Scenarios of Training Courses on Digital Modeling on the Example of Modeling a Route Network

Alexander V. Baldin* and Ivan D. Eroshok

1Bauman Moscow State Technical University, 2nd Baumanskaya str., 5/1, 105005, Moscow, Russia

Abstract. The article deals with the creation of training courses scenarios. The main objective of the development of software components of the training system is the ability to create an algorithmic structure of training fragments representing test tasks, workshops and just information support. On the basis of this instrumental environment a practical work on the topic of statistical analysis, modeling and forecasting of passenger flows in the urban route network has been developed. The study of passenger flows is considered as one of the important stages of designing and organizing the route network of the urban passenger transport. The purpose of this article is to develop a training scenario aimed at identifying hidden patterns of passenger flows at stopping points of bus routes with the subsequent development stage of associated models of random flows with specified autocorrelation properties. In addition to the auto-correlation estimation and spectral analysis, the main components analysis which allows to reduce significantly the dimensionability of the multidimensional time series of passenger flows at the stopping points of the route is carried out.

1 Introduction – Methods of collecting and structuring applications of the training system

The instrumental environment of the educational scenario formation whose volume is determined by the number of heterogeneous educational fragments [1] including multimedia [2], launching of mathematical packages applications [3,4], interactive [5] test tasks fragments [6,7] etc is given in this article.

One of the versions of scenario descriptions in the software environment of the training system is the track of applications \{hi\} of training elements (Figure 1). They represent some linear sequence of elementary applications and can be just information frames, extended workshops with quite complicated interactions.

* Corresponding author: bal@bmstu.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
The universal scheme for describing software training applications with the transfer of the completion code at the end (for example, the percentage of the solved test task) is proposed in the paper. A software mechanism of conditional transitions between parameterized applications is proposed. This allows to create a training scenario with an algorithmic structure quite simply but effectively. (Figure 2).

The universal scheme for describing software training applications with the transfer of the completion code at the end (for example, the percentage of the solved test task) is proposed in the paper. A software mechanism of conditional transitions between parameterized applications is proposed. This allows to create a training scenario with an algorithmic structure quite simply but effectively. (Figure 2).

![Algorithmic structure of applications.](image1)

There is a possibility in the developed application of workshops creation to include projects of OLE-objects [8] of software mathematical packages [9-11].

![Integrated structure of the training methodology scenario.](image2)

Such scenario designing is determined by abilities of the instrumental environment [12,13], namely, a formal description of the functionality of the basic application (a training fragment) of the scenario algorithmiation environment and the environment of the formation of scenarios hierarchy whose functional description is given below.

![Software and computing algorithms](image3)
1.1 Formulated description of the elementary application

In general, the training scenario is a set of elementary applications and data decoupling. The Fragment has the following structure:

\[ F_i = (t_i, d_i, a_i, s_i, r_i, p_i), \]

where \( t_i \) — a fragment type (information, calculation, selection etc.); \( d_i \) — a difficulty level (for test control); \( a_i \) — an access level to the fragment; \( s_i \) — time of the compulsory termination of the presentation; \( r_i \) — a subset of features connected with this fragment; \( p_i \) — a parametrization at activation.

\[ a_i \subseteq (r_p, r_n, r^6, r_u, r_s, r_v, r_o), \]

where \( r_p \) — a sign of transition prohibition from this fragment to the previous one in the sequence; \( r_n \) — a sign of transition prohibition from this fragment to the next one in the sequence; \( r_b \) — a sign of backtracking prohibition one step back on the track of fragments presentation; \( r_t \) — a sign of prohibition ion of the random access possibility to this fragment; \( r_s \) — a sign of prohibition of the atoppage of the fragment presentation; \( r_o \) — a sign of prohibition of the possibility of the transition from this fragment to the other random fragment.

Matrix of fragments adjacency (transitions):

\[ T = ||C_{ij}||, \]

where \( C_{ij} \) — a condition determining a transition from \( i \) fragment to \( j \) one;

\[ C_{ij} = \begin{cases} \left( r_y a_y \varphi (F_i) \right) \gamma_y \left( t_y \beta_y \psi (F_i) \right), & i = 1 \ldots N, j = 1 \ldots N, \text{ if a transition is envisaged;} \\ 0, & \text{otherwise}, \end{cases} \]

where \( N \) — a number of fragments in the structural element; \( r_{ij} \) — a result of the fragment presentation; \( a_{ij} \) — a comparison operation of the actual and predetermined results of the fragment presentation (\( \neq, <, \leq, =, \geq, > \)); \( t_{ij} \) — a duration of the fragment presentation; \( \beta_{ij} \) — a comparison operation of the actual and predetermined duration of the fragment presentation (\( \neq, <, \leq, =, \geq, > \)); \( \gamma_{ij} \) — a logical operation of the condition combination at the result and the duration of the fragment presentation \( (Y, \Lambda) \); \( \varphi (F_i) \) — a function definition of the actual result of the fragment presentation \( F_i \); \( \psi (F_i) \) — a function definition of the actual duration of the fragment presentation \( F_i \).

A condition of the transition correctness is defined as \( \forall i = 1..N, \sum_{j=1}^{N} C_{ij} = 0 \). The structural element is presented by

\[ S = \langle (F_i)_{i=1..N}, T \rangle, \quad i = 1..N, \]

where \( F_i \) — \( i \) fragment; \( T \) — a matrix of the fragments adjacency.

Fragment completion options can be the following:

- was completed by the user (the user pressed the button),
- Completed by itself (completed all its mm-flows),
- was completed by an actor (the forced end of acting time has expired).
The access level determines the enclosure of the scenario structure. This allows to create a hierarchy of scenarios and the usage of locking mechanisms to implement the structure of the enclosed processes. The parametrization of the application enables no only to make but also to solve the issues of coordination according to the data of different applications included in one training scenario.

1.2 Scenario of practical work on urban route network modeling

In general, the developed instrumental means include designers of educational test tasks (ETT), tests and structural elements (SE) (Figure 4).

![Instrumental means of practical work creation](image)

As a result of using of the test tasks designer is a multimedia fragment that recovers closed-type tasks for compliance, order, clustering and other tasks. The structural elements designer combines all the information fragments and test tasks created for this practical work into a single algorithmic structure where after solving the test task the next training fragment is selected depending on the result of solving their task. It can be said from the terminological point of view that the lecture contains only information fragments and is a linear track, and the test consists only of test tasks. In addition to the general scenario the test designer is used to create a test for practical work which is based on adaptive algorithms [14] based on Markov chains [15], stochastic approximation methods [16].

Further the task of formation practical work of the urban route network modeling is set. The scenario of work includes all range of statistical methods, methods of multidimensional analysis [17], simulation [18], gravitational and entropy methods of correspondence matrix construction [19,20] and other methods and models.

For all of each methods included in the work is given first an information fragment containing the basic theoretical information. Then the test task is given. If the task is not completed, the return to studying theoretical information is undertaken. If the task is completed, the route is randomly selected and time series of passenger flows are formed for the trainee. According to these data the trainee should perform calculations in appropriate mathematical packages. The results of the performance are perceived by the system as a test task. If the results of the comparison are negative, the calculations are made again, if positive the transition to another method is performed.

All the statistics of answers and results of calculations is saved in the form of the report on practical work which the teacher can view at any convenient time. The description of some educational fragments is given below.
2 Training fragment – statistical analysis of passenger flows

A detailed statistical analysis of many bus routes in the city of Krasnoyarsk is presented in the article. The main results with examples of analysis of passenger flows of bus route №85 (Figure 5), which is the first in popularity, are given in this article.

Fig 5. Scheme of bus route № 85.

Bus route №85 passes through the Historic Center of the city, connecting the city from the West to the East, passing through the October Bridge. The temporary series of passenger flows (Figure 6) at the stopping points (SP) has a clearly expressed cyclical character with a superimposed increasing trend.

Fig. 6. A temporary series of passenger flows of bus route №85.

The graph shows the volume of the passenger flows in the autumn months at SP with ordinal numbers 14, 15, and 16. These SP are quite close geographically, but as it will be shown above, such a close relationship is typical for geographically separated BS.

In addition to the main route characteristics, such as irregularity in hours, irregularity in directions and others, a correlation and a spectral analysis of passenger flows is carried out. (Figure 7,8).
Thus, all passenger flows have a clearly expressed seasonal component at a frequency of 0.144 that corresponds to the ideal cycle. In addition, there is another seasonal component at a frequency of 0.28 that corresponds to the half-weekly component. The contribution of this component is smaller, but taking it into account can improve the accuracy of the simulation results.
\[ \eta_i = \sum_{j=1}^{p} \alpha_{ij} \xi_j, ..., \eta_p = \sum_{j=1}^{p} \alpha_{pj} \xi_j \]  

(4)

which satisfy the conditions of orthogonality (cov(\(\eta_i, \eta_j\))=0, i,j=1..p) and monotonicity of the dispersion (D\(\eta_1\) ≥ D\(\eta_2\) ≥ ... ≥ D\(\eta_p\) and \(\sum_{i=1}^{p} D\eta_i = \sum_{i=1}^{p} \sigma_{ii}\)). At the same time the linear combination \(\eta_i = \alpha_{i1} \xi_1 + ... + \alpha_{ip} \xi_p\) is called the first main component if \(\{\alpha_{il}\}_{l=1}^{p} = \arg \max_{\alpha_{il}} D(\eta_i), \sum_{l=1}^{p} \alpha_{il} = 1\). The dispersion of the first principal component is \(D(\eta_1) = \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_{il} \alpha_{jj} \sigma_{ij}\). Similarly, the rest of the main components \(\eta_i = \alpha_{i1} \xi_1 + ... + \alpha_{ip} \xi_p\) is sought from the solution of the orthogonality condition to all the previous ones [21]. The geometric interpretation of this linear transformation is shown in Figure 9.

If the correlation analysis determines the measure of the relationship between two indicators, the purpose of the principal components method is to identify the overall relationship of all indicators at once (https://habr.com/ru/post/304214/). It is assumed that there are \(p\) indicators \(\{\xi_i\}_{i=1}^{p}\). With a vector of averages \(m=(m_1,...,m_p)\) and a covariance matrix \(D=(\sigma_{ij})\). The method of main components analysis determines the structural relationship between these indicators and the essence of the method is to seek the linear combinations of the initial variables.

**Table 1. Informativeness of main components.**

| Value | Eigenvalues (M_85) | Extraction: Principal components |
|-------|-------------------|---------------------------------|
|       | Eigenvalue | % Total variance | Cumulative Variance | Cumulative % |
| 1     | 7,291     | 81,010            | 7,291               | 81,010       |
| 2     | 0,644     | 7,156             | 7,935               | 88,166       |
| 3     | 0,306     | 3,395             | 8,240               | 91,561       |
| 4     | 0,248     | 2,755             | 8,488               | 94,316       |
| 5     | 0,184     | 2,043             | 8,672               | 96,359       |
| 6     | 0,144     | 1,600             | 8,816               | 97,959       |
The analysis of the table shows that three main components give more than 90% of informativeness i.e. by these three time variables, using the model of the multiple regression it is possible to restore passenger flows at all SPs with a loss of accuracy of 10% only. A priori the average value of the main components is 0, and СКО – 1 (Table 2).

**Table 2.** Descriptive statistics of time series of the main components.

| Variable | Valid N | Mean  | Minimum | Maximum | Std. Dev | Skewness | Kurtosis |
|----------|---------|-------|---------|---------|----------|----------|----------|
| F1       | 91,000  | -0.000| -1.294  | 1.993   | 1.000    | 0.827    | -0.861   |
| F2       | 91,000  | 0.000 | -2.232  | 2.515   | 1.000    | 0.240    | -0.231   |
| F3       | 91,000  | 0.000 | -2.417  | 2.273   | 1.000    | -0.182   | -0.069   |

In addition, the character of the autocorrelation functions of the first main components is quite interesting (Figure 10,11).

**Fig. 10.** Autocorrelation function of the 1-st component.
Fig. 11. Autocorrelation function of the 2-nd component.

Fig. 12. Autocorrelation function of the 3-rd component.
If the character of the autocorrelation function of the first main component (Figure 10) has a clearly expressed cyclic character, the second is monotonous and highly protracted (Figure 11) and the third autocorrelation is practically statistically insignificant, (Figure 12). Cross-correlation function is also statistically insignificant (Figure 13).

In connection with the obtained results of the statistical analysis, when two or three abstract time series can practically restore all the passenger flows, a model of generation of these flows, presented below, is proposed in this study.

4 Time series generation model with a given autocorrelation function

Since the simulation of route servicing requires the restoration of adequate associated passenger flows at all SPs the algorithm for generating these flows based on the inverse transformation of the main components is proposed in this study. By definition the main components are independent of each other which allows to solve adequately the problem of generating a sample trajectory of time series of each component since no other information for its generating is required. A stationary time series model with a given autocorrelation function is a basic generation model.

The following algorithm is proposed for generating main components with autocorrelation $R(m)$. Suppose that a sequence $x_i$ of length $N$. is considered. For $m=1,2,…$ the relation $R(m) = \frac{R(0)}{N - m} \sum_{i=1}^{N-m} x_i x_{i+m}$ which leads to the search of a solution of the systems of equations $R(m) = M(x_i, x_{i+m})$ with respect to $x_i$ is done. To generate the sequence a set of random equally distributed and independent random values $\eta_i i=1,2,…$ with zero
mathematical expectation is formed which is converted into the sequence \( \xi_j = \sum_{i=1}^{n} x_i \eta_{i+j} \), \( j=1,2,..... \)

From the condition \( M \eta_i = 0 \) follows that \( M \xi_j = M \sum_{i=1}^{n} x_i \eta_{i+j} = M \eta \sum_{i=1}^{n} x_i = 0 \), i.e. the mathematical expectation of the linear transformation is also 0. Herewith

\[
M \xi_j \xi_{j+l} = M \left[ \sum_{i=1}^{n} x_i \eta_{i+j} \cdot \sum_{i=1}^{n} x_i \eta_{i+j+l} \right] = \sum_{i=1}^{n} \sum_{i=1}^{n} x_i x_k M (\eta_{i+j} \cdot \eta_{i+j+l}) ,
\]

where \( M (\eta_{i+j} \cdot \eta_{i+j+l}) = \begin{cases} \{0, & \text{if } i \neq k + l \\ M \eta^2, & \text{if } i = k + l \end{cases} \). Because for values \( l \geq 1 \) it is possible only when \( k \leq l, u k = i-l \) i.e. \( M \xi_j \xi_{j+l} = M \eta^2 \sum_{i=1}^{n} \sum_{i=1}^{n} x_i x_k \), so based on the substitution of the summation of variables is obtained \( M \xi_j \xi_{j+l} = M \eta^2 \sum_{i=1}^{n} \sum_{i=1}^{n} x_i x_k = M \eta^2 R(l) \), i.e. covariance function \( R(m) \) of the time series \( \xi \) has a given under the conditions the formation of the random tie series, for solving a system of algebraic equations the Seidel iterative procedure is proposed to be used.

5 Restoration results

After generating sample series of the main components of passenger flows on the basis of the multiple regression model the linear transformation to form passenger flows at a certain SP (Figure 5) with the obtained regression coefficients of passenger flows at all SPs is implemented as the next stage (Table 3). Figure 14 shows an example of restoration for SP №30.

Fig. 14. Original and restored time series of passenger flows.
Table 3. - Regression parameters of passenger flows restoration at SP.

| N=91 | Regression Summary for Dependent Variable: 30 (M_85) | R=0.97338091  | R?=0.94747040  | Adjusted R?=0.94565904 | $F(3,87)=523.07 \ p<0.0000 \ \text{Std. Error of estimate: 8.7433}$ |
|------|-----------------------------------------------------|----------------|----------------|-------------------------|------------------------------------------------------------------|
|      | Intercept                                           | 176.780        | 0.917          | 192.876                 | 0.000                                                             |
|      | F1                                                 | -0.830         | 0.025          | -31.113                 | 0.922                                                             |
|      | F2                                                 | 0.487          | 0.025          | 18.264                  | 0.922                                                             |
|      | F3                                                 | 0.149          | 0.025          | 5.595                   | 0.922                                                             |

In this case the accuracy of restoration which can be expressed in terms of a multiple correlation coefficient (R=0.97) is quite high.

Thus, the efficiency of the proposed approach to simulate passenger flows of specified statistical characteristics based on the results of the main components analysis with the subsequent inverse conversion into passenger flows at some SPs is shown. The algorithm itself is based on the obtained data of the statistical analysis which made it possible to restore all the passenger flows at all SPs by two or three time series. This approach provides a more adequate model of the associated time series of passenger flows necessary to be included in the simulation model of bus route servicing.

6 Conclusion

Thus, instrumental means of a hybrid training system allow to form an algorithmic structure of software applications due to transitions between applications under the conditions of its completion using a standardized interface, so a user scenario is created. All mechanisms aimed at an operational creation of techniques having a typical proven set of universal applications. In addition to the mechanism of scenarios creation, a model of structurization of scenarios which allows to implement an application synchronization is proposed.

For methods and algorithms given in the article an appropriate software applications included in the hybrid training system have been developed. As a result the trainee has a possibility to perform a statistical analysis of passenger flows at the selected stopping points and analyze cross-correlation between them. On the base of the given analysis it is possible to launch an application of the model range generation, passenger flow and a comparative analysis with the original series of passenger flows whose characteristics were used for the series generation.

References

1. P. Drijvers, Digital technology in mathematics education: Why it works (or doesn’t), in Selected regular lectures from the 12th international congress on mathematical education, pp. 135-151, Springer, Cham (2015)
2. F.G. Alessio, M.C. Brambilla, A. Calamai, C. de Fabritiis, L. Demeio, M. Franca & M. Petrini, New Multimedia Technologies as Tools for a Modern Approach to Scientific Communication and Teaching of Mathematical Sciences, in The First Outstanding 50 Years of “Università Politecnica delle Marche”, pp. 393-402, Springer, Cham (2019)
3. A. Brenner, M. Shacham, & M.B. Cutlip, *Applications of mathematical software packages for modelling and simulations in environmental engineering education*, in Environmental Modelling & Software, **20**(10), pp. 1307-1313 (2005)

4. D. Wick (2009), *Free and open-source software applications for mathematics and education*, in Proceedings of the twenty-first annual international conference on technology in collegiate mathematics, pp. 300-304, Louisiana New Orleans (2009)

5. K.J. Kachashvili, D.G. Gordeziani, D.Y. Melikdzhanian, V.I. Khuchua & V.A. Stepanishvili, *Software packages for automation of environmental monitoring and experimental data processing*, in Geocology and Computers, pp. 273-278, Routledge (2018)

6. G. Wei, H. Shen & R. Xuehua, *Case Study on the Design and Teaching of MOOC: English Grammar*, in 4th International Conference on Economics, Management, Law and Education (EMLE 2018), December 2018, Atlantis Press (2018)

7. T. Haendler, G. Neumann & F. Smirnov, F., *An interactive tutoring system for training software refactoring*, in Instructor, **1**, 4 (2019)

8. H. Yang & Q. Jia, *Automatic synchronization technology of report data based on OLE*, in AIP Conference Proceedings (Vol. 1890, No. 1, p. 040065), AIP Publishing, October (2017)

9. A.S. Semenov, V.M. Khubieva & Y.S. Kharitonov, *Mathematical Modeling of Static and Dynamic Modes DC Motors in Software Package MATLAB*, in 2018 International Russian Automation Conference (RusAutoCon), pp. 1-5, September 2018, IEEE 2018

10. H.Y. Morokhovets, M.S. Sainenko, Y.N. Lysanets & O.V. Silkova, *THE USE OF MAPLE MATHEMATICAL SOFTWARE IN TEACHING MEDICAL AND BIOLOGICAL PHYSICS*, in *The Medical and Ecological Problems*, **22**(1-2), 63-65 (2018)

11. D. Ahmetovic, T. Armano, C. Bernareggi, M. Berra, A. Capietto, S. Coriasco, & E. Taranto, E., *Axessibility: a LaTeX Package for Mathematical Formulae Accessibility in PDF Documents*, in Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 352-354, October 2018, ACM (2018)

12. M. Shyshkina, U. Kohut & M. Popel, *The Systems of Computer Mathematics in the Cloud-Based Learning Environment of Educational Institutions* (2018) arXiv preprint arXiv:1807.01770.

13. G. Greefrath, C. Hertleif & H.S. Siller, *Mathematical modelling with digital tools—a quantitative study on mathematising with dynamic geometry software*, ZDM, **50**(1-2), pp. 233-244 (2018)

14. A. Benveniste, M. Métivier & P. Priouret, *Adaptive algorithms and stochastic approximations* (Vol. 22), Springer Science & Business Media (2012)

15. A.S. Poznyak, K. Najim & E. Gomez-Ramirez, *Self-learning control of finite Markov chains*, CRC Press (2018)

16. H. Robbins & S. Monro, *A stochastic approximation method*, in The annals of mathematical statistics, pp. 400-407 (1951)

17. A.A. Dezin, *Multidimensional analysis and discrete models*, CRC Press (2018)

18. I.V. Bychkov, G.A. Oparin, A.G. Feoktistov, I.A. Sidorov, V.G. Bogdanov & S.A. Gorsky, *Multiagent control of computational systems on the basis of meta-monitoring and imitation simulation*, in Optoelectronics, Instrumentation and Data Processing, **52**(2), pp. 107-112 (2016)

19. J. Solomon, G. Peyré, V.G. Kim & S. Sra (2016), *Entropic metric alignment for correspondence problems*, in ACM Transactions on Graphics (TOG), **35**(4), p. 72 (2016)

20. A. Andronov, *On some approach to an estimation of correspondence matrix of transport network* (2009)
21. L. Markauskaite, *Exploring the structure of trainee teachers' ICT literacy: the main components of, and relationships between, general cognitive and technical capabilities*, in Educational Technology Research and Development, 55(6), pp. 547-572 (2007)