Identification of Axle-Box Bearing Faults of Freight Cars Based on Minimum Entropy Deconvolution and Squared Envelope Spectra

Serhii Mykhalkiv1*, Vasyl Ravlyuk1, Andrii Khodakivskyi1, Viktor Bereznyi1

1Ukrainian State University of Railway Transport
*Corresponding author E-mail: svm_m@ukr.net

Abstract

Purpose: To improve the performance of vibration spectral methods in identification of bearing element faults of freight car axle-boxes.

Approach: An algorithm for simulating the expected vibration signal of outer race bearing was implemented. The autoregressive filter and minimum empirical deconvolution method was applied to identify the ball pass outer-race fault frequency and its harmonics on the envelope spectra and squared envelope spectra which were extracted in the proper frequency band.

Results: The simulated vibration signal of a faulty bearing shows suitability of the autoregressive filter and minimum empirical deconvolution method, envelope and squared envelope spectra for outer race fault identification. There were observed a lower amount of features and their impulse sharpness in outer race faults in the bearing test rig than on the spectra in the wheelset test rig.

Conclusions: The deterministic components are removed in the residual signal after using the AR filter and the impulse and noise components that decrease the kurtosis value remain in it. The MED technique additionally enhances the magnitude of increased BPFO components after using the AR filter and, together with it, provides satisfied performance and increases the efficiency of vibration diagnostics.

Keywords: Axle-box, Bearing; Deconvolution; Diagnostics; Envelope; Modelling, Spectra.

1. Introduction

Axle-box bearings are one of the most important elements of the underframe of freight cars. While in operation, axle-boxes are under the long-time influence of heavy loads, negative influence due to the interaction wheel-rail and torque loads, which results in the appearance of different faults of rolling bearing elements.

These faults can undergo unpredictable fast development and cause an accident. That is why it is important to implement diagnostic tools for the technical state of axle-boxes of freight cars. Because of the high sensitivity of vibration signals to the technical state of bearings, the vibration analysis based on methods of signal processing has become widespread. Methods of the vibration analysis can be concentrated on time domain, frequency domain signal or time-frequency domain signal when the signal under research is non-stationary. Time and frequency domain signals in most cases do not make it possible to search fault efficiently on the account of high noise and, even more, current diagnostic tools demand high practical experience of a diagnostician. Researches on the vibration-based diagnostics of bearing and gearbox faults are developing fast in recent years. Trustworthy diagnostics needs efficient signal processing generated by a faulty bearing.

In the furtherance of this goal, in recent years the appropriate methods have been developed, which are concentrated on the removal of noise from the faulty bearing signals, extraction of fault features from the bearing signals, enhancement of the impulsive features, high resolution in time frequency plane, disjointing of the interference terms due to other vibration sources such as gears, identification of the fault characteristic frequencies and its separation from rotational frequency [1]. A local fault occurring in bearing elements leads to the appearance of impulses in the vibration signal, which take form of repetition taking into account shaft rotation and loads [2]. Kurtosis is one of most important means of obtaining signatures for machinery faults, since it can detect the peaks induced by faults and is often chosen as a direct measure of the impulsiveness of the signal. However, the fault-induced repetitive transients are hidden by noise, which complicates diagnostics and has impact on a low kurtosis value. It was offered to make filtration of the signal that has diagnostic information in the vicinity of resonant frequency band which is called informative frequency band (IFB), that is why the search of the IFB is a key issue for the diagnostics of rotating machines [3]. An ability to detect impulse components of bearing and gearbox faults was demonstrated by the autoregressive model that generates a remaining signal and it is efficient only when the fault amplitude is much higher than the noise level [4]. The minimum entropy deconvolution (MED) method can eliminate drawbacks of the autoregressive model. MED technique is designed to counteract the effect of the transmission path. Assuming that the original excitation was impulsive and, thus, making the kurtosis maximum by finding an inverse filter to minimize the entropy of the filtered signal, the trustworthiness of the diagnostics increases [5]. Many articles [6—9] have proved the effectiveness of the MED method and its improvement techniques in separating the shock signal component from a complex signal that contains various components.

This paper focuses on the MED technique performance for enhancement of impulse components that are fault features of an axle-box bearing. The proposed algorithm of the vibration signal simulation [10] is validated on the real datasets from the bearing test rig and from the wheelset test rig. The experimental data of two rigs are also compared to define the best performance of bearing outer race fault identification.
2. Theoretical Background

When a rolling surface contacts the fault, this produces a short duration impulse which excites some structural resonance of the bearing or of the vibration transducer itself. The repetition of these impacts, when the bearing is operating, results in a series of impulse responses whose temporal spacing depends on the type of fault and on the geometry of the bearing. The series of impulse responses produced by an incipient fault are possibly amplitude modulated due to the passing of the fault into and out of the load zone. For a stationary outer race and in the presence of a radial load, an outer race fault would experience a uniform amplitude modulation.

The vibration signal of a fault in a rolling bearing element can be modelled as

\[ x(t) = \sum_{i=0}^{\infty} h(t - iT - \tau_i) q(iT) + n(t) \]  

(1)

where \( h(t) \) is the impulse response to a single impact, \( q(t) = q(t + P) \) is the periodic modulation of period \( P \) due to the load distribution, \( T \) is the inter-arrival time between two consecutive impacts on the fault, index \( i \) denotes the occurrence of the \( i \)th impact and \( n(t) \) accounts for an additive background noise that embodies all other vibration sources [11].

Diagnostic techniques developed for vibration-based condition monitoring involve processing of the signals through different filters. An autoregressive (AR) model-based filter isolates the impulse-like effect of localised faults in gears and bearings [6]. AR filter for linear prediction

\[ x_k - \sum_{i=1}^{k} a_i x_{k-i} + e_k \]  

(2)

where \( k \) is the time index, \( a_i \) are the AR coefficients and \( e_k \) is the residual [12]. The residual of the AR-based linear prediction filter comprises the non-predictable part of the signal and is a mixture of impulses and coloured noise. The AR model has no zeros, and cannot represent any maximum phase properties and it is based on autocorrelation measurements that are insensitive to phase relationships (which differentiate white noise from impulses). Moreover, the fault-induced impacts should be partly deconvolved (if the response can be assumed as having the minimum phase), and they become slightly sharper. It is achieved by application of an approach based on inverse filtering also known as the MED technique which also can be useful for inverting the effect of the transmission path of the system and recovering the impulses generated by the fault [13]. The deconvolution problem aims to reconstitute the fault signal \( \dd \) by applying filter \( \tilde{f} \) with \( L \) samples to measured machine acceleration \( \dd \)

\[ \dd = \tilde{f} \ast \dd = \tilde{f} \ast (\dd_0 + \dd_1) \]  

(3)

\[ \tilde{f} = [f_1 f_2 ... f_L] \]  

(4)

It is desired that the resulting filtered signal \( \dd \) approximates fault signal \( \dd \) and this is achieved by selecting filter \( \tilde{f} \) to minimize the noise effect \( \tilde{f} \ast (\dd_0 + \dd_1) \rightarrow \dd \), while closely cancelling the unknown input sequence of the system \( \tilde{f} \ast (\dd_0 + \dd_1) \rightarrow \dd \) and extracting a shifted approximation to the fault impulse train signal \( \tilde{f} \ast (\dd_0 + \dd_1) \rightarrow \dd \). The fault signal \( \dd \) is expected to be impulse-like (a signal of very high kurtosis) while competing signals \( \dd_0 \) and \( \dd_1 \) are of very low kurtosis. As a result of this kurtosis difference between the signals, the filter can be selected to reach a maximum in kurtosis.

The term kurtosis for deconvolution method was chosen because it is commonly used to quantify the impulse-like fault level of a vibration signal. Assuming \( y_n \) is zero-mean, \( \mu_y = 0 \)

\[ \text{max Kurtosis} = \max \frac{\sum_{i=1}^{\infty} (y_n - \mu_y)^4}{(\sum_{i=1}^{\infty} (y_n - \mu_y)^2)^2} \]  

(5)

By taking the derivatives of (5) with respect to filter coefficients \( \tilde{f} \) and solving it equal to zero, an iteratively converging local-maximum solution can be derived as

\[ \tilde{f} = \frac{\sum_{n=1}^{\infty} y_n^2}{\sum_{n=1}^{\infty} y_n^4} (X_0 X_0^\ast)^{-1} X_0 [y_1^2 ... y_N^2]^\ast \]  

(6)

\[ X_0 = \begin{bmatrix} x_1 & x_2 & x_3 & ... & x_N \\ 0 & x_1 & x_2 & ... & x_{N-1} \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & x_{N-L+1} \end{bmatrix} \]  

(7)

where \( \tilde{f} \) is iteratively selected. The general procedure is as follows:

1) Assume initial filter as a centred impulse, \( \tilde{f} = [0 0 0 ... 1 0 0]^T \).
2) Calculate \( X_0 \) and \( (X_0 X_0^\ast)^{-1} \) from input signal \( \dd \).
3) Calculate \( y \) as \( \dd = X_0 \tilde{f} \).
4) Determine new filter coefficients by solving for \( \tilde{f} \) in (6).
5) Repeat from step 3 for a specified number of iterations or until the change in kurtosis between iterations is below a specified small value.
6) The final deconvolved signal is calculated as \( \dd = X_0^{\ast} \tilde{f} \) [4].

3. Results and Discussions

3.1. Outer Rolling Bearing Fault Simulation

One important precondition to applying these signal processing techniques is that the basic fault generation mechanism model has to be correct and practical. An algorithm to simulate the expected vibration signal of outer race bearing fault [10] is based on the bearing model [11]. For the numerical implementation of the model (1) the axle-box bearing dimensions of freight car were selected. The bearing roller diameter (\( d_{na} \)) — 32 mm, outer race diameter — 250 mm, number of rolling elements (\( Z_{na} \)) — 15, cage diameter (\( d_c \)) — 190 mm. The vibration signal model main parameters are:

- sample frequency \( f_s \), Hz — 46666;
- carrier component of the shaft speed \( f_c \), Hz (rpm) — 5.5 (332);
- frequency deviation \( f_d \), Hz — 0.08f_s;
- modulation frequency \( f_m \), Hz — 0.5f_s;
- Single Degree of Freedom (SDOF) spring stiffness \( k \), N/m — 2·10^13;
- SDOF damping coefficient \( \zeta \), % — 5.

Figure 1 depicts the simulated time signal in case of outer-race localized fault with series of impulse responses.
The Fast Fourier Transform (FFT) spectrum (Figure 2) of time signal consists of the resonance frequency excited by the bearing impulses in the IFB 4 — 6 kHz.

In general, signal-to-noise ratio is used as an indication of the frequency band where the modulated signal should be filtered. After filtering, the selected frequency band is shifted at low frequencies in the spectrum and padded with zeros to double the length in order to obtain a one-side spectrum. When computing the inverse Fourier transform of this one-side spectrum, an analytic signal is obtained, such that, its imaginary part is the Hilbert transform of the real part. In this way, envelope corresponds to the modulus of real and imaginary parts [14]. Envelope spectrum (Figure 3) makes it possible to identify harmonics of the ball pass outer-race fault frequency (BPFO = \( \frac{Z_{rot} \cdot f_{rot}}{2} \cdot (1 - \frac{d_{in}}{d_{out}}) \)) 35 Hz, 2xBPFO (70 Hz), 3xBPFO (105 Hz).

At the next stage the proposed fault detection method MED [6] was applied to the resulting prediction residual of simulated signal after the AR filter. The expected results for the MED technique have a high kurtosis of the faulty bearing signal, unlike the AR method results that in the best-case scenario comprise a fault signal plus white noise [4]. The ARMED technique produces improved performance over traditional spectral methods (Figure 4) by appearing 4xBPFO (143 Hz) and 5xBPFO (178 Hz).

However, it is useful to analyze the squared envelope spectrum (SES), since it can improve signal-to-noise ratio by removing extraneous components in practical situations. Eventually, the SES method, which is the benchmark technique for bearing diagnostics [15], has replaced the envelope spectrum technique. The SES method consists of decreasing harmonics of BPFO where the first frequency is the strongest one (Figure 5).

Such a simulated vibration signal of a faulty bearing shows suitability of the ARMED, envelope, SES techniques for outer race fault identification.

3.2. Outer Rolling Fault Identification in the Bearing Test Rig

The experimental research was carried out in a freight car depot in the rolling bearing repair shop. After repair, bearings are installed in the test rig (Figure 6) for the verification of repair quality and identification of element faults. For the experiment the assembly of a rolling bearing with faulty outer race was performed (Figure 7).

The bearing inner race spins up to the frequency 735 rpm, and the outer race is subjected to the load 1900 N. An accelerometer with the frequency range 0 — 9000 Hz was mounted on the surface of the loading device. The revolutions were measured by an inductive sensor. The \( f_t \) of the digital recorder was 46666 Hz. The kurtosis of the raw signal (Figure 8) equals 3.76. To select the proper frequency band for further envelope spectrum extraction, the narrowband 4500 — 6000 Hz was chosen (Figure 9). The extracted envelope spectrum (Figure 10) comprises a rotor harmonic (\( f_t = 12 \) Hz), strong BPFO (76 Hz) and a weak 2xBPFO (152 Hz).

The features of outer race faults, acquired in optimal carrier frequency 5 kHz, remain the same (Figure 11) on the SES (BPFO, 2xBPFO).

For further feature enhancement of outer race faults (appearing in the series of BPFO harmonics) the AR technique was applied with AR order = 200. The vibration signal after using the AR filter (Figure 12) is full of noise components. The whitening of the AR-residual causing the lower kurtosis (3.56) decreases the signal-to-noise ratio and does not enhance the impacts comparing to the signal without pre-processing (Figure 8) [13]. At the next stage the MED technique with filter length 400 was applied to the AR-residual signal (Figure 12). The signal after using ARMED technique (Figure 13) has much higher signal-to-noise ratio and full of impulse components with higher kurtosis 7.79.
The FFT spectrum of the vibration signal (Figure 14) after using the ARMED technique has the clear IFB 4.5—6 kHz covered with resonance impulses caused by the bearing vibration.

Additional weak third harmonic 3xBPFO (228 Hz) appeared on the envelope spectrum after using the ARMED (Figure 15) and SES (Figure 16) techniques.

The MED technique designs a filter to approximately extract a single impulse even from a white noise signal. The MED definition assumes zero data for \( x_n = 0, n < 1 \), which creates a discontinuity between assumed zero sample \( x_0 \) and the first sample \( x_1 \). In application to rotating machine vibration this can cause a significant disturbance between samples observed at \( x_0 \) and \( x_1 \) causing a spurious impulse to be deconvolved at this location [7]. Further increase of the length of MED filter up to 800 and unchanged order 200 of AR model causes the appearance of spurious impulse and increase of kurtosis up to 32.29 (Figure 17). The idea of the AR model application for residual preprocessing before applying the MED technique to the vibration signal (ARMED) [6, 7] makes it possible to mitigate this discontinuity partially. Comparing the performance of the ARMED method to the MED technique (Figure 18) with the same length of MED filter (800) one can see the mitigation of the spurious impulse (0.053 m/s² vs 0.064 m/s²). Complete mitigation include selecting smaller filter lengths as it was done earlier (Figure 13).
Fig. 18: Vibration signal in the bearing test rig with outer race faults after using MED800 (kurtosis=33.45)

Fig. 19: Envelope spectrum of the vibration signal after using AR200+MED800 extracted in the IFB 4.5 — 6 kHz

Fig. 20: SES of the vibration signal after using AR200+MED800 filtered in the optimal carrier frequency 5 kHz

3.3. Outer Rolling Fault Identification in the Wheelset Test Rig

At the next stage of the research the same bearing with outer race faults was assembled together with a faultless one in an axle-box and the wheelset was mounted in the test rig (Figure 21). The speed of the wheelset was 346 rpm and the bearing load was 3500 N. The accelerometer was mounted on the inner surface of the axle-box. Time domain vibration signal (Figure 22) is full of strong background noise and no features of the bearing faults can be detected.

Faults in the rolling bearing elements produce a series of impacts which are usually of low amplitudes and buried in high levels of additive noise so that vibration signals should be first pre-processed by band-pass filtering in some frequency bands that maximize the signal-to-noise ratio [16]. Fortunately, such frequency band was successfully detected during simulation procedure (Figure 2), and was verified during the experimental research in the bearing test rig (Figure 9) and in the wheelset test rig (Figure 23). There was also applied the fast kurtogram (Figure 24) which displays the spectral kurtosis values as a function of frequency and windows length, and this technique defines the filter center frequency ($f_c = 4739.52$ Hz) and bandwidth ($B_w = 729.16$ Hz) with the highest value of kurtosis (3.2) applicable for envelope spectra extraction. The envelope spectrum of the vibration signal (Figure 25) comprises BPFO (36 Hz), 2xBPFO (72 Hz), 3xBPFO (108 Hz), 5xBPFO (180 Hz), 7xBPFO (252 Hz), which demonstrates much better performance than on the envelope spectrum in the bearing test rig (Figure 10).

The enhancing MED technique (filter length = 400) without the AR filter gives a vibration signal (Figure 26) with numerous sharp impulses and the kurtosis 9.91. The FFT spectrum (Figure 27) after using the MED technique has the strongest resonance in the IFB 4 — 6 kHz, and the extracted envelope spectrum (Figure 28) has the same amount of BPFO harmonics as the envelope spectrum without inverse filtering (Figure 25), which proves the ability of MED technique to enhance impulse components.

Then the AR filter before using the MED technique (AR order =200, filter length = 400) was applied to the vibration signal (Figure 22). The kurtosis of the filtered signal (Figure 29) has reached 9.79, and sharp impulses appeared like after using only the MED technique (Figure 26). The FFT spectrum after using the ARMed technique (Figure 30) has the strongest resonance in the IFB 4.5 — 6 kHz. Envelope spectrum (Figure 31) shows an excellent performance due to the appearance of the series of harmonics BPFO — 7xBPFO that comprise additional 4xBPFO (144 Hz) and 6xBPFO (216 Hz) harmonics. SES comprises all BPFO harmonics which decrease gradually and also demonstrates an excellent performance (Figure 32).
of a raw vibration signal, approximately by 50 % on the envelope spectrum after using the MED technique with the filter length 400 and did not improve the performance of the envelope spectrum by appearing the desired amount of BPFO harmonics.

The AR filter improves the performance of envelope spectrum dramatically by increasing the amount of BPFO harmonics. The deterministic components are removed in the residual signal after using the AR filter, but the impulse and noise components decreasing the kurtosis value remain in it. The MED technique additionally enhances the magnitude of increased BPFO components after using the AR filter, and together with it, provides satisfied performance. Moreover, the AR preprocessing procedure partly mitigates a spurious single impulse that appears after selecting a high length of MED filter due to the maximization problem.

At MED filter widths ranging from 2 to 800 in the bearing test rig and ranging from 2 to 1000 in the wheelset test rig, the sharpness of impulses of the first four BPFO harmonics on the envelope vibration spectra remains constant and the rest of BPFO harmonics gradually increase depending on increase of the filter width. The single spurious impulse becomes visible at widths more than 800 in the bearing test rig, and more than 1000 in the wheelset test rig. It was established that the increase of the filter width has an effect on the kurtosis increase, and high values of the kurtosis are not a feature of the improved performance of BPFO harmonics on the envelope spectra.

It was established that the performance of the envelope spectra very strongly depends on the load on the bearing and depends less on the rotation frequency. Under the load in the bearing test rig 1900 N and speed 735 rpm on the envelope vibration spectra after using the ARMed technique and SES, 4 BPFO harmonics can hardly be identified. However, despite the lower speed 346 rpm but the increased load up to 3500 N on the bearing in the wheelset test rig the performance of the envelope vibration spectra and SES after using the ARMed technique shows a great improvement.

4. Conclusion

According to the experimental data results, the verification of the bearing model can be obtained by extracting BPFO harmonics on envelope spectrum and SES. The fault impulses are obvious and are spaced between each other at the same width which equals to the BPFO value.

The MED technique has increased the amplitudes of all BPFO harmonics, which where earlier detected on the envelope spectrum

References


