ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse183942020.pdf
https://doi.org/10.30534/ijatcse/2020/183942020

Neural Networks in Statistical Analysis and Monitoring System Telecommunication Network

Borisov V. V.[1], Borisova S. V.[2], Lupinovich S. N.[3], Vihor V. G.[4], Torubara A. N.[5]
[1]Khortytsia national training and rehabilitation Academy (Zaporozhye, Ukraine)
[2]Donbass state pedagogical University (Slavyansk, Ukraine)
[3]Khortytsia national training and rehabilitation Academy (Zaporozhye, Ukraine)
[4]National aviation University (Kiev, Ukraine)
[5]National University "Chernihiv Collegium named after T. G. Shevchenko" (Chernihiv, Ukraine)

ABSTRACT

A promising approach to organizing the processing of implicit forms of knowledge representation has been developed, which is based on the use of neural network structures technology. It is proved that neural networks and their analog models can be successfully used to solve the problem of approximating continuous functions of many variables and predicting processes occurring in telecommunications networks over time.
In recent years, there has been a significant increase in interest in research on the application of applied intelligent technologies, their development and implementation in the industrial and industrial spheres. Today, we can talk about the formation of a new scientific direction — the theory of integrated management of complex distributed communication networks. Currently, fundamental and applied work on the creation of intelligent control systems is actively carried out in many branches of technology. This was facilitated by a long period of theoretical research in the field of artificial intelligence theory, situational management and simulation modeling.

Key words : : information and telecommunication network, intelligent technologies, neuron, neural network, traffic.

1. INTRODUCTION

Today, management based on the analysis of external situations (events) remains one of the key ideas of intelligent management. Intellectual systems have recently become a fairly common commercial product, which finds a wide demand in the most diverse areas of engineering and scientific and technical spheres of activity [3-5].

In the management systems that have intellectual in General, this property is manifested in such aspects as management in the conditions of uncertainty, self-learning and adaptation. These are complex systems with a multi-level hierarchical structure, capable of forming decisions that are adequate to the situation that has developed. As noted in [1, 2], the entire history of the development of artificial intelligence is mainly associated with attempts to develop the most modern methods and controls in conditions of uncertainty.
One of the most promising approaches to organizing the processing of implicit forms of Z-nan representation is related to the use of neural network structures technology, which accumulates and reproduces the main functional features of biological prototypes. One of the most important features of neural network structures is their high performance, which is achieved due to the economy of parallelism of information processing in their hardware implementation [6].

2. MATERIALS AND METHODS

Analysis of the operation of telecommunications networks shows that at this stage of development, it is quite difficult to ensure their effective operation. The practice of using heterogeneous telecommunications systems and computer networks is associated with their lack of transparency, complexity, organizational limitations and specificity, which determines the need for a broader and scientifically based introduction of statistical methods for their analysis and monitoring based on open streaming information [4-7], especially when solving complex tasks and emergency situations [8].
The conducted analysis of the works [9, 10] shows that it is advisable and necessary to use intelligent technologies to solve the tasks set.
Also, in today's time, rapidly evolving technologies create a neural network to stroke net the quality of the prediction (classification).
The structure of the neural network. The statistical system of telecommunication network analysis uses networks with several ordered layers of neurons. In this case, there is no interaction between neurons that belong to the same layer [12]. Neurons of each layer receive data (signals) from neurons of the previous layer, process them, and transmit the result of processing to the next layer. The exception is the input layer neurons. The number of neurons in the input layer is equal to the number of variables selected for solving the forecast problem or classifying it, so that each neuron corresponds to one of the variables. Thus, the signals coming

to the input layer represent the values of these variables [13-17].

Signals at the output of the last (output) layer of neurons are the result of the operation of a neural network. Therefore, if a neural network is supposed to be used to classify objects into one of M groups, then the number of neurons in the output layer is equal to M.

Processing of signals by neurons of intermediate layers. The input of each neuron of any intermediate layer receives signals from all the neurons of the previous layer. Signal processing consists of weighted summation of the received signals first. If the total weighted sum exceeds a certain threshold, then the output signal of the neuron is equal to 1, otherwise-0.

Formalize the above statement. Let zj^{l} , ..., $zjnk^{-1}$ signals received on the input of the j -th neuron k of layer from nk–1 neurons of the previous layer, and $w()kj^{l}$, ... , $w(\ )kjnk^{-1}$ — weight used by the neuron to form the sum: sk()k = w()jllk z jl + ... $w()jnkk^{-1}z\ jnk^{-1}$. (1)

Let t()jl be the limit value. The output signal of a given neuron is defined as a value (), where the jump function θ( )x =1, if x >> 0; and 0, if x ≤ 0, that is, if s()jk >> t()jk .

In practice, the jump function θ()x is replaced by some function [16].

Since the input of each neuron in *the k* -th layer receives signals from all the neurons of the previous (*k* – 1) - th layer, the number of weighting factors and thresholds for processing input signals by all neurons is equal to $(n_k+1)n_{k-1}$, where n_k is the number of neurons in *the k* -th layer. The set of weight coefficients of all the neurons *of the k* -th layer forms a matrix of communication $W^{()k}$ between *the k* -th and (*k* – 1)m layers.

To create a neural network that can be used to classify multidimensional objects or to predict the values of an independent variable (in the case of regression analysis or time series prediction), which is especially important in the case of statistical analysis of a telecommunications network, it is necessary:

— set the network architecture, i.e. set the number of layers and the number of neurons in each of them — estimate the weight coefficients for all neurons in the network (weights in connection matrices $W^{(\ )k}$).

A neural network must contain at least two layers: input and output. The number of neurons in the input layer is determined by the number of variables used. If all variables are continuous quantitative variables, then the number of neurons is simply equal to the number of variables. If this is the nominal variable, then for each such variable, for example, a variable *in*, is given (*l* – 1) input neurons, where *l* — the number of levels (categories) of the variable *y* in *and* m of the neuron (these (*l* – 1) neurons) and is assigned the value 1 if the variable takes *the i*e valuation, and 0 otherwise [18-20].

Therefore, the number of neurons in the input layer is uniquely determined as soon as the active variables are selected to solve the classification, regression, or prediction problem.

The number of neurons in the output layer is determined by the type of problem being solved, when solving problems of classifying objects into one of *M* groups, the output layer contains *M* neurons. When solving a prediction (regression) problem, the number of neurons is equal to the number of dependent and numerous ones. The number of intermediate layers and the number of neurons in each of them is set by the researcher before the weighting step.

To estimate the weight coefficients in the statistical system of analysis of the telecommunications network, the procedures of mad optimization by the method of conjugate gradients are applied. To solve the problem of local minima, we use the generation of a certain number of starting points [19].

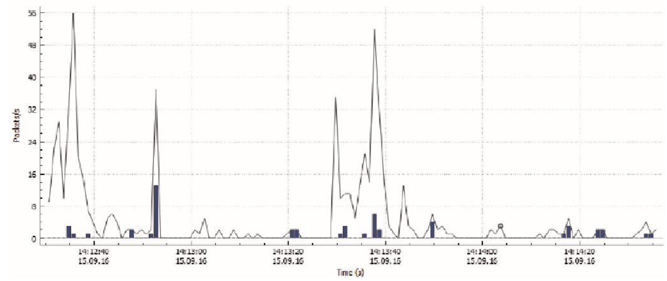


Figure 1: Independence of channel loading (packets/s) on time (tame(s)) between components of the DocFlow system

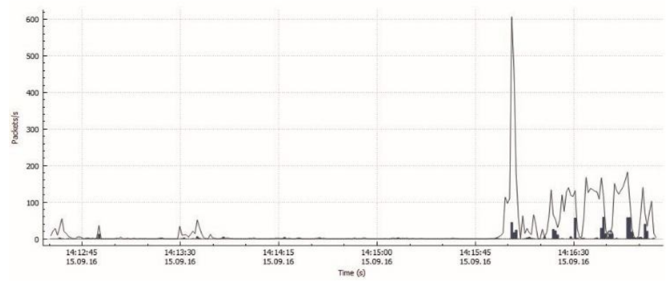
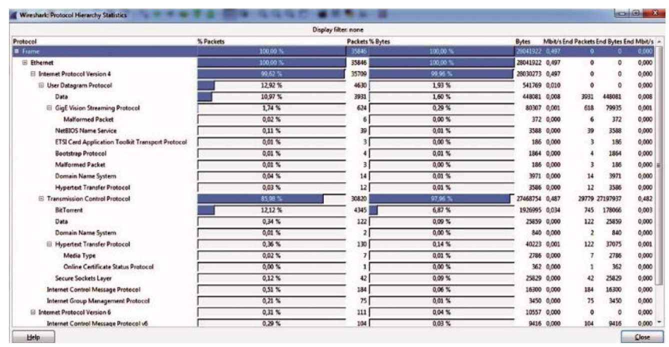


Figure 2: Dependence of channel loading (packets/s) on time (tame(s)) between SAP ERP system components


Wireshark: Protocol Hierarchy Statistics

Display filter: none

Protocol	% Packets	Packets	% Bytes	Bytes	Mbit/s	End Packets	End Bytes	End Mbit/s
Frame	100,00 %	35846	100,00 %	28041922	0,497	0	0	0,000
Ethernet	100,00 %	35846	100,00 %	28041922	0,497	0	0	0,000
Internet Protocol Version 4	99,62 %	35709	99,96 %	28030273	0,497	0	0	0,000
User Datagram Protocol	12,92 %	4630	1,93 %	541769	0,010	0	0	0,000
Data	10,97 %	3931	1,60 %	448081	0,008	3931	448081	0,008
GigE Vision Streaming Protocol	1,74 %	624	0,29 %	80307	0,001	618	79935	0,001
Malformed Packet	0,02 %	6	0,00 %	372	0,000	6	372	0,000
NetBIOS Name Service	0,11 %	39	0,01 %	3588	0,000	39	3588	0,000
ETSI Card Application Toolkit Transport Protocol	0,01 %	3	0,00 %	186	0,000	3	186	0,000
Bootstrap Protocol	0,01 %	4	0,01 %	1864	0,000	4	1864	0,000
Malformed Packet	0,01 %	3	0,00 %	186	0,000	3	186	0,000
Domain Name System	0,04 %	14	0,01 %	3971	0,000	14	3971	0,000
Hypertext Transfer Protocol	0,03 %	12	0,01 %	3586	0,000	12	3586	0,000
Transmission Control Protocol	85,98 %	30820	97,96 %	27468754	0,487	29779	27197937	0,482
BitTorrent	12,12 %	4345	6,87 %	1926995	0,034	745	178066	0,003
Data	0,34 %	122	0,09 %	25859	0,000	122	25859	0,000
Domain Name System	0,01 %	2	0,00 %	840	0,000	2	840	0,000
Hypertext Transfer Protocol	0,36 %	130	0,14 %	40223	0,001	122	37075	0,001
Media Type	0,02 %	7	0,01 %	2786	0,000	7	2786	0,000
Online Certificate Status Protocol	0,00 %	1	0,00 %	362	0,000	1	362	0,000
Secure Sockets Layer	0,12 %	42	0,09 %	25829	0,000	42	25829	0,000
Internet Control Message Protocol	0,51 %	184	0,06 %	16300	0,000	184	16300	0,000
Internet Group Management Protocol	0,21 %	75	0,01 %	3450	0,000	75	3450	0,000
Internet Protocol Version 6	0,31 %	111	0,04 %	10557	0,000	0	0	0,000
Internet Control Message Protocol v6	0,29 %	104	0,03 %	9416	0,000	104	9416	0,000

Help Close

Figure 3. Tree structure of protocols with an approximate percentage of total traffic

Figure 1 shows us independence of channel loading (packets/s) on time (tame(s)) between components of the DocFlow system, figure 2 shows us dependence of channel loading (packets/s) on time (tame(s)) between SAP ERP system components and figure 3 shows us the tree structure of protocols.

Of the above-mentioned characteristics of a telecommunications network, the most informative parameter can be singled out in particular — channel loading. The number of channels created directly affects the stability and reliability of communication channels and, accordingly, the quality of TCM functionality [21].

The developed procedures allow us to consider and analyze in more detail the dynamics of changes in information flows circulating in networks, and to determine the characteristic features of random sequences [22-25].

Of the above-mentioned characteristics of a telecommunications network, the most informative parameter can be singled out in particular — channel loading. The number of channels created directly affects the stability and reliability of communication channels and, accordingly, the quality of TCM functionality.

The sum of elements in column Pi'(t j), consisting of 0 and 1, gives the number of counter "zeroes" for the entire time period. The analysis of the distribution of units on the time axis allows us to establish numerical estimates of their distribution over certain time intervals Tk. To monitor the distribution of reset moments of the counter of the variable P ti(j) over the series Pi'(t j) a new table is formed, the elements of which record the moment when the "unit" is registered and the time interval between neighboring resets [26].

Using the above procedure convert the data from the storage type in a random sequence were obtained for other characteristics of the network traffic: download Kahn the crystals according to layers of protocols (interfaces), Fast Ethernet 1/0, IP Protocol, TCP Protocol, and a tree structure (Figure 4) protocols approximate percentage in the total traffic (number of packets, number of bytes).

On the basis of results, the features of statistical monitoring of the telecommunications network were determined:

— non-stationarity;

— heterogeneity;

— frequency (uneven loading of channels);

— complex periodic waveforms; - waveforms closer to din trapezoids with a pronounced "plateau" in the area of maximum loads;

— the amount of noise is higher at maximum downloads.

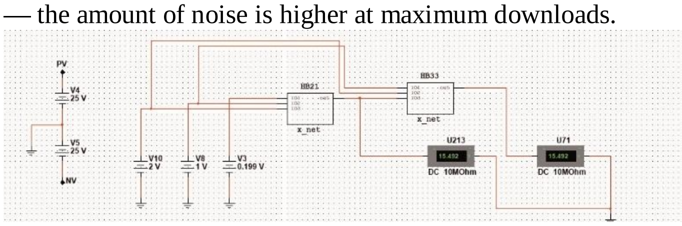


Figure 4: Scheme of parallel forecasting

The developed procedures allow us to consider and analyze in more detail the dynamics of changes in information flows circulating in networks, and to determine the characteristic features of random sequences.

3. CONCLUSION

When solving the tasks of analysis and monitoring of networks in the first place is considered the primary production of INFdeformations and solves the following problems of approximation of functions, prediction, optimization etc. To solve such problems, you need to use neural networks.

After reviewing the information on the work of the telecommunication network using a data transmission technology ATM 1/0, Fast Ethemet l/0, Fast Ethernet 4/0 found that the efficiency of the network depends on the following characteristics: the bandwidth of input and output (of bytes); the number of packets input and output; the number of errors in their Registration; CPU load (%); the amount of free processor memory and system I / o for routes

This may indicate that different weights need to be set for the y coordinate at different parts of the trajectory, meaning that the neural network needs to be trained again.

REFERENCES

1. Ameri, S, Fard, MJ, Chinnam, RB and Reddy, CK. 2016. **Survival Analysis Based Framework for Early Prediction of Student Dropouts**. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (CIKM), 16: 903–912. New York, NY, USA: ACM. DOI: https://doi.org/10.1145/2983323.2983351
2. Aulck, L, Aras, R, Li, L, Heureux, CL, Lu, P and West, J. 2017. **STEM-ming the Tide: Predicting STEM attrition using student transcript data**. In: Proceedings of ACM Knowledge Discovery and Data Mining Conference. Nova Scotia, Canada.
3. Aulck, L, Velagapudi, N, Blumenstock, J and West, J. 2016. **Predicting Student Dropout in Higher Education**. In: ICML Workshop on #Data4Good: Machine Learning in within the Open Polytechnic of New Zealand, relying Social Good Applications. New York, NY, USA.
4. Chen, Y, Chen, Q, Zhao, M, Boyer, S, Veeramachaneni, K and Qu, H. 2017. **DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction**. 2016 IEEE Conference on Visual Analytics Science and Technology, VAST 2016 – Proceedings, 111–120.
5. Fei, M and Yeung, D-Y. 2015. **Temporal Models for Predicting Student Dropout in Massive Open Online Courses.** 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 256–263. DOI: https://doi.org/10.1109/ICDMW.2015.174
6. Hu, Q and Rangwala, H. 2017. **Enriching Course-Specific Regression Models with Content Features for Grade Prediction**. In: Proceedings of ACM SIGKDD. Nova Scotia, Canada. DOI: https://doi.org/10.1109/DSAA.2017.74
7. Iam-On, N and Boongoen, T. 2017. **Generating descriptive model for student dropout**: A review of clustering approach. Human-centric Computing and Information Sciences, 7(1): 1. DOI: https://doi.org/10.1186/s13673-016-0083-0
8. Iqbal, Z, Qadir, J, Mian, AN and Kamiran, F. 2017. **Machine Learning Based Student Grade Prediction**:A Case Study
9. Jordan, MI and Mitchell, TM. 2015. **Machine learning: Trends, perspectives, and prospects**. Science, 349(6245): 255–260. DOI: https://doi.org/10.1126/science.aaa8415
10. Kumar, M, Singh, AJ and Handa, D. 2017. **Literature Survey on Educational Dropout Prediction**. I.J.

Education and Management Engineering, 2(March): 8–19. DOI: https://doi.org/10.5815/ijeme.2017.02.02

11. Liang, J, Li, C and Zheng, L. 2016. **Machine learning application in MOOCs: Dropout prediction**. ICCSE 2016 – 11th International Conference on Computer Science and Education (ICCSE), 52–57.
12. Młynarska, E, Greene, D and Cunningham, P. 2016. **Time series clustering of Moodle activity data**. CEUR Workshop Proceedings, 1751: 104–115.
13. Nunn, S, Avella, JT, Kanai, T and Kebritchi, M. 2016. **Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review**. Online Learning, 20(2): 13–29. DOI: https://doi.org/10.24059/olj.v20i2.790
14. Rovira, S, Puertas, E and Igual, L. 2017. **Data-driven system to predict academic grades and dropout**. PLOS ONE, 12(2): 1–21. DOI: https://doi.org/10.1371/journal.pone.0171207
15. Wang, W, Yu, H and Miao, C. 2017b. **Deep Model for Dropout Prediction in MOOCs**. Proceedings of the 2nd International Conference on Crowd Science and Engineering – ICCSE'17, 26–32. DOI: https://doi.org/10.1145/3126973.3126990
16. Xu, J, Moon, KH and van der Schaar, M. 2017. **A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs**. IEEE Journal of Selected Topics in Signal Processing, 11(5): 742–753. DOI: https://doi.org/10.1109/JSTSP.2017.2692560
17. Yang, D, Piergallini, M, Howley, I and Rose, C. 2014. **Forum Thread Recommendation for Massive Open Online Courses**. Proceedings of the 7th International Conference on Educational Data Mining (EDM), 257–260.
18. Yudelson, MV, Koedinger, KR and Gordon, GJ. 2013. **Individualized Bayesian Knowledge Tracing Models**.
19. Abramov, R. A. (2016). **Regional economic policy based on industrial sector clustering in the context of sustainable development**. Research Journal of Pharmaceutical, Biological and Chemical Sciences, 7(2), 2100–2106.
20. Abramov, R. A., Koshkin, A. P., Sokolov, M. S., & Surilov, M. N. (2018). **Transformation of the public administration system in the context of integration of the national innovation systems of the Union State**. Espacios, 39(14).
21. Belitskaya, M. (2018). **Ecologically adaptive receptions control the number of pests in the ecosystems of transformed at the forest reclamation**. World Ecology Journal, 8(2), 1-10. https://doi.org/https://doi.org/10.25726/NM.2018.2.2.001
22. Gribust, I. (2018). **Regulation of the state of plantings in the anthropogenically transformed territories: the principle of dendrological diversity**. World Ecology Journal, 8(2), 11-21 https://doi.org/https://doi.org/10.25726/NM.2018.2.2.002
23. Kondrasheva, N. K., Eremeeva, A. M., & Nelkenbaum, K. S. (2018). **Development of domestic technologies of producing high quality clean diesel fuel**. Izvestiya vysshikh uchebnykh zavedenii khimiya khimicheskaya tekhnologiya, 61(9-10), 76-82. https://doi.org/10.6060/ivkkt.20186109-10.5651
24. Burov A.G., Agüero D. **Implementation of the Principles of Innovative Entrepreneurship in the Field**. Academy of Entrepreneurship Journal. 2019. Vol. 25. Issue 1S. pp. 1-5.
25. Shashkova, A., Verlaine, M., & Kudryashova, E. (2020). **On modifications to the constitution of the russian federation in 2020. Russian Law Journal**, 8(1), 60–83. https://doi.org/10.17589/2309-8678-2020-8-1-60-83
26. Rajaallah, E. M., El Hissi, Y., & Haqiq, A. (2019). **Information system governance in moroccan universities**: Case of student affairs. International Journal of Advanced Trends in Computer Science and Engineering, 8(1). https://doi.org/10.30534/ijatcse/2019/2281.12019
27. Rajamanickam, L., Saat, N. A. B. M., & Daud, S. N. B. (2019). **Software testing: The generation tools.** International Journal of Advanced Trends in Computer Science and Engineering, 8(2), 231–234. https://doi.org/10.30534/ijatcse/2019/20822019