NEURAL NETWORKS APPLICATION FOR ANALYSIS OF FLIGHT INFORMATION IN AIRCRAFT ENGINE DIAGNOSTIC SYSTEM

Eugene Kopytov[1], Vladimir Labendik[1], Arvid Osis[2], Alexey Tarasov[1]

[1]*Transport and Telecommunication Institute*
Lomonosov Str. 1, Riga, LV-1019, Latvia
Ph.: (+371)7100590. Fax: (+371)7100660. E-mail: kopitov@tsi.lv

[2]*Aviation Institute of Riga Technical University*
Lomonosov Str. 1, Riga, LV-1019, Latvia
Ph.: (+371) 7207255, (+371) 7751589. E-mail: arvids.osis2003@navigator.lv

To increase the diagnostic systems reliability in flight in the case of partial information loss the using of artificial intellect systems in the form of neural networks as on-board software is suggested. Neural networks have property of training in the case of information flow changes. In the paper diagnostic matrices developed for turboprop engine PW125B of Fokker-50, which is exploited in the "Riga" airport, are used as a training tool for neural networks. The examples of neural networks training based on data obtained with the help of failures simulation using table of influence coefficients and diagnostic matrices are considered.

Keywords: *turboprop engine, flight information, thermo gas dynamic parameters, neural networks*

1. INTRODUCTION

For the diagnostics of the turboprop engine flowpath the usage of the diagnostic (localizing) matrices (DM) based on the modified linear mathematical model (LMM) of the aircraft engine adequately considering the changes of the gas generator condition during the exploitation process is possible while the number of measurement points is sufficient [1], though this method also has disadvantages. For example in the case of at least one sensor measuring the parameters' changes or its carrier failure the DM stop working, although deviations of other measurable parameters still carry the information about the perspective defect.

To raise the aircraft engines diagnostic efficiency during the flight in the case of partial information loss the usage of the systems with the elements of artificial intelligence in the form of neural networks (NN) as the on-board software is appropriate [2]. NN have property of training in the case of information flow changes. At the same time the DM can be used as the learning facility for the NN in the complex automated system of the engines ground diagnostics. But the DM can't be formed for many engines because the number of measurement points is not sufficient [3]. Unfortunately it is also applied to the turboprop engine PW125B of the Fokker-50 aircraft exploited in "Riga" airport (company "AirBaltic") [4]. For the developing of the DM for such engine's gas generator 9 parameters are needed, but only 4 parameters are measured during the flight (shown by dashed lines in the Figure 1). That's because the table of the qualitative estimation of the measurable parameters deviations is used (see Figure 2). As it can be seen some defects have the identical qualitative presentation and the localization of the defect node is not possible.

In this paper the task of raising the depth of aircraft engine diagnostics by the usage of the new methods of data processing and existing information analysis (apparatus of the aircraft engines theory and programming perspective technologies) has been formulated. Because of the small exploitation term and the absence of statistics about flowpath defects development the modelling of the defects indications with the help of the table of influence coefficients and the DM were carried out on the first stage. The DM is formed especially for the NN training in the condition that all sensors, which are necessary for the DM functioning, are present. It can be realized in the case of the ground-based aircraft engine diagnostic process. In the future only trained NN will be used during the flight exploitation. As it was said before, the DM loses their possibility of the defects localizing in the case of at least one sensor (or data carrier) failure. Properly trained non-linear NN can solve such tasks also in the case of incompleteness of input information. The information received after the transformations of the LMM of the engine in form of the table of influence coefficients is used as the training set.

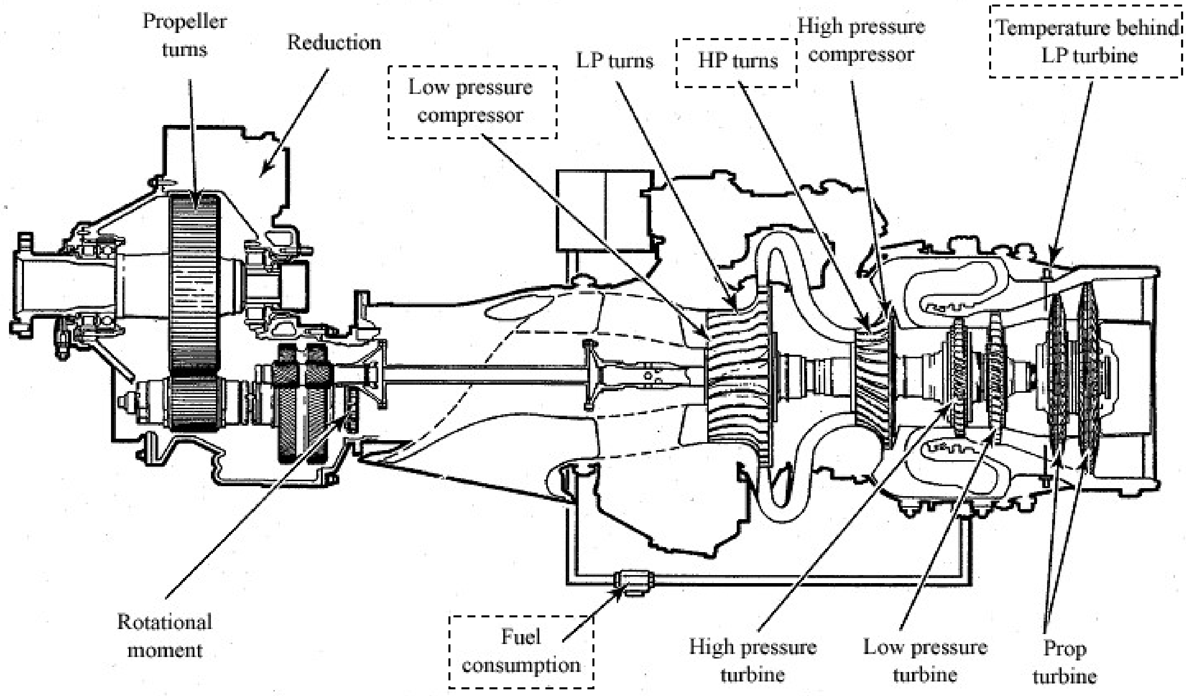

Figure 1. *PW125B engine measurable parameters*

HOT SECTION PERFORMANCE

COMPONENTS:	SYMPTOMS: AT CONSTANT SHP
HP VANE RING – INCREASE AREA	NL → Nh ↓ WF ↑ T6 ↑
LP VANE RING – INCREASE AREA	NL ↓ Nh ↑ WF ↑ T6 ↑
PT VANE RING – INCREASE AREA	NL ↑ Nh → WF ↑ T6 ↓
HP TURBINE – INCREASE TIP CLEARANCE	NL ↑ Nh ↓ WF ↑ T6 ↑
LP TURBINE – INCREASE TIP CLEARANCE	NL ↓ Nh ↑ WF ↑ T6 ↑
PT TURBINE – INCREASE TIP CLEARANCE	NL ↑ Nh ↑ WF ↑ T6 ↑
HP VANE SEALING RING – LEAKAGE	NL ↑ Nh ↓ WF ↑ T6 ↑
LP VANE SEALING RINGS – LEAKAGE OF FIRST SEALING RING	NL ↑ Nh ↓ WF ↑ T6 ↑
LP VANE SEALING RINGS – LEAKAGE OF SECOND SEALING RING	NL ↓ Nh ↓ WF ↑ T6 ↑
LP VANE SEALING RINGS INTERSTAGE TURBINE CASE SEALING RING – LEAKAGE	NL ↑ Nh ↑ WF ↑ T6 ↑
LP VANE SEALING RINGS SECOND PT VANE SEALING RING – LEAKAGE	NL ↑ Nh ↑ WF ↑ T6 ↑
LP VANE SEALING RINGS	

↑: UP
↓: DOWN
→: SAME

Figure 2. *Table of the qualitative estimation of deviations of the measurable parameters*

The DM is used for calculation of the criteria indications of the defects development during the exploitation process (length and direction of the vector characterizing the flowpath defect). The offered approach of the theoretical DM usage in the diagnostic systems software development is especially significant for the new engines, when statistics of the flowpath defects development is not available yet.

2. THE NEURAL NETWORKS TRAINING ALGORITHM

For considered task solution two-layered perceptrons are offered. This perceptron is the most developed and popular topology of the artificial NN. The classification task in this case can be described as follows: let us consider the set of input vectors $X_i = (x_1, x_2, ..., x_m)$ and according the set of output vectors $Y_i = (y_1, y_2, ..., y_n)$, where m is the number of input variables, n is the number of classification groups, and $y_j, j = \overline{1,n}$ are Boolean variables. Let the group to which the observation fed to the inputs of the NN applies has number j, then we have

$$y_j = 1 \text{ and } y_k = 0 \text{ for } k = 1, 2, ..., j-1, j+1, ..., n \ . \tag{1}$$

The scheme of the two-layered perceptron is given in the Figure 3. Its outputs are defined by formulas:

$$o_i^{(H)} = f^{(H)}\left(\sum_{j=1}^{m} w_{ij}^{(HI)} x_j\right), \ i = \overline{1,h}\ ; \tag{2}$$

$$y_k = f^{(T)}\left(\sum_{j=1}^{h} w_{kj}^{(TH)} o_j^{(H)}\right), \ k = \overline{1,n}$$

where $o_i^{(H)}$ is i-th output of the hidden layer; y_k – k-th output of the perceptron; $f^{(.)}$ – the activation function; h – number of hidden neurons; $w_{ij}^{(HI)}$ – weight of the connection between i-th hidden neuron and j-th input neuron; $w_{kj}^{(TH)}$ – weight of the connection between k-th output neuron and j-th hidden neuron.

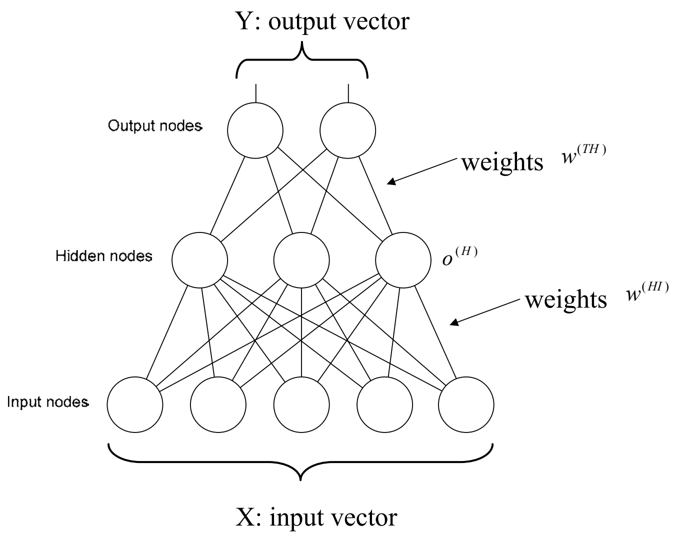

Figure 3. *The scheme of the two-layered perceptron*

It is necessary to take into account that the main advantage of multi-layered perceptrons (MLP) is their non-linearity that can be accomplished by the usage of the non-linear activation function. It can be shown that in such case the two-layered perceptron can represent any function with the finite number of break points if the size of hidden layer is sufficient. The most commonly used activation function is sigmoid curve because of its simple derivative:

$$f(x) = \frac{1}{1+e^{-\lambda x}} \tag{3}$$

where λ is the shape parameter.

In some cases the using of the hyperbolic tangent can be appropriate.

As the artificial NN training algorithm the back-propagation is offered. It consists of iterative corrections of weights in output and hidden layers:

$$\begin{aligned} & w_{ij}^{(HI)} = w_{ij}^{(HI)} + \alpha \cdot x_j \cdot \Delta w_{ij}^{(HI)}; \\ & \Delta w_{ij}^{(HI)} = f^{(H)'}(hidden_i) \cdot \sum_{k=1}^{n} \Delta w_{ki}^{(TH)} \cdot w_{ki}^{(TH)}; \\ & w_{ij}^{(TH)} = w_{ij}^{(TH)} + \alpha \cdot o_j^{(H)} \cdot \Delta w_{ij}^{(TH)}; \\ & \Delta w_{ij}^{(TH)} = (o_i - y_i) \cdot f^{(T)'}(output_i). \end{aligned} \tag{4}$$

where $output_i$, $hidden_i$ are i-th output of perceptron and i-th output of hidden neuron correspondingly before applying to them the sigmoid activation function, α is the learning rate.

The given algorithm is the kind of gradient descent on the error surface. It means that it doesn't guarantee the discovering of the global minimum of the error function (which depends on weights) or it's convergence in the acceptable terms. Nevertheless many researchers have reported about successful back-propagation using in the solution of the number of applied tasks. The weight correction occurs for each observation and in the most cases tens and hundreds training sets are needed for the reaching of acceptable results.

The stages of two-layered perceptrons training are as follows:

1. Forming the training and verification sets (as it will be shown below).
2. Defining the activation function for the hidden and output neurons.
3. Defining the number of hidden neurons.
4. Consecutive applying of the back-propagation algorithm during some amount of time or until the error of the network becomes permissible. After this stage weights never are changed.
5. Testing the network performance on the verification set.

The high efficiency of the network on the training set doesn't guarantee that network has correctly understood the dependence between input and output data. If the errors on the training and verification sets differ significantly, then the network is working incorrectly. In such case one of the given action has to be executed: repeated training (if it possible it must be longer), increasing the size of the training set or change of the network parameters defined on stages 2 and 3. Such situation means that network has *over-learned*.

When the training process has finished NN can be used for solving the real tasks and the weights are not changed anymore. According the set of input variables is fed to the network and almost instantly the calculated values of output variables appear at output layer neurons.

3. NEURAL NETWORKS TRAINING

Let us consider the PW125B engine diagnostic system in the case of the ground-based diagnostics using the DM when nine sensors are set up (see Figure 1). In general case the diagnostic apparatus has to localize defect in one of the following four diagnosable nodes of engines gas generator:

- low pressure compressor;
- high pressure compressor;

- low pressure turbine;
- high pressure turbine.

On the first stage using random-number generator the set of cases of defect development for the each diagnosable nodes (as the field of possible defects criteria) is created. In training process part of them makes up the training set, and part – the verification set.

It was supposed that single defects are the most probable. According to (1) the output vector of the diagnostic network contains four Boolean variables, each of them determines single defect of one diagnosable node ("true" or "false" correspondingly).

The function (3) with λ = 1 is taken as the activation function; the graph of it is shown in the Figure 4.

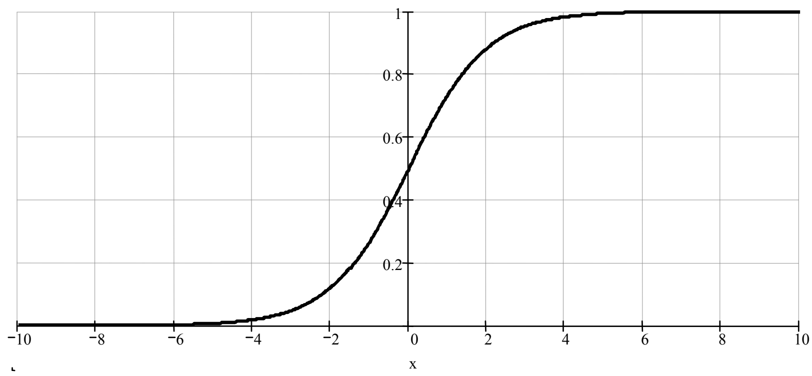

Figure 4. *Sigmoidal activation function with λ = 1*

Using STATISTICA Neural Networks 4.0 package the training of the NN was carried out. The stages 3, 4 and 5 of training process were executed automatically. Using the back-propagation method (4) the NN with 20 hidden neurons was created.

It is necessary to note that Neural Networks package automatically divides the initial set of defect development into three subsets: training, verification and test sets. The last one is used to compare received variants of networks. Anyway it only raises the effectiveness of the idea about training set division for the over-learning effect detecting. Such division and according classification errors (type 1 and 2 both) are shown in the Table 1. As it can be seen the initial set of defect development has 2000 cases.

Table 1. Classification errors in network with 9 inputs

Name of the set	*Size of the set*	*Average classification error, %*
Training set	1000	0,5
Verification set	500	0,4
Test set	500	0

From the received results it can be seen that over-learning hasn't occurred. Also the effectiveness of the network is practically ideal.

Next problem has to be solved is the training of considered NN when some sensors, which are necessary for the DM functioning, are absent or defective. For solution of this problem two approaches are offered.

1st approach. If the data from a single sensor (channel) are absent, then zero should be fed to the according input, but other input variables remain unchangeable. Then the process of the output variables calculation is carried out using (2). It is necessary to note that this formal approach is applied in practical task solving. As a rule, NN can perform correct classification in such cases too, but the error of classification can be significant.

2nd approach. Each NN should be trained using part of input variables and all output variables. In the case, when k inputs are absent, the set of C_m^k networks having $m–k$ inputs each will be created, where m is the total number of input variables. Thus primordial each network will be trained for the classification task in the case of some variable absence. As a rule, this approach is more effective than the first one, but it makes the whole system more complicated.

Let us consider examples of the offered approaches when some input variables are absent.

In given example one sensor is absent. On the first stage single universal NN with 9 inputs was trained. Then nine NN with 8 inputs were trained too, and there were considered 9 cases of one variable absence given in the first column of Table 2. Thus, the first network with 8 inputs was trained on all input variables except NH, the second – on all except NL, etc.

After the NN training the verification of created networks in the case of input information incompleteness was completed. The results of the first and second approaches are compared. The according errors of classification are shown in the Table 2, the graph of errors is presented in Figure 5.

Table 2. Errors of classification in universal and eight-inputs NN

Absent variable	*Average error of NN with 9 inputs (first approach), %*	*Average error of NN with 8 inputs (second approach), %*
NH	3.33	0.98
NL	1.06	0.69
T_LPT	0.7	0.65
WF	2.83	0.29
P_LPC	1.18	0.53
P_HPC	0.53	0.19
T_HPC	2.91	0.26
T_LPC	1.98	0.31
P_LPT	0.83	0.28

From these results the following conclusion could be made. It is necessary to avoid the universal network giving the preference to the set of networks with less input variables, because in this case the localization error will be smaller.

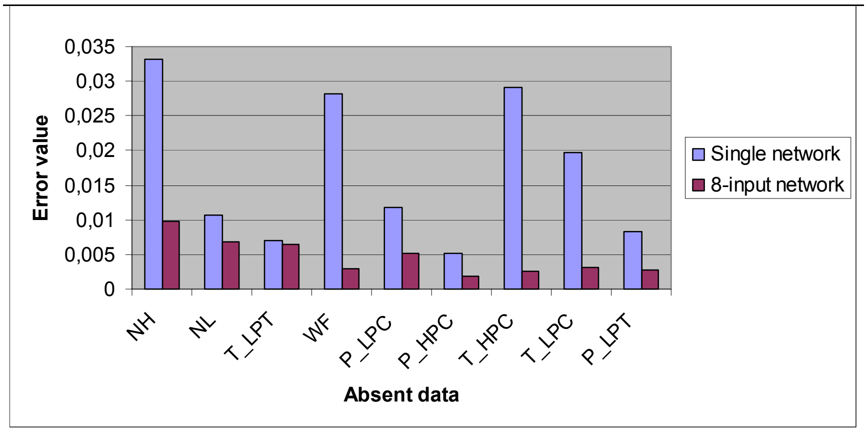

Figure 5. *Errors comparison of the universal and eight-inputs NN*

As it was mentioned before the additional five sensors in engine PW125B diagnostic system can be placed only on the ground because of the difficulties associated with the sensors setting up, as well as with additional information storing on board. Thus, the information only from four sensors is available during the flight. Using the approach applied before, let us train NN on some subset of the basic training set, namely only on four input variables which accord to sensors set up by a manufacturer for the engine flight model.

After the training process occurred analogically to the described above network with four inputs (NH, NL, T_LPT and WF), four outputs, and fourteen hidden neurons were received. Its effectiveness is illustrated in Table 3. From the higher values of errors (see above Table 1) it can be seen that the gathering of the information about the defects basing on four sensors was more complicated for the perceptron. But training time increasing accordingly can raise the efficiency of the network.

Table 3. Classification errors in network with 4 inputs

Name of the set	*Size of the set*	*Average error, %*
Training set	1000	2.5
Verification set	500	2.0
Test set	500	1.6

Let's consider the behaviour of the network in the case of the information incompleteness. For that purpose we will train the set of networks analogically to the method described above. The errors for such estimation are given in Table 4, the graph of errors is presented in Figure 6.

Table 4. Errors of classification in universal four-inputs NN

Absent variable	*Average error of universal network, %*	*Average error of the NN with one sensor absence, %*
NH	4.96	7.33
NL	10.6	1.18
T_LPT	5.93	1.98
WF	3.83	3.3

As it can be seen, the efficiency of the four-input network is evidently insufficient for the engine diagnostics when one sensor signal is absent and replaced by zero. Here especially it becomes apparent that the definite combination of the measurable parameters' deviations is in correlation with some defect. The decreasing the number of measurable parameters increases the weight of each measurable value according to its influence coefficient. The ponderability of the errors of parameters measurements also increases. It is because the change of absent signal to zero disturbs the "melody" of the defect, what excludes the DM usage. As a whole the carried calculations show that the set of NN trained on the cut DM information with less than 9 input channels allows to continue the diagnostic process during the flight, though in the case of the two-rotor gas generator four measures are the minimum when diagnostics preserves credibility.

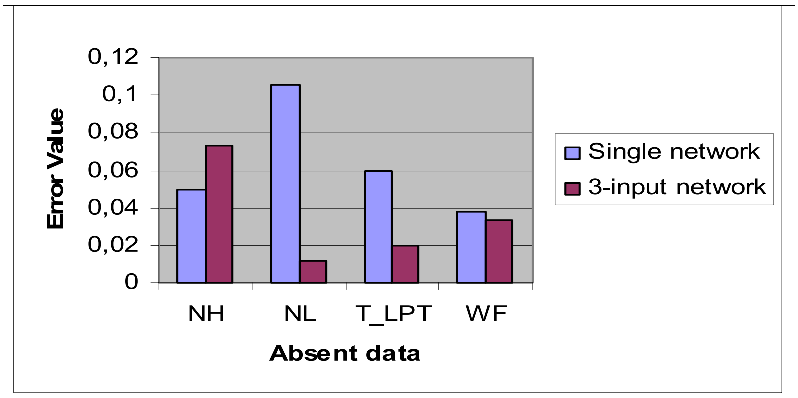

Figure 6. *Errors comparison of the four-input universal NN*

4. CONCLUSIONS

To increase the diagnostic systems reliability in flight in the case of partial information loss the using of artificial intellect systems in the form of neural networks as on-board software is suggested. NN have property of training in the case of information flow changes. In the paper we have considered the examples of NN training based on data obtained with the help of failures simulation using table of influence coefficients and DM. Diagnostic matrices developed for turboprop engine PW125B of Fokker-50, which is exploited in the "Riga" airport, are used as a training tool for NN. On the first stage of training the imitation model is used. It shows the NN usage possibilities in the diagnosis of aviation engines. On the second stage the real statistics of hazardous failures will allow developing of the real diagnostic model, which can be introduced in maintenance.

On the base of the suggested complex diagnostic tools the new generation of automatic engine diagnostic systems can be developed.

References

[1] Kuznetcov N.S., Labendik V.P. Peculiarities of Diagnostics Matrices Forming for Control of Gas Turbine Aeroengine Setting's Condition. In: *Proceedings of Institute of Higher Education "Aeronautical engineering "*. Kazan, KAI, 1993, № 3, pp. 97-101. (In Russian)

[2] Kopytov E., Labendik V., Kabelev N. Neural Network Application in Aircraft Engine Diagnostics. In: *Scientific Proceedings of the Scientific-technical Union of Mechanical Engineering*. *Second International Congress "Mechanical Engineering Technologies MT'01", June 24-26, 2001*, Sofia, Bulgaria, 2001, Vol.2 (57), pp.558-560.

[3] Kopytov E., Labendik V., Osis A., Tarasov A. Method of Aviation Engine Diagnostics in the Case of Partial Loss of Information. In: *Abstracts of II International Scientific and Technical Conference "Aero Engines of XXI Century"*, *December 6-9, 2005*. Moscow, Russia: CIAM, p. 266-267. (In Russian)

[4] Kopytov E., Labendik V., Osis A. Influence of Turbo-Prop Engine PW125B Parameters Measuring Errors on the Diagnostics Reliability of its Setting Conditions: Proceedings of the International Conference "RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION" (RelStat'03), 16-17 October 2003, Riga, Latvia, *Transport and Telecommunication,* 2004, Vol. 5 (2), pp. 112-117. (In Russian)