DIAGNOSING THE TECHNICAL CONDITION OF PLANETARY GEARBOX USING THE ARTIFICIAL NEURAL NETWORK BASED ON ANALYSIS OF NON-STATIONARY SIGNALS

Krzysztof POPIOLEK, Paweł PAWLIK

AGH University of Science and Technology, Faculty of Mechanical Engineering and Robotics
Department of Mechanics and Vibroacoustics
al. A. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: popiolekkrzysztof@gmail.com, pawlik@agh.edu.pl

Summary

This paper deals with the problem of diagnosing the technical condition of a planetary gearbox operating at variable load. The severity of the subject and related difficulties were discussed. Theoretical basis of analysis of non-stationary signals (order analysis) and its use in signal resampling was also presented. The paper tests the functionality of the planetary gearbox diagnostics method. The Multilayer Perceptron Network was used to identify and classify the damage. The network’s learning vectors were built on the basis of order analysis results and measurements of the planetary gearbox load. The functionality of two-layer and three-layer unidirectional artificial neural network was also analysed for potential use in diagnosing the technical condition of planetary gears.

Keywords: vibroacoustic diagnostics, order analysis, neural networks, planetary gearbox.

1. INTRODUCTION

Ensuring reliable operation of machinery and prevention of serious breakdowns means that procedures are needed which allow an early detection of faults and evaluation of technical condition of operating parts without the necessity to stop them. In addition, effective diagnostics systems can reduce the costs of routine maintenance and minimize the risk of an unplanned downtime. The main reason, accounting for 60% failures of rotating machines, is the gear train damage, of which 24% is caused by ineffective maintenance [1][2]. In addition, the gear train is a serial member of the reliability structure [3] where a failure of any member affects the reliability of the whole system. Consequently, there is a growing demand for monitoring the technical condition of large gear trains. The variety of requirements and operating conditions has resulted in a dynamic growth and development of newer and newer diagnostic methods.

Due to their numerous advantages, planetary gearboxes find application in complex mechanical structures, such as mining machines, helicopters, wind turbines, ships. Large ratios and favourable distribution of forces in planetary gearboxes allow to transmit much higher torques while maintaining a more compact size than traditional gear trains. Planetary gearboxes are often used in very difficult conditions and at high loads.
A particularly challenging task is the monitoring of technical condition of planetary gearboxes based on the vibration signal at operating conditions which differ in time (varying load of the gearbox). The varying load results in changes of vibration signal amplitude and also of the rotational speed. Absence of connection between the diagnostic features and the value of drive system load may result in a failure to detect a damage or in an incorrect detection [4][5][6].

Artificial neural network based on vibroacoustic signals has been so far successfully implemented, showing good precision for failure detection and condition monitoring of spur gears [7], bearings [8][9] and other complex mechanical systems [10][11]. There are also several other researches regarding planetary gearbox diagnostics using various pattern recognition methods [12][13][14].

The paper presents the order analysis based on the resampling of the signal relative to the signal from the rotational speed sensor (tachometer). This method allows diagnosing the technical conditions of machines operating at variable rotational speed. The analysis includes also the braking torque in order to account for the impact of the load on the values of diagnostic features. As for each type of machine, the load impact is different and sometimes difficult to define a priori, the neural networks as the artificial intelligence methods were used to detect and identify the damage.

This paper is structured as follows. In Section 2 the diagnostic experiment is described. Section 3 contains the methods of signal analysis used in the preparation of training vectors. Review of learning vectors correctness is presented in section 4. The architectures of neural networks were discussed in Section 5.

2. DIAGNOSTIC EXPERIMENT

The tests were conducted at the Department of Mechanics and Vibroacoustics of AGH University of Science and Technology. The diagnosed object was a Rexnord Mercury 1-A planet gear with ratio of 3.75. The laboratory setup is shown in Figure 1.

The acceleration was measured with the PCB 356B08 tri-axial piezoelectric sensor installed on the gearbox housing. The directions of axes are presented in Figure 2 (X – horizontally, perpendicular to the input shaft, Y – parallel to the input shaft, Z – vertically).

The input shaft speed was recorded with the Brüel&Kjær MM0360 tachometer; the signal from the tachometer allowed to synchronize the vibroacoustic signal with the rotational speed. The setup also allowed to control the powder brake by sending voltage signals proportional to the braking torque. The controller also allows to measure the current in the powder brake coil which is proportional to the braking torque.

The laboratory measurements were performed in two series, for selected gearbox operation settings and all faults which could be simulated in the test setup: input shaft misalignment - misalignment angle 1º) and flexible coupling unbalance (unbalance weight - 21g). The first series of measurements, or – to be more precise – the vectors built on the basis of its results, served as a learning sequence used in the process of teaching the neural networks how to recognize the operation conditions of the gearbox. The second series of measurements was used to build validation vectors which checked the correct operation of the artificial neural network. All gearbox settings tested during the experiment are given in Tables 1. Measurements were carried out for three operational states: good, misalignment and unbalance. The measured parameters were the same in both series of measurements. Additionally, the second series included measurements for two simultaneous faults, a so-called double damage (misalignment and unbalance).

Tab. 1. Planetary gearbox operation settings in the first series of measurements
Parameters of voltage signal sent to the brake

<table>
<thead>
<tr>
<th>Speed [rpm]</th>
<th>Signal type</th>
<th>Frequency [Hz]</th>
<th>Amplitude [Nm]</th>
<th>Constant (DC) component [Nm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>Constant</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>26.5</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>Constant</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>26.5</td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>Constant</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Sinusoidal</td>
<td>1.1</td>
<td>26.5</td>
<td></td>
</tr>
</tbody>
</table>

3. SELECTION OF DATA FOR LEARNING VECTORS

In order to determine the gearbox diagnostic parameters, the vibration acceleration signals were subjected to order analysis. The order spectrum is obtained by means of resampling of the vibration time signal relative to the input shaft rotational speed. Figure 2 presents a diagram of the order analysis algorithm. First, the signal from tachometer is interpolated using a cascaded integrator-comb (CIC). Then, the filtered signal from tachometer is used to resample the vibration signal in order to determine the even angle signal. Such resampled signal can be subjected to the Fourier Fast Transform (FFT) which converts the frequency to order numbers corresponding to multiples of the input shaft rotational frequency [15].

Using the order analysis we can determine the amplitude and phase of a chosen order in time. Rotational frequency of the input shaft corresponds to order No. 1. Shaft misalignment is a source of vibration with double and triple rotational frequency [16][17], hence the input shaft misalignment will produce changes in the amplitude of order No. 2 and 3. Unbalance will be observed in the amplitude of order No. 1. Figure 3 presents the amplitude of order No. 1 vs. time.

In addition to the technical condition, the changes of the order spectrum amplitude are also affected by variations of the system load. Consequently, also taken into consideration was the value of brake coil current which is proportional to the load. The signal from the brake (Figure 4) and amplitudes of individual orders are similar (Figure 3).

The vibration amplitude values and voltage values were averaged for 1-second intervals to reduce the amount of input data and obtain better separation between clusters corresponding to the gearbox operation conditions. The obtained averaged vectors, for speed 1500 rpm and sinusoidal load with frequency 1.1 Hz, are presented in figures 5, 6, 7 and 8.
The learning vectors were built using the data from the vibration acceleration signal on axes X and Y, focusing on orders 1 and 2, and the load torque values. The final form of learning vectors is shown in table 2. For the purposes of article were built eight training vectors for each combination of motor speed and load type, where each of them corresponds to one second measurement. As a result, were built 72 learning vectors for each damage.

| Tab. 2. Learning vectors consisting of 5 elements for each time instant |
|---------------------------------------------------|------------------|------------------|------------------|------------------|
| Learning vector 1 | Learning vector 2 | Learning vector 3 | Learning vector n |
| Load on brake [Nm] | 26.65 | 26.59 | 26.65 | ... |
| Amplitude of order 1 – X axis [m/s²] | 0.129 | 0.130 | 0.131 | ... |
| Amplitude of order 1 – Y axis [m/s²] | 0.054 | 0.052 | 0.054 | ... |
| Amplitude of order 2 – X axis [m/s²] | 0.024 | 0.019 | 0.018 | ... |
| Amplitude of order 2 – Y axis [m/s²] | 0.064 | 0.059 | 0.057 | ... |

4. REVIEW OF LEARNING VECTORS CORRECTNESS

The $k$-means clustering analysis was performed in order to verify the correctness of learning vectors built by the author. The analysis was performed in the MatLab environment with addition of Statistics and Machine Learning Toolbox™. The cluster analysis finally leads to grouping data into clusters so that the elements in the same group are as similar to each other as possible, and the elements from various groups are as dissimilar as possible. The measure of similarity depends on the area of application [18]. Consequently, the cluster analysis helps find structure in the data set without interpretation (without justification of their occurrence). The difference between the cluster analysis and statistical tests (methods) lies in the fact that the former is used without any “a priori hypotheses, during the exploratory phase of research.”

A silhouette plot was made using the group indices in order to check the quality of separation between clusters. The plot (Figure 9) presents the proximity of each point of one cluster to the neighbouring clusters. The measure of proximity (silhouette value) is in the $<-1,1>$ range, where 1 represents points very distant from other clusters, 0 represents points not included distinctly in any cluster, and -1 indicates points which are probably assigned to a wrong cluster.

The elements of cluster 1 have values from 0.3 to 0.55, elements of cluster 2 – from 0.2 to 0.4, elements of cluster 3 – from 0.7 to 0.85. This means that the elements in clusters separate in the learning vectors space. Large silhouette values were obtained in cluster 3. The group of these elements has the best separation. In remaining two clusters the separation is not so distinct, but still satisfactory. In addition,
no element was grouped wrongly, hence we can say that the choice of learning vectors is correct.

![Silhouette plot obtained by cluster analysis of learning vectors space with k-mean clustering](image)

**Fig. 9.** Silhouette plot obtained by cluster analysis of learning vectors space with k-mean clustering (1, 2, 3 – indices of successive clusters)

5. CHOOSING THE ARTIFICIAL NEURAL NETWORK ARCHITECTURE

5.1. Two-layer Network

We tested the multilayer unidirectional artificial neural network. The initial analysis was performed on the two-layer network (hidden layer and output layer). The output layer comprises two neurons. The desirable outputs from the network, corresponding to individual gearbox operation conditions, are presented in Table 3.

**Tab. 3.** Desirable outputs from the multilayer unidirectional network and corresponding gearbox operation conditions

<table>
<thead>
<tr>
<th>Neuron in output layer</th>
<th>Gearbox condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>neuron 1</td>
<td>neuron 2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The initial number of neurons in the hidden layer in the first tested network was specified as the number of inputs and outputs divided by 2. In this case, there were 4 neurons.

Sigmoid bipolar neuron activation functions were used in both layers. Training of the networks was performed using the Levenberg–Marquardt algorithm which is usually the fastest method for training the unidirectional networks. The parameter verifying the network performance was a mean-square error between the network outputs and defined target vector. The learning with the *train* function takes place in the batch mode, meaning that all learning vectors are sent to the network inputs before the synaptic weights are changed. This learning mode is significantly faster and generates smaller errors than the mode in which the weights are updated after each vector is sent to the network [19]. The input data are normalized so that they are included in the <1,1> range. In addition, a popular practice during the training of neural networks is to divide the data into three subsets. The first of them, the training subset, is used to calculate the gradient and change of weights and biases. The second, the validation subset, is used to monitor the error during the training. Usually, the validation error decreases in the beginning of the learning process just as the training subset error. However, a possible increase of the validation error may indicate the network overfitting. The values of weights and biases in the network are recorded for the least validation subset error. And finally, the third test subset is not used during the learning process; it can be used for comparing various models [19].

The division into subsets in our network was according to the default values: 70% - training subset, 15% - validation subset, 15% - test subset.

The first architecture of the two-layer unidirectional network is shown in Figure 10.

**Fig. 10.** The first unidirectional two-layer network created in MatLab. Where: 5 – number of input vector elements, 4 – number of neurons in hidden layer, 2 – number of output neurons (network outputs), w – weights vector, b – biases vector

As on the output we obtain the values from the <0,1> interval, three intervals were created corresponding to the quality of classification of gearbox operation condition by the network. The classification correctness criterion was absolute value of the difference between the target value and the value obtained on the output of both neurons $d_1$ and $d_2$:

$$d_1, d_2 \in <0, 0.25> - \text{accurate classification};$$

$$d_1, d_2 \in (0.25, 0.5) - \text{uncertain classification};$$

$$d_1, d_2 \in <0.5, 1> - \text{wrong classification}.$$  

The network was trained ten times using the data from the first series of measurements. The correctness of network performance was verified by applying to the network inputs the vectors from the second series of measurements and also the vectors with double damage.

Every ten learning and validation processes the network architecture was changed by adding neurons in the hidden layer and the process was repeated so that each architecture was trained ten times. As we
can see in Figure 11, with increased number of neurons in the hidden layer, the number of epochs needed to train the network decreased.

Fig. 11. Number of epochs to train the multi-layer network vs. number of neurons in the layer

Along with the increased number of neurons, the mean-square error was reduced too (Figure 12) and the classification correctness was increased.

Fig. 12. Mean-square error vs. number of neurons in the layer

The unidirectional network correctly classified 99.86% of validation vectors with as few as four neurons in the hidden layer. The biggest problem for the network was generalization of knowledge and correct classification of two damages at the same time (misalignment and unbalance). 13.26% of correct classifications were obtained with 4 neurons in the hidden layer, and the result was improved to 57.53% with 22 neurons.

Networks containing from 4 to 300 neurons in the hidden layer were tested in order to determine the number of neurons needed for maximum classification for this type of network. Trailing was repeated ten times for each tested number of neurons, and then the average value was calculated. The performance obtained for all two-layer networks depending on the number of neurons in the hidden layer is presented in Figure 13.

As we can see in Figure 13, the trend value of correct classification of double damage reached its maximum at 64% with 83 neurons. The decrease of correct classifications may have been caused by network overfitting, the phenomenon in which the neural network with too many neurons in relation to the learning data adapts itself to accidental errors on the learning sequence, thus losing its ability to generalize and apply to other similar data.

Fig. 13. Correct classifications vs. number of neurons in the hidden network layer (sigmoid bipolar activation functions)

The ratio of uncertain classification for this architecture was 6.38%, and of erroneous classifications - 14.71%. 100% of accurate classifications were obtained for validation vectors for single operation conditions.

5.2. Three-layer network – case I

Further tests were performed on a three-layer network with two hidden layers and one output layer. The second hidden layer had half of the neurons from the first layer, rounded down. We tested networks from 4 neurons in the first hidden layer, 2 neurons in the second hidden layer, and 2 neurons in the output layer, up to 100, 50 and 2 neurons, respectively. We used network models with different neuron activation functions.

Case I included only sigmoid bipolar functions – Figure 14.

Fig. 14. Architecture of unidirectional three-layer network (all neuron activation functions are sigmoid bipolar). Where: 5 – number of network inputs, 100 – number of neurons in hidden layer 1, 50 – number of neurons in hidden layer 2, 2 – number of output neurons (network outputs), w – weights vector, b – biases vector

This network architecture was also tested for performance depending on the number of neurons in the hidden layers. The best result (57.39%) was obtained for 70 neurons in the first hidden layer and 35 neurons in the second hidden layer, with 7.67% of uncertain classifications and 34.94% of erroneous classifications. 100% of accurate classifications were obtained for validation vectors for single operation conditions.
Generally, the results for the three-layer network were worse than for the two-layer network. In addition, the learning time was significantly longer.

5.3. Three-layer network – case II

In successive layers we used the following activation functions: linear, sigmoid bipolar, sigmoid bipolar – Figure 15.

The best average results of double damage classification correctness for this network were as follows: 77.17% of accurate classifications, 6.52% of uncertain classifications, and 16.31% of erroneous classifications. All validation vectors with a single damage were classified correctly.

![Fig. 15. Architecture of unidirectional three-layer network](image)

Fig. 15. Architecture of unidirectional three-layer network (neuron activation functions in successive layers: linear, sigmoid bipolar, sigmoid bipolar).

Where: 5 – number of network inputs, 80 – number of neurons in hidden layer 1, 40 – number of neurons in hidden layer 2, 2 – number of output neurons (network outputs), w – weights vector, b – biases vector

6. SUMMARY

The analysis and processing of non-stationary signals measured on a real object at the planetary gearbox test stand in the Department of Mechanics and Vibroacoustics of AGH University of Science and Technology was performed using the order analysis based on the signal resampling. A multithreaded program was developed in the LabVIEW environment to conduct the order analysis, extract the diagnostic features and create the learning vectors.

The correct choice of the learning vectors space was verified using the k-means clustering in the MatLab environment. A good separation was obtained for three clusters.

A functional analysis was performed on the artificial neural network with two-layer and three-layer unidirectional architecture, trained with the Levenberg–Marquardt algorithm.

A program was written in the MatLab environment in order to select the correct number of neurons and network layers. The program trained a given architecture ten times and recorded the results in a matrix. If the target operational correctness was not achieved, the network architecture was modified by adding more neurons and the learning process was repeated. The program proved helpful in finding the structure which fits the task the best. We obtained 100% of correct classifications of vectors from the second (validation) series of measurements. On the other hand, the tested multilayer networks showed different capability of generalizing the knowledge (testing with double-damage vectors) depending on the number of layers and the activation functions. The best results were obtained for the two-layer network with sigmoid bipolar neuron activation functions: 78.91% of accurate classifications, 6.38% of uncertain classifications, and 14.71% of erroneous classifications with 83 neurons in the hidden network hidden layer. In addition, the learning process in those networks was the fastest.

The performance was worse for three-layer networks. With one linear activation function and two sigmoid bipolar functions in the successive layers, the results were: 77.17% of accurate classifications, 6.52% of uncertain classifications, and 16.31% of erroneous classifications. The least classification correctness was obtained in case of three-layer networks with sigmoid bipolar activation functions in all layers: 57.39% of accurate classifications, 7.67% of uncertain classifications, and 34.94% of erroneous classifications. On top of that, the learning time of three-layer networks was much longer than of two-layer ones.

The obtained results prove the effectiveness of proposed methods in detecting and diagnosing the faulty operation conditions of a planetary gearbox operating at variable load. The proposed methods can be a valuable part of the diagnostic and monitoring systems, contributing to the increased quality of work, extended life and better safety of machinery and equipment.

ACKNOWLEDGMENT

The study has been developed under the statutory investigation of the Department of Mechanics and Vibroacoustics of the AGH University of Science and Technology.

REFERENCES


[5] Zimroz R. Modelowanie sygnałów drganiowych generowanych przez przekładnie planetarne w warunkach zmiennego obciążenia [Modelling of...
Vibration Signals Generated by Planetary Gearboxes under Variable Load [Diagnostyka'3 (39)/2006


[16] Cempel C. Diagnostyka Wibroakustyczna Maszyn [Vibroacoustic Diagnostics of Machines], Warszawa 1989 PWN.


Krzysztof POPIOŁEK, received Master of Science degree at the Faculty of Mechanical Engineering and Robotics, AGH University of Science and Technology, Cracow, Poland, in 2015. He currently works at GE Aviation, developing his passions: Machine Learning and Data Science.

Paweł PAWLIK, received PhD degree in Faculty of Mechanical Engineering and Robotics from AGH University of Science and Technology, Cracow, Poland, in 2013. In the scientific work he builds vibro-diagnostic monitoring systems and develops methods of signals processing. He is also involved in the development of non-classical methods of uncertainty assessment.