THE STUDY OF ANOMALOUS MACHINE SOUND DETECTION BASED ON CYCLOSTATIONARITY MODEL

Technical Report

Vladimir Klychnikov

Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
klyychnikovvv@student.bmstu.ru

Dmitriy Lapin[*]

Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
lapindv@bmstu.ru

Mark Hubbatulin

Bauman Moscow State Technical University
2-nd Baumanskaya St., 5, bld 1, Moscow, 105005, Russia
khmea15u755@student.bmstu.ru

ABSTRACT

In industrial predictive maintenance, one of the most important direction in Industry 4.0, machine monitoring and diagnostics is critical part of its operation. Non-contact acoustic data gathering is particular interest because of high ergonomics and low costs. The general method for processing such type of data is anomalous sound detection. This method allows express diagnostics of machines and units with minimum integrations. Based on DCASE 2020 Challenge, the study of the proposed method was presented. Problem description with physical interpretation and model elements review was conducted. Model based on Winger-Ville transform with architecture improvement and ensemble score calculation was developed. Model results on provided development dataset were calculated. Discussion of model results with assumptions for further research and development was shown. Conclusion about present study and future work was received.

***Index Terms*—** anomalous sound detection, machine condition monitoring, cyclostationarity processes, pseudo Winger-Ville transform, deep learning, autoencoder

1. INTRODUCTION

One of the most important direction in Industry 4.0 is predictive maintenance based on monitoring and diagnostic of machines and tools [1, 2]. Non-contact data mining [3] is particular interest because it provides non-destructive testing on important industrial units. Non-contact vibroacoustic data gathering is particular interest because of high ergonomics, low cost and ability to collect data from one or more objects simultaneously [4-10]. The general method for processing such type of data is anomalous sound detection [11-13], that allows express diagnostics of machines and units with minimum integrations and without complete knowledge of how the units work. Methods of machine health monitoring are presented on figure 1.

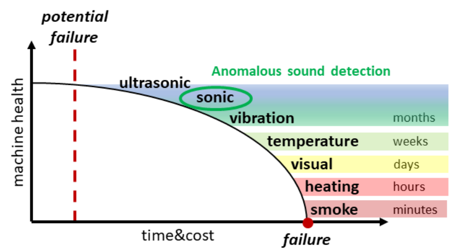


Figure 1: Methods of machine health monitoring.

The goal of this study, based on DCASE 2020 Challenge [14], is to analyze and synthesize anomalous machine sound detection model on provided datasets [15-17]. To achieve this goal, the following tasks need to be accomplished:

- to analyze problem description;
- to detail model elements and merge them;
- to calculate results of model;
- to discuss calculated result for suggestions;
- to draw conclusions about future work.

2. PROBLEM DESCRIPTION

Challenge task was explained in detail in the enclosed materials [14]. Provided datasets were descripted in accompanying papers [15-17]. This analysis is aimed at the physical interpretation of data mining methods and model elements review, that partially entered in the final model.

2.1. Physical interpretation

Based on the description of the experiments and our experience in diagnostics on machine sound data mining, a number of peculiarities are identified that determined the model concept.

Mechanical systems in this challenge – e.g. fan, slider, valve, pump, toy car, toy conveyer and their parts —sustain periodic motion of their components which in turn periodically modulate the vibration or noise they radiate. The cyclostationary property covers a rich statistical typology of signals including periodic signals and stationary random signals as particular cases. In particular,

[*] Thanks to LCC «Ctrl2GO» for funding.

because it is intuitive and physically meaningful, the approach relies as much as possible on the key concept of the energy conveyed by a signal: this will make easier the introduction of the various types of densities that characterize cyclostationary signals in the time, the frequency and the cyclic domains [18].

However, in order to use the cyclostationarity theory, it is necessary to know more about experiments on recording datasets, in particular the microphone parameters and technical specification of machines and their modes of operation.

2.2. Model elements review

The following hypotheses are considered individually and together according to numbered elements of general pipeline, shown on figure 2.

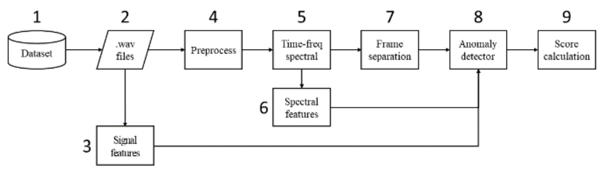


Figure 2: General pipeline of anomalous sound detection.

Data augmentation for expansion datasets on machines and on id (1 and 2 fig. 1). Since the operation sounds are presented for different machines, it may not be correct to create additional spectrograms in different modes of operation. Different augmentations showed completely different results for different machines and id. And since the estimated dataset will have an id that the model has not yet seen, it has been decided to reject the augmentation.

Features for more information from signal (3 fig. 1). An attempt to use signal features, such as – min, max, mean, 25 and 75 quartiles and fundamental frequency estimation [19] – does not improve the quality of anomaly detection, but on the contrary, worsens it. It is caused by the complex data structure of the source data and different states of machines on the records.

Sound data denoising to extract useful signal (4 fig. 1). The sound data of the target machine was noisy with sounds from the factory, so it was suggested that data cleansing can improve the score of the model. To do this, we have removed and tried to clean up areas where there is plant noise using correlation and REPET-SIM methods as shown in figure 3 [20]. But the adoption of these methods decreased the scores, perhaps because the defects are manifested exactly in such noise of the object, which this method considers the background noise of the factory.

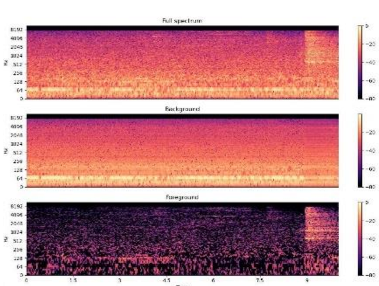


Figure 3: Example of spectrograms before and after pre-processing on FAN machine.

Time-frequency spectral transform on signal (5 fig. 1). In the process of work several types of transformations were applied: wavelets [21], constant-Q transform [22], MFCC [23], etc. After analysis of the received spectrograms it was found out that the nature of the abnormal signals does not allow to localize them at low resolution of the spectral picture, the selection of the target signal became impossible. To solve this problem, the window pseudo Winger-Ville transform was used, which currently allows us to obtain a better detailed spectral picture.

Features for more information from spectral transform (6 fig. 1). Using spectral descriptors [24] also does not improve the quality of anomaly detection, but on the contrary, worsens it for the similar cause as for signal features.

Different types of autoencoder anomaly detectors (8 fig. 1). Convolutional, variational and denoising autoencoders, pre-train encoder VGGish [25] were used. But in the current study these solutions did not allow to increase the final result.

After analysis, the highest result was achieved by window pseudo Winger-Ville transform together with mel-scaled spectrogram in concatenated frames on reshape baseline autoencoder.

Next, the selected elements and their merge will be discussed in more detail.

3. MODEL ARCHITECTURE

3.1. Feature extraction

For this study of anomalous machine sound detection of cyclostationarity processes it is important to have high resolution on spectrogram. both in time and frequency. Therefore, the Window Pseudo Wigner-Ville distribution (WPWVD) [26] was chosen for the basic transformation. WPWVD spectrograms are visually distinctly different than SFFT spectrograms. WPDV spectrograms are too slow for streaming audio compared to SFFT ones: they take about 50 times longer to compute. WPWVD is a better choice than SFFT when studying audio in one detail, where the highest quality TF graph is required. To generate a sample-accurate (1024 band) WPWVD spectrogram in real time would require about 16 cores. We did not have enough computing resources to calculate the full spectrograms. So, we had to downsample it to 2 kHz. For this reason, in the final model, we also used Mel-scaled spectrogram to find defects at frequencies above 2 kHz. Final configurations of acoustic features is shown of table 1.

Table 1: Configurations of acoustic features.

Feature	WPWVD	Mel spectrogram
Window length	20	1024
Hop length	-	512
Low Frequency	0 Hz	0 Hz
High Frequency	2 kHz	8 kHz
Feature dim	900	650

3.2. Architecture improvement

To improve the overall quality of the model it is necessary to perform a tuning of baseline autoencoder.

Experimentally, a rational number of neurons in layers has been established. For example, the dependence of average AUC on the number of neurons in dense and bottleneck layers on figure 4 is presented.

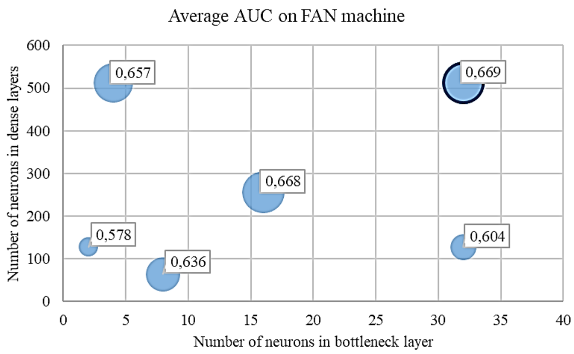
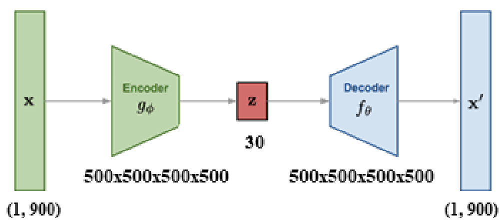
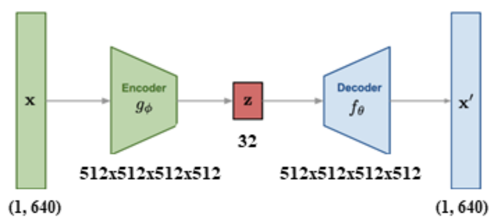


Figure 4: Average AUC on FAN machine.

Values of reshape layers, presented on figure 5, is chosen for current study.



a)



b)

Figure 5: Reshape autoencoders for: a) WPWVD; b) Mel spectrogram

3.3. Score calculation

The idea of the ensemble is to use two models that individually give common results, but together can improve it. The scheme is shown on figure 4.

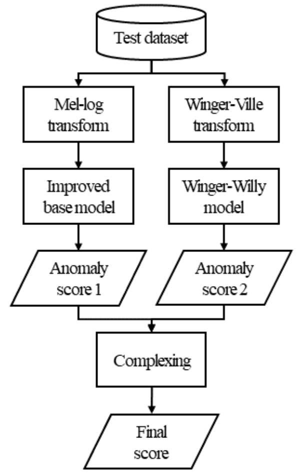


Figure 4: Scheme of score calculation.

Several complication algorithms were used. Logistic regression did not give statistically significant increase in clustering accuracy, as well as selection of the largest of the pair, the smallest of the pair, the middle pair. The sub-optimal algorithm is the addition of the anomaly scores and tuning coefficients, which were selected for each type of machine, by solving the optimization problem, the purpose of which was to maximize the average value of ROC_AUC.

4. RESULTS

Based on the model developed, the following results are obtained, as presented in Table 2.

Table 2: Score of ensemble model.

		Mel spectrogram		Winger-Ville		Final score	
Machine	ID	AUC ROC	pAUC ROC	AUC ROC	pAUC ROC	AUC ROC	pAUC ROC
PUMP	00	0,66042	0,51123	0,41720	0,51196	0,66916	0,50092
	02	0,62613	0,56377	0,52514	0,49692	0,65766	0,56377
	04	0,96350	0,83053	0,31755	0,50526	0,97070	0,86842
	06	0,75882	0,53818	0,31755	0,48194	0,74275	0,51135
	Aver.	0,75222	0,61093	0,42672	0,49902	0,76007	0,61112
VALVE	00	0,81134	0,54401	1,00000	1,00000	0,98370	0,91420
	02	0,87292	0,58509	0,20038	0,48202	0,86817	0,58289
	04	0,84425	0,53114	0,15617	0,47807	0,82325	0,51009
	06	0,66533	0,48509	0,63388	0,67193	0,72100	0,54430
	Aver.	0,79846	0,53633	0,49760	0,65800	0,84903	0,63787
FAN	00	0,51941	0,48959	0,62007	0,59472	0,61172	0,51442
	02	0,73476	0,54200	0,53474	0,51048	0,70348	0,53775
	04	0,62520	0,51664	0,57690	0,49350	0,62721	0,53962
	06	0,79825	0,52253	0,57199	0,53929	0,80773	0,59294
	Aver.	0,66941	0,51769	0,57593	0,53450	0,68754	0,54618
SLIDER	00	0,97570	0,87655	0,13292	0,47368	0,97174	0,86384
	02	0,82360	0,64794	0,50243	0,48847	0,84154	0,66154
	04	0,97287	0,85719	0,30787	0,47635	0,97416	0,86399
	06	0,84551	0,53105	0,51191	0,48019	0,84382	0,52277
	Aver.	0,90442	0,72818	0,36378	0,47967	0,90781	0,72803
TOYCAR	01	0,75677	0,58037	0,42690	0,49231	0,71471	0,57149
	02	0,77050	0,58502	0,35170	0,48486	0,71596	0,58122
	03	0,58717	0,52493	0,53209	0,50399	0,57383	0,52583
	04	0,63641	0,56760	0,55433	0,51573	0,80726	0,60777
	Aver.	0,68771	0,56448	0,46625	0,49922	0,70294	0,57158
TOY CONVEER	01	0,82207	0,64872	0,50703	0,50678	0,82169	0,68071
	02	0,68994	0,56994	0,50656	0,51116	0,67186	0,53293
	03	0,82798	0,65097	0,49606	0,50268	0,77355	0,59575
	Aver.	0,77999	0,62321	0,50322	0,50687	0,75570	0,60313

5. DISCUSSION

A number of the following assumptions can be made based on the results. These assumptions will make it possible to further improve the concept.

5.1. Harmonics distribution as definition of anomalies

The increase in model quality by high resolution spectral picture together with reshape of layers means the efficiency of cyclostationary theory application. It also means criticality of recovery process from input data of harmonic distribution. Combinations of linear dependent harmonics on fundamental frequency for selected machine parts give more information about anomalies. This allows a significant increase in the interpretability of anomaly detection. However, it imposes a number of requirements for data collection experiments and additional machine research.

5.2. Approach to weak classification

Based on previous assumption, the transition from anomaly detection to a weak classification model should be considered. In the framework of the conducted research it became obvious that it is difficult to generalize the problem of detecting anomalies from different machines. Therefore, in further research the concept will be modified to solve the problem of weak classification.

6. CONCLUSION

The following conclusions can be drawn from the study:

- the problem description made it possible to establish a basic approach in the model and to carry out an effective review of the methods;
- the model architecture allowed, with limited computing resources, to implement the required calculation methods;
- the results shown that the model have also high scores on development dataset;
- the discussion led to the following steps in the development of the solution.

Future work will be aimed on improvement of anomalous sound detection for machine condition monitoring by using harmonics distribution for weak anomaly classification.

7. ACKNOWLEDGMENT

Our team thanks LCC «Ctrl2GO» for funding the current study.

8. REFERENCES

[1] Lee, Jay, Hung-An Kao, and Shanhu Yang, "Service innovation and smart analytics for industry 4.0 and big data environment," *Procedia Cirp*, vol. 16.1, pp. 3-8, 2014.

[2] Zhalnin, V. P., Zakharova, A. S., Uzenkov, D. A., and Vlasov, A. I., "Configuration-making algorithm for the smart machine controller based on the internet of things concept*," In Int. Review of Electrical Engineering*, vol. 14.5, pp. 375-384, 2019.

[3] Chen, Xian-Yi, Zhi-Gang Jin, "Research on key technology and applications for internet of things," *Physics Procedia*, vol. 33, pp. 561-566, 2012.

[4] Cai, Chao, Rong Zheng, and Menglan Hu., "A survey on acoustic sensing," arXiv preprint arXiv, 1901.03450, 2019.

[5] Nazolin, A. L., V. I. Polyakov, and S. G. Gnezdilo, "Diagnostics of Shaft Trains of Large Turbine Sets for Integrity Using Torsional Vibration Measurements," *Thermal Engineering*, vol. 67.1, pp. 25-35, 2020.

[6] Serizawa, Y., Y. Shomura, "A Flexible Acoustic Sensing System and its Application to IIoT–Manufacturing Field Site," in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, 2019.

[7] Kiselev, M. I., V. I. Pronyakin, and A. K. Tulekbaeva, "Technical diagnostics functioning machines and mechanisms," in *IOP Conference Series: Materials Science and Engineering*, 2018, vol. 312.1.

[8] Delvecchio, Simone, Paolo Bonfiglio, and Francesco Pompoli, "Vibro-acoustic condition monitoring of Internal Combustion Engines: A critical review of existing techniques," *Mechanical Systems and Signal Processing*, vol. 99, pp. 661-683, 2018.

[9] Atamanov, V. V., et al, "Control of the State of Extreme Stacks of Active Steel of a Turbogenerator Under Load," *Power Technology and Engineering*, vol. 39.4, pp. 248-251, 2005.

[10] Nazolin, A. L., 'Evaluation of the Possibility of Flaw Detection in Turbine Generator Stator According to Spectrum of Vibroacoustic Signal," *Inorganic Materials*, vol. 54.15, pp. 1537-1542, 2019.

[11] Koizumi, Yuma, et al, "Unsupervised detection of anomalous sound based on deep learning and the Neyman–Pearson lemma," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27.1, pp. 212-224, 2018.

[12] Khan, Muhammad Salman, et al., "An unsupervised acoustic fall detection system using source separation for sound interference suppression," *Signal processing*, vol. 110, pp. 199-210, 2019.

[13] Kawaguchi, Yohei, et al., "Anomaly detection based on an ensemble of dereverberation and anomalous sound extraction," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019.

[14] http://dcase.community/challenge2020/task-unsupervised-detection-of-anomalous-sounds

[15] Yuma Koizumi, Yohei Kawaguchi, Keisuke Imoto, Toshiki Nakamura, Yuki Nikaido, Ryo Tanabe, Harsh Purohit, Kaori Suefusa, Takashi Endo, Masahiro Yasuda, and Noboru Harada, "Description and discussion on DCASE2020 challenge task2: unsupervised anomalous sound detection for machine condition monitoring," in *arXiv e-prints*, 2006.05822, pp. 1–4, June 2020.

[16] Yuma Koizumi, Shoichiro Saito, Hisashi Uematsu, Noboru Harada, and Keisuke Imoto, "ToyADMOS: a dataset of miniature-machine operating sounds for anomalous sound detection," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, November 2019, pp. 308–312.

[17] Harsh Purohit, Ryo Tanabe, Takeshi Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi,

"MIMII Dataset: sound dataset for malfunctioning industrial machine investigation and inspection," in *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, November 2019, pp. 209-213.

[18] Antoni, Jérôme, "Cyclostationarity by examples," *Mechanical Systems and Signal Processing*, vol. 23.4, pp. 987-1036, 2009.

[19] Lake, D. E. (1998, September), "Tracking fundamental frequency for synchronous mechanical diagnostic signal processing," in *Ninth IEEE Signal Processing Workshop on Statistical Signal and Array Processing (Cat. No. 98TH8381)*, September 1998, pp. 200-203.

[20] Rafii, Z., Pardo, B., "Music/Voice Separation Using the Similarity Matrix," in *ISMIR*, October 2012, pp. 583-588.

[21] Su, Hang, Quan Liu, and Jingsong Li, "Alleviating border effects in wavelet transforms for nonlinear time-varying signal analysi," *Advances in Electrical and Computer Engineering*, vol. 11.3, pp. 55-60, 2011.

[22] Schörkhuber, C., and Klapuri, A., "Constant-Q transform toolbox for music processing," in *7th Sound and Music Computing Conference*, July 2010, pp. 3-64.

[23] Głowacz, A., and Głowacz, Z., "Diagnostics of DC machine based on analysis of acoustic signals with application of MFCC and classifier based on words," *Archives of metallurgy and materials*, vol. 57, pp. 179-183, 2012.

[24] https://www.mathworks.com/help/audio/ug/spectral-descriptors.html

[25] https://github.com/tensorflow/models/tree/master/research/audioset/vggish

[26] Wang, Zaiyi, Bo Yin, and He Yan., "Weak Signal Detection Based on Pseudo Wigner Ville Distribution," *Journal of Physics: Conference Series*, vol. 1176.6, 2019.