Chapter 2
Trend Prognosis of Aero-Engine Abrupt Failure Based on Affinity Propagation Clustering

Limin Li, Zhongsheng Wang, Zhenbao Liu and Shuhui Bu

Abstract This paper presented a new method for trend prognosis of aero-engine abrupt failure which was based on clustering analysis. We used affinity propagation method to cluster vibration data from aero-engine rotating machinery to find the failure trend by means of features extracted from it. This system could remind rapidly which failure will happen next. It is possible to be applied to trend prognosis of aero-engine abrupt failure in airlines.

Keywords Trend prognosis · Aero-engine abrupt failure · Clustering analysis · Affinity propagation clustering

2.1 Introduction

In recent years, aero-engine catastrophic accidents happen occasionally and show an upward tendency [1–3]. Among these accidents, some were caused by abrupt failure. Because the occurrence time of aircraft abrupt failure is very short and highly random, it is very difficult to predict and prevent the accident. Although previous works [4–6] have focused on predicting gradual change failure of aero-engine, to our knowledge, there is little discussion about trend prognosis of abrupt failure. In order to enhance protection ability against abrupt failure, we attempt to address the important problem.

Project 51075330 Supported by National Natural Science Foundation of China.

L. Li (✉) · Z. Wang · Z. Liu · S. Bu
School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, People's Republic of China
e-mail: liliminxiaomi@mail.nwpu.edu.cn

J. Wang (ed.), *Proceedings of the First Symposium on Aviation Maintenance and Management-Volume II*, Lecture Notes in Electrical Engineering 297,
DOI: 10.1007/978-3-642-54233-6_2, © Springer-Verlag Berlin Heidelberg 2014

Methods of failure prognosis can be divided into two categories, model-based and data-driven methods [7]. Model-based methods require accurate process models. For data-driven approaches, only the availability of a large amount of process data is assumed. Because of the difficulty on modeling aero-engine system, in this paper we adopt data-driven method as our abrupt failure trend prognosis method. This can be achieved by intelligent data analysis based on features extracted from raw signals. Clustering data by identifying a subset of representative exemplars is important for handling sensory signals and detecting patterns in data. Generally, such 'exemplars' can be found by randomly choosing an initial subset of data points and then iteratively refining it. There are some methods which have been used in failure prognosis, such as k-means [8] and k-centers [9, 10]. Because these clustering methods are quite sensitive to initial selection of exemplars, it is usually rerun many times with different initialization in an attempt to find a good solution, whose cost is that it would take more time and obtain less accuracy rate. Another problem is that these algorithms have seldom been applied in abrupt failure trend prognosis. Affinity propagation [11] we adopted is an innovative clustering algorithm that identifies exemplars quickly and successfully. The speed of this method is fast and has smaller error rate; hence, it can satisfy the real-time requirement of abrupt failure trend prognosis.

The rest of this paper is organized as follows: in Sect. 2.2, we introduce our scheme of abrupt failure trend prognosis, as well as the principle of affinity propagation clustering. Then, in Sect. 2.3, we present the experiment results of our method. Finally, we conclude this paper in Sect. 2.4.

2.2 Trend Prognosis of Aero-Engine Abrupt Failure Based on Affinity Propagation Clustering

2.2.1 Trend Prognosis Module for Aero-Engine Abrupt Failure

We first collect vibration signals of abrupt failures generated from our test bench of aero-engine rotor and extract time-domain signals before failures to build a database of abrupt failure trend as shown in Fig. 2.1.

Then, we use affinity propagation algorithm to cluster these data in the database, find corresponding exemplars, and label them. When new data come, we cluster them and find their corresponding exemplar. After computing distance between new exemplar and old exemplars, choose the type which has the shortest distance to the result. Moreover, we set a threshold to decide whether the new data come from a new abrupt failure trend. If it is an unknown abrupt failure trend, we will use expert knowledge to analyze it and lastly enlarge the database of abrupt failure trend by increasing the number of classes of failures.

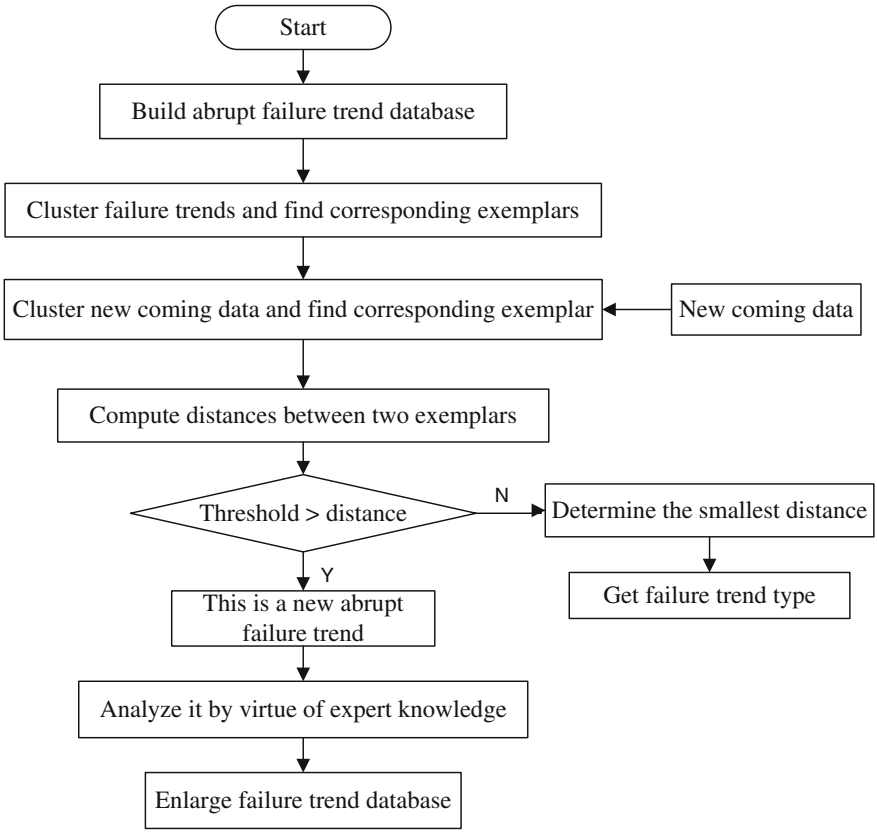
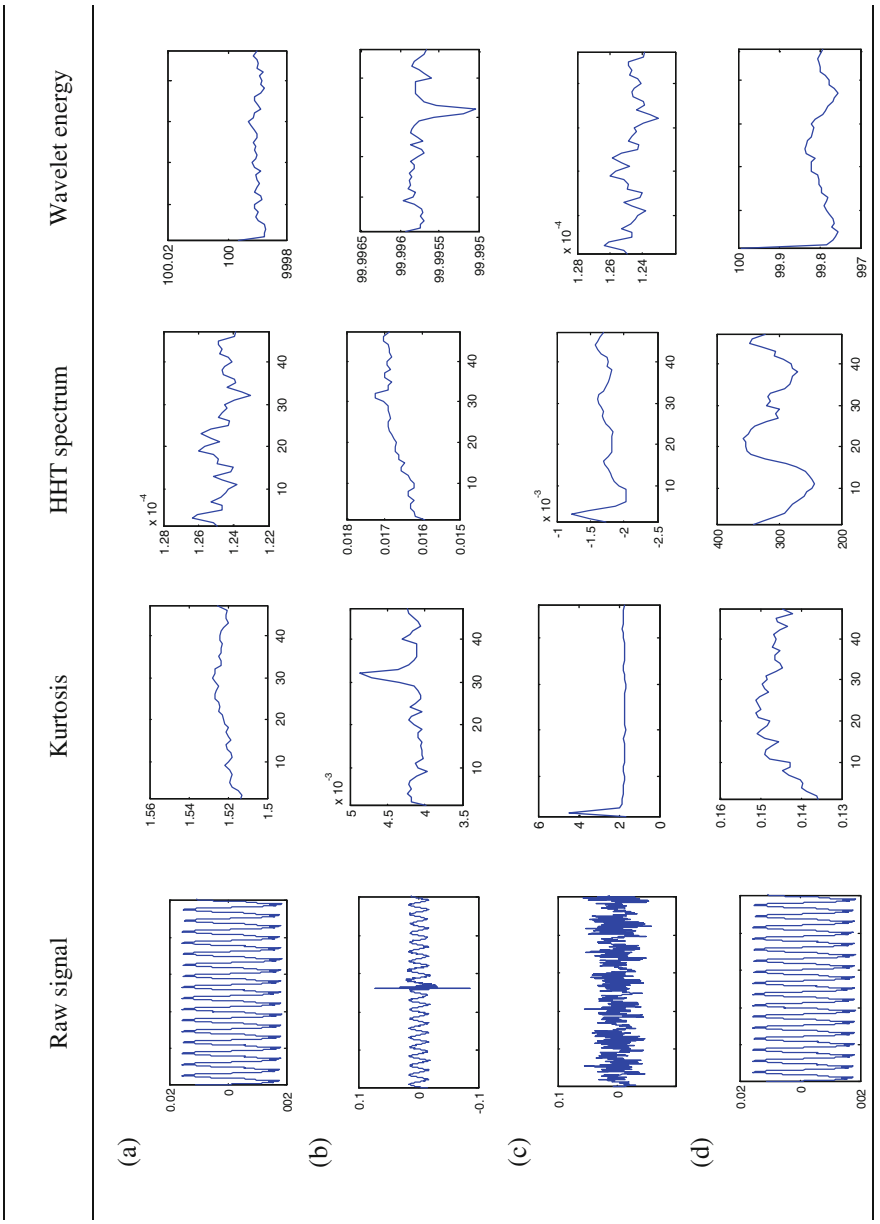


Fig. 2.1 Trend prognosis module for aero-engine abrupt failure

2.2.2 Feature Extraction

We adopted different feature extraction methods for failure signals. Fifteen basic features are extracted from raw signals, which include mean, median, variance, standard deviation, skewness, minimum value, maximum value, kurtosis, range, standard error, sum, maximum gradient, Hilbert-Huang Transform (HHT) spectrum, and first and second levels of energies in wavelet package decomposition. In order to choose discriminative features, reduce computational complexity, and add generalization performance, we use affinity propagation clustering to automatically select three features with high discrimination from these basic features. The resulting features are, respectively, kurtosis, HHT spectrum, and wavelet energy. Table 2.1 lists comparison between different failures in the feature space.

Table 2.1 Different failure features. (**a**) normal (**b**) abrupt failure of blade fracture (**c**) inner race failure of bearing (**d**) outer race failure of bearing



2.2.3 Affinity Propagation Clustering for Trend Features

The process of trend prognosis of aero-engine abrupt failure contains two steps, which contain finding exemplars of trend features in the database and prognosing failure trend by exemplars matching.

In the first step, according to the theory of affinity propagation, we need to determine the relationship between each pair of data points, through obtaining the index of each data point to determine which points are exemplars of failure trend database. A detailed process is as follows.

1. Calculate Euclidean distance d_{ij} between each pair of data points in X using Eq. 2.1.

$$d_{ij} = -\left\|x_i - x_j\right\|^2, \quad (1 \le i,j \le N) \tag{2.1}$$

2. Use binary variables $c_{ij}(1 \le i, j \le N)$ to represent the relationship of each pair of data points, such that $c_{ij} = 1$ if the exemplar for point i is point j; in this notion, $c_{jj} = 1$ indicates that j is an exemplar. c_{ij} is obtained by maximizing the sum of similarities between data points and their exemplars, which is shown in Eq. 2.2.

$$S(c_{ij}) = \sum_{i,j} s_{ij}(c_{ij}) + \sum_{i} I_i(c_{i1}, \ldots, c_{iN}) + \sum_{j} E_j(c_{1j}, \ldots, c_{Nj}). \tag{2.2}$$

In which

$$S_{ij}(c_{ij}) = \begin{cases} d_{ij} & \text{if} \quad c_{ij} = 1 \\ 0 & \text{otherwise.} \end{cases}, \tag{2.3}$$

$$I_i(c_{i1}, \ldots, c_{iN}) = \begin{cases} -\infty & \text{if } \sum_j c_{ij} \ne 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

$$E_j(c_{1j}, \ldots, c_{Nj}) = \begin{cases} -\infty & \text{if } \sum c_{jj} = 0 \ \ and \ \ \exists i \ne j \ s.t. \ c_{ij} = 1 \\ 0 & \text{otherwise} \end{cases}, \tag{2.5}$$

$S_{ij}(c_{ij})$ function nodes incorporate the input similarities d_{ij} between each data points and their potential exemplars. $\{I_i\}_{i=1}^N$ and $\{E_j\}_{j=1}^N$ are two constraint functions, function I_i operates on the ith row of c_{ij}, enforcing that x_i is assigned to only one cluster, the function E_j operates on the jth column of c_{ij}, ensuring that x_j chooses itself as an exemplar if another data point chooses x_j as its exemplar. We use graph theory to optimize the result of c_{ij}. We execute the max-sum algorithm

[12] on the graph; there are five message types passed between variable nodes and function nodes, update rules are Eqs. (2.6, 2.7).

$$\mu_{x\to f}(x) = \sum_{\{l|f_l\in ne(x)\backslash f\}} \mu_{f_l\to x}(x) \tag{2.6}$$

$$\mu_{f\to x}(x) = \max_{x_1,\ldots,x_M}\left[f(x,x_1,\ldots,x_m) + \sum_{\{m|x_m\in ne(f)\backslash x\}} \mu_{x_m\to f}(x_m)\right] \tag{2.7}$$

We use this algorithm to calculate the five messages, as shown in Eqs. (2.8–2.12).

$$s(i,j) = d_{ij} \tag{2.8}$$

$$\alpha_{ij} = \begin{cases} \sum_{k\neq j} \max[\rho_{kj},0], & i = j \\ \min\left[0, \rho_{jj} + \sum_{k\notin\{i,j\}} \max[\rho_{kj},0]\right], & i \neq j \end{cases} \tag{2.9}$$

$$\rho_{ij} = d_{ij} - \max_{k\neq j}(d_{ik} + \alpha_{ik}) \tag{2.10}$$

$$\beta_{ij} = d_{ij} + \alpha_{ij} \tag{2.11}$$

$$\eta_{ij} = -\max_{k\neq j} \beta_{ik} \tag{2.12}$$

Five messages are passed there iteratively, and the algorithm is said to converge once the message values no longer change. The final assignment of variable c_{ij} is based on the sum of all incoming messages to c_{ij}:

$$c_{ij}^* = \arg\max_{c_{ij}} \sum_{f_l\in ne(c_{ij})} \mu_{f_l\to c_{ij}}(c_{ij}) \tag{2.13}$$

where f_l represents S_{ij}, I_i, E_j. Till now, we get the result of c_{ij}, so we determine exemplars of failure trends in the database.

In the second step, we use the same algorithm to obtain the exemplar of new test data, and by matching the new exemplar with the exemplars in the first step by calculating Euclidean distances between them, as shown in Eq. 2.14.

$$D_i = \|X_i - X_n\| \tag{2.14}$$

where i is the index of exemplars in the database, and n represents the index of exemplar of new test data. If we set the threshold of distance as d, judge whether D_i are smaller than d; if it is so, choose the smallest value of D_i, the corresponding index is the failure trend class, inversely we can conclude that the new test data belongs to a new kind of failure trend.

2.3 Experiments

2.3.1 Physical Equipment Description

The data used in this paper is sampled from our test bench of aero-engine rotor as shown in Fig. 2.2, which is built for simulating failures including several types of abrupt failures of aero-engine.

As illustrated in Fig. 2.2, our rotor equipment contains blade, bearing, axle, motor, accelerometers, and eddy current sensors. Rated voltage of the motor is 220 V, rated current is 2 A, rated power is 300 W, and max rotation speed is 10,000 rpm. In this paper, we use two accelerometers to sample vibration signals, with a sampling frequency equal to 6 kHz, and set the motor speed to 1,200 rpm. After sampling data, we will test our method on these signals.

2.3.2 Trend Prognosis Results

In this paper, we simulated three kinds of failures in our experimental equipment. They contain blade fracture, inner race failure of bearing, and outer race failure of bearing. Features of failures are extracted, and a trend database is built for storing different types of failures. We sampled 47 feature points from each failure and clustered all the three kinds of points. Their exemplars are computed, and each point's class is determined. These points are labeled using different markers, and blank squares represent exemplars of failure trends in Fig. 2.3.

In order to test our prognosis scheme, we only extract exemplars of each type of failure trend, as shown in Fig. 2.3b. Point A is the exemplar of outer race failure trend, point B is the exemplar of abrupt blade fracture failure trend, and point C is the exemplar of inner race failure trend, respectively. The coordinates of three points A, B, and C are (1.5, 301.5, 99.9), (1.7, 214.7, 99.9), and (2.5, 104.6, 99.8), respectively. Then, we use another new sample of abrupt blade fracture failure as the test data, and Fig. 2.3c shows its exemplar. We first try to set the threshold to 1 according to the test data range. The coordinate of new test data exemplar N is (1.9, 127.7, 98.1); then, we calculate Euclidean distance between N and A, B, C. The distances d1, d2, and d3 are 0.41, 0.01, and 0.12, respectively. It is easy to see that d2 is smaller than two other values; moreover, it is smaller than the threshold value. Therefore, we can conclude that this new data belong to failure trend of abrupt blade fracture. Figure 2.3d shows that using our algorithm, the abrupt failure trend is able to be prognosed.

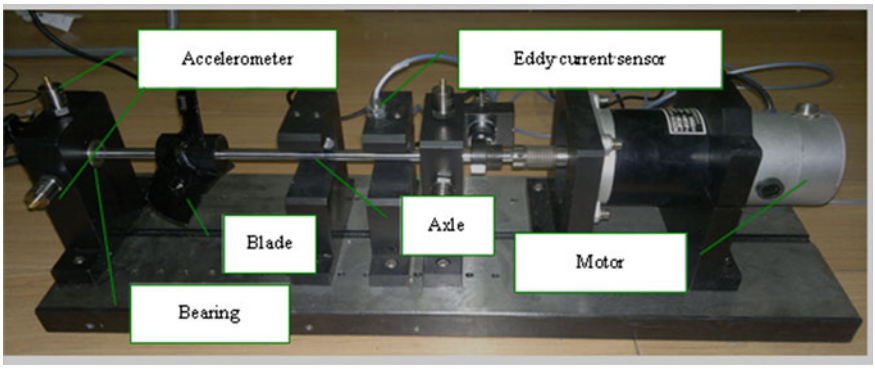


Fig. 2.2 Test bench of Aero-engine rotor

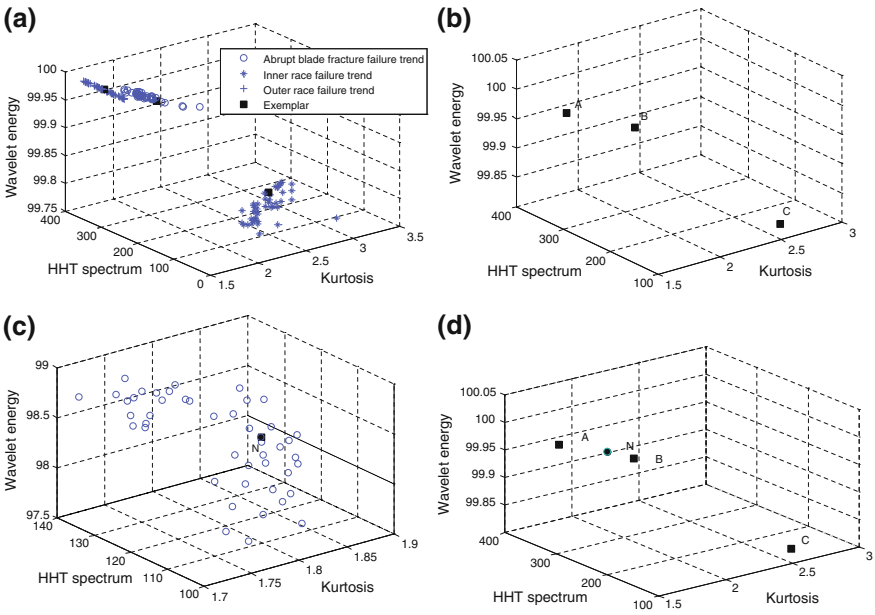


Fig. 2.3 Process of trend prognosis for aero-engine abrupt failure

2.3.3 Comparison with Other Methods

In order to check the performance of our scheme in aspects of time and accuracy, we compare clustering results on the same failure trend database to prognosis results on k-means [8] and k-centers [9, 10].

In order to reduce the random error, we run each program three times and got the average value of clustering time. From Table 2.2, we can see that while using

Table 2.2 Comparison of clustering time between methods based on affinity propagation, k-means, and k-centers

Time	Affinity propagation	k-means	k-centers
First	1.0168	3.0564	5.0513
Second	1.0021	3.0045	5.0141
Third	1.0007	3.0012	5.0002
Average	1.0215	3.0207	5.0219

Table 2.3 Comparison of error rate of failure trend prognosis between methods based on affinity propagation, k-means, and k-centers

	Affinity propagation	k-means	k-centers
Error (%)	0	16.31	46.09

affinity propagation to cluster failure trends, it took less time than k-means and k-centers methods.

The error of a clustering algorithm is the expected difference between its labels and the labels generated by the labeled point process Ξ, which was calculated by Eq. 2.15.

$$\varepsilon(S^{\Xi}, S^{\Phi}) = \frac{\left|\{x : I_{\Xi}(S;x) \neq I_{\Phi}(S;x)\}\right|}{|S|} \tag{2.15}$$

$I_{\Xi}(S;x)$ and $I_{\Phi}(S;x)$ denote the label of x for S^{Ξ} and S^{Φ} respectively, and $|\bullet|$ indicates the number of elements of a set. After running each algorithm for clustering failure trend one time, we obtained the error comparison results.

As shown in Table 2.3, while adopting affinity propagation algorithm to prognose failure trend, it produced less error than methods based on k-means and k-centers. The reason can be listed as follows:

1. Affinity propagation clustering determines relationship between points by adopting information exchange among all points, and this method belongs to global clustering partition. Moreover, it takes into account near regional relationship, which would produce less error, even no error at all.
2. K-means and k-centers clustering first divide data randomly into several parts and then adopt the iteration process to optimize some object functions.

   However, the setting of iteration itself is a source of error.

2.4 Conclusion

This paper presented a new method for trend prognosis of aero-engine abrupt failure which was based on clustering analysis. We used affinity propagation method to cluster vibration data from aero-engine rotating machinery to find the

failure trend by means of features extracted from it. This system could remind rapidly which failure will happen next. It is possible to be applied to trend prognosis of aero-engine abrupt failure in airlines.

References

1. Silveira E, Atxaga Irisarri G (2010) Failure analysis of two sets of aircraft blades. Eng Fail Anal 17(3):641–647
2. Etemadi AR, Behjati P, Madaah Hosseini HR, Kokabi AH (2010) Failure analysis of brazed air passages of an aircraft fuel system. Eng Fail Anal 17(6):1495–1499
3. Ukil A, Zivanovic R (2007) The detection of abrupt changes using recursive identification for power system fault analysis. Electr Power Syst Res 77(3–4):259–265
4. Baker A, Rajic N, Davis C (2009) Towards a practical structural health monitoring technology for patched cracks in aircraft structure. Appl Sci Manuf 40(9):1340–1352
5. Sathiya Naarayan S, Pavan Kumar DVTG, Chandra S (2009) Implication of unequal rivet load distribution in the failure the and damage tolerant design of metal and composite civil aircraft riveted lap joints. Eng Fail Anal 16(7):2255–2273
6. Smith Cary, Akujuobi Cajetan M, Hamory Phil, Kloesel Kurt (2007) An approach to vibration analysis using wavelets in an application of aircraft health monitoring. Mech Syst Signal Process 21(3):1255–1272
7. Pecht M, Jaai R (2010) A prognostics and health management roadmap for information and electronics-rich systems. Microelectron Reliab 50(3):317–323
8. Yiakopoulos CT, Gryllias KC, Antoniadis IA (2011) Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. Expert Syst with Appl 38(3):2888–2911
9. Ypma A, Pajunen P (1999) Rotating machine vibration analysis with second-order independent component analysis. In: Proceedings of the first international workshop on independent component analysis and signal separation, pp 37–42
10. Luo M (2006) Data-driven fault detection using trending analysis. PhD Dissertation, Louisiana State University
11. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315(5814):972–976
12. Kschischang FR (2001) Factor graphs and the sum-product algorithm. IEEE Trans Inf Theory 47(2):498–519

http://www.springer.com/978-3-642-54232-9

Proceedings of the First Symposium on Aviation Maintenance and Management-Volume II

Wang, J. (Ed.)

2014, XVI, 714 p. 303 illus., Hardcover
ISBN: 978-3-642-54232-9