International Journal of Rotating Machinery
2000, Vol. 6, No. 1, pp. 19–27
Reprints available directly from the publisher
Photocopying permitted by license only

An Efficient Method of Vibration Diagnostics for Rotating Machinery using a Decision Tree*

BO SUK YANG[a,†], CHUL HYUN PARK[b] and HO JONG KIM[b]

[a]*School of Mechanical & Automotive Engineering, Pukyong National University, Pusan, Korea;*
[b]*Research & Development Center, Korea Heavy Industries & Construction Co., Changwon, Korea*

(Received 15 May 1998; In final form 17 July 1998)

This paper describes an efficient method to automatize vibration diagnosis for rotating machinery using a decision tree, which is applicable to vibration diagnosis expert system. Decision tree is a widely known formalism for expressing classification knowledge and has been used successfully in many diverse areas such as character recognition, medical diagnosis, and expert systems, etc. In order to build a decision tree for vibration diagnosis, we have to define classes and attributes. A set of cases based on past experiences is also needed. This training set is inducted using a result–cause matrix newly developed in the present work instead of using a conventionally implemented cause–result matrix. This method was applied to diagnostics for various cases taken from published work. It is found that the present method predicts causes of the abnormal vibration for test cases with high reliability.

Keywords: Vibration, Diagnostics, Expert system, Decision tree, Rotating machinery

1 INTRODUCTION

A number of rotating machines are installed in power plants and petroleum refinery plants. Malfunction of a rotating machine due to some defects may cause shutdown of the plants resulting in high maintenance cost. It has been considered very significant to detect the defects of the machines at an early stage. The defects can be detected early by using vibration signals, since vibration signals contain current dynamic characteristics of the rotating machines. Moreover, the abnormal vibration is critical to operation of high speed rotating machines. So the necessity of vibration diagnostics for rotating machinery is gradually increasing. Vibration diagnosis require deep knowledge on dynamics of the rotating machines and operation principle, because vibration signals are very complicated and the causes of abnormal vibration are interrelated to one another. However, operators in the plants are usually not experts on the vibration of the rotating machines. Thus a lot of research has been done to make it possible for even an operator who is not an expert can diagnose the cause of

* This paper was originally presented at ISROMAC-7.
† Corresponding author. Tel.: +82 51 620 6450.

abnormal vibration using automatic diagnosis system. An expert system is a system that represents the knowledge of a special domain expert. Every expert system is composed of an inference engine and a knowledge base.

In this paper, the decision tree will be introduced to acquire the structured knowledge in the form of concepts and to build the knowledge base which is indispensable for the vibration expert system. The decision tree is a technology that builds the knowledge-based system by the inductive inference from cases and the decision tree itself can play the role of a diagnosis tool (Mui and Fu, 1980). The induction task is to develop classification rules that can determine the class of any case from its values of the attributes. These classification rules will be expressed as a decision tree. Since the decision tree is induced from cases, a set of cases relevant to the special domain is needed for decision tree (Safavian and Landgrebe, 1991). This set is called the training set. Each case in the training set is described in terms of a collection of attributes and has known classes. Therefore, considering induction tree for vibration diagnosis, we first have to construct a training set relevant to the vibration diagnosis, and then induce the decision tree. The training set is generated using a result–cause matrix newly developed in the present work, instead of using a conventionally implemented cause–result matrix. Of algorithms for induction of decision tree, the algorithm developed by Quinlan (1986; 1987; 1992) is basically adopted in this paper. The method developed is applied to a variety of cases taken from rotating machines to verify its reliability.

2 DECISION TREE

This technology for building knowledge-based systems by inductive inference from case histories is a typical algorithm used for construction of a model of the knowledge used by a human expert.

Because a decision tree executes supervised learning, classes and attributes must be established beforehand. The class is a category to which each case belongs. Each attribute measures some important features of a case, and may have either discrete or numeric value. All information about one case must be expressible in terms of a fixed collection of attributes. A set of training cases with known classes is called a training set. Training cases from which a classification rule is developed are known only through their values of a set of attributes, and the decision trees in turn are expressed in terms of these same attributes. In order to correctly classify each case in the training set, we start at the root of the tree, then evaluate the test, and take a branch appropriate to the outcome. The process continues until a leaf is encountered, at which time the case is asserted to belong to the class named by the leaf, i.e. the successive division of the set of training cases proceeds until all the subsets consist of cases belonging to a single class. Thus the classification rules are assigned to each node.

The above process is represented graphically in Fig. 1. Let T be a training set. If T is empty or contains only cases of one class, the simplest decision tree is just a leaf labeled with the class. Otherwise, if X is any decision rule on a case with possible outcomes $O_1, O_2, \ldots, O_n$, each object in T will give one of these outcomes for X. Thus, X produces a partition $\{T_1, T_2, \ldots, T_n\}$ of T with T_i containing those cases having outcome O_i. If each subset T_i could be replaced by a decision tree for T_i, the result would be a decision tree for all of T.

2.1 Selection of Test

The structure of the decision tree highly depends on selection of test (X) as a root. As a criterion for

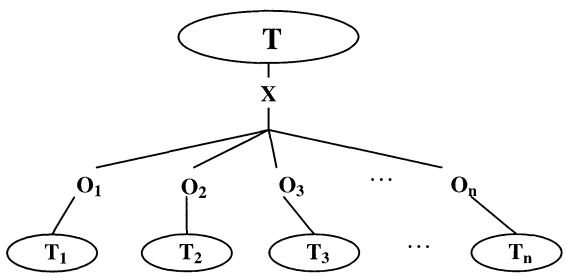


FIGURE 1 A decision tree structured with case in T.

selection, we use the information–entropy evaluation function (IEEF) based on the information theory (Quinlan, 1986) proposed by Shannon. The calculation procedure of a IEEF is as follows:

Step 1: Calculate the average amount of information, info (T), needed to identify the class of a case in a training set T.

$$\text{info}_X(T) = -\sum_{j=1}^{k}\left\{\frac{\text{freq}(C_j, T)}{|T|} \times \log_2\left(\frac{\text{freq}(C_j, T)}{|T|}\right)\right\} \tag{1}$$

where $|T|$ is the number of cases in T, C_j is the j-th class, k is the number of classes, $\text{freq}(C_j, T)$ is the number of classes belonging to C_j in T.

Step 2: Calculate the expected information, $\text{info}(T)$, required for a test X to classify T.

$$\text{info}_X(T) = -\sum_{i=1}^{n}\left\{\frac{|T_i|}{|T|} \times \text{info}(T_i)\right\} \tag{2}$$

where n is the number of outcomes of X, T_i is a subset of T which is divided due to the i-th outcome.

Step 3: Calculate the mutual information, $\text{gain}(X)$, obtained from classification of T depending on X.

$$\text{gain}(X) = \text{info}(T) - \text{info}_X(T). \tag{3}$$

Step 4: Calculate the split information, split $\text{info}(X)$, obtained from classification of T into n subsets.

$$\text{split info}(X) = -\sum_{i=1}^{n}\left\{\frac{|T_i|}{|T|} \times \log_2\left(\frac{|T_i|}{|T|}\right)\right\} \tag{4}$$

where n is the number of outcomes of X.

Step 5: Calculate the ratio of the mutual information to the split information, $\text{GR}(X)$.

$$\text{GR}(X) = \text{gain}(X)/\text{split info}(X). \tag{5}$$

The gain ratio, $\text{GR}(X)$, makes up for a shortcoming of the gain, $\text{gain}(X)$, which biases to an attribute of larger attribute value. The gain ratio also represents quantity of information provided by X in a training set. Thus, we select an attribute or test of the highest GR value as the root of the decision tree.

2.2 Evaluation of Unknown Attribute Values

In the previous section, if an attribute value is unknown, the outcome of a test cannot be determined since every test is based on a single attribute. In real world, occasionally one may not obtain necessary data, which requires special treatment. Problems coming from the unknown attribute values can be classified into three categories, and are solved through the following processes.

2.2.1 *Generalization of the Gain Ratio*

Selection of a test on which the training set is partitioned is made on the basis of heuristic criteria such as the gain or the gain ratio. If the attribute values are unknown in the tests, one may not know relative desirability of the attributes due to shortage of information. Therefore, generalized gain ratios should be computed taking into account the shortage of information.

For an attribute, let F be the ratio of a cause with known attribute values to total cases. The generalized $\text{gain}(X)$ is calculated by Eq. (6). The number of outcomes becomes $n+1$ considering that a case of an unknown attribute value be another set of attributes. Thus, the generalized gain ratio is obtained from Eq. (8).

$$\text{gain}(X) = F \times \{\text{info}(T_{\text{known}}) - \text{info}_X(T_{\text{known}})\}, \tag{6}$$

$$\text{split info}(X) = -\sum_{i=1}^{n+1}\left\{\frac{|T_i|}{|T|} \times \log_2\left(\frac{|T_i|}{|T|}\right)\right\}, \tag{7}$$

$$\text{GR}(X) = \text{gain}(X)/\text{split info}(X) \tag{8}$$

where T_{known} is a set of cases with known attribute values.

2.2.2 *Partitioning of the Training Set*

Once a test has been selected, training cases with unknown values of the relevant attribute cannot be associated with a particular outcome of the test, and so cannot be assigned to a particular subset T_i.

We introduce a weighting factor, w, representing probability of a case to belong to a subset. If the case with seen outcome O_i is assigned to a subset T_i, probability of belonging of the case to T_i is 1, while probability to remaining subsets is 0. However, if the case is with an unseen outcome, the probability will be between 0 and 1. In this case, the subset T_i becomes a set of probabilistic cases.

2.2.3 *Classification of Unseen Cases*

When the decision tree is used to classify a case, the outcome will not be decided if the case has an unknown value for the attribute tested in the current decision node. Similarly as in the previous subsection, the system explores all possible outcomes and combines the resulting classifications arithmetically. That is, since there can be multiple paths from the root of a tree or subtree to leaves, a classification is a class distribution rather than a single class. A class with the highest probability becomes the class of the unseen case.

2.3 Pruning

Pruning is a simplification process of the decision tree. For a set with unseen cases, capacity of the classification is decreased and evaluation is not feasible. Pruning strengthens the classification for the population, and evaluates the decision tree to classify unseen cases, with reasonable degree of reliability. This evaluation presents reliability of the decision tree, and is called predictive accuracy. The inverse of the predictive accuracy is called predictive error rate. The following way (Quinlan, 1987) in which the recursive partitioning method can be modified to produce simpler trees is used in the present work.

Let S and $L(S)$ be a subtree of a decision tree and the number of leaves belonging to S, respectively. And let $\sum N$ and $\sum E$ be the sum of total leaves of S and the sum of errors, respectively. In this case, the pruning assumes that, for unseen cases of $\sum N$, the number of misclassifications would be $\sum E + L(S)/2$, following the continuity correction for a binomial distribution. The standard error, for the errors $\mathrm{se}(\sum E)$, is defined by the following:

$$\mathrm{se}\left(\sum E\right) = \sqrt{\frac{\sum E \times (\sum N - \sum E)}{\sum N}}. \quad (9)$$

If the leaf has error not larger than $\sum E + \mathrm{se}(\sum E)$ after S has been replaced by the best leaf, the subtree is replaced with the leaf. This procedure makes the decision tree the simplest and the lowest error for given training sets.

3 APPLICATION AND RESULTS

Based upon past experiences on vibration of the rotating machines considered, we divided causes of vibration into 14 classes, and divided vibration phenomena into 20 attributes, which are summarized in Tables I and II. These classes and attributes are enough to represent vibration characteristics of

TABLE I Class of the decision tree

No.	Class (cause of vibration)	Symbol
1	Mechanical unbalance	Unbalance
2	Misalignment	Misalignment
3	Partial rub	Partial rub
4	Crack	Crack
5	Mechanical looseness	Looseness
6	Ball bearing damage	Ball bearing damage
7	Foundation distortion	Foundation distortion
8	Critical speed(1× resonance)	Critical speed
9	Subharmonic resonance	Subharmonic resonance
10	Oil whip/oil whirl	Oil whip/whirl
11	Vane passing vibration	Vane passing vibration
12	Clearance induced vibration	Clearance induced vib.
13	Static eccentricity of air gap or stator damage	Static air gap
14	Dynamic eccentricity of air gap or rotor damage	Dynamic air gap

TABLE II Attribute and attribute value of the decision tree

No.	Attribute	Symbol	Attribute value	Symbol
1	What is the predominant frequency?	Predominant frequency	1× component	1×
			2× component	2×
			1× and 2×	like same
			harmonics of 1×	multiples
			higher components more than 1×	higher
			lower components less than 1×	lower
2	Is there a natural frequency?	Natural frequency	yes	yes
			no	no
3	Is 0.4–0.48× component predominant?	0.4–0.48×	yes	yes
			no	no
4	Is 0.5–1× component predominant?	0.5–1×	yes	yes
			no	no
5	Is bearing damage frequency predominant?	Zf_{bearing}	yes	yes
			no	no
6	Is vane passing frequency predominant?	Zf_{vane}	yes	yes
			no	no
7	Is subharmonic predominant?	Subharmonic	yes	yes
			no	no
8	Is there harmonics of 1/2× component?	1/2-multiple	yes	yes
			no	no
9	Is there intense noise at the high frequency area?	High frequency	yes	yes
			no	no
10	Is there line frequency?	f_0	yes	yes
			no	no
11	Is there two times frequency as large as line frequency?	$2f_0$	yes	yes
			no	no
12	Is there a pulsation component, $2sf_0$	Pulse	yes	yes
			no	no
13	Do phase and amplitude of 1× component change?	1× -change	yes	yes
			no	no
14	Do phase and amplitude of 2× component change?	2× -change	yes	yes
			no	no
15	Does runout vector change?	Runout change	yes	yes
			no	no
16	Is axial amplitude larger than lateral amplitude?	Axial amplitude	yes	yes
			no	no
17	Is orbit shape leaning to one side or eight shape?	Orbit shape	yes	yes
			no	no
18	What is direction of orbit?	Orbit direction	forward direction	forward
			backward direction	backward
19	How is amplitude change during shut-down?	Shut-down	almost constant	stays same
			decrease at the same ratio	decrease
			temporary pause	temporary pause
			drop out suddenly	rapidly drop
20	What is predominant location of vibration?	Location	shaft	shaft
			bearing	bearing
			pipe system	piping
			casing	casing
			coupling	coupling

most rotating machines. The classes taken can be divided into three groups, depending upon sources of the vibration. The first group includes mechanical unbalance, misalignment, partial rub, crack, mechanical looseness, ball bearing damage, foundation distortion, and critical speed. These are mechanical faults. The second group is related with fluid flow in the machines, and includes subharmonic resonance, oil whirl/whip, vane passing vibration, and clearance induced vibration. The third group is related with electrical faults, and includes static eccentricity of air gap or stator damage and dynamic eccentricity of the air gap or rotor damage.

The attributes taken are frequency, amplitude, phase, trend, and location of vibration. Since vibration spectra usually contain such information for diagnostics, the frequency is divided into several components as shown in the tables. The classes and attributes are represented with corresponding symbols for implementation in the expert system.

We utilized the cause–result matrix (Sohre, 1980) in order to construct training set for decision tree. Sohre's cause–result matrix has been utilized in most automatic vibration diagnosis system. But this decision tree using Sohre's cause–result matrix was short of reliability. So, both attribute and class were defined again, and a new result–cause matrix improving its reliability was developed to induce vibration diagnosis decision tree. Table III shows a part of new result–cause matrix. The values in the table implies probability of frequency components for the vibration causes. Figure 2 shows the decision tree inducting throughout this training set.

The decision tree was applied to diagnostics for various rotating machines. Sample cases were taken from published work. Results of the diagnostics are summarized in Table IV. The table compares diagnostic result from using the present method with those obtained from the conventional cause–result matrix proposed by Sohre (1980). As shown in the table, the present method is found to predict causes of vibration more reasonably than the conventional method.

TABLE III Result–cause matrix of classes to attributes

	Predominant frequency					
Vibration cause	1×	2×	Almost same	Multiples	Higher	Lower
Unbalance	30%		30%			
Misalignment		50%	70%	30%		
Partial rub	13%	8%		13%	10%	20%
Crack		30%		13%		
Looseness				38%		
Ball bearing damage	13%				30%	
Foundation distortion		12%				15%
Critical speed	13%					
Subharmonic resonance						20%
Oil whip/whirl	5%			6%		25%
Vane passing vibration	13%				30%	
Clearance induced vibration						20%
Static air gap					30%	
Dynamic air gap	13%					

For an example, we take the case #2 in the table. This case represents an excessive vibration of a centrifugal fan with 24 vanes running at 3360 rpm. The vibration spectrum is shown in Fig. 3. From the spectrum, we see that the vibration is higher at 1344 Hz than at the running speed (56 Hz). The vane passing frequency is the product of the running speed times the number of vanes, which is 1344 Hz ($56\,\text{Hz} \times 24 = 1344\,\text{Hz}$). For this case, the decision tree first searches for condition of the predominant frequency, resulting in "higher". Then, the next step decides if the high frequency is originated from the ball bearing damage or the vane passing vibration. In this case, information on ball bearings are not available. Thus, the final diagnostics results in 90% probability of the vane passing vibration instead of 100% probability.

For another example, we take the case #4 in the table. This case is an excessive vibration of a centrifugal pump due to misalignment. The pump runs at 3600 rpm. In Fig. 4, vibration amplitudes

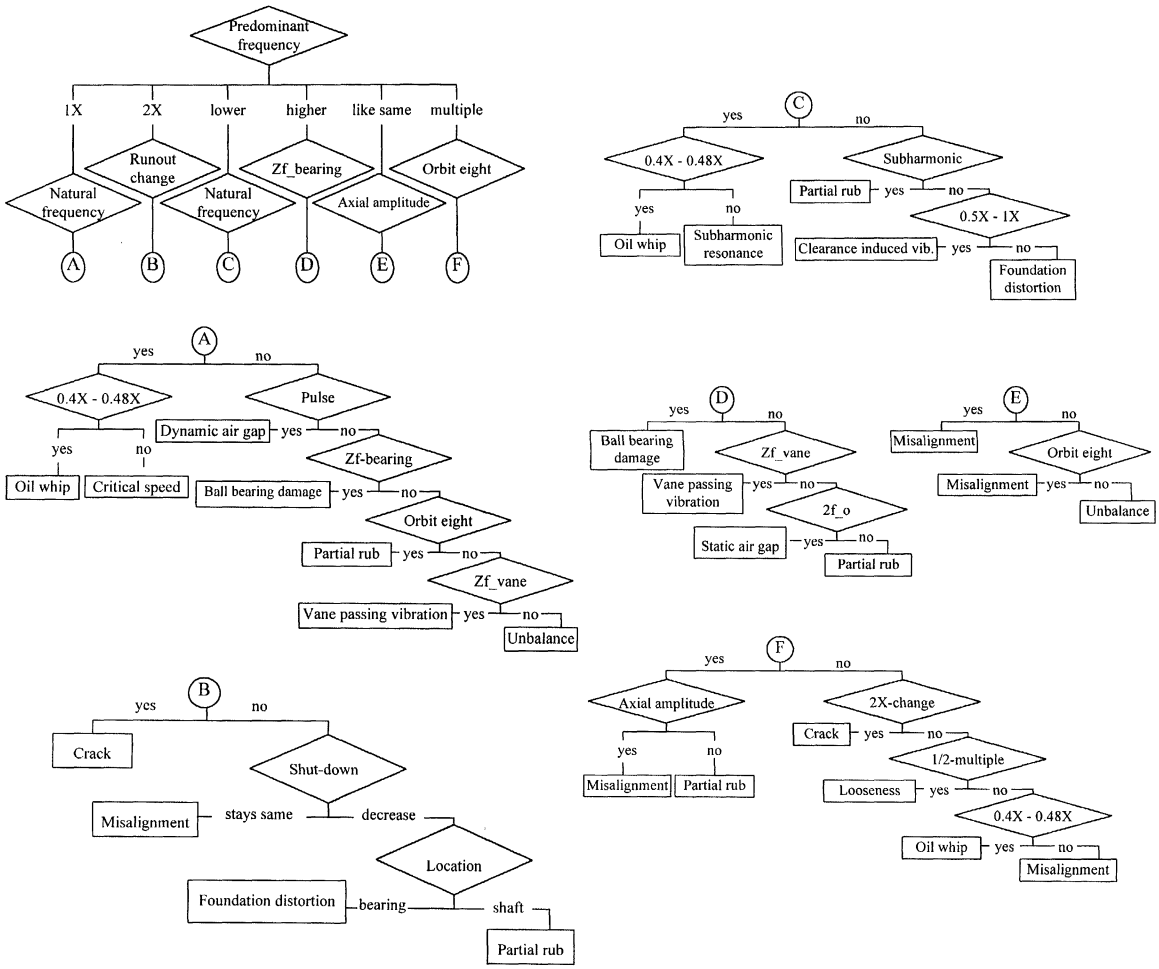


FIGURE 2 Decision tree for vibration diagnostics.

are summarized for each location of measurement. The vibration amplitude is the highest at the inner bearing, with 40.665 mm/s. The highest frequency component is 120 Hz, i.e., 2 × revolution vibration, as shown in Fig. 4. In diagnosing the vibration, the present method first asks the frequency component and then 2 × is selected. In the next step, the run-out vector is checked. For this case, since data on the run-out vector is not known, the more probable branch is selected, based on probabilistic branching by the branching algorithm of the present method. In the next step, amplitude change during shut down is checked. Since data is not known, the same principle is applied, as in the previous step. In the final step, the location of the most severe vibration is checked. The diagnostic result shows that misalignment is the primary source of vibration with probability of 47%. In this sample case, since known data are very restricted, the probability is not so high although the primary source of vibration is correctly diagnosed. Thus, when one diagnoses the vibration of the rotating machines, one has to acquire as much data as possible, including not only steady state and transient vibration data but also data on operation of the machines.

TABLE IV Diagnostics result of rotating machinery

Machines		Vibration cause	Diagnosis result			
			Result–cause matrix		Cause-result matrix	
1	Induction motor, Wowk (1991)	Looseness	Looseness	50%	Unbalance	36%
			the others	50%	the others	64%
2	Centrifugal fan, Wowk (1991)	Vane passing vibration	Vane passing vibration	90%	Bearing damage	67%
			Ball bearing damage	10%	Seal rub	33%
3	Centrifugal pump, Mitchell (1993)	Vane passing vibration	Vane passing vibration	90%	Vane passing vibration	75%
			Ball bearing damage	10%	the others	25%
4	Centrifugal pump, Miguel (1996)	Misalignment	Misalignment	47%	Misalignment	45%
			Foundation distortion	53%	the others	55%
5	Draft fan, Wowk (1991)	Misalignment	Misalignment	54%	Vane passing vibration	90%
			the others	46%	Seal rub	10%
6	Fan, Wowk (1991)	Unbalance	Unbalance	29%	Unbalance	36%
			the others	71%	the others	64%
7	Steam turbine, Mariano (1993)	Misalignment	Misalignment	100%	Misalignment	100%
8	Turbo-blower, Maegawa (1990)	Oil whip	Oil whip/whirl	72%	Oil whip	59%
			the others	28%	the others	41%
9	Air compressor, Maki (1993)	Oil whip	Oil whip/whirl	53%	Oil whip	59%
			the others	47%	the others	41%
10	Centrifugal pump, Maki (1993)	Misalignment	Misalignment	51%	Foundation distortion	80%
			the others	49%	Misalignment	20%

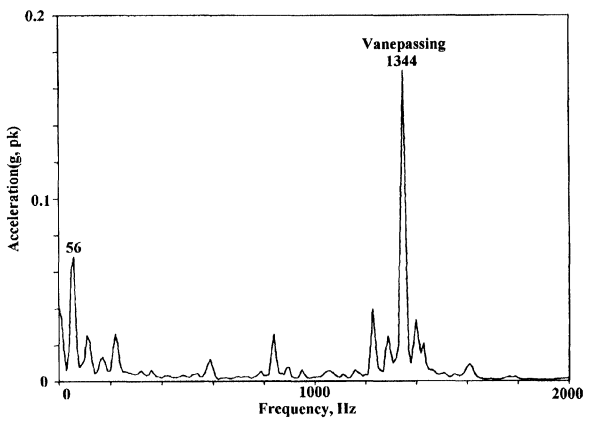
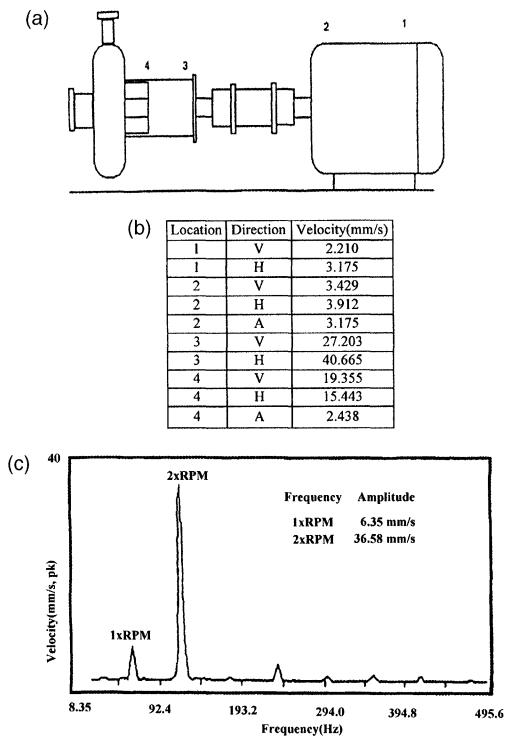


FIGURE 3 Vibration spectrum of centrifugal fan (Wowk, 1991).

4 CONCLUSION

A method of vibration diagnostics for rotating machines has been developed, using a decision tree with a result–cause matrix. The method was applied to a variety of rotating machines. It is found that the present method diagnoses vibration causes of the rotating machines relatively well.



Location	Direction	Velocity(mm/s)
1	V	2.210
1	H	3.175
2	V	3.429
2	H	3.912
2	A	3.175
3	V	27.203
3	H	40.665
4	V	19.355
4	H	15.443
4	A	2.438

FIGURE 4 Vibration spectrum of centrifugal pump (Wowk, 1991). (a) Measuring points; (b) Overall vibration level; (c) Vibration spectrun at the location 3H.

NOMENCLATURE

C_j	j-th class
F	ratio of a cause with known attribute values to total cases
freq(C_j, T)	number of classes belonging to C_j in T
gain(X)	mutual information
GR(X)	ratio of the mutual information to the split information
info$_X$(T)	expeted information
k	number of classes
n	number of outcomes of test X
O_i	possible outcome of T_i
split info(X)	split information
T	training set
T_i	subset of T
$\|T\|$	number of cases
T_{known}	set of cases with known attribute values
X	any decision rule with possible outcomes $O_1, O_2, \ldots, O_n$

References

Maegawa, K. (1990) *Plant Diagnosis Technique*, Japanese Plant Maintenance Association, Tokyo, p. 107.

Maki, S.L. (1993) *Practice of Machine Diagnosis Technique by Vibration Method*, Japanese Plant Maintenance Association, Tokyo, p. 268, 288.

Mariano, S.C. (1993) Misalignment and balancing problems at a refinery, *ORBIT*, pp. 20–22.

Miguel de Priego, J.C. (1996) Vibration analysis of a centrifugal pump, *Vibration*, **12**(3), 14–15.

Mitchell, J.S. (1993) *Introduction to Machinery Analysis and Monitoring*, Pennwell Books, Oklahoma, p. 203.

Mui, J.K. and Fu, K.S. (1980) Automated classification of nucleated blood cells using binary tree classifier, *IEEE Trans. PAMI*, **2–5**, 429–443.

Quinlan, J.R. (1986) Induction of decision tree, *Machine Learning*, **1**, 81–106.

Quinlan, J.R. (1987) Simplifying decision trees, *International Journal of Man-Machine Studies*, **27**, 221–234.

Quinlan, J.R. (1992) *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, California, 17–44.

Safavian, S.R. and Landgrebe, D. (1991) A survey of decision tree classifier methodology, *IEEE Transactions on Systems, Man, and Cybernetics*, **21**(3), 660–674.

Sohre, J.S. (1980) *Turbomachinery Problems and Their Correction, Sawyer's Turbomachinery Maintenance Handbook, Vol. 2*, Gas Turbine Publications, Inc., Connecticut, Ch. 7.

Wowk, V. (1991) *Machinery Vibration, Measurement and Analysis*, McGraw-Hill, New York. Ch. 5.