# Forecasting network traffic in the information and telecommunication system of railway transport by means of a neural network

*Igor* Zhukovytskyy, and *Victoria* Pakhomova*

Ukrainian State University of Science and Technology, Chair of electronic computing machines, 49010 Dnipro, Lazaryan Street 2, Ukraine

**Abstract.** Network traffic is one of the most important actual indicators of the information and telecommunication system (ITS) of railway transport. Recent studies show that network traffic in the ITS of railway transport is self-similar (fractal), for the study of which the Hirst indicator can be used. One of the possible solutions is a method of network traffic forecasting using neural network technology, which will allow you to manage traffic in real time, avoid server overload and improve the quality of services, which confirms the relevance of this topic. The method of forecasting the parameters of network traffic in the ITS of railway transport using neural network technology is proposed: for long-term forecasting (day-ahead) of network traffic volume based on network traffic volumes for the previous three days using the created multilayer neuro-fuzzy network; for short-term prediction (one step forward, which takes five minutes) of network traffic intensity based on network traffic intensities for the previous fifteen minutes using the created multilayer neural network. The corresponding samples are formed on the basis of real values of network traffic parameters in the ITS of railway transport. Studies of optimal parameters of the created multilayer neural network, which can be integrated into specialized analytical servers of the ITS of railway transport, are carried out, which will provide a sufficiently high level of short-term forecasting of network traffic parameters (in particular intensity) in the ITS of railway transport at the stage of deepening the integration of the national transport network into the Trans-European Transport Network.

**Keywords:** railway transport, ITS, analytical server, forecasting, network traffic, parameter, neural means, error, epoch.

## 1 Entry

Ensuring the interoperability of rail transport at the stage of deepening the integration of the national transport network into the Trans-European Transport Network is possible only with a developed information structure. Today in Ukraine the information and telecommunication system (ITS) of railway transport is used. Until recently, the work of railway transport in Ukraine was an interaction of six railways: Lviv; Southwest; South; Odessa; Prydniprovska; Donetsk, each of which has implemented a corresponding ITS. Network traffic is one of the most important actual indicators of the information and telecommunication system (ITS) of railway transport. Recent studies show that network traffic for most types of services is self-similar (fractal), for the study of which the Hirst indicator can be used, which allows to determine the stochasticity of the process, but there is one major drawback - it takes considerable time to receive and process information about network traffic. The change in network traffic in the ITS of railway transport during the day leads to the need to determine in real time the overload in the computer network and monitor data flows. One of the possible solutions is the method of forecasting network traffic using neural network technology, which will allow you to manage traffic in real time, avoid server overload and improve the quality of services, which confirms the relevance of this topic.

### 1.1 Analysis of recent research and publications

It is known that forecasts can be long-term, medium-term and short-term.

In general, at the present stage there are methods for constructing forecasts with high accuracy based on the use of the following neural networks (NN) [1-4]: Multi-Layer Perceptron (MLP); Radial Basis Function Network (RBF); General Regression Neural Network (GRNN); Adaptive-Network-Based Fuzzy Inference System (ANFIS). Each NN has its advantages and disadvantages that affect the results of forecasting and require additional research.

Previously, the authors presented the results of studies of the use of intelligent network traffic routing tools in the prospective network MPLS ITS railway

* Corresponding author: viknikpakh@gmail.com

© 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/).

transport [6], but it is also necessary to conduct a study of the use of intelligent tools for predicting network traffic parameters.

Intelligent forecasting mechanisms in information and control systems of railway transport, in particular in the automated control system of freight transportation of railway transport, were described in [7, 8]. In these works, the use of specialized analytical servers was proposed, within which unified intellectual software systems are created, which can be used for a number of tasks, including forecasting tasks. The Figure 1 shows that the network structure of analytical servers (application server level) has a load balancing service [7, 9]. To implement this service, which can manage the load balance of individual network shares, it is necessary to have data on the load of these particles and the forecast of this load. This task can also be assigned to analytical servers.

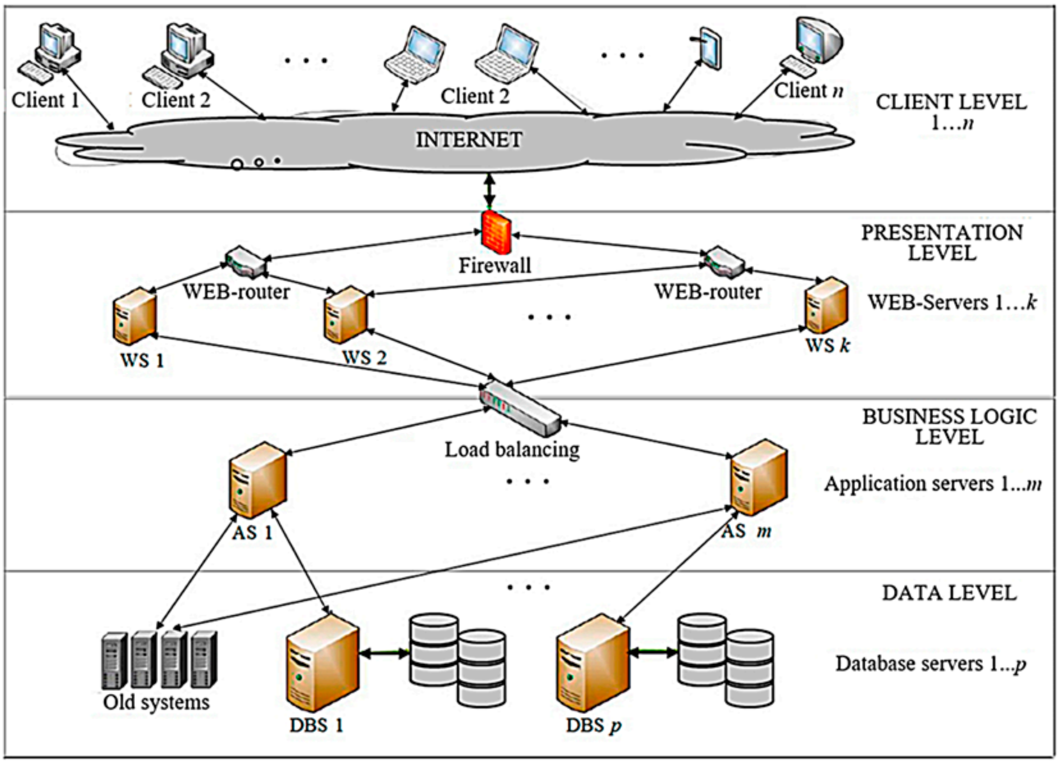


**Fig. 1.** Structural model of automated control system of freight transportation of railway transport [7].

### 1.2 The purpose of the work

The aim of this work is to develop a methodology for predicting network traffic parameters in the ITS of railway transport using neural network technology. In accordance with the goal, the following tasks were set: to analyze network traffic in the ITS of railway transport; when performing machine learning, identify the optimal parameters of the neural network, which will provide a sufficiently high level of forecasting the parameters of network traffic (in particular its volume and intensity) in the ITS of railway transport.

## 2 Main part

### 2.1 General characteristics of the methodology for forecasting network traffic parameters in the ITS of railway transport

To predict the parameters of network traffic in the ITS of railway transport using neural network technology, the authors propose method.

The method of forecasting network traffic parameters in the ITS of railway transport proposed by the authors provides: creation of a neuro-fuzzy network (NFN) for long-term forecasting of network traffic volume; creation of a neural network (NN) to obtain a short-term forecast of the intensity of network traffic. The use of NFN and NN to predict network traffic parameters in the ITS of railway transport requires the following steps: formation of a training sample; choice of configuration of NFN and NN; selection of the algorithm of NN training and the method of optimization of NFN learning.

### 2.2 Network traffic forecasting

In [5], an analysis of network (incoming and outgoing) traffic for the most important fragment of the Dnipro-Kyiv in the ITS of the Prydniprovska railway was performed. The main server is located in Kiev, according to which incoming traffic is the volume of information received per unit of time, and outgoing traffic is the volume of information transmitted per unit of time. As an example, Figure 2 presents network traffic (outgoing) for a time series length of 24 hours for different days of the week.

 https://doi.org/10.1051/matecconf/2023???02003

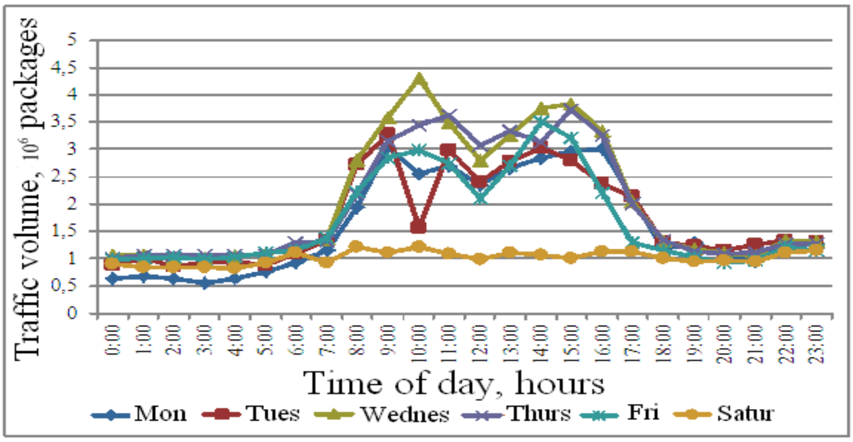


**Fig. 2.** The volume of network traffic (outgoing) in the ITS of railway transport [5].

In Figure 2, there is a tendency to behavior in the volume of network traffic for the week: it is approximately the same on Monday, Tuesday, Thursday and Friday; there are certain regular changes over certain periods. Thus, in particular, a smaller volume of traffic and more or less stable from 00:00 to 7:00, a significant and unstable traffic volume from 8:00 to 17:00, again a smaller and relatively unchanged traffic volume from 18:00 to 23:00. Network traffic is highest on Wednesday, and much less traffic on weekends than on weekdays. The Figure 2 shows that the volume of outgoing traffic on Wednesday compared to Monday, Tuesday, Thursday and Friday is about 1.3 times higher. To make a forecast of the volume of network traffic, the interval from 8:00 to 17:00 is chosen, where it fluctuates significantly, but on days of the week (Monday, Tuesday, Thursday, Friday), when the nature of traffic is approximately the same.

The analysis of the time series of network traffic volume in the ITS of railway transport was performed, the Hirst indicator was approximately $H$=0.89 (at $a=\pi/2$; $N$=144); Thus, the time series of network traffic is persistent, which is characterized by the effect of long-term memory.

The study of NFN configuration 3-6-8-8-1 (Fig. 3) was carried out, where 3 – number of neurons of the input layer; 6 – number of neurons of the inputmf layer; 8 – number of neurons of the rule layer; 8 – number of neurons of the outputmf layer; 1 – the number of neurons of the output layer created in MatLAB according to the Sugeno algorithm [5].

For linguistic estimation, each input variable had two terms (maximum and minimum values), the Gaussian function (gaussmf) is taken as the membership function, and the linear function (linear) is taken as the resulting variable. On the created NFN, studies were conducted on the average learning error from the number of its inputs, from the number of terms of input variables, from the number of training examples using the following learning optimization methods: Backpropa (error backpropagation method) and Hybrid (hybrid method that combines the method of reverse error propagation with the least squares method).

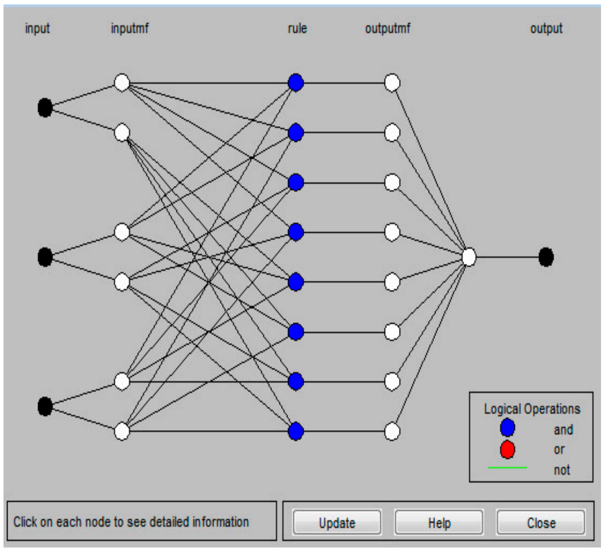


**Fig. 3.** Created in MatLAB NFN configuration 3-6-8-8-1 [5].

On the created NFN, studies were conducted on the average learning error from the number of its inputs, from the number of terms of input variables, from the number of training examples using the following learning optimization methods: Backpropa (error backpropagation method) and Hybrid (hybrid method that combines the method of reverse error propagation with the least squares method).

### 2.3 Network traffic intensity forecasting

The analysis of the intensity of network traffic in the ITS of railway transport was performed, the Hirst indicator was approximately $H$=0.96 (at $a=\pi/2$; $N$=96). As an example, Figure 4 shows a fragment of the time series of network input traffic intensity in the ITS of the Prydniprovska railway (the interval between adjacent values is five minutes, the fragment length is 8 hours). The corresponding interval was chosen due to the fact that in modern ITS the recommended monitoring interval is five minutes.

 https://doi.org/10.1051/mateccconf/2023???02003

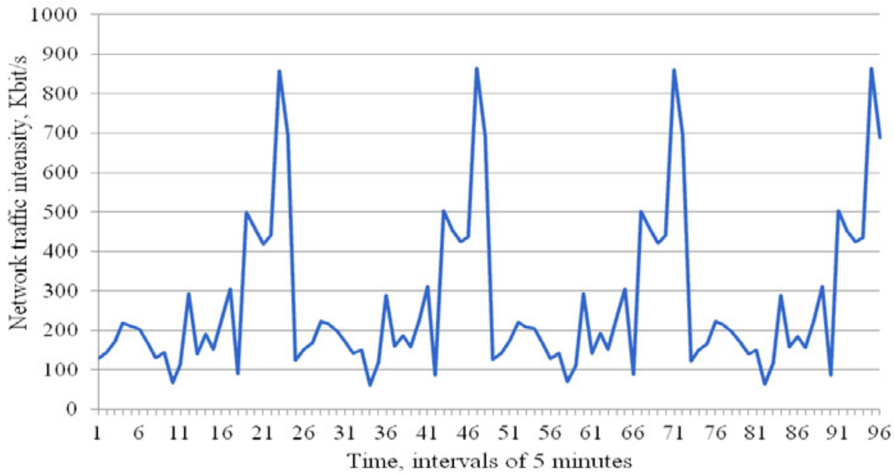


**Fig. 4.** Fragment of network (incoming) traffic in the ITS of railway transport.

For time (aggregated) series of network traffic intensity (incoming and outgoing) in the ITS of railway transport, some anomalous points have been adjusted, for the diagnosis of which the Irvine criterion is used; the anomalous values are replaced by the average of two adjacent (non-abnormal) values. As an example, Figure 5 shows a fragment of the time series of the intensity of aggregated network input traffic in the ITS of the Prydniprovska railway

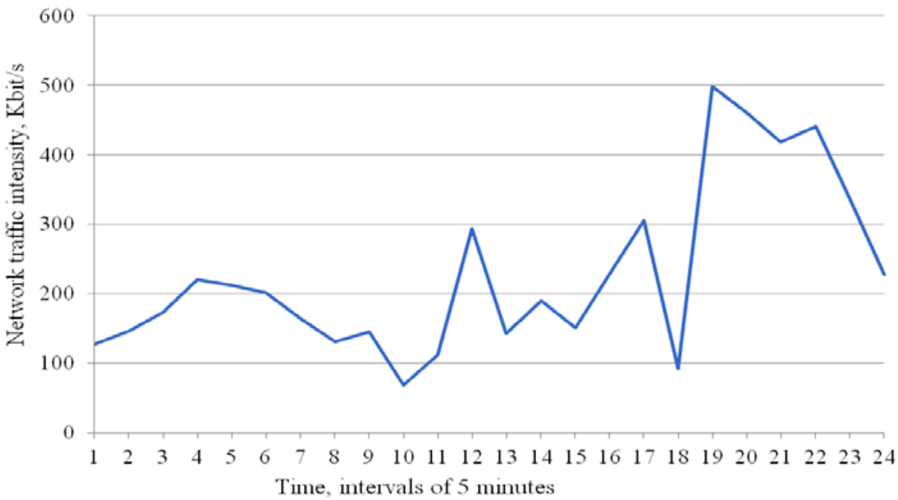


**Fig. 5.** Fragment of aggregated network (incoming) traffic in the ITS of railway transport.

Forecasting (one step ahead; five minutes are taken as a step) of network traffic intensity in the ITS of railway transport is performed by means of NN, the input of which is fed X(t-3), X(t-2), X(t-1), X(t) – network traffic intensity at times t-3, t-2, t-1, t, respectively; as the resulting characteristic X(t+1) – network traffic intensity at time t+1.

The study of NN configuration 4-2-16-1 (Fig. 6) was conducted, where 4 – number of input neurons; 2 – number of hidden layers; 12+4=16 – number of hidden neurons; 1 – the number of resulting neurons built in the MatLAB system. As a activation function of neurons of the first hidden layer is taken hyperbolic tangent, of the second hidden layer – sigmoid (logistic) function, resulting layer – linear function.

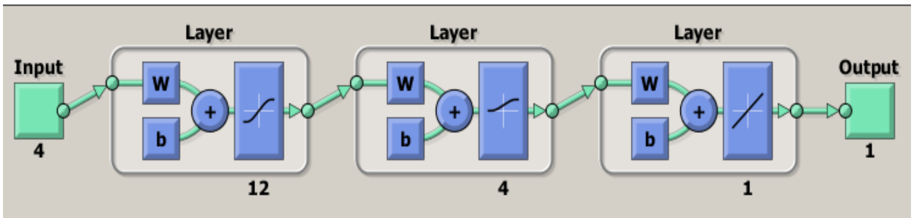


**Fig. 6.** Created in MatLAB system NN configuration 4-2-16-1.

On the NN configuration X-2-16-1, studies of mean square error (MSE) and the number of learning epochs with different numbers of inputs according to different learning algorithms were carried out (Fig. 7-8).

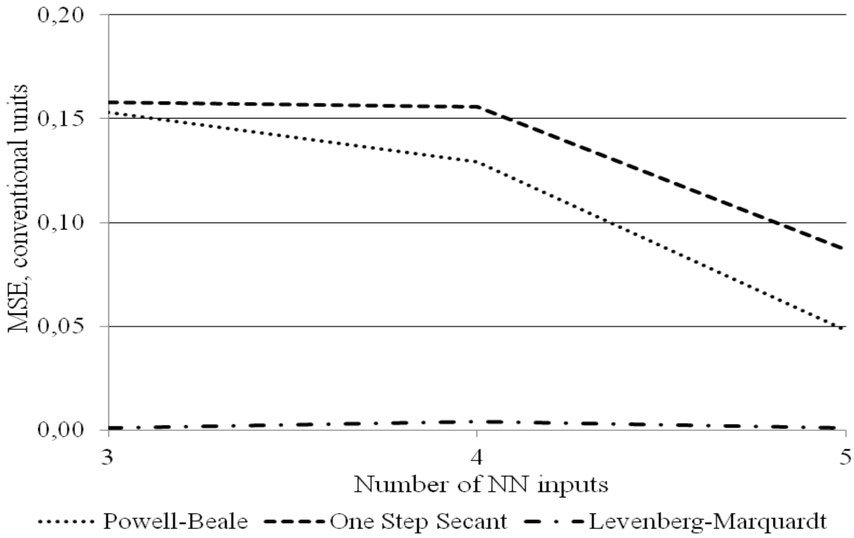


**Fig. 7.** The MSE NN configuration X-2-16-1 according to different training algorithms.

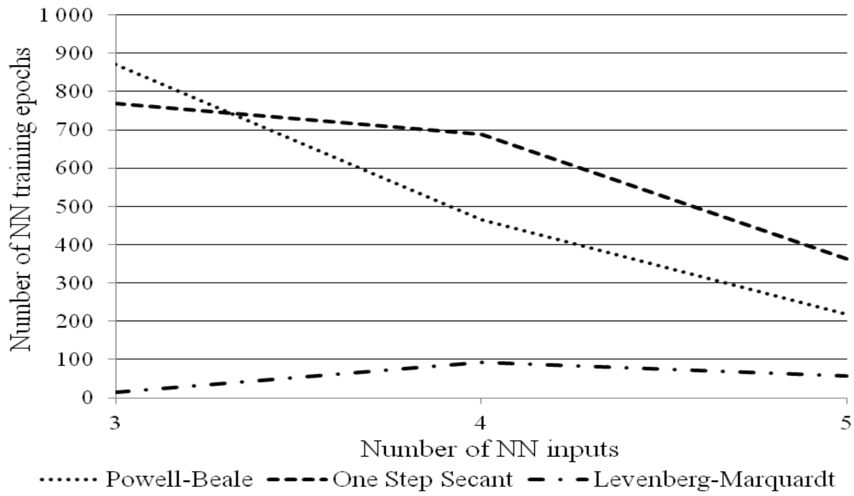


**Fig. 8.** The training time NN configuration X-2-16-1 according to different algorithms.

From Fig. 8-9 shows that the smallest values of the error and training time of the NN configuration X-2-16-1 were obtained according to the Levenberg-Marquardt algorithm, while it is enough to take 4 as the immersion depth.

On the NN configuration 4-2-X-1, studies of the MSE error and the number of learning epochs with different numbers of hidden neurons according to different learning algorithms were carried out (Fig. 9-10).

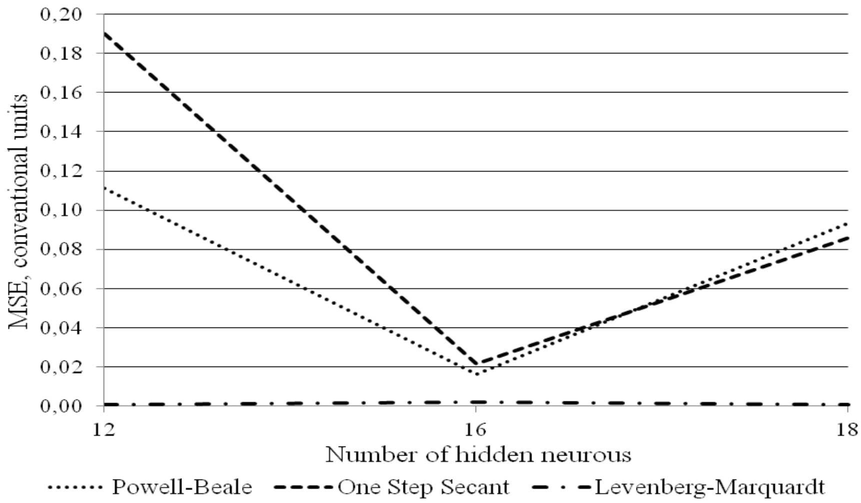


**Fig. 9.** The MSE NN configuration 4-2-X-1 according to different learning algorithms.

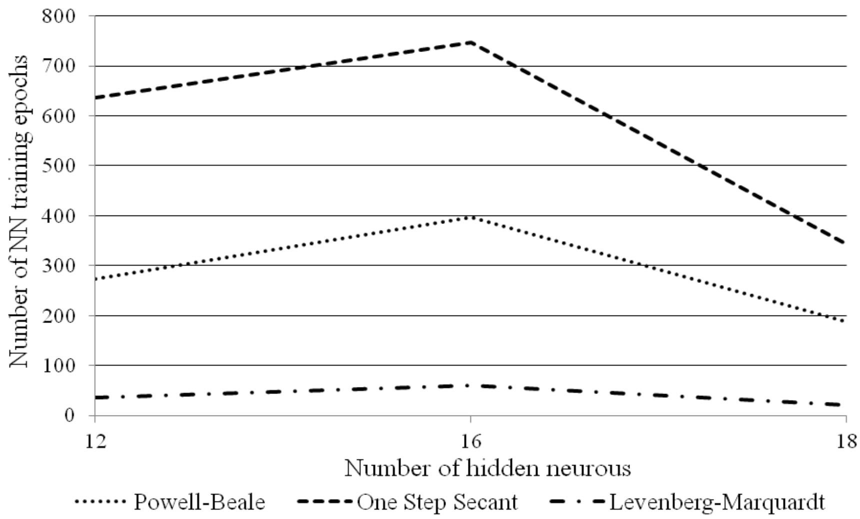


**Fig. 10.** The training time NN configuration 4-2-X-1 according to different algorithms.

From Fig. 9-10 it can be seen that the smallest values of the error and training time of the 4-2-X-1 configuration were obtained according to the Levenberg-Marquardt algorithm, while it is enough to have 16 hidden neurons.

## 3 Discussion

A long-term forecast is made in the ITS of railway transport using NFN, the input of which is supplied with the volume of network traffic for the previous three days, while the values of MAPE (Mean Absolute Persentage Error) were 6.9 % and 7.7 % for incoming and outgoing network traffic, respectively.

A short-term forecast (one step ahead; five minutes is taken as a forecasting step) of network traffic intensity in rail transport ITS based on the use of NN (immersion depth - 4) was made, while MAPE values were 2.43 % and 1.30 % for incoming and outgoing network traffic, respectively.

When executing long-term and short-term forecasts, MAPE values of less than 10 % and 5 %, respectively, were obtained, which indicates a sufficiently high quality of the performed forecasts of network traffic parameters (in particular volume and intensity) in the ITS of railway transport based on the use of created NFN and NN.

## 4 Further research

Until recently, the work of railway transport in Ukraine was an interaction of six railways: Lviv; Southwest; South; Odessa; Prydniprovska; Donetsk, each of which has implemented a corresponding ITS, which requires at the next stage the proposed methodology for forecasting network traffic parameters using real data on network traffic in the ITS of all relevant railways.

## 5 Conclusions

A methodology has been developed using tools for predicting network traffic parameters in the ITS of railway transport for long-term traffic forecast and short-term intensity forecast based on the use and study of optimal parameters of NFN and NN, respectively, which is important in the tendency to increase the nodes of ITS railway transport at the stage of deepening the integration of the national transport network into the Trans-European Transport Network, since at the present stage the maximum volume of network traffic already corresponds to a network load of more than 80 %.

## References

1. S. Chabaa, A. Zeroual. Identification and prediction of internet traffic using artificial neural networks. *Intelligent Learning Systems & Applications*. 147-155 (2010)
2. M. Crovella, A. Bestravos. Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes. *IEEE Transactions on Networking*. **Vol. 5**. No. 6. 835-836 (1997)
3. P. Cortez, M. Rio, M. Rocha, P. Sousa. Multi-scale Internet traffic forecasting using neural networks and time series methods. *Expert Systems*. **Vol. 29**. No. 2. 143-155 (2010)
4. S. Gowrishankar, P. Satyanarayana. A time series modeling and prediction of wireless network traffic. *International Journal of Interactive Mobile Technologies (IJIMT)*. **Vol. 4**. No. 1. 53-62 (2009)
5. V. Pakhomova. Network traffic forcasting in information-telecommunication system of Prydniprovsk railways based on neuro-fuzzy network. *Science and Transport Progress*. **No. 6 (66)**. 105-115 (2016)
6. I. Zhukovyts'kyy, V. Pakhomova, H. Domanskay, A. Nechaiev. Distribution of information flows in the advanced network of MPLS of railway transport by means of a neural model. *MATEC 294 (EOT-2019)*, 04007. 7 p. URL: https://doi.org/10.1051/matecconf/201929404007 (2019)
7. I. Zhukovyts'kyy, V. Skalozub, A. Ustenko. Intelligence zasobi upravlins parkami technical system zaliznychne transport: monography. *Dnipropetrs. natz. un-te zall. transp. im. akad. V. lazariana*. Dnipro: Standard-service. 190 p. (2018)

8. I. Zhukovyts'kyy, V. Skalozub, A. Ustenko, V. Klymenko. Formuvannya intelektualnykh informatsiynykh tekhnolohiy zaliznychnoho transportu na osnovi modeley analitychnykh serveriv ta ontolohichnykh system. *Informatsiyno-keruyuchi systemy na zaliznychnomu transporti*. Kharkiv. **No 6**. 3-11 (2018)

9. A. Kosolapov, I. Zhukovyts'kyy. Rezervy arkhitektury avtomatizirovannoy sistemy upravleniya gruzovymi perevozkami Ukrainskikh zheleznykh dorog. *Zalíznichniy transport Ukraíni*. **No 1**. Kiyev. 10-13 (2013)