unknowns. Possible variant of its solution is a trivial solution, i.e., all values of the attractiveness function are equal to zero or there are many linearly dependent solutions. In the first case, the rank of the coefficient matrix must be equal to the rank of the extended matrix; for a homogeneous system, this means that the coefficient matrix must be unexpressed. Since the trivial solution does not correspond to the problem statement, it is necessary to prove that the determinant of the coefficient matrix is equal to zero. In matrix form, the system looks as follows:

$$\begin{vmatrix} P_1 - 1 & P_1 & \ldots & P_1 \\ P_2 & P_2 - 1 & \ldots & P_2 \\ \ldots\ldots & \ldots\ldots & \ldots & \ldots \\ P_n & P_n & \ldots & P_n - 1 \end{vmatrix} * \begin{vmatrix} A_1 \\ A_2 \\ \ldots \\ A_n \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{vmatrix} \tag{5}$$

Or alternatively, in matrix form:

$$P \cdot A = Z \tag{6}$$

If we add all n-1 of other rows to the last row of the matrix P, then each column of this matrix will contain the expression.

$$\sum_j^n P_i - 1 = 0. \tag{7}$$

Thus, the n^{th} row will be zero, since such transformations do not change the value of the determinant, which means that $det\,P = 0$, i.e., the solutions of the system are linearly dependent and at least one of them can be specified arbitrarily.

If we specify a value for one attraction from the set $1\ldots n$, for example A_1, then other values A_i, $i \neq 1$ are determined from the statement.

$$A_i = A_1 \cdot P_i / P_1. \tag{8}$$

Therefore, for each questionnaire it is necessary to set one arbitrary value, however, the best approximation for it may be the value obtained not by calibration, but using the least-square method. That is, the A1 value for each questionnaire should be the variables or factors that determine the A1 value. For each i questionnaire, one can write down n equations of the following type, if a linear description is used.

$$\begin{cases} a_0 + a_1 \cdot F_{11} + a_1 \cdot F_{21} + \ldots + a_m \cdot F_{m1} - 1 \cdot A_i = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1 \cdot F_{1n_i} + a_1 \cdot F_{2n_i} + \ldots + a_m \cdot F_{mn_i} - \dfrac{P_{n_i}}{P_1} \cdot A_i = 0. \end{cases} \tag{9}$$

where a_m is the regression coefficient at m factors that determine the value of attractiveness; F_{kj} is the value of the k^{th} factor in the j^{th} variant of the movement path; A_i is the value of the attractiveness function for the 1^{st} variant of the movement path in the i questionnaire; P_1 is the probability of choosing the first variant of the movement path of the *i* questionnaire; n_i is the number of movement options in the *i* questionnaire; m is the number of factors in the regression equation (path parameters).

Thus, each questionnaire adds one variable A_i to the array of factorial features.

The total number of factorial features N_f in the set of information for regression analysis is determined from the expression:

$$N_f = m + N_a, \tag{10}$$

where N_a is the number of questionnaires based on the survey results.

However, the result for all experiments, the number of which N_o is determined from the following expression, is equal to zero.

$$N_0 = \sum_{i}^{N_a} n_i. \tag{11}$$

Most $n_i > 1$, and $N_a > \; > m$, i.e., the number of experiments significantly exceeds the number of factorial features:

$$N_o >> N_f. \tag{12}$$

One should not expect a linear relationship between parameters or observations, since they are the outcomes of a random selection of the results of the behavior of different people in the transport system. There are quite a lot of motives for such behavior so that even in the equal conditions different people make different decisions. Therefore, here one should expect not only one result, in accordance with the properties of a homogeneous matrix:

$$\forall a_k = 0,\ k = 1 \ldots m \ldots \tag{13}$$

It is possible to get out of this situation by abandoning the homogeneity of the data set, i.e., setting the value of one base variable. Let it be A = 1.

The search for regression coefficients for such a dataset must be performed with an intercept, since it has a completely definite physical meaning here - this is the attractiveness of a route with zero values of the travel path parameters. For most of the parameters, the regression coefficients should be negative if they characterize the cost of travel of one type or another. It should be borne in mind that these values of the regression coefficients are meaningful only for $m + 1$ of the first factors.

The next consideration concerns the MNL model. Since $A_i = exp(\Pi_i)$, for it, the regression coefficients can be obtained only after transformations based on (10). This expression, taking into account (5), is written as follows.

$$exp(A'_1)/P_1 = exp(A'_i)/P_i. \tag{14}$$

After taking the logarithm of this equation, by analogy with (15), we can obtain.

$$A'_i = A'_1 -\!-ln(P_1) + ln(Pi). \tag{15}$$

Thus, to obtain the type of the MNL model, it is necessary to first determine the natural logarithms of the probabilities of choosing the options for the travel routes.

Final consideration. For the convenience of using the obtained regression coefficients, it is not necessary to set the value of the base indicator A_1 equal to one. The higher this value is, the higher the value of the regression coefficients.

3 Preparing Data for the Route Selection Model for Further Processing by the Operational Analysis Method

After calculating the parameters of the questionnaires in the processing system, they were transferred to the STATISTICA package to obtain the coefficients of the regression models. Calculations using the linear model gave poor results, the total correlation coefficient did not exceed 0.6 for various parameters. Therefore, it was concluded that it was necessary to use the MNL model with its results which looked successful. The general characteristics of the model are quite high. However, a number of indicators such as path length, coefficient of transport fullness and number of transfers have not confirmed their significance. Therefore, they were not included in the final version of the model. However, the model for determining attractiveness cannot fully characterize the final result of the calculations - probability of selection the route by passengers. Therefore, an additional calculation was carried out in order to determine the accuracy of the resulting model with respect to the probability of selection the route by passengers. Here, theoretical probabilities were compared with probabilities that were obtained from the survey. Since the maximum accuracy of the results is 0.2 (five days of the survey), in each case the value of the qualitative indicator of the error was found, which is compared to one if the error is greater than 0.2 and exactly zero otherwise. The calculation results indicate a fairly high degree of model accuracy. Only in minor cases there is exceeding the permissible deviation of the probability of selection the route by a passenger. The maximum discrepancy between theoretical and actual values is 0.279.

Such deviations cannot be considered significant. Refinement of the parameters of the model can be carried out by increasing the survey period and increasing the sample size. When using the MNL model, the question arises of determining the options for the route, which should be considered as alternative when calculating the probability of choosing the route, since for any combination of significant parameters, the model will provide a positive value of the probability. The survey results confirmed the conclusions of [18] that it is enough to consider 3 alternatives that have the maximum attractiveness.

4 Implementation of Intelligent Monitoring Technology

Modern information and communication technologies without unnecessary software costs allow monitoring and evaluating the questionnaire, which proves the possibility of introducing such systems on the basis of existing computer networks in a large city. Compliance with the existing computer network and the HTTP protocol is sufficient. On the part of the respondent, it is necessary to have a regular Web browser, which will make it possible to carry out the survey on any modern computer, regardless of the operating system installed. The database server can be commercial solutions based on Microsoft SQL Server 2005/2008 and later, or a free version of MySQL supported by Oracle. An important aspect for solving monitoring problems is the implementation of registration of data provided by respondents in their responses to the electronic form of the questionnaire. Such a questionnaire, as a component of a transport Web portal, can be implemented using the ASP.NET MVC Framework technologies - a framework for creating web applications that implements the Model-View-Controller pattern.

To solve the assigned monitoring tasks and reduce the requirements for the qualifications of system administrators who will maintain the system, one of the optimal solutions will be the use of cloud computing technology - Platform as a Service. This model of providing cloud computing assumes that the entire information technology infrastructure, including computer networks, servers, storage systems, is entirely controlled by the provider, and the consumer is given the opportunity to use the platform to launch applications. The noted features of receiving and transforming monitoring data were used in Kharkiv National Automobile and Highway University (KNAHU) when creating a universal transport portal (Fig. 1).

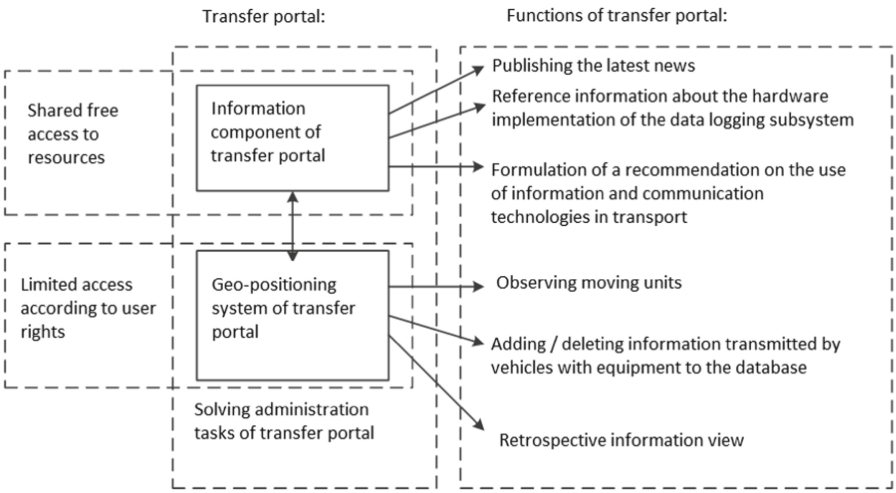

Fig. 1. The structure of the components of the transport portal.

It is universal distribution computer system for monitoring land transport and traffic environment. This is a two-level compute system that provides information and communication technology for the movement of land transport and consists, first of all, of a special website. The transport DBMS technology is proposed to be built on the basis of the "client-server" technology, which will have two parts - a client application (front-end) and a database server (back-end). In general, the developed monitoring system can be found in Fig. 2.

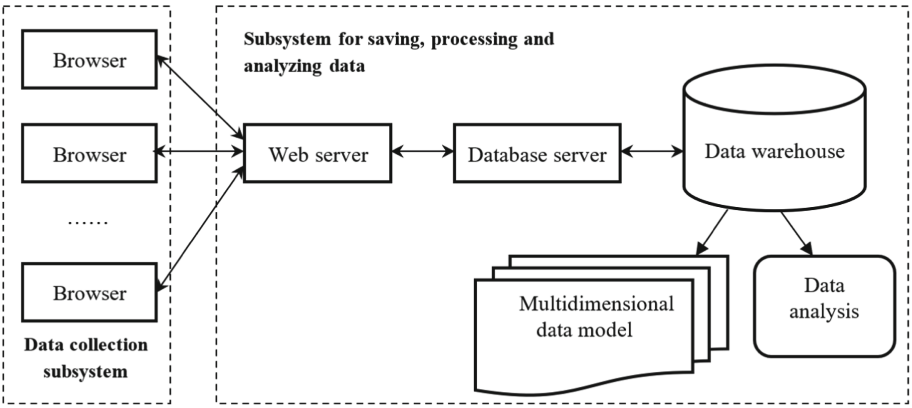

Fig. 2. General diagram of the monitoring system.

5 Technical Efficiency of the Monitoring System

By the technical efficiency of monitoring we mean the definition of the reliability of a given system.

In our case, we have client-server software. The server serves requests from N client programs. In software, errors are evenly distributed over the input domain, among which the most common are:

1. Software failures. If the software is not modified, then the rate of its failures remains constant.
2. Internal failures in the program. Such failures are due to fundamental limitations of the algorithm used in the software (for example, the use of heuristic algorithms can lead to random failures).
3. Failures due to real-time operation restrictions. In the systems under consideration, the environment can change dynamically.

The results of the analysis of the system allowed us to draw preliminary conclusions about the possible load. With the given system characteristics, channels with 10 remain unloaded. Average intensity of servicing requests equals 2.5. The average processing time for one request by the application server is 0.4 s. Comparison of the results of simulation and analytical modeling of the system allows us to draw conclusions about the correctness of the constructed simulation model Average service time of a request – 0.4 and 0.40 s; system load – 4 and 3.8; the average number of requests in the system is 3.9 and 3.8 for analytical and simulation models, respectively. The resulting model is adjusted taking into account the heterogeneity of the flow of requests, different types of input flows. The simulation results are given in Fig. 3.

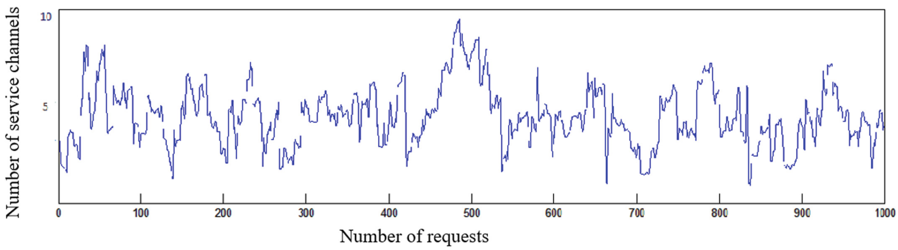

Fig. 3. System load simulation results.

6 Conclusions

The analysis of the development of urban passenger transport showed the existence and need to overcome the contradictions between the desired and the existing system of providing transport services to the population. Their detection and elimination is the key to the successful development of the corresponding transport infrastructure of cities and regions. The basis for such development is the updating and implementation of the latest information, intelligent and network technologies for monitoring the state of the passenger transportation system.

The expediency of creating a decision support system for the subjects of the passenger transportation market is shown. It is based on the creation of a single information space for traffic participants in transport systems.

A new method for assessing the passenger's choice of route based on the MNL (multinomial model) has been developed. The maximum discrepancy between the theoretical and actual values of the model is 0.279. Refinement of the parameters of the model can be carried out by increasing the survey period and sample size. The developed methodology allows one to obtain and predict the type of the attractiveness function of urban routes in any city for its further use in models of the formation of rational route networks.

The practical significance of the results obtained lies in a new approach to monitoring passenger traffic, proposed to be based on the use of information technologies and distributed multidimensional databases. The result of the application of appropriate information and communication technologies for monitoring the movement of passenger transport is useful for the development of the transport infrastructure of large cities, tested in the conditions of Dnepropetrovsk, Donetsk and Kharkov regions. This provides up to 10–15% of the profit from the rational distribution of resources for the maintenance of the route network of cities and regions. Instead of the usual costs of 10,000–15,000 UAH for the preparation of recommendations to assess the human factor for making one decision on organizing a new route, costs of 1–2,000 UAH are sufficient. In a big city, the savings from the implementation of the proposed monitoring technology are about 10,000 UAH per year. Theoretical regulations on the information development of transport systems have become the basis for teaching the following training courses: "Flexible computerized transport systems", "Modeling of transport machines and processes."

References

1. Gorobchenko, O., et al.: Intelligent locomotive decision support system structure development and operation quality assessment. In: 2018 IEEE 3rd International Conference on Intelligent Energy and Power Systems (IEPS), Kharkiv, pp. 239–243 (2018). https://doi.org/10.1109/IEPS.2018.8559487
2. Prizhibyl, P., Svitek, M.: Telematics in Transport. MADI (STU), Moscow (2003)
3. Volodarets, M., Gritsuk, I., Chygyryk, N., Belousov, E., et al.: Optimization of vehicle operating conditions by using simulation modeling software. SAE Technical Paper 2019–01–0099 (2019). https://doi.org/10.4271/2019-01-0099
4. Vychuzhanin, V.V., et al.: Analysis and structuring diagnostic large volume data of technical condition of complex equipment in transport. In: Machine Modelling and Simulations 2019 (2020). https://doi.org/10.1088/1757-899X/776/1/012049
5. Livshits, V.V.: The system concept of the city and mathematical modeling of the adaptive behavior of the urban population. The use of applied systems analysis in the design and management of urban development, p. 120 – 147. Moscow, Stroyizdat (1974)
6. Rogova, G.L.: Modeling the Choice of Routes for the Movement of Passengers in the Transport Systems of Cities. Moscow (1987)

7. Branovitskaya, S.V.: Economic and Mathematical Models of Forecasting Passenger Traffic in Cities. Kyiv (1971)
8. Vasiliev, V.M.: Investigation of Some Regularities in the Formation of Passenger Traffic on the Urban Route Network of Public Transport. Saratov (1978)
9. Krivosheev, D.P.: Methods of distribution of passenger traffic in transport calculations (review). CSTI on civil engineering and architecture. Moscow (1974)
10. Rickberg, G.S.: On the Method of Modeling Intracity Labor Relations. City and passenger, pp. 48–51. Stroyizdat, London (1975)
11. Soymina, E.Y.: Development of Dynamic Models of Passenger Traffic and their Application to Control Loads on Public Transport Networks. Almaty (1988)
12. Faience, O.G.: Mathematical models of the formation of passenger traffic flows. Mathematical methods in the management of urban transport systems. Institute of Socio-Economic Problems, pp. 621, Nauka (1979)
13. Shaposhnikov, S.V.: Methodology for Calculating Passenger Traffic in Citie. Moscow (1969)
14. Makarov, I.P., Yavorsky, V.V., Tuzovsky, A.F.: Prediction of the distribution of passenger traffic when changing the route network of passenger transport, Modeling of transport systems management processes, pp. 3941. Vladivostok (1977)
15. Ryzhenko, L.I.: Modeling the movement of passengers along the city route network. In: Regional scientific and practical conference "Increasing the efficiency of the transport complex", pp. 1214. Omsk (1989)
16. Sadikhova, O.S.: Choice of the route by passengers. Urban transport and engineering preparation of territories, pp. 3341. LISI (1991)
17. Alekseev, V.O.: Intelligent technology of organizing the movement of vehicles. Automobile transport, Vol.10, pp. 305311. KNAHU, Kharkiv (2002)
18. Kruglov, V., Dli, M.: Intelligent Information Systems: Computer Support for Systems of Fuzzy Logic and Fuzzy Inference. Fizmatlit, Moscow (2002)