

Keywords: Intelligent agents; transport choosing model; passenger satisfaction index; transport quality

Sergey VIKHAREV*, Maxim LYAPUSTIN, Danil MIRONOV,
Irina NIZOVTSEVA, Vladimir SINITSYN

Ural Federal University
19 Mira, Ekaterinburg, 620002, Russia

*Corresponding author. E-mail: sergey@viharev.com

MODELING OF PASSENGERS' CHOICE USING INTELLIGENT AGENTS WITH REINFORCEMENT LEARNING IN SHARED INTERESTS SYSTEMS; A BASIC APPROACH

Summary. The purpose of this paper is to build a model for assessing the satisfaction of passenger service by the public transport system. The system is constructed using intelligent agents, whose action is based on self-learning principles. The agents are passengers who depend on transport and can choose between two modes: a car or a bus wherein their choice of transport mode for the next day is based on their level of satisfaction and their neighbors' satisfaction with the mode they used the day before. The paper considers several algorithms of agent behavior, one of which is based on reinforcement learning. Overall, the algorithms take into account the history of the agents' previous trips and the quality of transport services. The outcomes could be applied in assessing the quality of the transport system from the point of view of passengers.

1. INTRODUCTION

This work is devoted to the study of relationships between transport companies and passengers in a system with shared interests. In order to formulate the concept of a system with shared interests [1-3], we need to introduce some notation. We will denote the set of regulators A (authorities), the set of executors E (executers), and the set of consumers (customer).

2^X is the set of all subsets of an arbitrary set X (Boolean of X). The set of real numbers is traditionally denoted by \mathcal{R} . A system with shared interests is a collection $(A, E, C, \alpha, \gamma, \tau)$, where $\delta: A \times 2^E \rightarrow \mathcal{R}$, $\gamma: E \times 2^C \rightarrow \mathcal{R}$, $\tau: A \times 2^C \rightarrow \mathcal{R}$ are the functions of interaction.

In systems with shared interests, there is a problem of determining the required financing and the stringency of regulation of the industry. There is also a problem of determining the functions of interactions between δ, γ, τ .

In this paper, two objectives are pursued: the first is to propose a model for the interaction δ between the set of regulators and the set of executors; the speed of response of carriers to changes in regulation is evaluated with the help of this function. The second objective is to propose a model for the interaction of γ between the set of executors and the set of consumers.

The approach to definition of function of interaction δ for a pair of sets of regulators and carriers is set on the basis of population model Lotka–Volterra. This approach allows us to give a short-term (for several years) forecast of the response of many carriers to changes in the rigidity of regulation and the amount of subsidization by a set of regulators. The application of a discrete series of queuing systems to the determination of the interaction function γ is used for a set of carriers and a set of passengers. The interaction function γ is compiled on the basis of data on the profitability of carriers and the quality of their transport services provided by the function δ . It is supposed to divide the range of

possible quality of services into a number of levels. The function γ calculates the proportion of passengers who have received a transport service of a certain quality level.

In the trivial case, the interaction function τ of authorities and passengers in the modeling of social tension caused by transport services can be calculated as the arithmetic mean of the quality of transport services measured in time and space. Such an approach, with obvious simplicity, has several significant drawbacks. First, the negative situation with transport services, for example, in one of the city's districts, can be veiled by the high quality of trips in other areas. It is clear that social tension in this region is high and requires the reaction of the authorities. Simple averaging will not allow to notice this hotbed of tension. Second, important for the satisfaction of passengers is not the absolute level of the quality of the transport service, but how much better or worse things are for the "neighbors", for example, in other areas of the city, regions of the country, or even neighboring countries. It is easy to see (for example, [4]) that the quality of transport services can vary significantly from country to country. The same absolute level of quality of transport services in one area can lead to a social explosion, whereas for another, it is absolutely natural and does not lead to social tension of the population.

Thus, simple averaging is not sufficient to determine the level of social tension resulting from the quality of transport services provided. It is necessary to propose a model that meets the following requirements.

1. The model should be sufficiently sensitive to identify foci of dissatisfaction with transport services.
2. The model should evaluate the satisfaction of passengers, taking into account the transport situation of their "neighbors".
3. The model should provide for a computer implementation that automates the calculation of the level of social tension and visualizes the results.
4. It seems useful if the model realizes the decision-making mechanism by the passenger, about the type of transport (for example, public transport, private car or taxi) to be used to achieve the maximum level of comfort.
5. It is necessary to ensure that in the model, the particular type of transport was used by a particular group of passengers, united by geographic and temporal characteristics.

Based on these requirements, it seems reasonable to consider the apparatus of agents' networks for the simulation purposes, when agents use reinforcement learning.

The basic idea of reinforcement learning is simply to capture the most important aspects of the real problem faced by a learning agent interacting with his/her environment to achieve a goal [5]. Clearly, such an agent must be able to sense the state of the environment to some extent and must be able to take actions that affect the state. The agent also must have a goal or goals relating to the state of the environment. The MDP formulation is intended to include just these three aspects – sensation, action, and goal – in their simplest possible forms without trivializing any of them. Any method that is well suited to solve such problems would be considered to be a reinforcement learning method.

Reinforcement training assumes that each agent interacting with the environment is trying to improve something for himself. All agents who use reinforcement learning have explicitly expressed goals; they can perceive the characteristics of the environment and also choose the actions that affect this environment.

2. MODELING

In order to display this approach on the presented system, the following mathematical model will be compiled and the main aspects of the real problem will be fixed.

We divide the entire set of passengers (transport service consumers, designated C) into subsets $[C^1, \dots, C^n]$ such that each subset is a group of people who live in some definite proximity to one another and somehow can communicate with each other.

Next, for simplicity, it is supposed that usually the choice consists of two actions to choose a trip by the car or by public transport. In general, a selection society may contain a greater number of

elements, corresponding, for example, to the selection of various types of public transport, pedestrian and bicycle routes. Randomly, the passengers are divided from set C into these two classes.

For each trip, regardless of the element of the multiple choices of transport, an assessment of the quality of services of the transport system is introduced and denoted by k . The works [6-8] give a detailed description and methodology of construction of the transport service quality assessment. This assessment is influenced by parameters such as the modern bus, traffic congestion, waiting time at the bus stop and others.

In the case of choosing a bus, we consider the set of all trips. Depending on the quality assessment (see Fig. 1), we divide it into 3 classes: high-quality, normal, bad. To correctly identify these classes, they have been studied in detail in article [9]. Then, passengers are randomly assigned to bus routes, each of which corresponds to one of the three classes. Thus, for each passenger, we will get an assessment of the quality of services provided.

For all car owners, the assessment of the quality of personal transportation depends on traffic congestion, that is, on the activity of using private vehicles by neighbors. This simplified approach cannot claim the title of evaluation of the quality of personal transportation, but it is sufficient to demonstrate the work of the model. We introduce the notation k_a , the evaluation of the quality of personal transport, then to calculate the quality for car owners, we obtain the following formula:

$$k_a = 1 - \frac{\text{count}_a(c_i)}{n}, \text{ where } \text{count}_a(c_i) \text{ is the number of agents who chose personal transport.}$$

For all car owners at the time of initialization, we set the coefficient k equal to a certain constant.

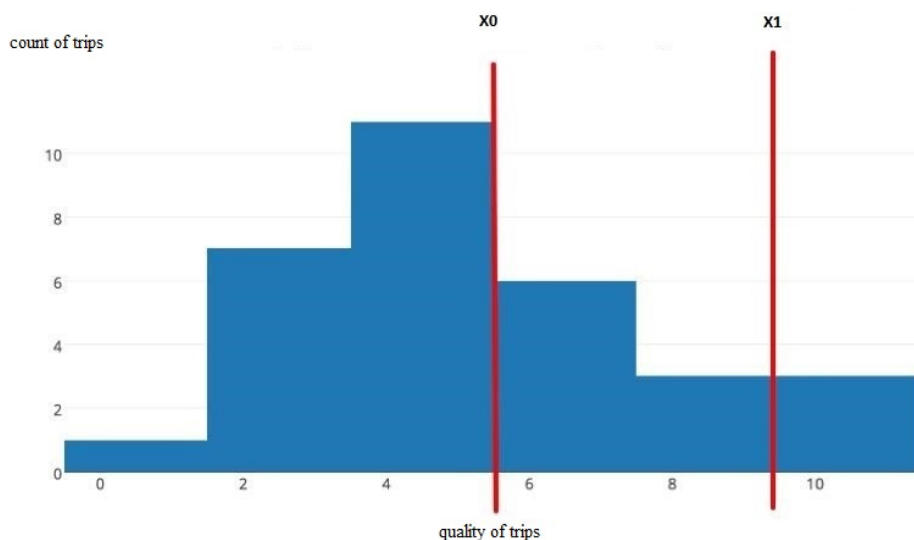


Fig. 1. Quality assessment, X0 – acceptable quality, X1 – high quality

Another indicator is introduced, S , the level of satisfaction, which determines the passenger's satisfaction with the quality of services provided by the transport system. The approach to considering the satisfaction rate as a perception of actual quality has been discussed for example in [10-11]. If the level of satisfaction of a passenger is below a certain limit, then he/she is transplanted to another mode of transport (for simplicity, every person in our system is capable of buying a car and of driving it at any moment). [12]

Thus, at the initialization stage, there are subsets $[C^1, \dots, C^n]$, in which each person either drives a car or a bus, and he/she can assess the level of satisfaction depending on the quality of the services provided to him/her. It should be highlighted that people communicate with each other and are in approximately the same conditions. This means that each person can determine his/her level of satisfaction with other people, which means he/she can decide whether to change to another mode of transport on the next day.

The transport selection algorithm is directly considered for the next trip. It is reasonable to note that the choice of transport is affected by the satisfaction ratio. Therefore, for simplicity, we split the algorithm into two steps (a mini-algorithm):

1. Determination of the level of satisfaction, depending on the quality of service delivery and personal travel history.
2. Deciding the next trip on the basis of the level of satisfaction and information about neighbors.

Thus, there is a model of the system, which at the time of initialization has arbitrary parameters, and after some iterations, it stabilizes somehow.

The following examples can be considered:

Example 1. If most people transplanted to cars realize that they spend a lot of time in traffic jams, they are transplanted back to public transport. Then some people are transplanted to the cars again, and some continue to use the services of buses. After a number of iterations, the system comes to a state of rest, so the level of mobility and social tension can be assessed.

Example 2. People in some area constantly drive cars and are not transplanted to buses, which means that the quality of services provided by buses is extremely low, thus revealing a hotbed of dissatisfaction with transport services.

Let us proceed to the formalization of this system in the form of an agent network wherein each intelligent agent uses the method of reinforcement learning.

Let each person be an agent and the environment for any agent is one of the subsets $[C^1, \dots, C^n]$. The purpose of the agent is to raise his/her level of satisfaction, which is a reward function that is needed to maximize. For this, the agent needs to interact with the environment. The interaction is reduced to the fact that the agent, analyzing the level of satisfaction of his/her neighbors (the same agents), as well as his/her own, makes a decision on what transport should he/she opt for the next day. Based on the decisions of all agents, the compensation function will calculate the reward, depending on the level of satisfaction that each agent will receive while choosing the transport for the next trip. After a few iterations, the agent will accumulate his/her own experience and, based on this experience, will eventually be able to improve the level of satisfaction [8] until this level stabilizes. Thus, we can expect that the system will come to a more or less stable state.

3. TRANSPORT SELECTION ALGORITHM

To expand the research capabilities, we will consider that it is possible to influence how the level of satisfaction will be built and how, on the basis of this level, transport for the next trip will be chosen.

The concept of transfer function is introduced – this implies how from the current assessment of the passenger's satisfaction they get their choice of transport for the next trip.

An algorithm obtained in the course of reinforcement learning can be an excellent transfer function. An agent will come to the optimal choice of transport with time. But for the initial illustration of the choice of mode of transport personal satisfaction, the application of an empirically selected transfer function is considered.

Possible algorithms are considered without the participation of machine learning, based on our intuitive understanding of the operation of the transport system and the behavior of passengers, and also based on the ideas of the article [13].

In order to make a choice, passengers need to know two things:

- Their own level of satisfaction
- Level of satisfaction of neighbors and their mode of transport.

Algorithm No. 1 (simple comparison with average)

Determining the level of satisfaction

Let each passenger have his/her own level of satisfaction, which depends on the personal qualities of the person. This level has an effect on satisfaction. Also, each passenger stores information about

the quality of the last five trips. Satisfaction is considered by the following formula: (current trip quality - average for the last five trips + 1) / 2 + personal satisfaction.

Transport selection for the next trip (transfer function)

Every passenger knows about those neighbors who "surround" him/her. So, when the passenger decides to choose how he/she will commute the next day, he/she follows this algorithm:

- 1) divides his/her neighbors into groups by the type of transport
- 2) for each group, he/she considers the average level of satisfaction
- 3) transplants to the mode of transport whose middle level is the largest

In the event that a passenger has chosen a car, he/she is 85% likely to remain on it, taking into account the possibility of a car breakdown, problems with gasoline and other unforeseen circumstances.

Pseudocode:

Input:

- Neighbors (array[passengers]) – array of neighbors of a particular passenger
- Satisfaction – the level of satisfaction of the passenger
- CurrentTransportType – type of passenger transport

Output: TransportType – new type of transport

1. Groups = Neighbors .GroupBy(TransportType);
2. for group in Groups
3. averageSatisfaction = getAveregeSatisfaction(group);
4. if (CurrentTransportType!= group.TransportType and Satisfaction < averageSatisfaction)
5. CurrentTransportType = group.TransportType
6. return CurrentTransportType;

Algorithm No. 2 (choice of deviation from a given level)

1. Algorithm for determining the level of satisfaction
The same as in Algorithm No. 1.
2. Transport selection algorithm for the next trip (transfer function)

Consider the behavior of the average user of the transport system, if the neighbors level of satisfaction on the same transport is higher, that is, the meaning of transplanted to an alternative mode of transport.

Since the passenger satisfaction values are random values, it would be reasonable to count the difference more than one standard deviation for significant differences. However, for simplicity, we assume that there is a certain border, at the intersection of which, there is a need to change the mode of transport.

Pseudocode:

Input:

- Neighbors (array[passengers]) – array of neighbors of a particular passenger
- Satisfaction – the level of satisfaction of the passenger
- CurrentTransportType – type of passenger transport
- DeviationValue – value significant variance

Output: TransportType – new type of transport

1. neighborsSameTransport = getNeighborsWithType(CurrentTransportType);
2. averageSatisfactionSameTransport = getAveregeSatisfaction(neighborsSameTransport);
3. if (averageSatisfactionSameTransport > Satisfaction)
4. CurrentTransportType = ChangeTransportType();
5. neighborsOtherTransport = getNeighborsWithType(OtherTransportType);

6. $\text{averageSatisfactionOtherTransport} = \text{getAverageSatisfaction}(\text{neighborsOtherTransport});$
7. $\text{if} (\text{averageSatisfactionSameTransport} > \text{Satisfaction})$
8. $\text{CurrentTransportType} = \text{ChangeTransportType}();$
9. $\text{return CurrentTransportType};$

Algorithm No. 3 (based on reinforcement learning)

1. Algorithm for determining the level of satisfaction
The level of satisfaction will be measured as the quality of the journey, which is derived from the transport system and multiplied by certain coefficient characterizing personal qualities. This does not take into account the history of satisfaction, but the history of travel will be taken into account in reinforcement learning.
2. Transport selection algorithm for the next trip (transfer function)
We introduce the following concepts:
 - S – the set of all states in which the agent can be located
 - a – action that can be performed by the agent
 - A_s – set of actions that can be performed from state s
 - $Q(S, a)$ – the agent's subjective assessment function of the quality of action a , which can be selected from states [5].

Example (chess):

s_0 is the initial state of the field

a_0 is moving pawn in the center of the field

a_1 is moving pawn in the edges of the field

It is clear that the first pawn move in the center of the field gives more profit than the first move in the edges of the board, which means $Q(S, a_1) > Q(S, a_0)$.

$Q(S, a)$ evaluates how beneficial it is to perform a specific action from a particular state.

The task of reinforcement learning is to construct $Q(S, a)$.

Indeed, if the function $Q(S, a)$ is known, it is enough to always take an action that maximizes this function for the optimal strategy.

The algorithm of constructing $Q(S, a)$ can be considered.

Q-learning algorithm

Init :

1. for each s and a do $Q[s, a] = \text{RND} // 0..1$
- Observe:
2. $s' = s$
3. $a' = a$
4. $s = \text{FROM_ENVIROMENT} // \text{get current state from enviroment}$
5. $r = \text{FROM_ENVIROMENT} // \text{get reward for previous action}$

Update:

6. $Q[s', a'] = Q[s', a'] + \text{LF} * (r + \text{DF} * \text{MAX}(Q, s) - Q[s', a'])$

Decision:

7. $a = \text{ARGMAX}(Q, s)$
8. $\text{TO_ACTIVATOR} = a$
9. GO TO 2

Notations:

- LF is a factor of learning. The higher it is, the stronger the agent trusts the new information.
- DF is a discount factor. The smaller it is, the less agent thinks about the benefits of his/her future actions.

Function $\text{MAX}(Q, s)$.

1. $max = minValue$
 2. for each a of $ACTIONS(s)$ do
 3. if $Q[s, a] > max$ then $max = Q[s, a]$
 4. return max
- Function $ARGMAX(Q,s)$
1. $amax = First\ of\ ACTION(s)$
 2. for each a of $ACTION(s)$ do
 3. if $Q[s, a] > Q[s, amax]$ then $amax = a$
 4. return $amax$

Using this approach, we describe the choice of the mode of transport agent for the next trip:

- State is the object that characterizes the agent at a given time. The object contains the following:
 - current transport of the agent itself
 - transport information for each neighbor
 - information for each neighbor on more or less (1 or -1) the level of satisfaction of a neighbor compared to an agent.
- Actions:
 - Select car
 - Select bus
- Reinforcement function that returns the reward r :
 - if after the action, the current level of satisfaction of the agent has become higher, then the agent should be praised, and $r = 1$
 - if it becomes worse, on the contrary, the agent should be punished, $r = -1$
 - if remained the same, then $r = 0$
- $Q(S, a)$ is a matrix that stores for each state and action from that state a value that will be calculated the last time in step 3 of the Q – learning algorithm.

Agents will be trained as follows: multiple agents can be specified with different initial conditions, running them on the system. Then, a sufficient number of iterations (to be established empirically) should be conducted, after which the most successful agent should be selected with the highest level of satisfaction, and its function $Q(S, a)$ should be assigned to all the others. After several such "games", the optimal function $Q(S, a)$ will be obtained.

This means that the transport selection algorithm itself is reduced to a comparison between the values $Q(S, \text{"select machine"})$ and $Q(S, \text{"select transport"})$. The action whose function value is greater is selected.

4. AN EXAMPLE OF WORK MODELS

Table 1 presents an example of work agents' networks, when agents use reinforcement learning for passenger transport.

Consider three agents c_1, c_2, c_3 , where c_1, c_2 use machines, and c_3 is moved on the bus, k is the coefficient of the service quality and S is the satisfaction. See Table 1.

The first iteration can be considered in which agents will choose the mode for the next day. It can be established that c_3 will be always completely satisfied with the services of the bus and will never be transplanted to car. It can be established that c_1 will switch to the bus, as his comfort level is the least, whereas c_2 will remain with the car option.

Then, we will see the following picture, because the coefficients will change (see Table 1 Iteration 1).

Considering the second iteration, it can be observed that c_1 and c_3 always travelled by bus; c_2 perceived that neighbors had a higher level of comfort compared to his comfort, so he also decided to take the bus.

Table 1

States of Agents on different iterations

Stage	c_1			c_2			c_3		
	Decision	k	S	Decision	k	S	Decision	k	S
Initiation	Car	0,23	0,30	Car	0,61	0,35	Bus	0,51	0,40
Iteration 1	Bus	0,57	0,43	Car	0,63	0,35	Bus	0,51	0,40
Iteration 2	Bus	0,57	0,43	Bus	0,49	0,33	Bus	0,51	0,40
Iteration 3	Bus	0,57	0,43	Car	0,63	0,35	Bus	0,51	0,40

For some reason, the comfort level of c_2 was lower than it was before, so c_2 decided to move again to the car. Thus, our simplified system comes to balance.

In a real problem, there will be more neighbors, and thus more iterations, but it is important to note that the system will tend to equilibrium wherein it will be possible to estimate the number of people satisfied with services provided.

5. SCIENTIFIC WEB SERVICE

In the course of current work, a scientific web service was developed that allows checking all possible hypotheses of passenger behavior under different initial conditions and with different transfer functions described above (Algorithm No. 1, Algorithm No. 2, Algorithm No. 3).

The work of this web service can be considered, for example 3 passengers, with the transfer function using reinforcement learning (Algorithm No. 3).

It can be supposed that an initial state is set wherein passenger No. 0 is moved to the bus and the other passengers are moved to cars. (Fig. 2.)

After the first iteration, passenger No. 0, based on his experience and knowing the satisfaction level of neighbors, decides to take the car, whereas the other two decide to take the bus as an experiment (Fig. 3).

Then, passenger No. 1 notices that the neighbor on the left who chose the car has a higher satisfaction than the neighbor on the right who chose the bus, so he decides to take the car (Fig. 4).

In the end, the passenger No. 2 also decides to choose the car, because his satisfaction level is the lowest, but with the roads being very busy, the quality of transport deteriorates, hence worsening the satisfaction from the trips (Fig. 5). Therefore, in the next iteration, passenger No. 1 and passenger No. 2 are transferred back to the bus (Fig. 6). System comes to balance (as in Fig. 3) and to the maximum satisfaction with transport services.

6. CONCLUSIONS

Returning to the main task of this work, the definition of macro-satisfaction of the population with the quality of their transport services, the proposed model allows to determine the centers of tension, due to the division of the set of passengers into subsets. Now each agent can analyze the level of satisfaction of his/her "neighbors" and make the decision about the mode to choose on the next day on the basis of these facts. Moreover, the initial grouping can be made not only on a geographical basis, but also on any other basis, for example, on time. Now there are only two modes of transport in the model, but this list can be easily expanded. The proposed method allows to calculate satisfaction based on the actual quality of transport services. The satisfaction rate can be further used to assess the mobility and social tensions depending on the transport service. The algorithm consists of the selected transfer function and the algorithm of determination of satisfaction. Algorithm results can be easily visualized. Of the above advantages, we can safely say that the approach can be used to calculate the function τ .

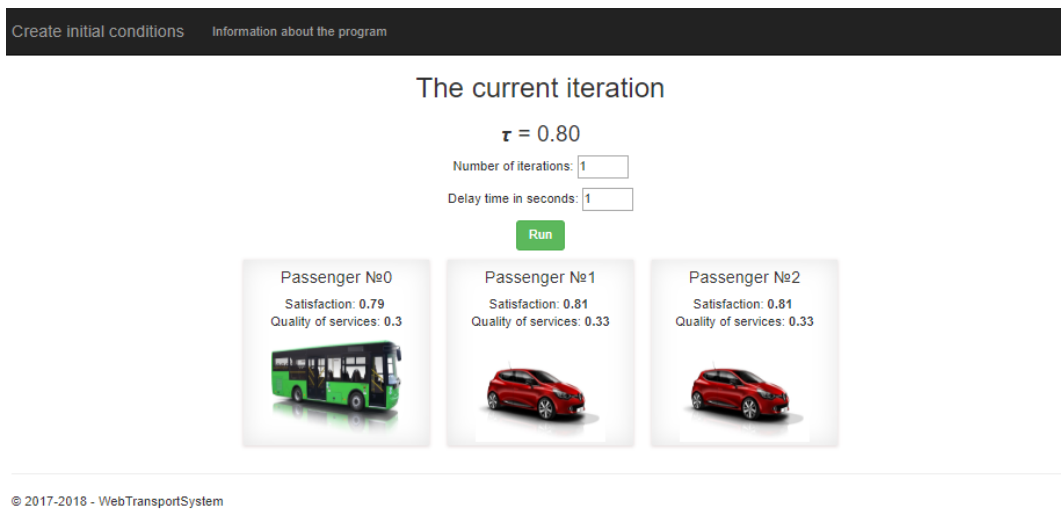


Fig. 2. Initial state

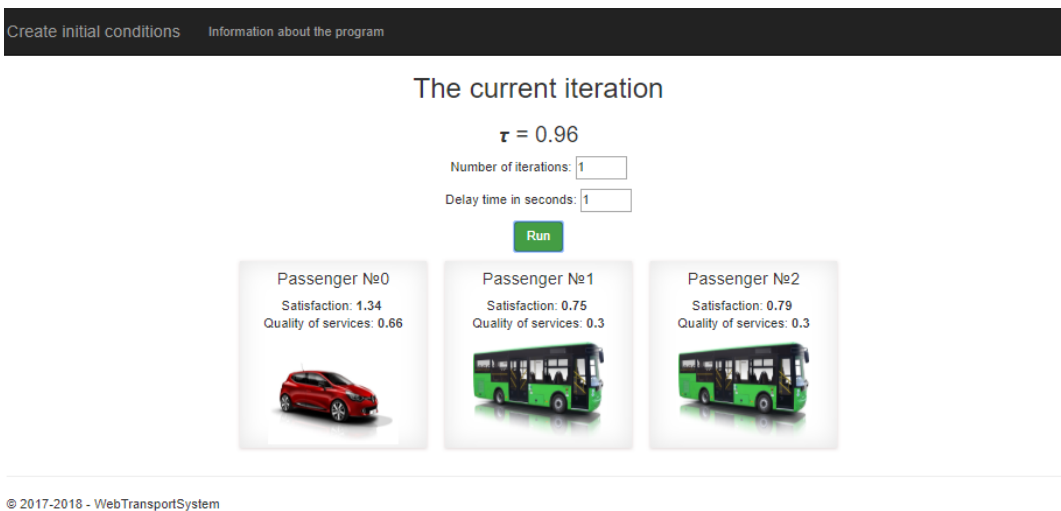


Fig. 3. After the first iteration

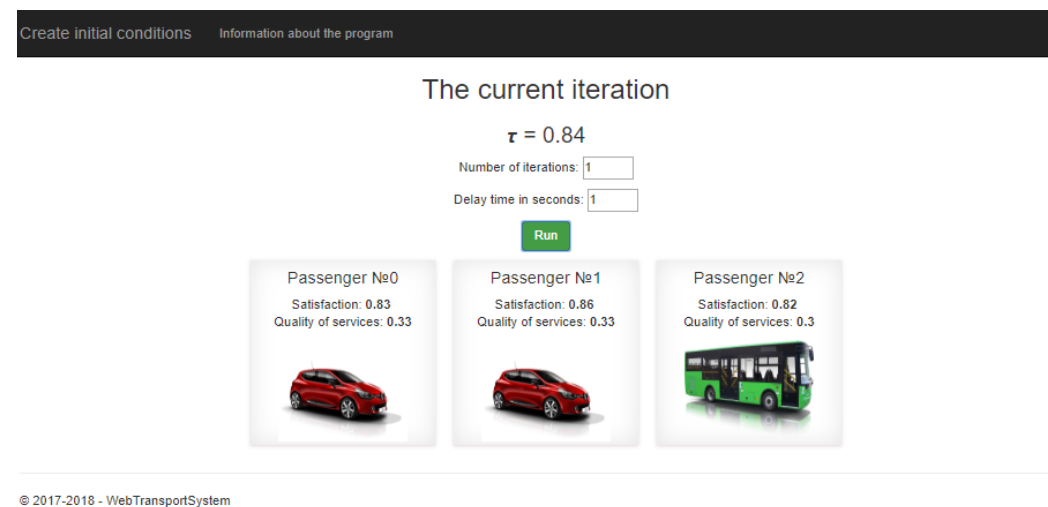


Fig. 4. After the second iteration

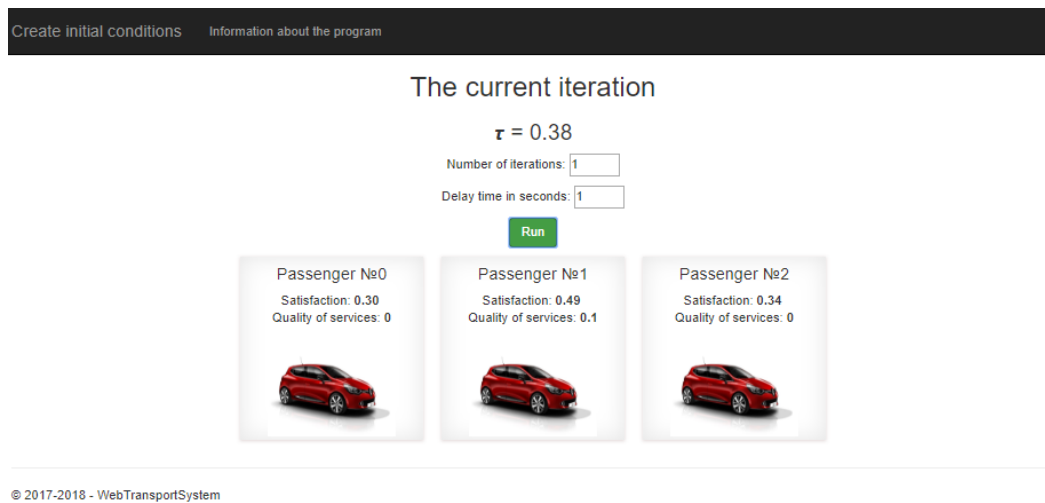


Fig. 5. After third iteration

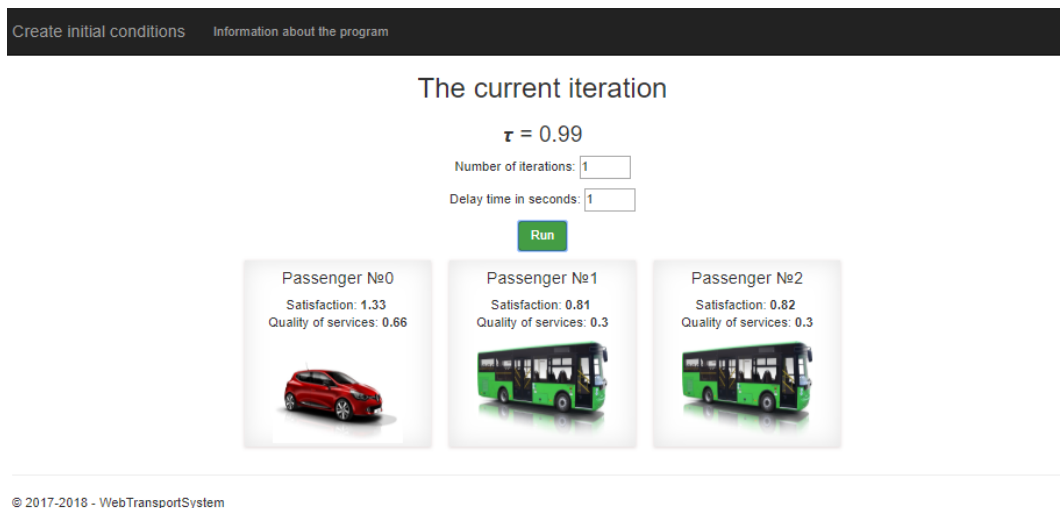


Fig. 6. After fourth iteration

In further studies, it is necessary to develop a «docking» algorithm for γ and τ functions. A method of representing residential agglomerations in the form of two grids will be considered. The first grid should reflect the transport topology. The second grid should reflect the relationship of information connectivity between neighbors in the broad sense of the word. The configuration of such grids could be used as input to the functions γ and τ . So the output data of the γ function could be used as additional data for τ including, for example, time-detailed forecasts based on the actual quality of transport services received by passengers. The use of common initializing grids for both functions will allow us to go to the core-term prediction of the choice of modes of transport.

Acknowledgements

The authors acknowledge the support from the Russian Science Foundation (project No. 17-71-20108)

References

1. Sizii, S. & Shichko, A. & Vikharev, S. Organizational processes in networks with shared interests: relevance, statement of the problem, research plan. *Bulletin of USURTU*. 2009. No. 1-2. P. 34-42.

2. Сай, В. *Планетарные структуры управления на железнодорожном транспорте*. Москва: ВИНТИ РАН. 2003. 336 с. [In Russian: Sai, V. *Planetary management structure on railway transport*. Moscow: VINITI RAN, 2003. 336 p.].
3. Сай, В. & Сизый, С. *Образование, функционирование и распад организационных сетей*. Екатеринбург: УрГУПС. 2011. 270 с. [In Russian: Sai, V. & Sizii, S. *The formation, functioning and dissolution of organizational networks*. Ekaterinburg USURTU. 2011. 270 p.].
4. *Logistics performance index: Quality of trade and transport-related infrastructure*. The World Bank Group. Available at: <https://data.worldbank.org/indicator/LP.LPI.INFR.XQ>.
5. Sutton, R. & Barto, A. *Reinforcement Learning: An Introduction*. London: The MIT Press. 2017.
6. Olivkova, I. Evaluation of public transport criteria in terms of passengers' satisfaction. *Transport and Telecommunication*. 2016. Vol. 17. No. 1. P. 18-27.
7. Ismail, R. & Hafezi, M.H. & Nor, R.M. & et al. Passengers preference and satisfaction of public transport in Malaysia. *Australian Journal of Basic and Applied Sciences*. 2012. Vol. 6(8). P. 410-416.
8. Hwe, S.K. & Cheung, R.K. & Wan, Y. Merging bus routes in Hong Kong's central business district: Analysis and models. *Transportation Research Part A: Policy and Practice*. 2006. Vol. 40. No. 10. P. 918-935.
9. Van Lierop, D. & Badami, M.G. & El-Geneidy, A.M. What influences satisfaction and loyalty in public transport? A review of the literature. *Routledge*. 2018. Vol. 38. No. 1. P. 52-72.
10. Wang, C. & Weng, J. & Chen, Z. & et al. A method of building bus satisfaction evaluation index system based on passengers' perception. *American Society of Civil Engineers (ASCE)*. 2018. Vol. 2018-January. P. 4675-4683.
11. Weng, J. & Di, X. & Wang, C. & et al. A bus service evaluation method from passenger's perspective based on satisfaction surveys: A case study of Beijing. China. *Sustainability MDPI AG*. 2018. Vol. 10. No. 8. P. 1-15.
12. Fitzsimmons, E.G. Downside of ride-hailing. More gridlock. *The New York Times*. March 8, 2017.
13. Crawford-Brown, D.J. The changing influences on commuting mode choice in urban England under Peak Car: A discrete choice modelling approach. *Transportation Research Part F: Traffic Psychology and Behavior*. October, 2018. Vol. 58. P. 167-176.

Received 09.12.2017; accepted in revised form 05.06.2019