Advances in Intelligent Systems Research, volume 174
Proceedings of the 8th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2020)

Threats Complex Distributed Systems Parrying Based on their Development Prognostication

Evgeny Palchevsky
Department of Geographic Information Systems
Ufa State Aviation Technical University
Ufa, Russia
teelxp@inbox.ru

Olga Khristodulo*
Department of Geographic Information Systems
Ufa State Aviation Technical University
Ufa, Russia
o-hristodulo@mail.ru

Sergey Pavlov
Department of Geographic Information Systems
Ufa State Aviation Technical University
Ufa, Russia
psvgis@mail.ru

***Abstract*—This article proposes a method for countering threats in complex distributed systems based on predicting their development. Initially, the relevance of the topic under study is justified: it seems promising to use approaches that have found application in solving problems associated with the management of complex systems. Further, an artificial neural network is proposed for forecasting: its structure is shown, as well as a mathematical model of self-learning, which allows achieving more accurate (with less error) results in the framework of threat prediction (in this case, the level of water rise at gauging stations) in complex distributed systems. Testing was carried out, the purpose of which is to confirm the effectiveness of the proposed solution in the framework of forecasting threats in complex distributed systems: the error of the predicted values from real varies from 9% to 10%, which allows us to predict the flood situation in a few days.**

Keywords—Water level forecasting, flood situation, neural networks, neural network for forecasting

I. INTRODUCTION

In the modern world, complex systems include qualitatively new systems, including physical (including technical), biological (including people and environmental objects), and digital (including computers, software, data) components [1-3]. In the context of this article, by complex distributed systems we will mean a set of technical objects (for example, oil or gas pipelines, potentially dangerous objects, etc.) distributed in the study area and vulnerable to a certain threat.

Today, for such systems there are many threats that pose a danger to them, the consequence of which can be substantial material damage. One of these threats, for example, for the Republic of Bashkortostan, is the spring flood, the consequences of which often threaten complex distributed systems, thereby making them vulnerable to flooding and flooding. To counter such threats, it is proposed to forecast the level of water rise in advance in order to plan further measures to prevent the negative impact of the flood on complex distributed systems falling into the zone of its spread for further assessment of the vulnerability of these systems.

This problem is addressed by many scientists in various fields of science [4-28]. However, due to the lack of work containing a description of the method for early detection of threats in order to counter them in complex distributed systems, it seems promising to use approaches that have found application in solving problems related to the management of complex systems: tasks of monitoring and predicting the state and parry threats; teaching methodology with and without a teacher of neural networks of various generations.

Thus, it seems important and necessary to develop a method for early detection of threats in order to counter them in complex distributed systems using the example of early prediction of the level of water rise during the flood period, which will allow the special services to give the necessary time to carry out flood control measures to prepare for the protection of the above complex distributed systems.

II. DEVELOPMENT OF A RECURRENT NEURAL NETWORK FOR FORECASTING THE WATER LEVEL AT HYDROPOSTS

The main characteristic of the possible impact of the flood situation in a certain territory (for example, the Republic of Bashkortostan) can serve as A – the level of water rise in water bodies, which is measured daily at n-posts of the hydrometer, so A^{n}_{ijt}– i the measured value of this level at *n* post. To counter the threat, it is necessary at some fixed point in time *t* to determine the value of the water level at all posts, that is, to determine $A^{n}_{ij(t+1)}, j = \overrightarrow{1,n}$.

Accordingly, to solve the problem, a recurrent self-learning was implemented (the scheme of work with the training stages is shown in fig. 4) neural network.

It is worth noting that a recurrent neural network is one of the types of high-speed (second generation) ANNs, in which the connections between neurons together form a directed output sequence of stream-processed information. The structure of this neural network is shown in fig. 1.

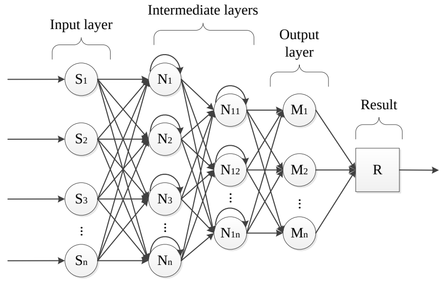

Fig. 1. The structure of RANN

The proposed structure of a recurrent neural network for predicting water levels is shown in fig. 2. Initially (fig. 2), data (gauging stations codes, dates, water levels) enter the input layer for further processing in the RANN intermediate link (layers) for training purposes. Teaching without a teacher is based on the integration of the Hebb methods and the back propagation of the error for the shortest and more accurate forecast of the water level at gauging stations. At the exit, we get a ready forecast for a given period of days.

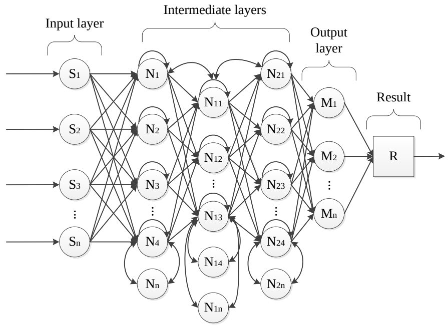

Fig. 2. RANN structure for predicting water levels

The neuron itself has a sigmoidal activation function (fig. 3) with a range of values (0, 1), described by the equation:

$$\psi = \frac{1}{1+e^{-Q}}. \quad (3)$$

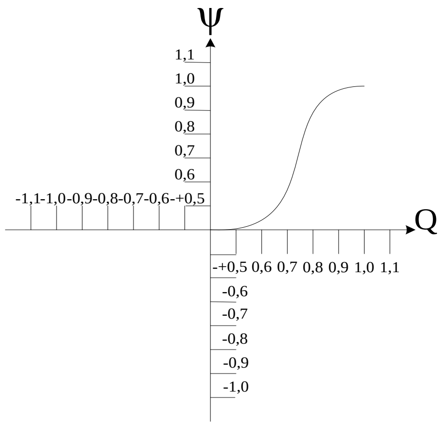

Fig. 3. The activation function in the developed neural network

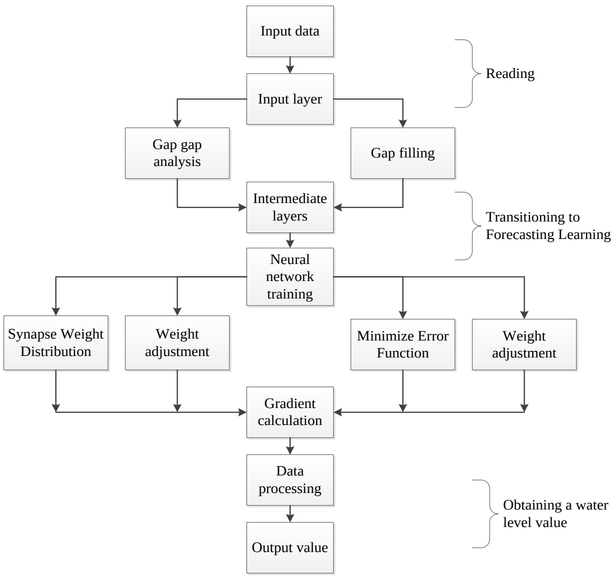

Fig. 4. Scheme of work of RANN with training stages

From a mathematical point of view, this network is characterized by the union of signals in the input layer (node, vector) $C(i_1)$. In general, the system of input, intermediate, and output layers is expressed as follows:

$$C(i_1+1) = f((C(i)) \cdot C((i_1-1)) \cdot C(i_1-(N-1)) \cdot \\ \cdot C_1(i_1-1) \cdot C_1(i_1-P)), \quad (1)$$

where $N-1$ – is the delay of the input signal (quantity), P – is the delay of the output signal (quantity), i_1 – the number of neurons in the intermediate layers. Thus, in this case, a recurrent neural network can be characterized by a set of numbers $\{N, P, I_1\}$. Consequently, the vector input to the network $C(i_1)$ has the following form:

$$C(i_1) = [1 \cdot (C \cdot (i_1)) \cdot (C \cdot (i_1-1)) \cdot \ldots \cdot (C \cdot (i_1-(N-1))), \\ ((C_1 \cdot (i_1-P)) \cdot (C_1 \cdot (i_1-P+1)) \cdot \ldots \cdot (C_1 \cdot (i_1-1)))]^T. \quad (2)$$

Denote U_i as the sum of the signals of each i-th neuron of all j-th intermediate layers, and g is the sum of the signals of each i-th neuron in the output layer, then

$$U_i = \sum_{j=0}^{N+P} w_{ij} \cdot C_j \,, \quad (4)$$

where $w_{ij} = f(U_i)$. Respectively,

$$g = \sum_{i=0}^{i_1} w_i \cdot C(i_1), \quad (5)$$

where $C(i_1) = f(g)$. Thus, the output is the result in the form of an output data signal (I_1).

The read values (fig. 4) are transferred to the input layer for further processing and analysis, followed by filling in the gaps. The fact is that at gauging stations occasionally misses measuring the water level for certain reasons. The developed neural network allows them to be restored, which makes the forecast even more accurate. In the end, all the data is sent to training, after which the output value for each day is displayed on a graph.

The mathematical model of training is presented in section III.

III. Integration Of Teaching Methods Without A Hebb Teacher And Reverse Error Distribution For Threat Forecasting

The advantage of this integration (association) is to obtain a more accurate forecast of the water level at gauging stations in the shortest possible time. Partially, the first version of this approach was used to neutralize cyber threats [29].

Hebb's method. This method of self-training of a neural network is based on tuning the weights of synapses. There are two varieties of Hebb's teaching method: signal and differential. Hebb's signaling method of training is based on strengthening the connections between excited neurons (this is the main reason for choosing within the framework of integration). Thus, synapse weights will be distributed as follows:

$$w_{ij}(t+1) = \frac{(w_{ij}(t) + a \cdot y_i^{[n-1]} \cdot y_j^{[n]}) \cdot A_{ij(t+1)}^n}{A}, \quad (6)$$

where a is the learning speed (coefficient), $y_i^{[n-1]}$ is the value at the output of each i-th neuron of the layer n-1, $w_{ij}(t+1)$ and $w_{ij}(t)$ is the weights of the synapse that connects the neurons in iterations t and t+1, $y^{[n]}$ is the output value of each j-th neuron of the layer n.

In the case of the backpropagation training method, everything is different, since the main idea is to propagate error signals from the ANN outputs to its inputs. The initial task is to minimize the error function:

$$E = \frac{1}{2}\sum_{j=1}^{p}(y_j - d_j)^2, \quad (7)$$

where y_j is the obtained value of the j-th output of the ANN and d_j is the reference value of the j-th output of the neural network. Accordingly, the minimization of E is determined by the gradient descent method. At the first stage, there is an automatic adjustment of the weight coefficients of the synapses:

$$\Delta w_{ij} = -m \cdot \frac{dE}{dw_{ij}}, \quad (8)$$

where w_{ij} is the synapse weight, $-m$ ($0<m<1$) is the neural network learning rate expressed by a coefficient. Next, it is necessary to disclose (8) for more accurate training of the neural network. Thus,

$$\frac{dE}{dw_{ij}} = \frac{dE}{dy_j} \cdot \frac{dy_j}{ds_j} \cdot \frac{ds_j}{dw_{ij}}, \quad (9)$$

where s_j is the sum of the input signals of each ANN neuron and y_j – is the output of the j-th neuron. Respectively,

$$\frac{dE}{dy_j} = \sum_k \frac{dE}{dy_k} \cdot \frac{dy_k}{ds_k} \cdot \frac{ds_k}{dy_j} = \sum_k \frac{dE}{dy_k} \cdot \frac{dy_k}{ds_k} \cdot w_{jk}^{[n+1]}, \quad (10)$$

where k is the number of neurons in the n+1 layer. Also for disclosure (8) we introduce a new variable:

$$\delta_j^{[n+1]} = \frac{dE}{dy_j} \cdot \frac{dy_j}{ds_j}. \quad (11)$$

After entering the variable, we obtain recursive formulas for the output (12) and input (13) layers of the ANN:

$$\delta_i^{[n]} = \frac{dy_i}{ds_i} \cdot (y_i^{[n]} - d_i), \quad (12)$$

$$\delta_j^{[n]} = \frac{dy_j}{ds_j} \cdot \sum_k \delta_k^{[n+1]} \cdot w_{jk}^{[n+1]}. \quad (13)$$

Thus, the disclosed formula (14) for automatically adjusting the weight coefficients of synapses:

$$\Delta w_{ij} = -m \cdot \delta_j^{[n]} \cdot y_i^{[n-1]}. \quad (14)$$

Next, it is necessary to integrate the error minimization function (7), the automatic adjustment of the weight coefficients of the synapses (8) and the distribution of the weights (6) for faster training and adjustment of the weight coefficients in order to increase the accuracy of calculations:

$$T_1 = \frac{\sum A_{ij(t+1)}^n \cdot w_{ij}(t+1)}{\left(\frac{1}{2}\sum_{j=1}^{N}(y_{ij} - d_{ij})^2\right) \cdot \Delta w_{ij}}, \quad (15)$$

$$\Delta w_{ij}^N(t) = -m \cdot \delta_j^{(n-1)} \cdot y_i^{(n-1)} + \Delta w_{ij}(N+1). \quad (16)$$

The last step is to calculate the gradient for the output and intermediate layers in order to adjust the received data:

$$\frac{dE}{dw_{ij}} = \left(\sum_{n=1}^{n} \delta_j^{[n]} \cdot y_i^{[n-1]}\right) \cdot \left(\sum_{n=1}^{n} \delta_{ij}^{n}\right), \quad (17)$$

$$\frac{dE}{dw_{ij}} = \left(\sum_{n=1}^{n} \delta_j^{[n]} \cdot y_i^{[n-1]}\right) \cdot \left(\sum_{n=1}^{n}\sum_{ij} \delta^{n-1}\right), \quad (18)$$

The next step is the testing presented in section IV. In the future, it is planned to finalize the neural network: add for more accurate prediction and improvement of the mathematical model of training the parameters (in the form of additional data) “rain”, “snow cover”, “soil moisture”, “air temperature”.

IV. Estimation Of The Effectiveness Of The Proposed Solution For Forecasting The Water Level At Hydroposts

The purpose of testing is to identify the effectiveness of the developed solution within the framework of the mathematical problem posed in the second section. Initially, a water level forecast was made for ten days, which is presented in table 1. For the reliability of the forecast, a check was made on old data.

TABLE I. FORECASTING THE WATER LEVEL AT HYDROPOSTS

Date	Hydropost number	Actual water level, cm.	Predicted water level ($A_{ij(t+1)}^n$), cm.
01.04.19	76289 (Ufa, Belaya River)	-11,00	0,00
02.04.19		8,00	20,00
03.04.19		39,00	34,00
04.04.19		63,00	50,00
05.04.19		74,00	86,00
01.04.19	3000014 (microdistrict Shaksha, Ufa)	124,00	150,00
02.04.19		131,00	147,00
03.04.19		134,00	160,00
04.04.19		134,00	161,00
05.04.19		160,00	180,00
01.04.19	76288 (Iglinsky district, the village of Okhlebinino)	291,00	300,00
02.04.19		270,00	300,00
03.04.19		281,00	305,00
04.04.19		296,00	312,00
05.04.19		323,00	330,00

Based on the data of table 1 it is seen that the predicted data do not differ much from the real ones (a difference of 9-10%). An important difference, compared with other methods, is the speed with which a forecast is obtained and its correctness (more accurate) for long-term forecasting.

V. Conclusion

A method for the early detection of threats (in this case, the water level at the gauging stations of the Republic of Bashkortostan) with the aim of countering them is proposed. An artificial neural network is proposed for prediction. Testing of this artificial neural network shows the effectiveness of its work: the difference between the predicted and real values varies from 9 to 10%. This allows you to give the necessary time to special services for flood control measures to prepare for the protection of complex distributed (including technical) systems.

Acknowledgment

The results of the research presented n this article were supported by grant RFBR no. 20-08-00301 A – «Methodological and theoretical foundations of vulnerability management based on threat repression in complex distributed systems».

References

[1] O.E.J. Wing, C.C. Sampson, and P.D. Bates, “A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model,” Journal of Hydrology X,vol. 4, pp. 1–17, 2019.

[2] D.S. Gundrov, “Automated system for monitoring and forecasting the flood situation at water bodies of the Krasnodar Territory and the prospects for its development,” Problems of ensuring safety in responding to emergency situations, vol. 1, pp. 155–157, 2018.

[3] M.Ya. Zdereva, V.F. Bogdanova, N.A. Khluchina, “Evaluation of the possibility of using model precipitation forecasts for forecasting rain floods in the mountain rivers of Altai,” Transactions of the Hydrometeorological Research Center of the Russian Federation, No. 359, pp. 128–141, 2016.

[4] Y.V. Grebnev, A.V. Spring, “Flood monitoring and forecasting in the Krasnoyarsk Territory using neural network algorithms,” Siberian Fire and Rescue Bulletin, no. 3 (10), pp. 13–16, 2018.

[5] J. Noymanee, T. Theeramunkong, “Flood Forecasting with Machine Learning Technique on Hydrological Modeling,” Procedia Computer Science, vol. 156, pp. 377–386, 2019.

[6] Y. Zhou, S. Guo, F. Chang, “Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts,” Journal of Hydrology, vol. 570. pp. 343–355, 2019.

[7] S. Han, P. Coulibaly, “Bayesian flood forecasting methods: A review,” vol. 551, pp. 340–351, 2017.

[8] A.E. Tekeli, H. Fouli, “Evaluation of TRMM satellite-based precipitation indexes for flood forecasting over Riyadh City,” Saudi Arabia, vol. 541, Part A, pp. 471–479, 2016.

[9] K. Kim, P. Pant, E. Yamashita, “Integrating travel demand modeling and flood hazard risk analysis for evacuation and sheltering,” International Journal of Disaster Risk Reduction, vol. 31, pp. 1177–1186, 2018.

[10] V.E. Gvozdev, A.E. Kolodenkova, “The method of extrapolation forecasting of the flood situation on the basis of retrospective and current data (on the example of the Republic of Bashkortostan),” Vestnik UGATU, vol. 9, no. 2 (20), pp. 119–126, 2007.

[11] S. Barbetta, G. Coccia, T. Moramarco, and E. Todini, “Real-time flood forecasting downstream river confluences using a Bayesian approach,” Journal of Hydrology, vol. 565, pp. 516–523, 2018.

[12] A.V. Wesemael, L. Landuyt, H. Lievens, and N. Verhoest, “Improving flood inundation forecasts through the assimilation of in situ floodplain water level measurements based on alternative observation network configurations,” Advances in Water Resources, vol. 130, pp. 229–243, 2019.

[13] T. Cowan, M.C. Wheeler, O. Alves, S. Narsey, and M.K. Hawcroft, “Forecasting the extreme rainfall, low temperatures, and strong winds associated with the northern Queensland floods of February 2019,” Weather and Climate Extremes, vol. 26, pp. 1–14, 2019.

[14] X. Zhang, P. Liu, L. Cheng, Z. Liu, and Y. Zhao, “A back-fitting algorithm to improve real-time flood forecasting,” Journal of Hydrology, vol. 562, pp. 140–150, 2018.

[15] O.A. Efremova, Yu.N. Kunakov, S.V. Pavlov, and A.H. Sultanov “Algorithm for constructing zones of actual water areas of area during spills rivers based on radar surface images and analysis of characteristic relief points,” Computer optics, vol. 42, no. 4, pp. 695–703, 2018.

[16] M.M. Morsy, J.L. Goodall, G.L. O’Neil, J.M. Salder, D. Voce, G. Hassan, and C. Huxley, “A cloud-based flood warning system for forecasting impacts to transportation infrastructure systems,” Environmental Modelling & Software, vol. 107, pp. 231–244, 2018.

[17] S.V. Pavlov, V.G. Krymsky, and O.I. Khristodulo, Information support for assessing and managing water resources based on geoinformation technologies, Moscow: Dada+, 2010, p. 284.

[18] S.V. Pavlov, R.R. Sharafutdinov, and O. I. Khristodulo, “Development of a geoinformation model of a river network taking into account cartographic, hydrological and morphometric information to determine the boundaries of flood zones when the water level in water bodies changes,” Vestnik of USATU, vol. 11, no. 1, pp.18–27, 2008.

[19] S.V. Pavlov, S.A. Abramov, R.A. Shkundin, and O.I. Khristodulo, “Geoinformation system for water resources management at the territorial level (on the example of the Republic of Bashkortostan),” Geoinformatics, no. 4, pp. 14–20, 2008.

[20] I.N. Zaitov, O.I. Khristodulo, and R.R. Sharafutdinov, “Automated information support system for flood control in the Republic of Bashkortostan,” Management of water resource systems in extreme conditions, Moscow, pp. 219–220, 2008.

[21] C.H. Chang, H. Lee, F. Hossain, S. Basnayke, S. Jayasinghe, F. Chishtie, D. Saah, Yu. Hanwen, K. Sothea, and D.D. Bui, “A model-aided satellite-altimetry-based flood forecasting system for the Mekong River,” Environmental Modelling & Software, vol. 112, pp. 112–127, 2019.

[22] H. Shafizadeh-Moghadam, R. Valave, H. Shahabi, and K. Chapi, “Shirzadi A. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping,” Journal of Environmental Management, vol . 217, pp. 1–11, 2018.

[23] H. Bonakadri, A.H. Zaji, A.D. Binns, and B. Gharabaghi, “Integrated Markov chains and uncertainty analysis techniques to more accurately forecast floods using satellite signals,” Journal of Hydrology, vol. 572, pp. 75–79, 2019.

[24] L. Alfieri, E. Zsoter, S. Harrigan, F.A. Hipra, C. Lavaysse, C. Prudhomme, and P. Salamon, “Range-dependent thresholds for global flood early warning”, Journal of Hydrology X, vol. 4, pp. 1-10, 2019.

[25] S. Berkhahn, L. Fuchs, and I. Neuweiler, “An ensemble neural network model for real-time prediction of urban floods,” Journal of Hydrology, vol. 575, pp. 743–754, 2019.

[26] S. Park, M. Berenguer, D. Sempere-Torres, “Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale,” Journal of Hydrology, vol. 573, pp. 768–777, 2019.

[27] A. Bukvic, J. Harrald, “Rural versus urban perspective on coastal flooding: The insights from the U.S. Mid-Atlantic communities,” Climate Risk Management, vol. 23, pp. 7–18, 2019.

[28] F. Policelli, A. Hubbard, H. Chul Jung, B. Zaitchik, and C. Ichoku “A predictive model for Lake Chad total surface water area using remotely sensed and modeled hydrological and meteorological parameters and multivariate regression analysis,” Journal of Hydrology, vol. 568, pp. 1071–1080, 2019.

[29] E.V. Palchevsky, O.I. Khristodulo, “Development of a self-learning method for a pulsed neural network to protect against DDoS attacks,” Software & Systems, vol. 32, no. 3, pp. 419–432, 2019.