Fault Prognosis and Diagnosis of an Automotive Rear Axle Gear Using a RBF-BP Neural Network

Yimin Shao[1* 3], Jie Liang[1], Fengshou Gu[2], Zaigang Chen[1], Andrew Ball[2]

1 State Key Laboratory of Mechanical Transmission, Chongqing University, PR China;
2 School of Computing and Engineering, University of Huddersfield, HD1 3DH, UK;
3 College of Automobile and Mechanical Engineering, Changsha University of Science and Technology, Changsha, PR China

Abstract: The rear axle gear is one of the key parts of transmission system for automobiles. Its healthy state directly influences the security and reliability of the automobiles. However, non-stationary and nonlinear characteristics of gear vibration due to load and speed fluctuations, makes it difficult to detect and diagnosis the faults from the transmission gear. To solve this problem a fault prognosis and diagnosis method based on a combination of radial basis function (RBF) and back-propagation (BP) neural networks is proposed in this paper. Firstly, a moving average pretreatment is used to suppress the time series fluctuation of vibration characteristic parameter tie series and reduce the interference of random noise. Then, the RBF network is applied to the pretreated parameter sequences for fault prognosis. Furthermore, based on self-learning ability of neural networks, characteristic parameters for different common faults are learned by a BP network. Then the trained BP neural network is utilized for fault diagnosis of the rear axle gear. The results show that the proposed method has a good performance in prognosing and diagnosing different faults from the rear axle gear.

1. Introduction

With the rapid development of automobile industry, the automotive safety and reliability have paid increasing attention in different fields. Rear axle power system is a critical device that transfers the power from the engine to the travel system. According to statistics, 60% failures of transmission system are caused by gear faults. If gear service life can be predicted timely before gear fault happened and effective measures are taken according to the fault types. It is of great significance for eliminating the occurrence of failures and reducing the maintenance and economic loss.

For gear life Prognosis and fault diagnosis, plenty of work has been done by scholars from different aspects with various methods. Li et al. [1] proposed an embedded model to predict remaining useful life of a gear with a fatigue crack. Zhan et al [2] developed a robust model-based technique for the detection and diagnosis of gear faults under varying load conditions using the gear motion residual life and a noise-adaptive Kalman filter-based auto-regressive (AR) model is fitted to the gear motion residual signals. Loutridis et al. [3] proposed energy-based features for gear fault diagnosis and Prognosis which obtained when defected teeth are engaged. Loutridis et al. [4] introduced multi-scale
local statistic tools for gear failure prognosis and established an empirical law that related to variance at various scales to crack magnitudes. Kramberger et al. [5] analyzed thin-rim gear fatigue life by using the finite element method and boundary element method, which leads to continuum mechanics based approach for the Prognosis of the fatigue process initiation phase. Khan et al. [6] made an effort to validate the competency of standard’s gear useful lifetime formulation which used for helical gear useful lifetime estimation under linear pitting fatigue conditions.

In the mean time, gear fault diagnosis study has also had a great development in the past few years. A number of effective diagnostic feature parameters have been developed based on classic vibration analysis methods including statistical analysis and Fourier transforms in the time and frequency domains. Neural networks, wavelet transforms, support vector machine and the fuzzy fault tree have been developed greatly in recent years [7]. Niranjan Subrahmanya [8] proposed regression and classification algorithm of samples choice based on algorithm of Bayesian framework recently.

However, these researches have addressed little in combining gear fault prognosis with fault diagnosis to show gear operating status timely. In this paper, a method unifying both fault prognosis and diagnosis together is proposed and applied to the fault diagnosis of a rear axle gear. In order to reduce temporal fluctuations, a smooth pretreatment is carried on by applying a moving-average to the time series of feature parameters extracted from vibration signals. A multi-step life forecast is them implemented using RBF neural network to predicted the future values of time series. An increase of the time series will indicate a likely fault of the gear and a back-propagation (BP) network will be initiated for fault diagnosis to find the origins and the severity of the faults.

In conventional threshold based fault detection, it is often difficult to specify a correct threshold. Especially a high threshold may need to be specified when a time series has high statistical fluctuations. This means that threshold based method has less sensitivity and also produce more false alarms. In contrast, a RBF-BP combination will perform fault detection based on both current measurements and future predication, which would be produce more reliable results. Moreover, the diagnostic results from BP can be further referenced to confirm the detection results. Although parametric models such as exponential model, polynomial models, auto regression models etc can be used for modeling the process. However, two issues: model selection and model parameter estimation need be solved to use these models properly. In contrast RBF requires a light effort in both model estimation and parameter estimation. BP neural network is involved in this study just because that it is an effective method for multiple fault classification.

The content of this paper is organized into 5 sections. Next section i.e. section 2 outlines the fundamentals of RBF and BP neural networks and the scheme of RBF-BP based fault prognosis and diagnosis. Section 3 evaluates the performance of the scheme in fault prognosis and diagnosis by using simulated time series. In section 4, fault prognosis and diagnosis results are discussed in applying the proposed method to an automotive rear axle gear. Finally, concluding remarks are presented in section 5.

2. Fault Prognosis and Diagnosis using RBF-BP neural network

Because of highly non-stationary and nonlinear characteristics in vibration signals from the rear axle gear, direct time-domain analysis will yield low precision for fault Prognosis and diagnosis. For these reasons, the present paper has proposed a model combining a RBF-BP neural network with moving-average method to solve this problem.

The flowchart of axle gear fault Prognosis and fault classification based on the RBF-BP network is illustrated in Figure1. Firstly, the characteristic parameters are extracted from vibration signals.
Secondly, the moving average method is used to reduce the variations of these parameters. Then, a RBF neural network Prognosis algorithm is employed to predict the fault of rear axle gear. Finally, if the gradient of forecasts point show ascendant trend, a BP network will be initiated for fault diagnosis.

Figure 1. Flowchart of fault prognosis and diagnosis scheme

2.1 Feature extraction
The collected gear vibration signals can be described as a discrete time series \( \{ s(i) \} (i=1, 2...N) \). In these time domain signals many feature parameters can be extracted for fault detection and diagnosis. In this study, RMS and kurtosis are selected to be the two inputs for monitoring gear healthy condition due to their combined stability and sensitivity in tracking condition changes[9]. The RMS of a vibration signal \( s \) is calculated by Equation 1.

\[
s_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} s_i^2}
\]

and kurtosis is calculated by Equation 2.

\[
Kurt = \frac{N \cdot \sum_{i=1}^{N} (s_i - \bar{s})^4}{\left( \sum_{i=1}^{N} (s_i - \bar{s})^2 \right)^2}
\]

where \( s_{\text{rms}} \) and \( Kurt \) denotes the RMS and kurtosis respectively, \( s \) is the velocity of the vibration signal; \( N \) is the number of points in signal \( s \) and \( s_i \) is the \( i^{th} \) member of signal \( s \).

The feature parameters obtained in different time instant constitute a time series \( x \). Because of
inevitable operating condition changes and noise contamination, the feature time series consists of not only the determined part due condition changes but also random part due to noises from measurement and environmental interferences. Thus the series can be denoted as \( x = x_{\text{status}} + e_{\text{noise}} \), in which the noise part indicates the statistical fluctuations.

2.2 Moving Average

A direct use of feature series will produce low precision for fault Prognosis and diagnosis due to the statistical fluctuations, so a moving average method is proposed to reduce the influence of temporal fluctuation and improve the smoothness of the time series. The pretreatment of moving average is defined in Equation 3[10].

\[
\begin{align*}
y_1 &= \frac{1}{5}(3x_1 + 2x_2 + x_3 - x_4)G \\
y_2 &= \frac{1}{10}(4x_1 + 3x_2 + 2x_3 + x_4) \\
& \vdots \\
y_i &= \frac{1}{5}(x_{i-2} + x_{i-1} + x_i + x_{i+1} + x_{i+2}); \\
& \vdots \\
y_{m-1} &= \frac{1}{10}(x_{m-3} + 2x_{m-1} + 3x_{m-1} + 4x_m); \\
y_m &= \frac{1}{5}(-x_{m-3} + x_{m-2} + 2x_{m-1} + 3x_m)
\end{align*}
\]

Where \( x \) is the raw time series; \( y \) is the results after smooth processing; \( m \) is the number of data; \( x_i \) is the \( i \)th element of \( x \). The processed time parameter sequence may be a better approximation of the determined part and described as \( y = x_{\text{status}} \).

2.3 Fault Prognosis

A RBF neural network is a linear network and is efficient for on-line training. Therefore, it is used widely for nonlinear modeling and classification. In the paper, the RBF neural network is used for modeling the time series and subsequently for fault prognosis.

Figure 2 shows the architecture of a typical RBF neural network. It consists of three layers: input, hidden and output layers [11]. For gear fault prognosis a rolling forecast model (RFM) is proposed based on the RBF neural network so that it allows the multiple step forecast of the vibration feature parameters on which fault severity is determined in real time. Figure 3 shows the data frame flow of the RFM. At time instant \( t \), RPM takes \( Y_{t-k+1}, Y_{t-k+2}, Y_{t-k+3}, \ldots, Y_t \) as the input of the RBF network trained by the dataset before \( t \) to predict \( \hat{Y}_{t+1} \) at time \( t+1 \). Then it uses \( Y_{t-k+2}, Y_{t-k+3}, \ldots, Y_t, \hat{Y}_{t+1} \) to predict \( \hat{Y}_{t+2} \) and so on. For multi-step prognosis, it adopts a recursive algorithm [12]. As the higher step prognosis by RPM is based on the updated information, it is likely to produce a more accurate prediction compared with that from the fixed data sets at time \( t \).
2.4 Fault Diagnosis

For fault diagnosis, a BP neural network is employed due to its high performance in modeling more complicated processes in comparison with the RBF network. A BP network has a back-propagation architecture with a nonlinear differential function for training. Network weights were adjusted by error feedbacks. By means of revising weights of the network, actual output is closer to the expected output. A typical architecture of a BP neural network is shown in Figure 4 [13]. The architecture of fault diagnosis is shown in Figure 5. EFM forecasted values \( \hat{Y}_{t+1}, \hat{Y}_{t+2}, \ldots, \hat{Y}_{t+k} \) are taken as the inputs of the BP network trained by known fault feature patterns. Its output \( \hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_k \) represents different fault cases. In this forecast feature based diagnosis, fault diagnosis result indicates possible fault natures in the future. With this information, machine operating conditions may be adjusted to prevent the fault from further deterioration. Moreover, the diagnostic result can be used to confirm the predicted value if the fault is diagnosed as a fault that happens commonly to the component.

3. Performance Evaluation

In order to validate the performance of RBF-BP neural network for fault prognosis and diagnosis, two time series: \( y_1(x) \) and \( y_2(x) \) are produced numerically and mixed with Gaussian white noise \( e(x) \). The processes of \( y_1(x) \) and \( y_2(x) \) are defined mathematically as follows:

\[
\begin{align*}
    y_1(x) &= 10 + 1.5 \sin(3\pi x) + e(x) \\
    y_2(x) &= 10 + 2 \sin(2\pi x) + 10^{-3} \exp(x/10) + e(x)
\end{align*}
\] (4)
In equation (4), $y_1(x)$ may simulate a normal running condition which remains a constant amplitude over time course but with statistical fluctuations whereas $y_2(x)$ may simulate a progression trend of a vibration system such as the axle gear in which its vibration feature exhibits a gradual increase in amplitude in early operation period and a fast increasing amplitude in late stage. As shown Figure 6, $y_2$ shows a clear increase trend during the time period from 75 to 100. However, both of the time series have clear local fluctuations, representing the inevitable uncertainties in measurement systems.

Obviously the fluctuations will also lead to high uncertainty in trend prediction which in turn results in poor performance for fault diagnosis. On the other hand, after moving average suggested in 2.2, the sequences become much smoother but maintain the basic trends of constant in $y_1$ and gradual increase in $y_2$. Based on these smooth time series a more accurate model can be developed for better prediction.

![Figure 6. Simulated time series of common vibration processes](image)

![Figure 7. Prognosis performance at time indices of 74 and 88](image)

To implement the prognosis of these simulated time series, a RBF network with 8 input neurons and 2 output neurons are used. To train this network at a specific time such as time index 1, the 1st 8 data points are taken as the 1st input pattern whereas the 9th data point is taken as the output of 1st pattern. By shifting data one point ahead, the 2nd input pattern and its output is obtained. In the same way, 15
patterns can be constructed from the time series as the 1st training dataset to obtain a network for prediction at time instant 8+15. This is also the effective data size used for training network and should be larger than the local fluctuation duration. Similarly, a new training dataset can be obtained by shift the whole data frame to a new training time instance, which updates the network with latest measurements and then produces more accurate prognosis results.

It is common to normalize the training dataset for better numeric performance in training process. In this study, each pattern \( y^k \) in the training dataset is normalized by

\[
y'^k_{ai} = \kappa [y^k_i - \min(y^k)] / [\max(y^k) - \min(y^k)]
\]

And the corresponding denomalization is performed by

\[
y_i^k = y'^k_{ai} [\max(y^k) - \min(y^k)] / \kappa + \min(y^k)
\]

Where \( \kappa \) is a parameter to adjust the slop of the time series with time.

During the training, the mean squared error (MSE) is set to 0.1 and a large spread parameter of 5 is used for a more generalized model and hence for better prognosis results.

**Figure 8.** Detection using forecasted trend

Figure 7(a) and (b) shows 2 typical cases of prognosis performed using the trained network. In the constant value period such as at time index 74, the forecasted values of steps 7 and 10 show a under-prediction, which helps to avoid too many false alert. However, in the increasing phase such as at time index 88, the forecasted values show an over-prediction, which allows to provide earlier warning of the monitored process. As shown in Figure 8 the forecasted trends show 10 steps earlier than the measured trend when it exceeds the 1st threshold that is built up based on the standard deviation of the 1st training dataset. Even in the increase phase, it also shows 2 steps earlier than the measured values to exceed the 2nd threshold. Those earlier alerts will provide more leading time to take corresponding actions to prevent sever damage to the system, which demonstrates significance of the detection scheme proposed.

Furthermore, original time sequences \( y_1(x) \) and \( y_2(x) \) is used to train a BP network for fault diagnosis. Based on the threshold values and engineering knowledge, the two data sequences is organized into 9 groups to indicate the healthy status of the system. The network is trained to have 10 input neurons and 2 output nodes, denoting as \([Y_1, Y_2]\), in which the input sequence of \( y_1(x) \) leads to \([1 \ 0]\) whereas the sequence of \( y_2(x) \) results in to \([0 \ 1]\). Classification results in Table 2 explain that BP network can identify signal types based on the forecast results and shows that the trained network performs correctly in differentiating different types of datasets.
Table 1. Predictive samples after normalization.

| Pattern | Predicted samples after normalization |
|---------|--------------------------------------|
| $y_1(x)$ | 0.691 0.984 0.990 0.920 0.951 0.852 0.476 0.323 0.211 0.010 |
| $y_2(x)$ | 0.010 0.081 0.160 0.249 0.346 0.451 0.563 0.689 0.830 0.990 |

Table 2. Pattern classification results.

| Pattern | Y1 | Y2 | Classification | Error |
|---------|----|----|----------------|-------|
| $y_1(x)$ | 0.9967 | 0.0127 | $y_1(x)$ | 0.0160 |
| $y_2(x)$ | 0.0000 | 1.0000 | $y_2(x)$ | 0.0000 |

4. Application Example

4.1 Monitored System

The proposed method was applied to monitoring a rear axle gear fatigue process which was tested in a system shown in Figure 7. The system consists of (1) a 570kW induction motor; (2) a transmission unit; (3) torque and speed sensors; (4) the rear axle unit; (5) a speed increaser; (6) a secondary speed increaser; and (7) a 150kW dynamometer.

To monitor the vibration of the real axle for further condition assessment, a data acquisition system consists of (8) a signal conditioner; (9) a Dewetron vibration acquisition unit; (10) a signal indicator; (11) Acceleration transducer 1 (12) Acceleration transducer 2; and (13) Acceleration transducer 3. More details of the system construction and performance parameter have been described in [12].

During the fatigue test the vibrations are measured continuously and common feature parameters including RMS, kurtosis, peak factors are calculated at a time interval of 5 min to monitor the healthy status of the gear and provide a reference base to terminate the test before a server damage to the test system by a unexpected failure of the gearbox.

4.2 Fault Prognosis and Diagnosis Results

During test it has found that RMS and kurtosis values produce a better trend to indicate the fatigue process. As shown in Figure 8 and Figure 9, RMS series exhibits a gradual increase trend start at time
instant 350 mintes whereas kurtosis increases rapidly starting at 550 minutes. These varying trends allow the use of RMS for indicating small fault at early stages and kurtosis for fault with high severity at late stages. However, the raw series has many local spikes, showing clear operating condition changes and random influences from measurements, which can be suppressed without loss of overall trend characteristics. As shown by the smoother plots in the Figures the smoothed trends have much less fluctuations and maintain the key increase behaviors of the time series.

![Figure 8. Original and pretreated RMS.](image1)

![Figure 9. Original and pretreated kurtosis factor.](image2)

In the rear axle gear practical operation, RBF network is used to learning parameters value in different stages with the shifty trends, so as to set up accurate prediction models according to different states. Number of input layer node is 15 and number of output layer node is 1, Target error is 0.001. Each online learning data length is 225min parameter values, when the error reach to the target error, the trained RBF neural network can be used to process ten steps rolling Prognosis. Moreover, mobile data learning length is 75min parameter values. Vibration parameters after preprocessing and normalization act as input of the network, predicted value is output value of the network after denormalization. Figure 12 shows error curve of RBF network model which learned the RMS value after preprocessing within 175min - 400min. It shows that RBF neural network could arrive at the error goal and meet the demand after 10 training epochs. Figure 10 shows RMS prognosis results in small fault period based on this trained network. In the same way, Figure 11 shows kurtosis 10 steps prognosis results in the same stage. The performance error RMSE (root mean square error) are 0.0265 and 0.0172 respectively.

According to actual operating condition of the rear axle gear, its running state can be organised into four periods: normal operation, small fault period, medium fault period and serious fault period. Original typical fault feature of gear is used for training BP neural network. Number of input neurons is 2 and number of output layer node is 2, network learning rate is 0.001. Data sets with 40 groups which used for training network was extracted from four typical fault feature. Input of the network is RMS and kurtosis value after normalization, Target output defined as two nodes: [Y1 Y2], normal operation correspond to [0 0], small fault correspond to [0 1], medium fault correspond to [1 0] and serious fault correspond to [1 1]. The training error curve is shown in Figure 13. It is obvious that network convergences after 90 steps and it is training error is about 0.00089. then the trained network can be used for identify the predictive points with rising gradient in different time, yet gear operating condition is obtained in corresponding moment. Here, several predictive points with rising gradient in different time are selected for fault diagnosis on BP network. The results are shown in table 3.
According to Table 3, fault diagnosis rate is nearly 100% for different conditions including normal situation and different degrees of fault. In addition, a lot of experiments have been done. Abundant results show that the proposed method is more efficient to normal operation, medium fault and serious fault. However diagnosis rate of small fault is not so high. The reason may be lie in that small fault characteristics are not obvious.
5. Conclusions
The present paper has proposed a fault diagnosis and prognosis scheme using RBF-BP neural network coupling with a moving average pretreatment. Simulation and experiment results show that the proposed method is effective for remaining life prognosis and fault diagnosis for an automotive rear axle gear. Moving-average pretreatment is useful in reducing the influence from random noise and caused by loads and speed fluctuation so that gear fault prognosis has high accuracy. Meanwhile, fault diagnosis results exhibit that RBF-BP network is available to forecast gear operating condition in advance and can be used for fault prognosis and diagnosis practically. Furthermore, the proposed method has higher accuracy and fault diagnosis capability compared with traditional methods.

Acknowledgement
This research was supported by the National Natural Science Foundation of China under Contact No.51035008, and supported by Visiting Scholar Foundation of the State Key Lab. of Mechanical Transmission in Chongqing University.

References
[1] Li C J and Lee H 2005 Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics Mech. Syst. Signal Process. 19 836-846
[2] Zhan Y M and Makis V 2006 A robust diagnostic model for gearboxes subject to vibration monitoring J. Sound Vib. 290 928-955
[3] Loutridis S J 2006 Instantaneous energy density as a feature for gear fault detection Mech. Syst. Signal Process. 20 1239-1253
[4] Loutridis S J 2008 Gear failure prediction using multiscale local statistics Eng. Struct. 30 1214-1223
[5] Kramberger J, Sraml M, Potrc I and Flaszer J 2004 Numerical calculation of bending fatigue life of thin-rim spur gears Eng. Frac. Mech. 71 647-656
[6] Khan M A, Cooper D, Starr A 2009 BS-ISO helical gear fatigue life estimation and wear quantitative feature analysis Strain 45 358-363
[7] Hong S Z, Wei H 2008 Remote fault diagnosis system of rotating machinery based on inspection model Modern manuf. Eng. 4 23-25
[8] Subrahmanya N 2010 A Bayesian machine learning method for sensor selection and fusion with application to on-board fault diagnostics Mech. Syst. Signal Process. 24 182–192
[9] Shao Y and Nezu K 2002 Prognosis of remaining bearing life using neural networks J. Syst. Control Eng. 214 217-230
[10] Wang J and Hu X 2005 Matlab applications in vibration signal processing Water conservancy and hydropower press
[11] Wei H K 2005 The structural design of neural network theory and method Beijing Defense Industry press
[12] Shao Y M, Li X X, Chris K, Mechefske and Chen Z G 2009 Rear axle gear damage prediction using vibration signal preprocessing coupled with RBF neural networks Main. Relia. 4 57-64
[13] Zhang X H, Lei Y 2008 Application of BP Neural Network in Mechanical Fault Diagnosis J. noise vib. control 10 95-97

Wealth