APPLICATION OF THE NARX NEURAL NETWORK FOR PREDICTING A ONE-DIMENSIONAL TIME SERIES

Time series data analysis and forecasting tool for studying the data on the use of network traffic is very important to provide acceptable and good quality network services, including network monitoring, resource management, and threat detection. More and more, the behavior of network traffic is described by the theory of deterministic chaos. The traffic of a modern network has a complex structure, an uneven rate of packet arrival for service by network devices. Predicting network traffic is still an important task, as forecast data provide the necessary information to solve the problem of managing network flows. Numerous studies of actually measured data confirm that they are nonstationary and their structure is multicomponent. This paper presents modeling using Nonlinear Autoregression Exogenous (NARX) algorithm for predicting network traffic datasets. NARX is one of the models that can be used to demonstrate non-linear systems, especially in modeling time series datasets. In other words, they called the categories of dynamic feedback networks covering several layers of the network. An artificial neural network (ANN) was developed, trained and tested using the LM learning algorithm (Levenberg-Macwardt). The initial data for the prediction is the actual measured network traffic of the packet rate. As a result of the study of the initial data, the best value of the smallest mean-square error MSE (Mean Squared Error) was obtained with the epoch value equal to 18. As for the regression R, its output ANN values in relation to the target for training, validation and testing were 0.97743. 0.9638 and 0.94907, respectively, with an overall regression value of 0.97134, which ensures that all datasets match exactly. Experimental results (MSE, R) have proven the method's ability to accurately estimate and predict network traffic

Keywords: one-dimensional time series, NARX model, forecasting, neural network, nonlinear autoregression

T a n s a u l e S e r i k o v
PhD*

A i n u r Z h e t p i s b a y e v a
Corresponding author
PhD*
E-mail: aigulji@mail.ru

S h a r a f a t M i r z a k u l o v a
PhD, Assistant Professor
Department of Radio Engineering, Electronics and Telecommunications
Turan University
Satpayeva str., 16a, Almaty, Republic of Kazakhstan, 050013

K a i r a t b e k Z h e t p i s b a y e v
PhD
LLP «NTS Design»
Zhakyp Omarov str., 100, Nur-Sultan, Republic of Kazakhstan, 010000

Z h a n a r I b r a y e v a
Master, Senior Lecturer
Department of Radio Engineering, Electronics and Telecommunications
International University of Information Technology
Manas str., 34/1, Almaty, Republic of Kazakhstan, 050000

A r a i T o l e g e n o v a
PhD, Head of Department*

L y u d m i l a S o b o l e v a
Master, Senior Lecturer*

B e r i k Z h u m a z h a n o v
Doctoral Student
School of Engineering and Digital Sciences
Nazarbayev University
Kabanbay Batyr ave., 53, Nur-Sultan, Republic of Kazakhstan, 010000
*Department of Radio Engineering, Electronics and Telecommunications
S. Seifullin Kazakh Agro Technical University
Zhenis ave., 62, Nur-Sultan, Republic of Kazakhstan, 010011

1. Introduction

Further evolution of the telecommunication network based on infocommunication with packet switching caused a sharp increase in the amount of data associated with information flows from various human activities.

Today, topical areas in the Republic of Kazakhstan (RK) are data center management, cloud and cognitive technologies, IT security, etc. There are projects of electronic document management, electronic government, and the main investments are made in the development of cognitive technologies. Artificial intelligence collects various data and develops through machine learning, as well as based on previous requests or statements, provides various information. All these prerequisites contribute to the growth of network traffic, and taking into account the implementation of the concept of the Internet of Things, data volumes will increase more and more in the cognitive infocommunication network. Analysts of the Republic of Kazakhstan report that based on the analyses carried out, the growth in network traffic volume, taking into account the implementation of the concept of the Internet of Things, will receive an astronomical growth in data volumes. Therefore, today the task of network management, including based on predicted future data, is necessary to make the right decision. Identifying

and quantifying the components of a complex structure – the presence/absence of a trend, periodicity, random component is the main task of time series analysis. Identifying a nonlinear function and predicting it is an urgent task. Predictive data provide the necessary information to solve the problem of managing information flows in the network.

Practice has shown that the measured time series are non-stationary. The peculiarity of non-stationary time series is that they contain a trend, seasonal and cyclical fluctuations and noise.

Predictive data provide the necessary information to solve the problem of managing information flows in the network and allow you to prevent packet loss on a control basis.

This work is relevant, because in order to predict non-stationary time series in the traditional way, additional research and processing of initial data are required. As for the models of the artificial neural network, they do not require any additional information and preliminary processing of initial data. Analyzing network traffic is an important step in developing successful proactive congestion control schemes and identifying normal and malicious packets. These schemes aim to avoid network congestion due to network resource allocation relative to predicted traffic. The predictability of network traffic is an important benefit in many areas such as dynamic bandwidth allocation, network security, network planning, predictive congestion control and so on [1].

In addition, the task of detecting and preventing abuse on the network becomes very difficult due to the growing volume of traffic and the complexity of networks. Anomaly detection is one of the measures to counter online abuse. Significant deviations from normal traffic behavior can be exploited to detect an attack. The efficiency of anomaly detection is directly related to the accuracy of traffic prediction [2].

2. Literature review and problem statement

In [3], the authors described that, according to the estimates of forecasters, there are already more than a hundred forecasting methods, which raises the problem of choosing methods that would give adequate forecasts for processes or systems under study. At the same time, to obtain predictive data, they used a neural network to identify complex patterns of a time series.

The authors of [4] propose several recurrent neural network (RNN) architectures (standard RNN networks, Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU)) to solve the problem of predicting network traffic. Analyze the performance of these models on three important network traffic prediction problems: volume prediction, packet protocol prediction, and packet distribution prediction. As a result, we achieved a solution to the problem of predicting volumes in publicly available datasets such as the GEANT and Abilene networks. It can also be seen that RNN architectures show promising results in all three areas of network traffic prediction, far exceeding standard statistical prediction models.

A variety of deep learning models have been proposed for solving time series forecasting problems. Recurrent neural networks (RNNs) such as LSTM and GRU have made outstanding advances in sequential modeling. These models have become the obvious choice for time series modeling and a promising tool for handling irregular time series data. RNNs have the unique ability to adapt to efficiently use missing value patterns, time intervals and complex time relationships in irregular univariate and multivariate time series data [5, 6].

Irregular time series data is becoming more common with the growth in the number of multi-sensor systems and the continued use of unstructured manual data recording mechanisms. Irregular data and resulting missing values severely limit the ability to analyze and model data for classification and forecasting tasks. Traditional methods used to process time series data often introduce bias and make strong assumptions about the underlying data generation process, which can lead to poor model predictions [7].

The work [8] provides a complete overview of the IoT traffic forecasting model using classical time series and an artificial neural network. Real network traces are used to predict IoT traffic. Experimental results show that LSTM and FNN based prediction models are very sensitive and therefore can be used to provide better performance as a time sequence prediction model than traditional traffic prediction methods.

It is this approach that was used in [9]. Several methods were proposed for predicting the traffic matrix (TM) based on neural networks (NN) and predicting the traffic matrix from three points of view: directly predicting the total TM, predicting each origin-destination (OD) flow separately and predicting the overall TM combined with the correction of key elements. In addition to the prediction accuracy, various prediction methods are evaluated in terms of traffic management performance as well as prediction time. Experimental results show that forecasting methods based on recurrent neural networks (RNNs) can provide better forecasting accuracy than methods using convolutional neural networks (CNNs) and deep belief networks (DBNs). Predicting each OD stream with RNN models can further improve prediction accuracy as well as performance. However, it takes more time to predict each OD stream sequence. In contrast, overall TM prediction, combined with key element correction, can provide a trade-off between traffic control results and prediction overhead, which is more suitable for dynamic traffic control.

In [10], the authors present the results of a study on the use of echo state networks (ESN) to predict network traffic. ESN is a type of reservoir learning algorithm whose internal state consists of a pool of randomly connected neurons, similar to a reservoir. The experiments compare the results obtained using ESNs with those obtained using other network prediction methods on two datasets obtained from a well-known public network trace repository. The authors show that ESNs provide significantly less training and prediction times at similar levels of accuracy compared to other methods.

In [11], the authors provided a methodology for predicting future power consumption using nonlinear autoregressive (NAR) and nonlinear autoregressive neural networks with exogenous inputs (NARX), respectively. The results show that NAR and NARX neural networks are suitable for performing energy predictions, but also that exogenous data can help improve prediction accuracy. Each network model has its own advantages and disadvantages. NAR methods are simpler than NARX. However, the latter model allows additional information to be used that can improve the forecast.

In the NARX time series closed-loop neural networks with multistage prediction, IoT traffic time series prediction models have been proposed. The forecast was evaluated using reliability assessment functions such as MSE, SSE, MAE, and mean absolute percentage of error (MAPE) [12].

The authors in their research believe that NARX is one of the best methods in the neural network (for example, much faster in convergence) [13, 14].

In [15], an automated system for forecasting time series based on a neural network apparatus is implemented. All this

allows us to assert that it is advisable to conduct studies of a time series devoted to actually measured data.

Taking into account the above review of scientific research, for predicting network traffic, it is proved that the NARX method learns network traffic efficiently with an acceptable result of the forecast accuracy obtained [16]. For this, nonlinear autoregressive networks with exogenous inputs (NARX) were chosen, which augment multivariate time series using external information to improve the performance of time series forecasting.

3. The aim and objectives of the study

The aim of the work is to predict the actually measured time series. From a practical point of view, forecasting the actually measured time series (packet intensity) solves the problem of identifying traffic for creating devices for managing information flows.

To achieve the aim, the following objectives were set:
– substantiation of the choice of a forecasting model based on a neural network;
– description of the architecture and modeling of the neural network;
– building a forecast of future values of the time series.

4. Materials and methods of the research

The original data was obtained by means of a sniffer program – Wireshark. Within five hours, 278.557 packages were tracked, including: 158 TCP (Transmission Control Protocol); 493 ARP (Address Resolution Protocol); 25733 MPEG (Moving Picture Experts Group); 250242 UDP (User Datagram Protocol), etc.

A time series was created based on the collected data. This row represents the intensity of User Datagram Protocol packets. At the same time, their number was calculated for every 10 seconds. User Datagram Protocol is used to transmit real-time traffic that does not guarantee packet delivery, so it is important to consider its structure for forecasting (Fig. 1). A fragment of the measured network traffic is as follows:

22313 36.346551 HewlettP_db:77:81 Broadcast ARP Who_has_192.168.0.49? Tell_192.168.0.76 42 bytes
22314 36.355642 192.168.172.1 239.2.1.8 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22315 36.355646 192.168.172.20 239.2.2.60 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22316 36.355652 192.168.172.1 239.2.1.8 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22317 36.355655 192.168.172.20 239.2.2.60 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22318 36.355658 192.168.172.1 239.2.1.8 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22319 36.355661 192.168.172.1 239.2.1.8 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes

22320 36.355664 192.168.172.20 239.2.2.60 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22321 36.355667 192.168.172.1 239.2.1.8 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22322 36.355672 D-Link_4e:ea:bf Broadcast ARP Who_has_192.168.0.76? Tell_192.168.0.1 60 bytes
22323 36.355674 Vmware_9f:38:18 HewlettP_db:77:81 ARP 192.168.0.49_is_at_00:50:56:9f:38:18 60 bytes
22324 36.356075 HewlettP_db:77:81 D-Link_4e:ea:bf ARP 192.168.0.76_is_at_18:a9:05:db:77:81 42 bytes
22325 36.356111 192.168.0.76 192.168.0.49 TCP 49362_>_icslap_[SYN]_Seq=0_Win=8192_Len=0_MSS=1460_WS=2 66 bytes
22326 36.404231 192.168.172.20 239.2.2.60 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22327 36.404236 192.168.172.20 239.2.2.60 UDP Source_port:_bre Destination_port:_cisco-sccp 1358 bytes
22328 36.404239 192.168.0.49 192.168.0.76 TCP icslap_>_49362_[SYN,_ACK]_Seq=0_Ack=1_Win=65535_Len=0_MSS=1460_WS=0 66 bytes
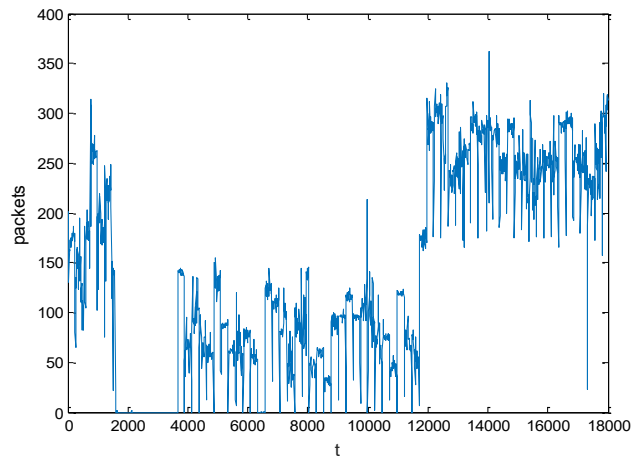

Fig. 1. User Datagram Protocol packet arrival rate

Program code was created for predicting the initial series (Fig. 2).

```
clc
clear all
close all

Data_Inputs=xlsread('Time series.xlsx'); % Import file
    %The training data sample are randmonized by using the function'randperm'
    Shuffling_Inputs=Data_Inputs(randperm(end),1:2); % integers (training sample)
    Training_Set=Data_Inputs(1:end,1);%specific training set
    Target_Set=Data_Inputs(1:end,2); %specific target set
    Input=Training_Set'; %Convert to row
    Target=Target_Set'; %Convert to row
    X = con2seq(Input); %Convert to cell
    T = con2seq(Target); %Convert to cell
    %%2. Data preparation
    N = 365; % Multi-step ahead prediction
    % Input and target series are divided in two groups of data:
    % 1st group: used to train the network
    inputSeries  = X(1:end-N);
    targetSeries = T(1:end-N);
    inputSeriesVal  = X(end-N+1:end);
    targetSeriesVal = T(end-N+1:end);
    % Create a Nonlinear Autoregressive Network with External Input
    delay = 2;
    inputDelays = 1:2;
    feedbackDelays = 1:2;
    hiddenLayerSize = 79;
```

Fig. 2. A fragment of program code for predicting the source series

MATLAB R2014a (8.3.0.532) was used as a research tool, which interactively performs analysis, modeling and forecasting of a time series.

5. Results of the research

5. 1. Justification of the choice of a forecasting model based on a neural network

Since 2007, a new generation packet-switched network has been deployed in Kazakhstan. It is currently developing in the direction of the cognitive infocommunication network. In the course of its evolutionary development, the volumes of transmitted information, the types of its presentation, the methods of transmission and storage, the number of sources and consumers, the distribution between users, the requirements for timeliness and reliability (quality) change.

Modern studies of real traffic show that it is necessary to take into account the «explosive» nature of traffic (pulsating over a wide range of time scales), which is combined with the Triple Play and Quadra Play service packages, etc. This fractal property implies a significant impact on the performance of network devices associated with a non-uniform packet arrival rate [17].

Based on the above, it can be seen that the measured network traffic is not stationary. And classical forecasting methods are not suitable for forecasting such series. Nonlinear identification with prediction is possible using neural network prediction. At the same time, neural networks and deep learning cope well with such tasks.

A non-stationary series always has a tendency, which is characterized by non-random factors in the processes represented by this time series. A time series is called non-stationary if its characteristics (mean value, variance and autocorrelation function) depend on time [18].

The importance of network traffic prediction is relevant in the following cases:
– when managing a network:
– detecting network congestion;
– data flow control.

In this case, the selected model for predicting network traffic is important. Practice shows that modern network traffic is not stationary. A functioning telecommunications network must be ready to serve a huge number of devices in accordance with the concepts of the Internet of Things (IoT) and Industrial Internet of Things (IIoT). It is about interchange through a network of sensors, actuators, controllers and human-machine interfaces.

When forecasting the time series, one takes into account that the series values in the past contain information about its future behavior. The stronger the future depends on the past, the better: it is possible to build a forecast – this is a key feature. The most common methods today are: methods of temporary extrapolation intersecting with autoregressive forecasting models, exponential smoothing models, neural network forecasting methods, etc.

In the present work, a forecast is made based on ANN (artificial neural network) associated with their universal approximating and predicting capabilities.

Back in 1943, foreign scientists introduced the concept of ANN and proposed a formal model of an artificial neuron.

Further in 1949, scientists described the basic principles of the network and created the first algorithm for training neurons [19]. And today, ANN is used as a model, mean and tool for approximating multidimensional functions, predicting and diagnosing processes, for searching by associations and searching for patterns in data arrays, for adaptive control and statistical analysis, for identifying and recognizing patterns, etc.

5. 2. Description of the architecture and modeling of the neural network

ANNs are mathematical models (software or hardware implementations), which are built on the principle of the organization and functioning of biological neural networks – networks of nerve cells of a living organism.

ANN is a set of artificial neurons connected to each other.

The work uses one of two types of ANN architecture – a network with direct signal propagation, a NARX nonlinear autoregression network, in which nonlinear coefficients are introduced into the difference autoregressive equation for the implementation of dynamic nonlinear systems using neural networks. At the same time, a network of this type is multilayer with direct signal transmission and output feedback, and the input signal is passed through a vector of time delays.

NARX networks come in two flavors:
– parallel;
– serial-parallel.
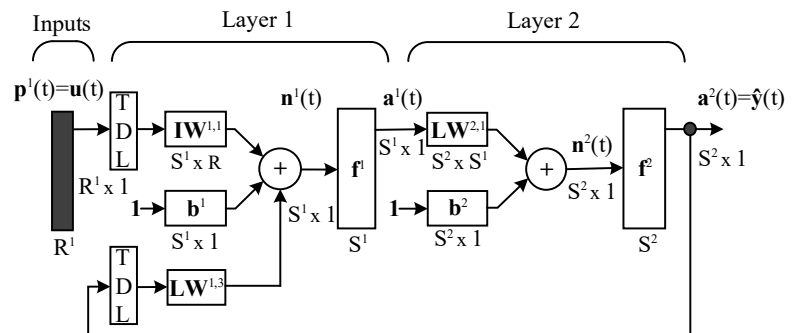
The work uses a parallel version (Fig. 3).

Fig. 3. NARX network architecture

The most common area of application of NARX networks is forecasting systems, that is, predicting the values of the output signal of the system under study based on the results of previous measurements.

One of the simplest dynamic neurons is the Hopfield neuron. As for the NARX non-linear autoregression network, it uses a modified Hopfield neuron, whose state is determined by a more distant background. The dynamics of the NARX model is described as follows:

$$y(n+1) = F\begin{pmatrix} y(n),...,y(n-q+1), \\ u(n),...,u(n-q+1) \end{pmatrix}, \tag{1}$$

where F – some nonlinear function of its arguments, which is approximated during the training; q – unit delay.

The NARX network is a multilayer network with direct signal transmission and output feedback, the output of which is passed through a time delay vector. As a function of activation, a sigmoid function is used.

Each network layer converts the input feature space into some other space, possibly with a different dimension. Such

a nonlinear transformation occurs until the classes turn out to be linearly separable neurons of the output layer. All layers of the ANN, except for input and output, give the network the ability to simulate nonlinear phenomena.

5. 3. Building a forecast of future values of a time series

For training the neural network, a nonlinear optimization method – the Levenberg-Macwardt algorithm was used, which is based on achieving the lowest mean square error and the mean square error (MSE) is used to evaluate the performance of the neural network.

Input (source data) and target data vectors are randomly divided into the following three sets:

– 70 % of vectors are used for training;

– 15 % vectors – to verify the reliability of the results and to avoid network overtraining;

– 15 % used for independent network testing.

Fig. 4 shows the nntraintool window, which allows you to get various graphs describing different processes and quality of training the NARX network. At the top of the nntraintool window is a model of a nonlinear autoregressive neural network with exogenous inputs. In this case, the forecast is made on the basis of the previous values of the studied data and exogenous input signals. This network has one Hidden layer with 52 neurons and one Output layer with one neuron.
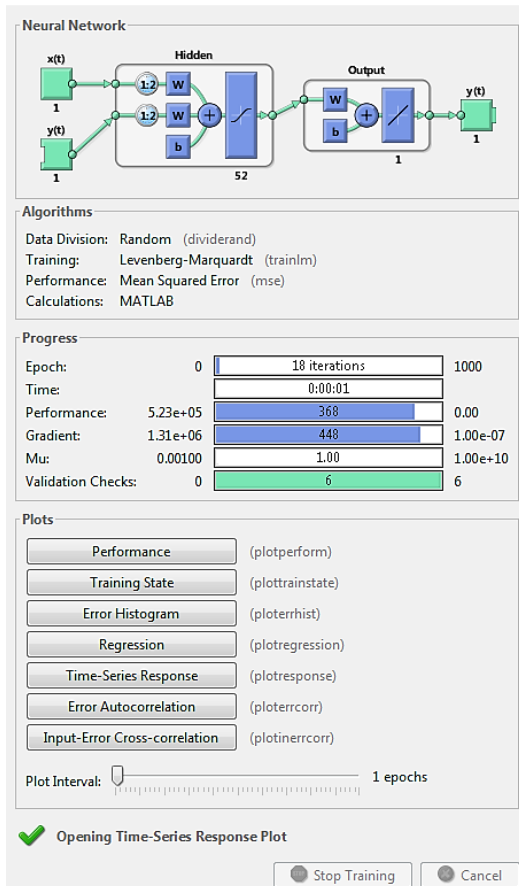

Fig. 4. NARX network learning outcomes

The principle of the processes of such learning is that a recurrent neural network «turns back in time» and is represented as a multilayer perceptron with a large number of layers, each layer corresponds to some past tact.

The learning process was completed when the target accuracy was achieved, while the studied network was trained in 18 epochs.

When the Performance button is activated, a graph of the mean square error (MSE) versus the number of epochs is obtained (Fig. 5). It can be seen that by the end of the learning process, the error becomes higher.
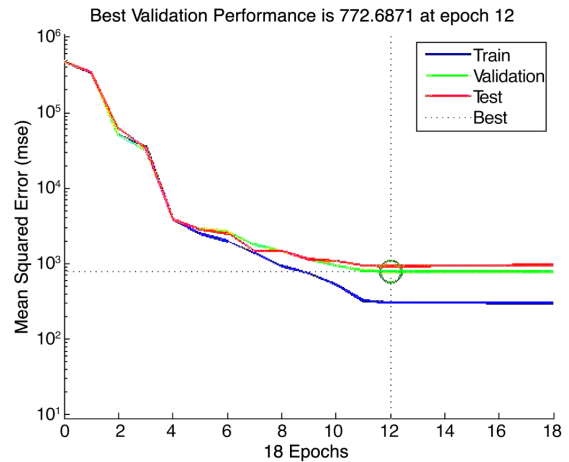

Fig. 5. Dependence of the standard error on the number of epochs

The above plot of the validation efficiency between the number of epochs and the root mean square error demonstrates that the green circle in the plot shows that training stopped when the validation error increased for twelve iterations. Both the validation set error and the test set error have similar characteristics, as indicated by the green and red lines. It is clear that with an increase in the number of epochs, the MSE decreases for all trained, validation, and test data with an insignificant change in slope. The best validation efficiency from the MSE perspective is 572.6426 in epoch 12.

As a result of the launch and training of ANNs, informative graphics were obtained:

– train – describes the training settings. At the same time, the value of the epoch (cycle) is 18, which the network shows, is represented by the introduction of a new adjustment pattern (backpropagation), coupled with a direct activation flow, occurred under the adjustment of the ANN weights with the total mean square error. A further increase in the epoch value leads to a decrease in the mean square error of the MSE, that is, the training lasted 18 epochs;

– validation – the studied data after the value of the epoch equal to 12 does not change its values, that is, the best performance is taken from the epoch with the smallest validation error (dot circled);

– test – shows the best learning performance;

– best – the best performance.

Fig. 6 shows the ANS training tools, which additionally display various parameters for predicting progress. The characteristics of the learning state of the network at epoch 18, when the gradient is 447.8859. The network stops the training session because its generalization stops improving.

A histogram of output errors of a nonlinear autoregressive network with exogenous NARX inputs is shown in Fig. 7.

For a detailed analysis of the approximation quality, a regression analysis of the network outputs has been obtained,

in which a linear regression of the network learning results is constructed on the three considered subsets and on the entire training set. For each result, the correlation coefficient R was calculated and graphs were plotted (Fig. 8).
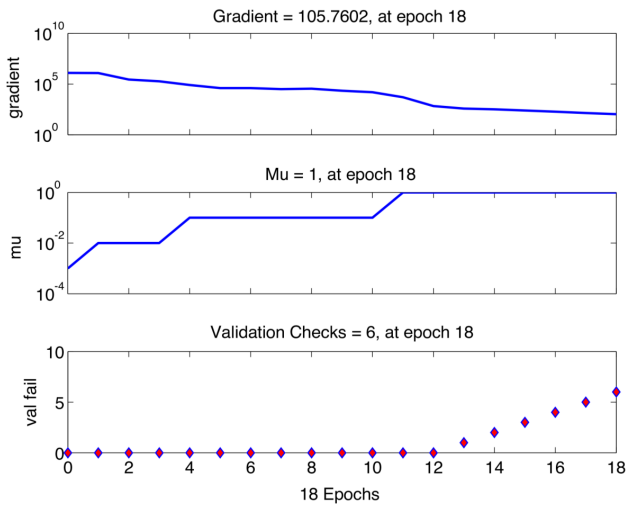

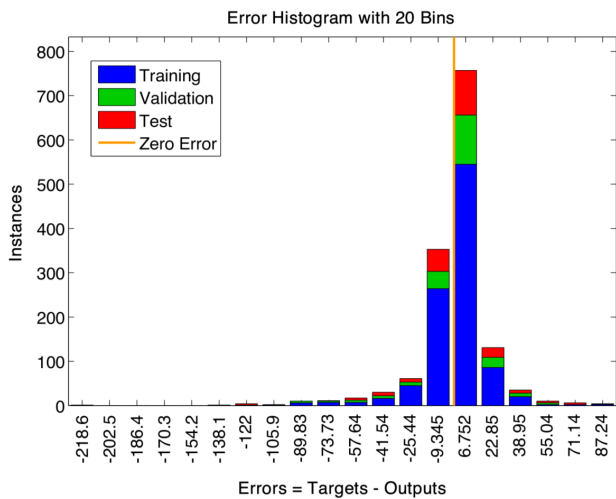Fig. 6. Artificial neural network training tools

Fig. 7. Histogram of output errors

When solving the problems of approximating dependencies, the quality of the network is determined by the differences between the known values of the function and the approximated network. Fig. 9 shows the approximation functions in the Neural Network Toolbox Matlab.

Fig. 10 shows a graph that displays an erroneous autocorrelation function.

Fig. 11 shows the cross-correlation function of the input error, which illustrates how errors correlate with the input sequence $x(t)$. Moreover, all correlations are within the confidence limit around zero.

Fig. 12 shows a graph of packet intensity prediction.

Is an error autocorrelation function that describes how prediction errors are related over time. An ideal prediction model contains only one non-zero autocorrelation function value, which should occur at zero latency, and is the root-mean-square error and indicates that the prediction errors were completely uncorrelated with each other (white noise). In this case, the autocorrelation function of the error has

a single non-zero value, which confirms that the forecast model is relatively perfect.
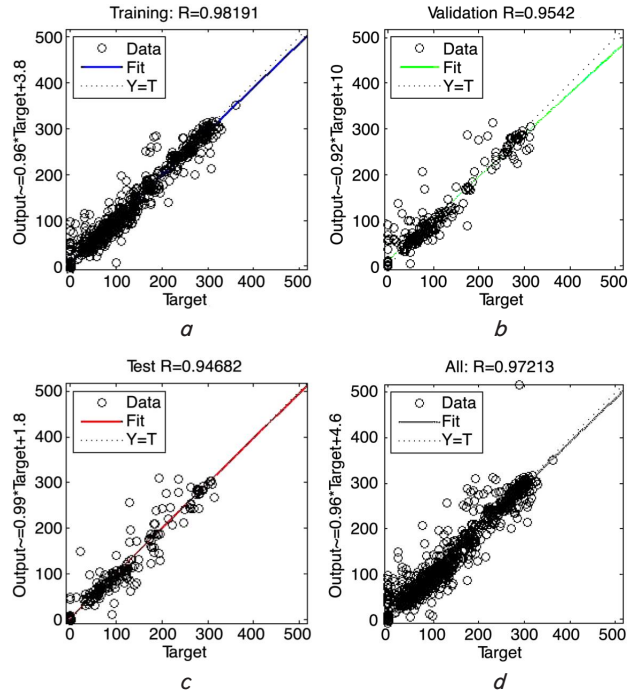

Fig. 8. Regression graphs:
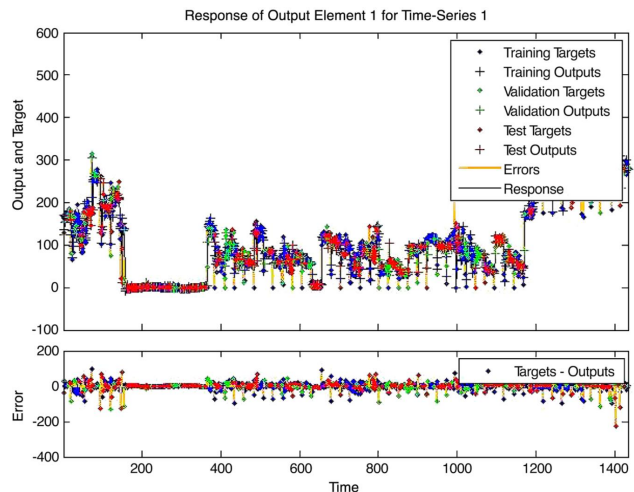a – training; b – validation; c – test; d – all: $R=0.97134$
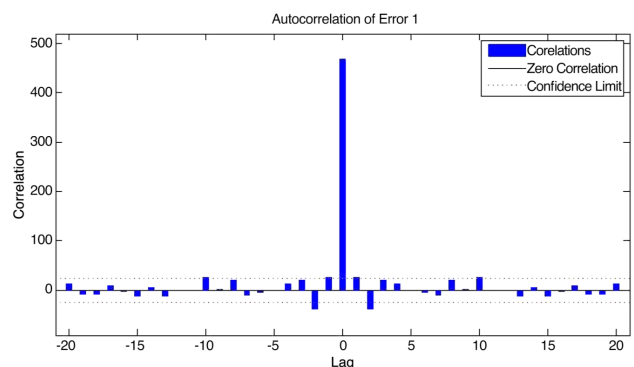

Fig. 9. Approximation function

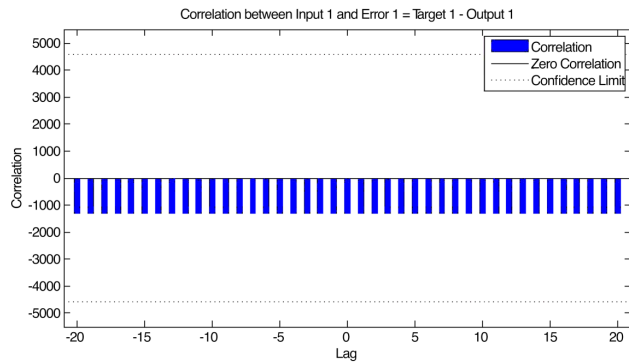Fig. 10. Erroneous autocorrelation function

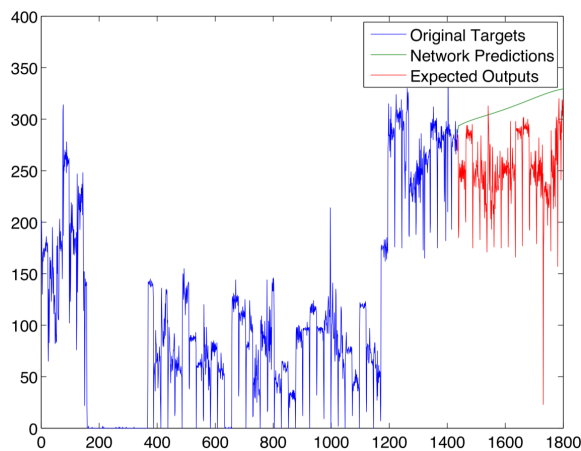Fig. 11. Cross-correlation function of the input error

Fig. 12. Prediction of packet intensity

6. Discussion of the results of one-dimensional time series forecasting

Fig. 8 shows the regression of the output (fitted) versus the target (actual) values fitted by the NARX model, the solid line on which represents the best fit linear regression line between the actual and the fitted network and shows how far from ideal. At the same time, regression analysis graphs were obtained with the value of the correlation coefficient R, which for the training set (Train) was 0.97743, for the validation parameter – 0.9638, for the best training performance (Test) – 0.9497, and for the entire training set – 0.97134. These data show good agreement.

After several experiments by the authors in [16] using three architectures, the results show that the NARX model gives the best result when using the parameters 189:31:94 or 60 %:10 %:30 %, in which the training performance value was <0.90. The overall performance of the NARX model is above 0.90, and between 0.86 and 0.90. In other words, the NARX architecture can be used as a model to predict a time series dataset.

In addition, the results show that the MSE and r values for the NARX architecture model are 0.006717 and 0.90764, respectively. Thus, the NARX architecture model with a configuration of 189:31:94 or 60 %:10 %:30 % can be used as an alternative model for predicting daily network traffic, because the MSE and r results obtained are very high.

In Fig. 9, 10, there is only one non-zero value of the autocorrelation function, which occurred at zero latency (standard error). This shows that the forecast model is relatively perfect (forecast errors do not correlate with each other, that is, they represent white noise). The rest of the values of the autocorrelation function are below the confidence interval and are approximately 95 % of the confidence interval near zero. Otherwise, you will have to retrain the network to improve the forecast.

The cross-correlation function plot shows how the errors correlate with the input sequence. Ideally, the forecasting model should have all correlations equal to zero. In this case, all correlations are within confidence limits below zero.

Based on the prediction rule, namely the error autocorrelation, cross-correlation of input errors and time series the reaction parameters that were observed. For an accurate forecast model, error autocorrelation function values there must be only one non-zero value and reach zero lag, this would mean that the errors in expectations were completely uncorrelated with each other [16]. The research results show that the NARX network with exogenous inputs provides a satisfactory forecast quality, but this model is relevant only for long-term network forecasting.

7. Conclusions

1. The choice of a forecasting model based on a neural network is primarily due to the fact that the measured traffic has an uneven intensity. This row is not stationary. Forecasting such series by traditional forecasting methods requires bringing the series to a stationary one using differentiation.

2. NARX model is a dynamic neural architecture commonly used to model nonlinear dynamic systems. When applied to time series forecasting, the NARX network is designed as a feedforward time-delay neural network.

3. The total root mean square error (MSE) is the root mean square error difference between the actual output values and the target values and is equal to 572.6429 with an epoch of 12. Regarding the regression values (R), they measure the correlation between the outputs. The difference in errors is very small, indicating a complete agreement between the values predicted by the NARX network and the experimentally measured values. The NARX network predictions show what is accurately modeled by the artificial neural network. Therefore, the artificial neural network is intended to be a useful tool for predicting packet intensity.

References

1. Oliveira, T. P., Barbar, J. S., Soares, A. S. (2016). Computer network traffic prediction: a comparison between traditional and deep learning neural networks. International Journal of Big Data Intelligence, 3 (1), 28. doi: https://doi.org/10.1504/ijbdi.2016.073903

2. Romirer-Maierhofer, P., Schiavone, M., D'Alconzo, A. (2015). Device-Specific Traffic Characterization for Root Cause Analysis in Cellular Networks. Lecture Notes in Computer Science, 64–78. doi: https://doi.org/10.1007/978-3-319-17172-2_5

3. Joshi, M., Hadi, T. H. (2015). A Review of Network Traffic Analysis and Prediction Techniques. arXiv.org. Available at: https://arxiv.org/ftp/arxiv/papers/1507/1507.05722.pdf

4. Ramakrishnan, N., Soni, T. (2018). Network Traffic Prediction Using Recurrent Neural Networks. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). doi: https://doi.org/10.1109/icmla.2018.00035

5. Cirstea, R.-G., Micu, D.-V., Muresan, G.-M., Guo, C., Yang, B. (2018). Correlated Time Series Forecasting using Multi-Task Deep Neural Networks. Proceedings of the 27th ACM International Conference on Information and Knowledge Management. doi: https://doi.org/10.1145/3269206.3269310

6. Tian, C., Ma, J., Zhang, C., Zhan, P. (2018). A Deep Neural Network Model for Short-Term Load Forecast Based on Long Short-Term Memory Network and Convolutional Neural Network. Energies, 11 (12), 3493. doi: https://doi.org/10.3390/en11123493

7. Weerakody, P. B., Wong, K. W., Wang, G., Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing, 441, 161–178. doi: https://doi.org/10.1016/j.neucom.2021.02.046

8. Khedkar, S. P, Canessane, R. A., Najafi, M. L. (2021). Prediction of Traffic Generated by IoT Devices Using Statistical Learning Time Series Algorithms. Wireless Communications and Mobile Computing, 2021, 1–12. doi: https://doi.org/10.1155/2021/5366222

9. Liu, Z., Wang, Z., Yin, X., Shi, X., Guo, Y., Tian, Y. (2019). Traffic Matrix Prediction Based on Deep Learning for Dynamic Traffic Engineering. 2019 IEEE Symposium on Computers and Communications (ISCC). doi: https://doi.org/10.1109/iscc47284.2019.8969631

10. Adeleke, O. A. (2019). Echo-State Networks for Network Traffic Prediction. 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). doi: https://doi.org/10.1109/iemcon.2019.8936255

11. Ruiz, L., Cuéllar, M., Calvo-Flores, M., Jiménez, M. (2016). An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings. Energies, 9 (9), 684. doi: https://doi.org/10.3390/en9090684

12. Abdellah, A. R., Mahmood, O. A. K., Paramonov, A., Koucheryavy, A. (2019). IoT traffic prediction using multi-step ahead prediction with neural network. 2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). doi: https://doi.org/10.1109/icumt48472.2019.8970675

13. Andalib, A., Atry, F. (2009). Multi-step ahead forecasts for electricity prices using NARX: A new approach, a critical analysis of one-step ahead forecasts. Energy Conversion and Management, 50 (3), 739–747. doi: https://doi.org/10.1016/j.enconman.2008.09.040

14. Li, X., Yang, C. (2014). Research and Application of Data Mining and NARX Neural Networks in Load Forecasting. International Journal of Database Theory and Application, 7 (2), 13–24. doi: https://doi.org/10.14257/ijdta.2014.7.2.02

15. Chandra, R. (2015). Competition and Collaboration in Cooperative Coevolution of Elman Recurrent Neural Networks for Time-Series Prediction. IEEE Transactions on Neural Networks and Learning Systems, 26 (12), 3123–3136. doi: https://doi.org/10.1109/tnnls.2015.2404823

16. Haviluddin, Alfred, R. (2015). Performance of modeling time series using nonlinear autoregressive with eXogenous input (NARX) in the network traffic forecasting. 2015 International Conference on Science in Information Technology (ICSITech). doi: https://doi.org/10.1109/icsitech.2015.7407797

17. Boussaada, Z., Curea, O., Remaci, A., Camblong, H., Mrabet Bellaaj, N. (2018). A Nonlinear Autoregressive Exogenous (NARX) Neural Network Model for the Prediction of the Daily Direct Solar Radiation. Energies, 11 (3), 620. doi: https://doi.org/10.3390/en11030620

18. Serikov, T., Zhetpisbayeva, A., Akhmediyarova, A., Mirzakulova, S., Kismanova, A., Tolegenova, A., Wójcik, W. (2021). City Backbone Network Traffic Forecasting. International Journal of Electronics and Telecommunications, 67 (3), 319–324. doi: https://doi.org/10.24425/ijet.2021.135983

19. McCulloch, W. S., Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5 (4), 115–133. doi: https://doi.org/10.1007/bf02478259