Wavelet decomposition algorithm for machine learning model in wind turbines

M Pukhova, E Lopatin, N Sokolinskaya and S Tarakanov
Financial University under the Government of the Russian Federation, 49, Leningradsky Avenue, Moscow, 125468, Russia

E-mail: evgeniy_lopatin_97@mail.ru

Abstract. The paper considers the wavelet decomposition for machine learning model to solve the problem of analyzing the shape of vibration signals in wind energy turbines. The questions of training and optimization of the parameters of the classmate method are highlighted. The principle of constructing a training sample is described. The algorithm for the instance recognition of circuit elements by vibration signals removed from wind turbines. Thus, the method of recognition of equipment elements based on sparse wavelet decomposition of vibration signals can be used in practical renewable energy. The research results will stimulate of the development of generating facilities based on wind energy with an installed capacity up to 15 kW will contribute to the stimulation of the development of wind power with a vertical axis in the urban environment.

1. Introduction
The recently adopted concept of industrial energy development “Industry 4.0” implies a shift from routine maintenance of industrial equipment to service. A prerequisite for the implementation of such a transition is there are fully automatic vibration diagnostic energy systems integrated with a single enterprise management system. The difference between energy automatic energy systems and expert oriented is the complete exclusion of the person (expert) from the decision-making process about the technical condition of the equipment [1, 2].

The reliability of the operation of energy automatic energy systems is determined by the methods for isolating static signs characterizing the state of industrial equipment. In many cases, insufficient and sometimes excessive diagnostic information makes acceptance difficult right decision [3,4].

There are two ways to increase the reliability of energy automatic energy systems for vibration diagnostics:

- development of auxiliary methods that increase information content, stability and authenticity of expert-oriented spaces of diagnostic features;
- development of fundamentally new methods and approaches for assessing the technical condition of mining, assessment of qualitatively new features, the creation of invariant feature spaces.
2. Materials and methods

This is an insufficiently generated training set for the recognition of gears, the formation of which did not take into account the models of impact processes of gears at kink teeth [5,6]. Moreover, if we assume that each type of equipment item has a set of its own, form templates that are unique to him, and then based on a compact description of signals in the temporal region [7, 8, 9]. A system of recognition of these elements can be constructed. This kind of energy system will be relevant in cases of vibrational diagnostics of equipment with complete or partial absence its kinematic scheme [10, 11]. An example of this is the process of the appearance of shock processes in electricity networks.

Except as indicated, the sensitivity of the developed recognition method items of equipment lies in the range of $TPR = [0.706, 1]$ with recognition accuracy $PPV = [0.733, 1]$. It should be noted that the sensitivity and recognition accuracy of a defect-free bearing rolling on average lower than defective, which is reflected in lower $TPR$ and the $PPV$ for a dataset «Bearing 6213 Norm / an Dataset OR» compared with «Bearing an Dataset 6213 OR» (Table 2). The reason for this is either the complete absence of shock processes in the signal of a defect-free bearing, or their insignificant energy contribution to the total energy of the signal.

It is highly reliable to define the signal as “bearing”. The average values of the $TPR$ and $PPV$ parameters for all the algorithm of the signals ($N = 1115$ pcs.) amounted to $TPR = 0.882$ and $PPV = 0.972$. This position was proved in researches about energy market [12, 13].

3. Results

In order to obtain the elements of equipment, experiments were conducted on sets of vibrational signals. For of each data set, several resonant frequencies $f_1 - f_P$ were selected. In this case, the desired element (true) for each frequency $f_1$ was set by expert analysis of spectral features in its vicinity.

Table 1 presents the results of recognition of a rolling bearing by single classic OCSVM fixtures based on various attribute spaces for the Bearing 6213 OR Dataset.

| Parameters | $f_1 = 600$ | Frequency, Hz | $f_2 = 1100$ | $f_3 = 3050$ | $f_4 = 5200$ |
|------------|------------|---------------|---------------|---------------|---------------|
|            | $TPR$ | $PPV$ | $TPR$ | $PPV$ | $TPR$ | $PPV$ | $TPR$ | $PPV$ |
| BFS        | 0.079 | 1 | 0.219 | 1 | 0.072 | 1 | 0.040 | 1 |
| BFS-C-I    | 0.158 | 1 | 0.610 | 1 | 0.247 | 1 | 0.182 | 1 |
| BFS-C-II   | 0.099 | 1 | 0.276 | 1 | 0.083 | 1 | 0.040 | 1 |
| BFS-F      | 0.436 | 0.917 | 0.741 | 1 | 0.701 | 1 | 0.717 | 1 |
| BFS-wn     | 0.079 | 1 | 0.219 | 1 | 0.052 | 1 | 0.040 | 1 |
| BFS-wn     | 0.168 | 1 | 0.591 | 1 | 0.165 | 1 | 0.182 | 1 |
| C-I        | 0.109 | 1 | 0.276 | 1 | 0.052 | 1 | 0.040 | 1 |
| C-II       | 0.396 | 1 | 0.687 | 1 | 0.588 | 1 | 0.667 | 1 |
| BFS-wn-cl  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| BFS-wn-cl  | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cl-C-I     | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cl-C-II    | 0.465 | 0.940 | 0.105 | 1 | 0.381 | 1 | 0.879 | 1 |
| cl-F       | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
Recognition was carried out for four selected frequencies $f_1 - f_4$, decomposition efficiency signal. In the vicinity it exceeded the threshold value. For classifiers based on each characteristic spaces TPR sensitivity and PPV accuracy parameters were calculated. The result contains an assessment of the quality of the algorithm for instance recognition a set of classifiers presented.

In accordance with the results presented in table 1, we can conclude that the joint the use of a set of independent classifiers can significantly increase the sensory consistency (TPR) and accuracy (PPV) of element recognition in comparison with single classifier (based on one attribute space). Therefore, for example, for frequency $f_3$ the sensor recognition of an element by individual classifiers varies in the range from 0 to 0.701, while the resulting value is $TPR = 0.897$.

The authors do not consider in detail each type of feature space presented in table 1, as this is not included in the subject of this article. It should only be noted that OCSVM classifiers based on the BFS-F attribute space have the highest sensitivity and its modifications describing the distribution of signal energy over families of basis functions MEXH, MEXP, MORL, and SINP. The space BFS-F is integral in nature, its dimension ($dim = 3-4$) is significantly lower than the dimension of the original BF spectrum ($dim = 30$).

For this reason, classifiers based on BFS-F have, on average, lower recognition accuracy than classifiers based on other types of spaces (Table 1), but greater sensitivity.

Table 2 summarizes the results of recognition of equipment elements for all used data sets. In the column "Dataset / Parameters" next to the name of each set, the data in parentheses indicates the type of element sought: (G) - gear, (B) – bearing rolling. Four significant frequencies $f_1 - f_4$ were allocated in each data set.

| Dataset / Parameters                  | Frequency |
|--------------------------------------|-----------|
|                                      | $f_1$     | $f_2$     | $f_3$     | $f_4$     |
|                                      | TPR       | PPV       | TPR       | PPV       | TPR       | PPV       | TPR       | PPV       |
| High Speed Gearbox Dataset (G)       | 0.589     | 0.786     | 0.923     | 1         | 0.706     | 0.733     | 0.804     | 0.945     |
| High Speed Bearing Dataset (G)       | 1         | 1         | 0.951     | 1         | 0.782     | 0.932     | 0.742     | 1         |
| Bearing 6213 OR Dataset (B)          | 0.921     | 0.960     | 1         | 1         | 0.897     | 1         | 0.950     | 1         |
| Bearing 6213 Norm/OR Dataset (B)     | 0.821     | 0.92     | 1         | 1         | 0.865     | 1         | 0.756     | 0.982     |

Source: Matlab.

In accordance with the results of experiments for various data sets, we can but conclude that the algorithm for instance recognition of equipment elements based on a set of one-class classifiers is able to recognize items of equipment with enough high reliability.
The least sensitivity of the method ($\text{TPR} = 0.589$) was recorded for vibrating signals. High-speed gearbox dataset is in the vicinity of frequency $f_1$.

4. Conclusion
The paper considers the space of informative features of the sectional algorithm. The wavelet decomposition to solve the problem of analyzing the shape of vibration signals. Proposed by algorithm for instance recognition of equipment elements based on sparse wavelet decomposition [14, 15, 16].

The questions of training and optimization of the parameters of the classmate method are highlighted. Features of One-Class SVM. The principle of constructing a training sample generator for the problem is briefly described.

The algorithm for the instance recognition of circuit elements by vibration signals removed from rolling bearings and gears [17, 18, 19].

The average values of the parameters of TPR and accuracy PPV recognition amounted TPR = 0.882 and PPV = 0.972, respectively. Thus, the method of recognition of equipment elements based on sparse wavelet decomposition of vibration signals can be used in practice energy.

Acknowledgements
The article is based on the results of research carried out at the expense of budget funds on the state task of the Financial University.

References
[1] Dzeroski S and Zenko B 2004 Is combining classifiers with stacking better than selecting the best one? Mach. Learn. 54 255–73
[2] Freund Y and Schapire R 1996 Experiments with a new boosting algorithm Proceedings of the 13th international conference on machine learning 148–56
[3] Friedman M 1940 A comparison of alternative tests of significance for the problem of m rankings Ann. Math. Stat. 11 56–92
[4] Goldberg D, Nichols D, Oki B M and Terry D 1992 Using collaborative filtering to weave an information tapestry Communications of the ACM 35(12) 61–70
[5] Dayong N, Mikhaylov A, Bratanovsky S, Shaikh Z A and Stepanova D 2020 Mathematical modeling of the technological processes of catering products production Journal of Food Process Engineering 43(2) e13340 https://doi.org/10.1111/jfpe.13340
[6] Demsar J 2006 Statistical comparisons of classifiers over multiple data sets J. Mach. Learn Res 7 130
[7] Mikhaylov A 2018 Volatility Spillover Effect between Stock and Exchange Rate in Oil Exporting Countries International Journal of Energy Economics and Policy 8(3) 321-326
[8] Dietterich T 2000 Ensemble methods in machine learning Lect. Notes Comput. Sci. 1857 115
[9] Dunn O 1961 Multiple comparisons among means Am Stat. Assoc. 56 52–64
[10] Meynkhard A 2019 Fair market value of bitcoin: halving effect Investment Management and Financial Innovations 16(4) 72-85
[11] Nyangarika A Mikhaylov A and Richter U 2019 Oil Price Factors: Forecasting on the Base of Modified Auto-regressive Integrated Moving Average Model International Journal of Energy Economics and Policy 9(1) 149-60
[12] Nyangarika A Mikhaylov A and Tang B-J 2018 Correlation of Oil Prices and Gross Domestic Product in Oil Producing Countries International Journal of Energy Economics and Policy 8(5) 42-48
[13] Meynkhard A 2020 Priorities of Russian Energy Policy in Russian-Chinese Relations International Journal of Energy Economics and Policy 10(1) 65-71
[14] Meynkhard A 2019 Energy Efficient Development Model for Regions of the Russian Federation: Evidence of Crypto Mining International Journal of Energy Economics and Policy 9(4) 16-21
[15] Abdi H 2007 The method of least squares. Encyclopedia of Measurement and Statistics (CA, USA: Thousand Oaks)

[16] An J Dorofeev M and Zhu S (2020) Development of Energy Cooperation Between Russia and China International Journal of Energy Economics and Policy 10(1) 134-9

[17] An J and Dorofeev M (2019) Short-term FX forecasting: decision making on the base of expert polls Investment Management and Financial Innovations 16(4) 72-85

[18] Denisova V 2019 Energy efficiency as a way to ecological safety: evidence from Russia International Journal of Energy Economics and Policy 9(5) 32-7

[19] Denisova V, Mikhaylov A and Lopatin E 2019 Blockchain Infrastructure and Growth of Global Power Consumption International Journal of Energy Economics and Policy 9(4) 22-9

[20] Mikhaylov A 2018 Pricing in Oil Market and Using Probit Model for Analysis of Stock Market Effects International Journal of Energy Economics and Policy 8(2) 69-73