Learning process

The step of traffic filtering / selecting features is desirable in order to limit the number of features actually used in training the classifier and thus create a classification model. The output signal in Figure 5 is a classification model.

Conclusion

In conclusion, it should be noted that the use of machine learning methods for classifying IP traffic makes it possible to solve the problems of classification and regression and to predict the output for new input streams of network traffic. The proposed model of classifier training with supervised learning allows to calculate the processing of traffic statistics and filter traffic to limit the number of attributes used in the learning process.

References

1. S. Zander, T. Nguyen, and G. Armitage, "Automated traffic classification and application identification using machine learning," in IEEE 30th Conference on Local Computer Networks (LCN 2005), Sydney, Australia, November 2005.
2. IETF RFC 7011. Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information, September 2013.
3. Gulomov Sherzod Rajaboevich, Xoshimova Charos Saidaminovna, Ganiyeva Toxira Irkinovna, Djurayeva Shoxista Tagirovna. Analysis of Methods for Measuring Available Bandwidth and Classification of Network Traffic. International Journal of Emerging Trends in Engineering Research. Volume 8. No. 6, June 2020. Indexed in Scopus. – P. 2753-2759.
4. Chen Z., Yang B., Chen Y., et al. Online hybrid traffic classifier for Peer-to-Peer systems based on network processors // Appl. Soft Comput. 2009. V. 9, N 2. P. 685–694.
5. J. Erman, A. Mahanti, and M. Arlitt, "Internet traffic identification using machine learning techniques," in Proc. of 49th IEEE Global Telecommunications Conference (GLOBECOM 2006), San Francisco, USA, December 2006.
6. Canini M., Li W., Moore A. W., Bolla R. GTVS: Boosting the collection of application traffic ground truth // Lecture Notes Comput. Sci. 2009. V. 5537. P. 54–63.

APPROACHES TO SOLVING PROBLEMS OF OPTIMIZATION OF SOLVING MONITORING PROBLEMS BASED ON NATURAL COMPUTING ALGORITHMS

Muhamediyeva D.T.

Professorof Scientific and Innovation Center of Information and Communication Technologies at TUIT named after Al-Kharezmi, Tashkent, dilnoz134@rambler.ru

Abstract: *To find approximate optimization solutions, many algorithms are used, seven of which are considered in this work: fuzzy sets, artificial neural networks, genetic algorithm, ant algorithm, particle swarm algorithm, DNA computation, and a new approach based on artificial immune systems (IIS). All these methods belong to the direction of "natural computing", ie. model certain biological processes, the algorithms of which nature has created for millions of years. It should be noted that the efficiency of one or another algorithm depends on the characteristics of the initial data of the problem, so it is impossible to unambiguously determine which of the algorithms is the most effective*

Keywords: *fuzzy sets, fuzzy models, neural networks, estimation, risk, knowledge base, membership function, decision making*

1. Introduction.

Consider a test problem of combinatorial optimization about sales agents and a routing problem. The trading company sells goods in *n* different cities, the purchasing power of residents of which is estimated in b_j conv. units, $j=\overline{1,n}$. For the sale of goods, the company has *n* sales agents, each of whom she sends to one of the cities. The professional level of agents varies; share of implemented i-th sales agent purchasing power is a_i, $i=\overline{1,m}$.

Let's introduce the parameter $r_{ij} = a_i b_j$, characterizing the value of the purchasing power realized i - th as a sales agent in j -th town.

Control variables z_{ij}, $i = \overline{1,n};\ j = \overline{1,n}$ determined by the formula

$$r_{ij} = \begin{vmatrix} 1, \ if \ \ i - th \, agent \, sent \ to \ \ j - th \ town, \\ 0, \ in \ other \, case. \end{vmatrix}$$

Mathematical model will be written in the following form:

$$R = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} z_{ij} \rightarrow \max. \tag{1}$$

$$\begin{cases} \sum_{j=1}^{n} z_{ij} = 1, \ i = \overline{1,n}, \\ \sum_{i=1}^{n} z_{ij} = 1, \ \ j = \overline{1,n}, \\ z_{ij} \in \{0;1\}, \ \ i = \overline{1,n} \quad j = \overline{1,n}. \end{cases} \tag{2}$$

First and second constraints formalize, respectively, the conditions that one sales agent is sent to each city, and one sales agent cannot work in two cities. Objective function *R(z)* - it is the sum of the realized purchasing power of all sales agents in all cities. It should be as high as possible.

To solve the sales agent problem is to find z_{ij}, satisfying (1) and minimizing functions (2).

There are three most efficient algorithms for finding the shortest path:

• Dijkstra's algorithm (used to find the optimal route between two peaks);
• Floyd's algorithm (for finding the optimal route between all pairs of vertices);
• Jen's algorithm (for finding k-optimal routes between two vertices).

Indicated algorithms are easily performed with a small number of vertices in the graph. With an increase in their number, the task of finding the shortest path becomes more difficult.

In this regard, modern methods for solving optimization problems are being investigated.

2. Fuzzy sets.

Consider the formation of a route matrix based on the theory of F-sets.

Let the source I and the receiver P be known a priori, the set of feasible route plans $L = \{l_j\}$, many route parameters $\pi_n = \Pi$ (route length, transmission delay, security, reliability, throughput, information aging and network element availability and downtime factors, etc.), possible situations of network damage $C_ц = \{c_k\}$ [1].

In real operating conditions, as a rule, there is no statistics on the distribution of information flows in extreme situations at a time T_i. However, in these conditions, it is necessary to choose an acceptable route l^* to establish the required connection with a satisfactory quality of information reception.

In this regard, at each switching node, fuzzy reflexive matrix of route preferences are constructed a priori M_n^l, based on standard membership functions by condition of the situation $c_k \in C_ц$, stored in the database B_{Mk}. Pre-database B_k is formed on the basis of the opinions of experts on the telecommunications network in the form of logical fuzzy rules.

If we take into account that the manager in the network management system is dominant at his level of distributed control (DC), then it is enough to form a routing F-matrix only for him.

Thus, taking into account the address (code) of the agent, route parameters $\pi_n = \Pi$ and possible situations $c_k \in C_ц$ control system M defines a subset of acceptable routes $L^* \subset L = \{l_j\}$. From the set L^* fuzzy preference matrix is constructed, which has the form:

$$M_n^l = \begin{array}{c|ccccc|} & l_1 & l_2 & l_3 & \dots & l_j \\ \hline l_1 & 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1j} \\ l_2 & \alpha_{21} & 1 & \alpha_{23} & \dots & \alpha_{2j} \\ l_3 & \alpha_{31} & \alpha_{32} & 1 & \dots & \alpha_{3j} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ l_q & \alpha_{q1} & \alpha_{q2} & \alpha_{q3} & \dots & 1 \end{array}, \tag{3}$$

where M_n^l - route relationship matrix; n - parameter index π, in relation to which the expert determines the degree of the preference relationship for routes in the network $l_j \in L$; α_{qj} - numerical value of the membership function

$$\mu_R(l_q, l_j) = < not\, bad\, l_j >, l_q, l_j \in L, \alpha_{qj} \in [0,1]. \tag{4}$$

A reflective preference matrix for route parameters is also built $\pi_n = \Pi$, based on a priori defined database subject to the situation $c_k \in C_ц$:

$$M_k^\pi = \begin{array}{c|ccccc|} & \pi_1 & \pi_2 & \pi_3 & \dots & \pi_j \\ \hline \pi_1 & 1 & b_{12} & b_{13} & \dots & b_{1j} \\ \pi_2 & b_{21} & 1 & b_{23} & \dots & b_{2j} \\ \pi_3 & b_{31} & b_{32} & 1 & \dots & b_{3j} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \pi_q & b_{q1} & b_{q2} & b_{q3} & \dots & 1 \end{array}, \tag{5}$$

where π_n - route parameter when situation c_k; b_{ji} - numerical value of the membership function

$$\mu_R(\pi_i, \pi_j) = < \pi_i\, not\, bad\, \pi_j >, i, j = \overline{1,n}, b_{ij} \in [0,1].$$

Next, it is necessary to correlate matrices (3) and (5) and decide on the most acceptable route for transmitting information to the recipient, i.e. get matrix

$$M_k^l = \begin{array}{c|ccccc|} & l_1 & l_2 & l_3 & \dots & l_j \\ \hline \pi_1 & \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1j} \\ \pi_2 & \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2j} \\ \pi_3 & \alpha_{31} & \alpha_{32} & \alpha_{33} & \dots & \alpha_{3j} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \pi_q & \alpha_{q1} & \alpha_{q2} & \alpha_{q3} & \dots & \alpha_{nj} \end{array}. \tag{6}$$

Each row of matrix (6) is a fuzzy set L_{kn} = < POSSIBLE ROUTE AREA BY PARAMETER CONDITION π_n > with membership function $\mu(l_j) \in [0,1]$, $\alpha_{nj} \in [0,1]$, where n, j - current indices, k - situation index c_k.

Each row of matrix (6) is a compression of matrices of type (3) according to the algorithm [1].

Step 1. Based on matrices of type (1), strict preference matrices are constructed by the expression

$$R^S = R^p - (R^p)^{-1}$$

with membership function

$$\mu_R(x,y) = \max_{x,y\in L}(0, \mu_R(x,y) - \mu_R(y,x)), \tag{7}$$

where R^p - fuzzy preference relation $R^p = M^d$; $(R^p)^{-1}$- inverse relation; R^S - fuzzy attitude of strict preference; x,y - values of variables identical l (routes), $l_j \in L$.

Step 2. Take the complement from (7)

$$\overline{R} = 1 - R^S$$

with membership function

$$\mu_{\overline{R}}(x,y) = 1 - \mu_{R^S}(x,\mathrm{y}). \tag{8}$$

Step 3. A fuzzy subset of non-dominated (n.d.) alternatives is determined, i.e. the matrix is convolved (3) M_n^l as a string

$$\mu_{R_i}^{НД}(y) = \min_{x\in L}\left[1 - \mu_R(x,y)\right],\ y \in L, \tag{9}$$

or

$$\mu_{R_i}^{НД}(y) = 1 - \max_{x\in L}\mu_R(x,y),\ y \in L,\ i = \overline{1,n}. \tag{10}$$

As a result of operations (7) - (10), a matrix of fuzzy relations (6) is formed.

Stage 2. Next, a matrix of fuzzy relations is formed on the basis of matrices (4) and (6) according to the algorithm.

Step 4. Take a max-min composition of matrices (6) and (7)

$$R_k = M_k^l \circ M_k^\pi \tag{11}$$

with membership function

$$\mu_R(l_i, 1_j) = \max\min_{x,y\in L}\left[\mu_{M^l}(l_i, \pi_j), \mu_{M^\pi}(\pi_i, \pi_j)\right]. \tag{12}$$

Step 5. From the matrix R_k (11) a fuzzy subset of non-dominant alternatives is determined similarly to (9) or (10). Thus, the subset is found $L_k \subset L$ possible routes of the system for the kth network situation with the membership function

$$\mu_{L_k}(l_j / \pi_1, \pi_2, ..., \pi_n), \tag{13}$$

by parameter condition $\pi_i,\ i = 1,...,n.$

Step 6. In accordance with expression (5), the most possible route in the control network is determined.

Next, matrix (6) is compressed into a row matrix according to the algorithm from [1] and the most acceptable route is determined according to the rule

$$l_k^* = \arg\max_{l_i\in L}\mu_L(l_j / \pi_1, \pi_2, ..., \pi_i), \tag{14}$$

where $\mu_L(l_j / \pi_1, \pi_2, ..., \pi_i)$- membership function <THE MOST POSSIBLE ROUTE AREA L ACCORDING TO THE PARAMETERS $\pi_1, \pi_2, ..., \pi_i$>; l_k^* - the most possible route for the kth situation in the network.

3. Neural networks.

To solve problem (1) - (2), a recurrent neural network [3] is proposed, which is described by the differential equation

$$\begin{aligned}\frac{\partial u_{ij}(t)}{\partial t} = -\eta\left(\sum_{k=1}^{n} z_{ik}(t) + \sum_{l=1}^{n} z_{lj}(t) - 2\right) + \\ + \lambda r_{ij}\exp\left(-\frac{t}{\tau}\right),\end{aligned} \tag{15}$$

where $z_{ij} = f(u_{ij}(t)), \quad f(u) = \frac{1}{1+\exp(-\beta u)}$. As in the Hopfield network, it uses a matrix of neurons of size $n \times n$, but neurons do not interact according to the principle "each with each", but in rows and columns.

The difference version of this equation has the form

$$u_{ij}^{t+1} = u_{ij}^{t} - \Delta t \cdot \left[\eta \left(\sum_{k=1}^{n} z_{ik}(t) + \sum_{l=1}^{n} z_{lj}(t) - 2 \right) - \right. \\ \left. - \lambda r_{ij} \exp\left(-\frac{t}{\tau} \right) \right]. \tag{16}$$

Where Δt - step by time. Parameters $\Delta t, \eta, \lambda, \tau, \beta$ are selected experimentally and significantly affect the speed of achieving a solution to the problem and the quality of this solution.

To speed up the solution of the system of equations (16), the principle "Winner takes all" is proposed [3]:

1. Matrix is generated $\left\| z_{ij}^{0} \right\|$ random values $z_{ij}^{0} \in [0,1]$. Iteration (16) continues until the inequality

$$\sum_{k=1}^{n} z_{ik}(t) + \sum_{l=1}^{n} z_{lj}(t) - 2 \le \varepsilon,$$

where ε - specified accuracy of fulfillment of constraints (2).

2. Transformation of the obtained solution matrix is performed $\left\| z_{ij} \right\|$:

2.1. $i = 1$.

2.2. The i-th row of the matrix finds the maximum element $z_{i,j_{\max}}$, $j_{\max}$ - column number with maximum element.

2.3. Conversion in progress $z_{i,j_{\max}} = 1$. All other elements of the first row and column vanish:

$$z_{ij} = 0, \; j \ne j_{\max},$$
$$z_{k,j_{\max}} = 0, \; k \ne i.$$

Next, there is a transition to the line $j_{\max}$.

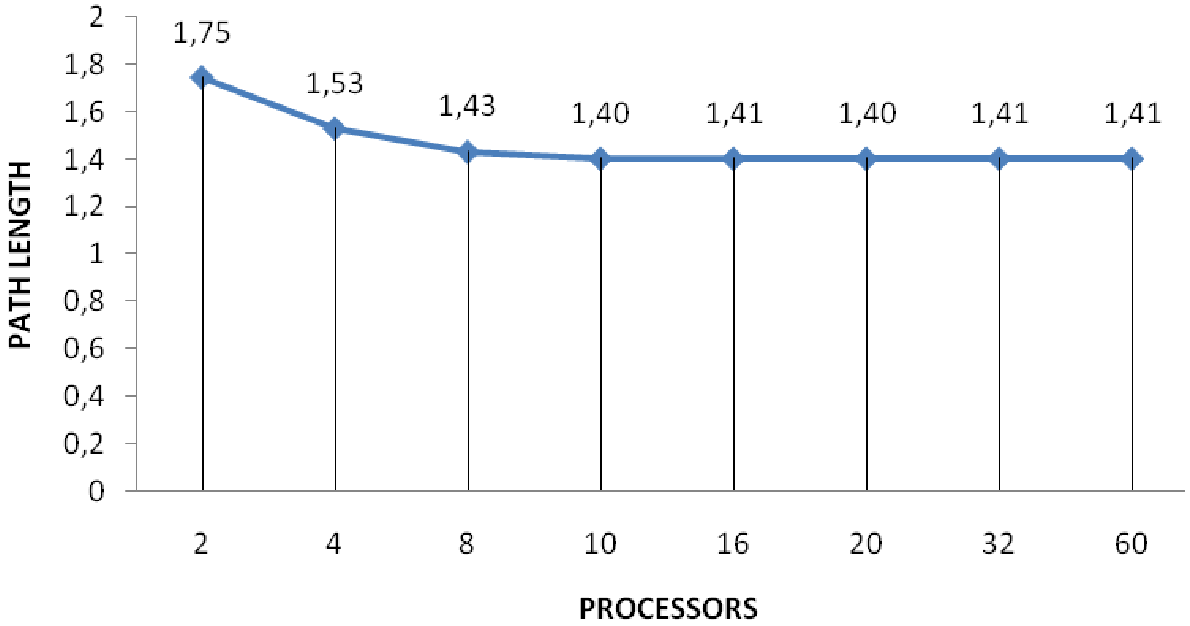

Figure 1. ***The result obtained using the Hopfield network - a graph of the dependence of the number of processors with the length of the optimal direction.***

Actions 2.2. and 2.3 are repeated until the return to the first line occurs, which will mean the completion of the construction of the loop

3. If the return to row 1 occurred earlier than in the matrix $\|z_{ij}\|$ value 1 received n elements, this means that the length of the constructed loop is less than n. In this case, steps 1 and 2 are repeated.

With this approach, the computational complexity of the algorithm for solving routing problems decreases from $O(n^4)$ до $O(n^2)$.

To solve this problem in the Java programming language, the FMPJ library is used. For the solution, 15 dual-processor and dual-core computers were used (60 processors in total).

Figure 1 shows a graph of the results obtained using the Hopfield network algorithms:

4. Conclusion.

The analysis of the results obtained showed that the results of the proposed algorithms of artificial neural networks, in comparison with the algorithms created based on Hopfield neural networks, are distinguished by low resource intensity and efficiency in terms of high speed of work. However, it should be noted that if the volume of tasks is very large, then neural network algorithms might become less efficient due to a longer computation time. Usually in such cases, it is advisable to use evolutionary algorithms.

References

1. Bychkov E. D. Application of the theory of fuzzy (Fuzzy) sets in mathematical models of communication systems. Research and materials: Supplement to the journal "Omsk Scientific Bulletin" / Bychkov ED, Salakhutdinov RZ, Lendikrey VV - Omsk: OGMA, 2000 .- 188 p.
2. Hopfield J.J., Tank D.W. “Neural” computation of decisions in optimization problems // Biological Cybernetics, 1985, vol. 52, no. 3, pp. 141-152.
3. Bychkov E. D. Application of the theory of fuzzy (Fuzzy) sets in mathematical models of communication systems. Research and materials: Supplement to the journal "Omsk Scientific Bulletin" / Bychkov ED, Salakhutdinov RZ, Lendikrey VV - Omsk: OGMA, 2000 .-188 p.
4. Holland J. H. Adaptation in natural and artificial systems. An introductory analysis with application to biology, control, and artificial intelligence.— London: Bradford book edition, 1994 —211 p.
5. Zadeh L.A. Fuzzy sets: Information and control. 1965 , -Vol.№8. - pp. 338-353.
6. Aliev R.A, Aliev R.R. The theory of intelligent systems. –Baku: Publishing House "Chashyolgy", 2001. -720 pAltunin A.E., Semukhin M.V. Models and algorithms for decision making in fuzzy environment. Tyumen: Publishing House of the Tyumen State University, 2000. - 352 p.
7. Muhamediyeva D.T. and Egamberdiyev N.A. Algorithm and the Program of Construction of the Fuzzy Logical Model //2019 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2019, pp. 1-4
8. Egamberdiev N., Mukhamedieva D. and Khasanov U. Presentation of preferences in multi-criterional tasks of decision-making // IOP Conf. Series:Journal of Physics: Conference Series 1441 (2020).
9. Bekmuratov T.F., Mukhamedieva D.T. A training algorithm of fuzzy inference system //International scientific and technical journal “Chemical technology. Control and management., № 3-4” and “Journal of Korea multmedia society” South Korea, Seoul – Uzbekistan, Tashkent. – 2015. –C.108-114.
10. Muhamediyeva D.K. Study parabolic type diffusion equations with double nonlinearity // IOP Conf. Series:Journal of Physics: Conference Series 1441 (2020) 012151. DOI: https://doi.org/10.1088/1742-6596/1441/1/012151.
11. Mukhamediyeva D.T., Niyozmatova N.A. Application of Immune Algorithm for Constructing a Model of Fuzzy Inference // American Journal of Mathematical and Computational Sciences 2016; 1(2): 74-78. http://www.aascit.org/journal/ajmcs.
12. Marakhimov, A.R., Siddikov, I.H., Nasridinov, A., Byun, J.-Y. A structural synthesis of information computer networks of automated control systems based on genetic algorithms // Lecture Notes in Electrical Engineering. 2015.