

UTILISATION DE L'ARBRE DE DECISION POUR LA DETECTION DES DEFAUTS DE ROULEMENTS

M. BOUMAHDJ¹, J.P. DRON², S. RECHAK³ et O. COUSINARD⁴

¹ Université de Médéa, LBMPT, Médéa, Algérie;

^{2,4} Laboratoire de mécanique appliquée, URCA/GRESPI Reims, France;

³ Laboratoire de génie mécanique et développement, ENP Alger, Algérie
jp.dron@univ-reims.fr

ABSTRACT

For vibratory follow-up of the rotating machinery, different techniques and tools are used for, on the one hand to detect a failure then to evaluate the degree of severity of this failure. Among the most sensitive machine parts, we can cite the bearing. The techniques of monitoring are currently well established and based on signals vibration analysis resulting from accelerometers. They allow statistical analysis of the signals (Crest factor, Kurtosis, root mean square RMS) which highlights a vibrations global level variation and the frequential contents analysis that allows to locate the origin of this level variation more precisely. Our objective is in the use of a decision tree approach for the monitoring and the diagnosis of the bearing defects. The construction of this tree is based on the application of an algorithm on dataset resulting from experiments realized on ball bearings. This approach allows to model the state of bearing and to generate the expert rules. It is based on criteria for the choice of a monitoring or diagnosis indicator which gives the maximum of information in relation to the other indicators. We show in this study the advantage of the decision tree to supervise and diagnose the bearings defects.

RÉSUMÉ

Pour le suivi vibratoire des machines tournantes, différentes techniques et outils sont utilisés pour, d'une part détecter une défaillance puis pour évaluer le degré de sévérité de cette défaillance. Parmi les organes mécaniques les plus sensibles, on peut citer le roulement. Les techniques de surveillance sont actuellement bien établies et sont basées sur l'analyse de signaux vibratoires issus d'accéléromètres. Elles permettent l'analyse statistique des signaux (Facteur de crête, Kurtosis, valeur RMS) qui met en évidence une variation du niveau global des vibrations et l'analyse du contenu fréquentiel (spectre, cepstre, détection d'enveloppe...) qui permet de localiser plus précisément l'origine de cette variation de niveau. Notre objectif réside dans l'utilisation de l'approche arbre de décision pour la surveillance et le diagnostic des défauts de roulements. La construction de cet arbre est basée sur l'application d'un algorithme de classification sur un ensemble de données issu des expériences réalisées sur des roulements à billes. Cette approche permet de modéliser l'état de roulement et de générer des règles expertes. Elle repose sur des critères pour le choix d'un indicateur de surveillance ou de diagnostic qui apporte le maximum d'informations par rapport aux autres indicateurs. Nous montrons dans cette étude l'intérêt de l'arbre de décision pour surveiller et diagnostiquer les défauts de roulements.

1. INTRODUCTION

Plusieurs éléments de machines tournantes produisent des vibrations lorsqu'ils sont partiellement ou totalement dégradés. La mise en place d'une maintenance conditionnelle basée sur l'analyse vibratoire permet d'obtenir une disponibilité accrue de ces machines et d'éviter des arrêts non planifiés dus aux pannes. Elle permet notamment de détecter et d'identifier les défauts se manifestant dans les roulements, l'un des éléments importants des machines tournantes. Il peut s'agir d'un défaut de cage, de bille ou de piste sur la bague intérieure ou extérieure. Ces défauts se traduisent par des chocs périodiques, et par une augmentation de la valeur efficace du signal et du facteur de crête [1]. Ils sont caractérisés par des fréquences de répétitions. Harris [2] présente les expressions de ces fréquences en fonction des paramètres géométriques du roulement. Cependant, pour détecter ou identifier un défaut, il faut utiliser le traitement de signal le plus approprié et choisir l'indicateur le plus pertinent. Dans le cas d'une détection des défauts de roulement, plusieurs études ont montré l'apport des indicateurs scalaires, tels que Kurtosis, Facteur de crête et valeur efficace. Néanmoins, le kurtosis semble plus sensible que le facteur de crête en terme de détection [3]. Dans le cas d'une identification des défauts de roulement, des études ont montré que l'analyse du contenu fréquentiel (spectre, cepstre, détection d'enveloppe...) est un outil essentiel pour les localiser avec précision. La référence [4] présente une synthèse des différentes techniques utilisées dans l'analyse vibratoire des défauts de roulement.

Dans le processus de la maintenance conditionnelle, l'interprétation des signaux vibratoires reste le domaine du spécialiste. Il ressort d'études antérieures [5] que les systèmes experts permettent d'envisager une aide appréciable à la décision durant le processus de la maintenance où les raisonnements ne sont plus bridés par un algorithme. Sachant que ces systèmes simulent le raisonnement d'un (ou plusieurs) spécialistes de ce problème et comportent une base de connaissances, sous forme de règles, une base de faits et un moteur d'inférence. Un système expert peut être envisagé si l'on dispose d'une expertise suffisante dans le domaine visé et si cette expertise est formalisable en termes de règles. Dans le cas où les experts n'arrivent pas à dégager les règles qu'ils utilisent. Elles peuvent être générées à partir des données historiques de la machine et par l'utilisation de l'une des méthodes d'apprentissage automatique. Cette approche est appelée datamining qui vise à extraire des connaissances à partir des données.

Dans cette étude, on utilise l'approche d'apprentissage par arbres de décision, une branche de l'apprentissage automatique, pour modéliser l'état de roulement graphiquement puis l'extraction des connaissances, sous forme de règles, à partir des données issues de l'analyse vibratoire. Des études récentes [6] [7] ont montré l'apport de cette approche dans la surveillance et le diagnostic des défauts dans une machine tournante.

Les règles extraites par cette approche peuvent être exploitées ultérieurement dans le développement d'un système expert. En outre, les arbres de décisions représentent un formalisme compréhensible par tout utilisateur et aussi un outil d'aide à la décision. Dans cette étude, pour construire cet arbre, on utilise l'algorithme de classification C4.5, incorporé dans le logiciel WEKA sous code J48.

C'est dans ce contexte que s'inscrit la présente étude à savoir d'appliquer l'approche d'arbre de décision sur un ensemble de données issues des essais expérimentaux sur des roulements à billes sans ou avec défauts (défaut sur la bague interne, défaut sur la bague externe et deux défauts, un

sur la bague interne et l'autre sur la bague externe.) pour différentes charges radiales et pour différentes vitesses de rotation.

2. ÉTUDE THEORIQUE

Analyse statistique des signaux

Divers indicateurs sont utilisés dans le suivi vibratoire des roulements, on peut citer à cet effet, la valeur de l'accélération efficace (ou RMS), le Kurtosis et le Facteur de crête.

L'accélération efficace Acc_{eff} , (ou RMS) est définie par l'expression suivante:

$$Acc_{eff} = RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (1)$$

x_i l'amplitude du signal pour l'échantillon i , [m/s^2]

N le nombre d'échantillons prélevés dans le signal.

Le Kurtosis est un paramètre statistique permettant d'analyser la répartition des amplitudes vibratoires continues dans un signal temporel. Défini comme suit :

$$Kurtosis = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^4}{\left[\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right]^2} \quad (2)$$

Avec

$$\bar{x} \text{ la valeur moyenne des amplitudes, tel que : } \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

Le facteur de crête FC correspond au rapport entre la valeur crête d'un signal (en valeur absolue) et sa valeur efficace. Défini comme suit:

$$FC = \frac{Acc_{crête}}{Acc_{eff}} = \frac{\sup |x_i|}{\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}} \quad (4)$$

La présence d'une anomalie peut être détectée si un indicateur dépasse un seuil prédéterminé. Les seuils, sont définis, soit par la norme (par exemple : ISO10816), soit par expérience. La norme ne donne qu'un ordre de grandeur car elle ne tient pas compte du type de machine et des conditions de fonctionnement. Avec une bonne expérience, la meilleure méthode consiste à fixer deux seuils d'alarme: un seuil de dégradation et un seuil d'intervention obligatoire.

Sans expérience, il est possible de se fixer une alarme sur la variation de niveau mesurée entre la première et la dernière mesure (une multiplication par deux du niveau reflète généralement d'un début de dégradation).

Dans ce contexte, nous utilisons le niveau de l'accélération efficace en décibels pour apprécier la sévérité des vibrations. Son expression est:

$$RMS(dB) = 20 \log_{10} (RMS / RMS_{réf}) \quad (5)$$

Avec : $RMS_{réf}$ l'accélération efficace de référence (ou de la première mesure ou de celle d'un roulement neuf).

En conséquence, si une mesure dépasse le niveau d'alerte, cela signifie que l'accélération relevée est deux fois plus importante que celle qui fait référence [8]. Et si une mesure dépasse le niveau de danger, cela signifie que l'accélération relevée est dix fois plus importante que celle qui fait référence (Tableau 1).

Tableau 1 RMS en décibels

Niveau	$RMS / RMS_{réf}$	$RMS(dB)$
Danger	100	40
	10.00	20
	3.16	10
	2.51	8
Alerte	2.00	6
	1.41	3
	1.12	1
Référence	1.00	0

Plusieurs études ont montré que le Kurtosis d'un signal sans défaut est égale à trois et que le facteur de crête du même signal peut varier de 3 à 6. De ce fait, ils sont indicateurs de la présence d'un choc si leurs valeurs dépassent ces seuils [5].

Analyse du contenu fréquentiel

Les défauts de roulements se traduisent par des chocs périodiques. Ils sont caractérisés par des fréquences de répétitions qui sont cependant faciles à mettre en évidence dans le signal d'enveloppe.

La procédure de détection d'enveloppe demande une série de traitements du signal temporel brut. Elle consiste à d'abord fenêtrer puis filtrer le signal en passe bande dans les régions de résonance afin de réduire l'effet de bruit. Puis on calcule l'enveloppe du signal par la transformée de Hilbert, il s'agit en quelque sorte d'un redressement du signal. Ensuite, on calcule le spectre par la transformée de Fourier, on obtient le spectre de l'enveloppe du signal qui est un graphe d'amplitude de vibration en fonction de la fréquence.

Arbre de décision

Les arbres de décision sont utilisés lorsqu'on traite un problème de classification (ou de décision). Le processus de la classification consiste à affecter une classe à des objets à l'aide d'un modèle entraîné sur un autre ensemble d'objets. Il s'agit de classer un ensemble de données en les représentant sous forme d'un arbre. Il s'agit de prédire avec précision les valeurs prises par une décision à partir d'un ensemble d'attributs prédictifs. Ainsi chaque ensemble de données est constitué d'une liste d'attributs prédictifs et des décisions à prédire. Un arbre de décision est composé de feuilles, de nœuds et de branches. Chaque nœud de l'arbre correspond à une propriété de l'objet à classer, appelé attribut. Chaque branche de l'arbre correspond à une valeur possible de l'attribut père et chaque feuille de l'arbre correspond à une classe (ou décision) Figure 1.

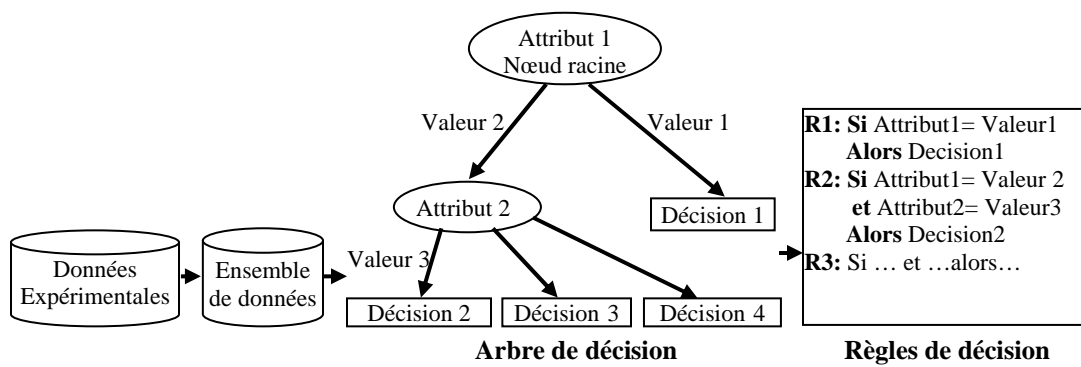


Figure 1 Schéma pour l'extraction des connaissances.

Le principe de la construction de cet arbre est basé sur l'extraction des informations contenues dans les données en utilisant des algorithmes de classification. Parmi ces algorithmes on cite : CART (Breiman, Friedman, Olshen & Stone, 1984), ID3 [9], C4.5 [10]. C4.5 est l'un des algorithmes d'induction d'arbres de décisions les plus connus et les plus utilisés aujourd'hui [11]. Dans le présent travail on utilise l'algorithme C4.5, incorporé dans le logiciel WEKA sous code J48.

2.3.1 Construction de l'arbre de décision

La construction de l'arbre de décision nécessite la définition des attributs et des classes (ou décisions) pour former un ensemble de données. Convenons de noter par :

- $X = \{ X_1, X_2, \dots, X_i, \dots, X_n \}$ l'ensemble des attributs,
- n , le nombre d'attributs,
- $C = \{ C_1, C_2, \dots, C_k \}$ l'ensemble des classes,
- K , le nombre de classes,
- $|C_j|$, $j=1,2,\dots,k$ le nombre des exemples qui appartiennent à la classe C_j ,
- T , l'ensemble des exemples,
- $|T|$, le nombre total des exemples.

La structure de l'arbre de décision dépend infiniment du choix du test X_i . Ce choix repose sur le critère entropie et le critère ratio du gain définis comme suit :

Entropie, c'est la quantité d'information nécessaire pour identifier la classe d'un élément de T , définie comme suit:

$$Info(T) = - \sum_{j=1}^k \frac{|C_j|}{|T|} \log_2 \left(\frac{|C_j|}{|T|} \right) \quad (6)$$

Si l'ensemble T est partitionné sur la base des valeurs d'un attribut prédictif X_i , en sous-ensembles T_i , avec $i=1, \dots, n$. Figure 2, alors on définit l'entropie conditionnelle.

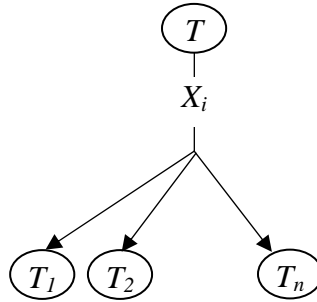


Figure 2 Structure de l'arbre de décision pour l'ensemble T .

Entropie conditionnelle, C 'est la quantité d'informations nécessaires pour identifier la classe d'un élément de T sachant les valeurs de l'attribut X_i , définie comme suit:

$$Info(X_i, T) = \sum_{i=1}^n \left(\frac{|T_i|}{|T|} \cdot Info(T_i) \right) = - \sum_{i=1}^n \left(\frac{|T_i|}{|T|} \cdot \sum_{j=1}^k \frac{|C_j|}{|T_i|} \log_2 \left(\frac{|C_j|}{|T_i|} \right) \right) \quad (7)$$

Gain d'Entropie pour l'attribut X_i . Défini comme suit :

$$Gain(X_i, T) = Info(T) - Info(X_i, T) \quad (8)$$

Cela représente la différence entre l'information nécessaire pour identifier un élément de T et l'information nécessaire pour identifier un élément de T après que la valeur de l'attribut X_i ait été obtenue.

Ce critère permet de choisir l'attribut le plus important. L'attribut choisi est celui qui a le gain le plus grand par rapport aux autres attributs.

Ratio du gain (Gain d'Entropie normalisée), défini comme suit :

$$GR(X_i, T) = \frac{Gain(X_i, T)}{Split\ info(X_i, T)} \quad (9)$$

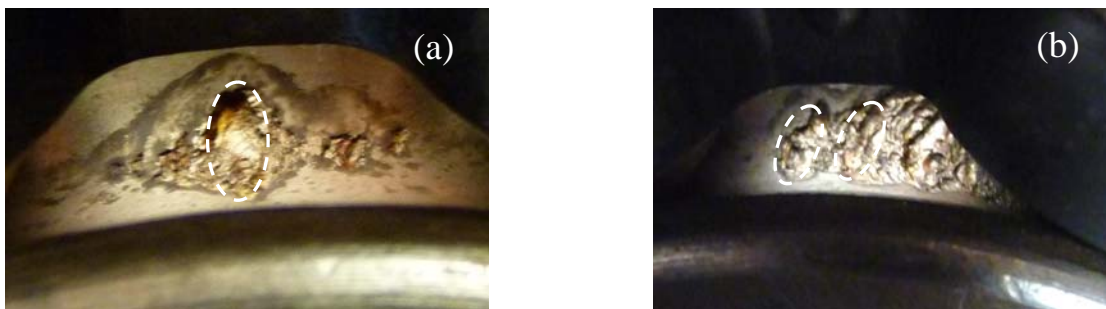
Avec

$$Split\ info(X_i, T) = - \sum_{i=1}^n \frac{|T_i|}{|T|} \log_2 \left(\frac{|T_i|}{|T|} \right) \quad (10)$$

Le ratio du gain est un critère permettant de corriger les points faibles du gain d'entropie en prenant en considération le nombre de valeurs d'un attribut et la proportion de ces valeurs dans les données. Il permet aussi d'éviter d'avoir un arbre mal équilibré (la majorité des nœuds d'un seul côté). L'attribut choisi est celui qui a le ratio du gain le plus grand par rapport aux autres attributs.

3. ÉTUDE EXPERIMENTALE

Plusieurs expériences ont été réalisées sur des roulements à une rangée de billes de type SKF6206 montés sur un banc d'essai de détection de défauts de roulement. On a effectué des mesures large bande [0 – 20 kHz] en accélération efficace. Des défauts de différentes tailles de 8 mm² à 20 mm², ont été réalisés par électroérosion sur la bague intérieure, sur la bague extérieure et sur les deux bagues à la fois. Les défauts de forme elliptique où le grand axe est dans le sens de la largeur de roulement. Le choix de cette forme est justifié par le fait que quand la bille roule sur la piste d'une bague et sous l'effet des charges appliquées, qui donnent lieu à des déformations élastiques dont les contraintes générées sont estimées par la théorie d'Hertz, un contact Hertzien elliptique est formé entre la bille et la bague. D'autre part, les défauts artificiels représentent réellement des défauts rencontrés dans l'industrie. Par exemple, les figures 3a et 3b illustrent des écaillages naturels de forme elliptique sur la bague intérieure d'un roulement à billes de type SKF6414, d'un broyeur dans l'industrie.



Figures 3a, 3b Défauts d'écaillage sur la bague intérieure du roulement de type SKF6414.

Les essais ont été effectués pour les vitesses de rotation (600 tr/mn, 1000 tr/mn et 2000 tr/mn) et pour trois charges radiales horizontales (3000 N, 5000 N et 7000 N) générées par un vérin hydraulique. Les détails des expériences sont résumés dans le tableau 2. L'acquisition des signaux vibratoires est effectuée par un système d'acquisition SIGLAB et par l'utilisation d'accéléromètres piézoélectriques.

La figure 4 représente quelques signaux délivrés par un accéléromètre fixé sur un palier selon une direction radiale horizontale et pour une même vitesse de rotation et pour un même chargement. On peut constater la présence des impulsions périodiques, caractéristiques des défauts de roulement, produites par le passage d'une bille sur une discontinuité.

Tableau 2 Caractéristiques du roulement et les détails des tests

<i>Dimensions du roulement</i>	<i>Fréquence de mesure</i>	<i>Fréquences de rotation tr/min</i>	<i>Charges radiales N</i>
Type SKF 6206	20 kHz	600	3000
Nombre de billes 9		1000	5000
Diamètre primitif 46 mm		2000	7000
Diamètre d'une bille 9.53 mm			

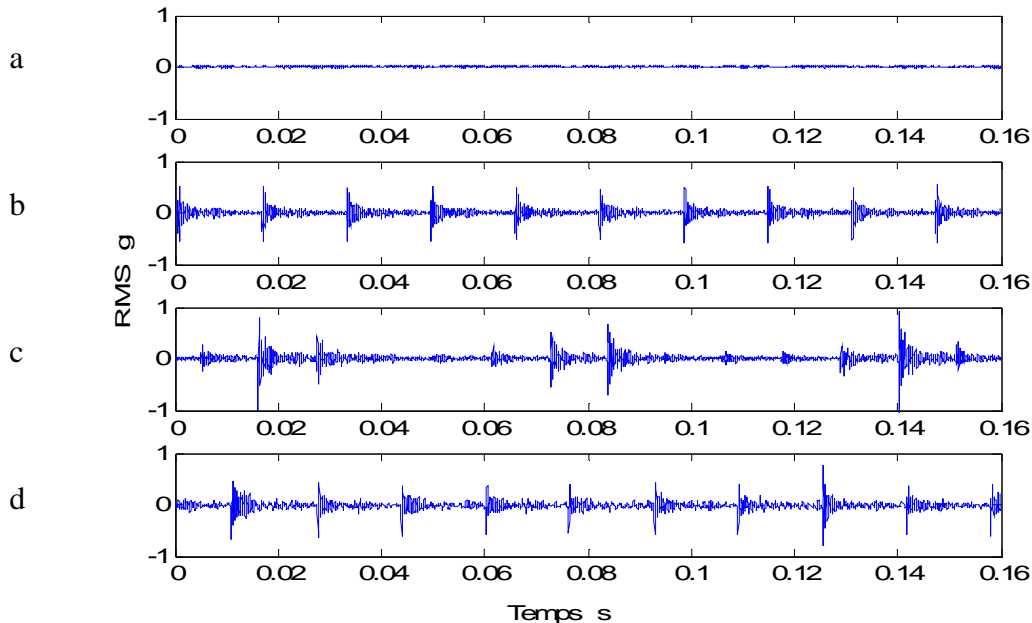


Figure 4 Signaux temporels pour un roulement a). Sans défaut, b) défaut sur la bague externe, c) défaut sur la bague interne, d) deux défauts l'un sur la bague externe et l'autre sur la bague interne.

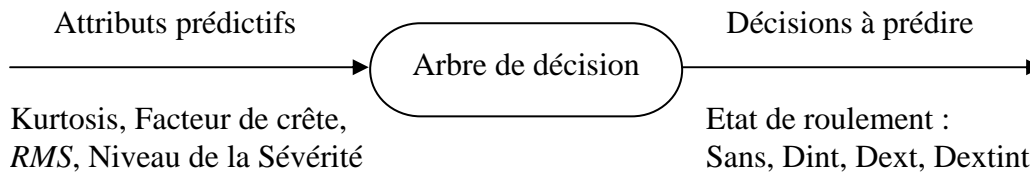
3.1 Application de l'arbre de décision

Pour chaque processus, détection ou identification de défauts, la construction de l'arbre de décision peut se décomposer en trois étapes principales:

1. Acquisition des signaux vibratoires bruts issus des accéléromètres et le traitement adéquat de ces signaux puis l'extraction des données à partir des indicateurs de surveillance et de diagnostic appropriés aux défauts de roulement.
2. Choix des attributs et des classes et mise en forme de l'ensemble de données.
3. Application de l'algorithme de classification J48 sur le l'ensemble de données

3.1.1 Détection des défauts de roulement

La liste des attributs choisis inclut les indicateurs de détection suivants : le Kurtosis, le Facteur de crête, le RMS et le niveau de la sévérité qui est fonction de la valeur *RMS* en décibels. Les classes représentent l'état de roulement : Sans, Dint, Dext, Dextint (figure 5). Les détails des attributs et des classes désignés sont résumés dans les tableaux 3 et 4.

**Figure 5 Attributs et décision****Tableau 3 Attributs de l'arbre de décision**

	<i>attributs</i>	<i>valeurs</i>
1	Kurtosis	numérique
2	Facteur de crête	numérique
3	RMS	numérique
4	Sévérité	Normale : si $RMS (dB) < 6$ Alerte : si $6 < RMS (dB) < 20$ Danger : si $RMS (dB) > 20$

Tableau 4 Classes de l'arbre décision

	<i>Classes</i>	<i>symbole</i>
1	Roulement sans défaut	Sans
2	Roulement avec un défaut sur la bague interne	Dint
3	Roulement avec un défaut sur la bague externe	Dext
4	Roulement avec deux défauts : Un sur la bague externe et l'autre sur la bague interne.	Dextint

L'Application de l'algorithme de classification J48 sur l'ensemble de données a permis de construire les arbres de décision illustrés par les figures 6, 7 et 8 qui correspondent respectivement aux trois cas suivants :

Cas1 : L'ensemble de données issues des signaux vibratoires non débruités,

Cas2 : L'ensemble de données issues des signaux vibratoires débruités par les ondelettes de Daubechies qui semble adaptées aux impulsions caractéristiques des défauts de roulement selon la référence [12]. La même référence présente les expressions pour faire le choix de niveau optimal et le nombre des niveaux pour ce type d'ondelettes.

Cas3 : L'ensemble des données issues des signaux vibratoires non débruités et ceux débruités par les ondelettes de Daubechies

Cas 1 : L'ensemble de données issues des signaux vibratoires non débruités,

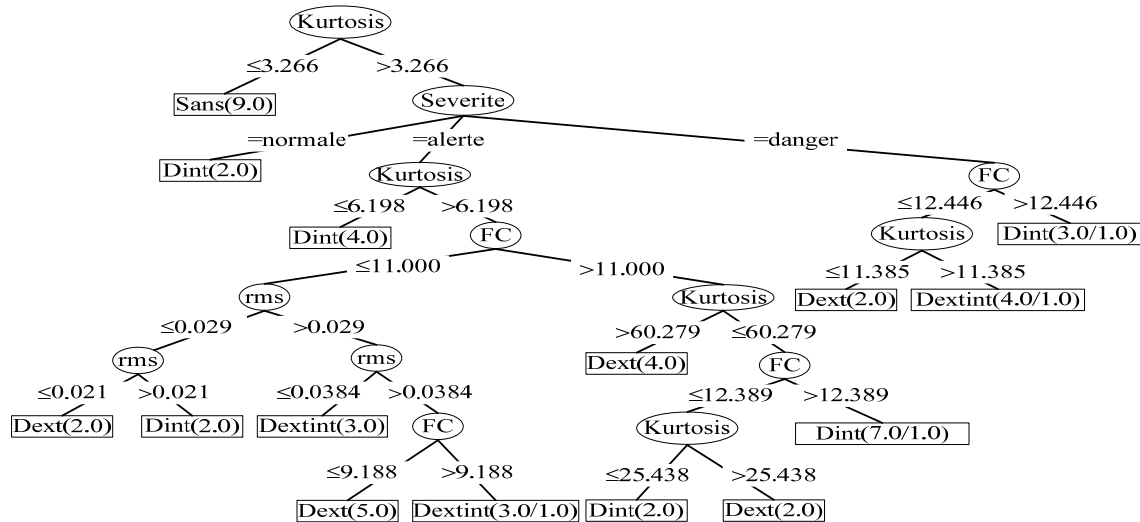


Figure 6 Arbre de décision issu des signaux vibratoires non débruités.

Tableau 5 Matrice de confusion, cas1

Classes	Sans	Dint	Dext	Dextint
Sans	9	0	0	0
Dint	0	18	0	0
Dext	0	1	15	2
Dextint	0	1	0	8

D'après le tableau 5, les éléments de la diagonale de la matrice de confusion indiquent le nombre des exemples bien classés. Le troisième élément de la troisième ligne indique le nombre d'exemples appartenants à la classe Dext et qu'ils ont la même classe selon l'arbre de décision. Par contre, le deuxième et le quatrième élément de la même ligne indiquent le nombre d'exemples appartenant à la classe Dext mais qui sont mal classés par l'arbre de décision. Les exemples du deuxième élément sont classés comme Dint et les exemples du troisième élément sont classés comme Dextint et ainsi de suite.

D'après les figures 6, 7 et 8 qui correspondent aux cas 1, 2 et 3 respectivement, on constate que tous les attributs choisis de l'ensemble de données étudié apparaissent dans l'arbre de décision. Cela confirme l'importance de ces attributs pour prendre une décision dans le processus de détection des anomalies. En outre, le Kurtosis défini par l'arbre, du nœud racine, semble plus proche de celui défini par la référence [5] et signifie qu'il est plus fiable que le facteur de crête et la valeur RMS.

Chaque chemin de l'arbre qui va de la racine à une feuille peut être traduit en règle sous forme : **Si conditions Alors Décision**. À partir de la figure 6, on peut extraire les règles de décision suivantes :

Règle 1 : **Si** kurtosis ≤ 3.266 **Alors** Roulement sans défaut,

Règle 2 : **Si** kurtosis > 3.266 **et** sévérité = normale, **Alors** Roulement avec un défaut sur la bague interne.

De la même manière, on peut traduire l'ensemble des chemins sous forme de règles de décision. Les règles obtenues peuvent être simplifiées par élimination de certaines de ces conditions.

Cas 2 : L'ensemble des données issues des signaux vibratoires débruités par les ondelettes de Daubechies,

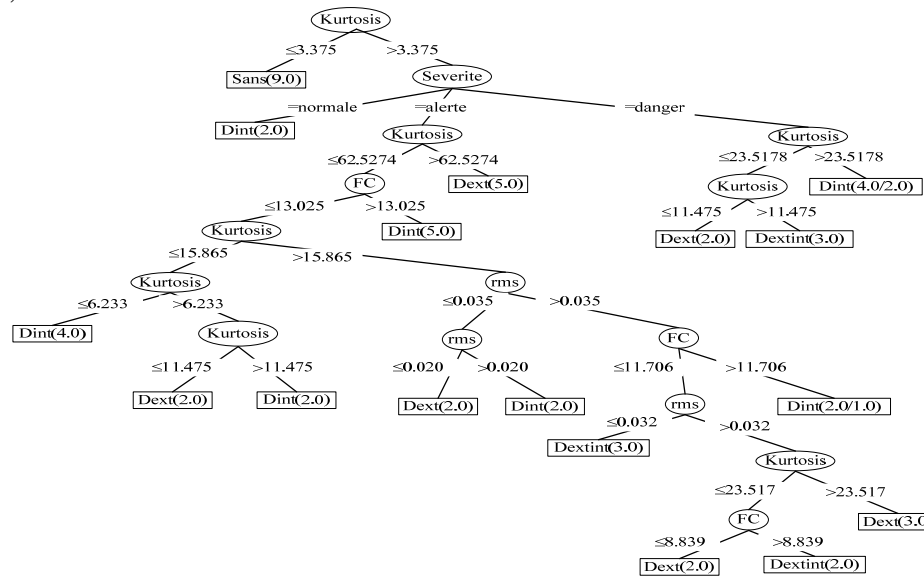


Figure 7 Arbre de décision issu des signaux débruités par les ondelettes de Daubechies.

Tableau 6 Matrice de confusion, cas 2

Classes	Sans	Dint	Dext	Dextint
Sans	9	0	0	0
Dint	0	18	0	0
Dext	0	2	16	0
Dextint	0	1	0	8

Cas 3 : L'ensemble des données issues des signaux vibratoires non débruités et ceux débruités par les ondelettes de Daubechies.

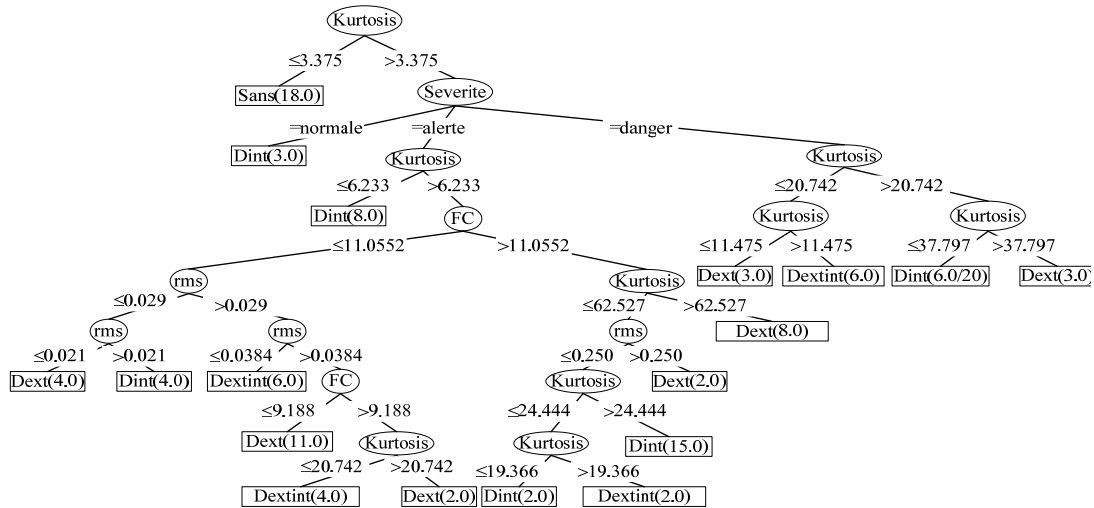


Figure 8 Arbre de décision issu des signaux vibratoires non débruités et ceux débruités par les ondelettes de Daubechies.

Tableau 7 Matrice de confusion, cas 3

Classes	Sans	Dint	Dext	Dextint
Sans	18	0	0	0
Dint	0	36	0	0
Dext	0	2	34	0
Dextint	0	0	0	18

Le tableau 8, présente quelques paramètres de performances pour l'évaluation d'un arbre de décision. Par exemple, il montre que l'amélioration du taux de classification dépend du débruitage des signaux vibratoires et du nombre d'exemples. Sachant que le taux de classification correspond au rapport entre le nombre des exemples bien classés par l'algorithme et le nombre total des exemples qui forment l'ensemble de données.

Tableau 8 Paramètres de performances pour la détection des défauts

	Taux de classification	Kappa statistique	Erreur absolue moyenne	Nombre de feuilles	Nombre d'exemples
Cas 1	92.59 %	0.90	0.0545	15	54
Cas 2	94.44 %	0.92	0.0278	17	54
Cas 3	98.15 %	0.97	0.0123	18	108

3.1.2 Identification des défauts de roulement

La liste des attributs choisis inclut les indicateurs de détection suivants : le Kurtosis, le FC et le RMS. Aussi, elle inclut l'inégalité entre l'amplitude de la première et de la deuxième fréquence caractéristique du défaut de roulement. On peut vérifier cette inégalité à partir le graphe du spectre d'enveloppe. Les classes représentent l'état de roulement suivant : Sans,

Dint_n, Dint_a, Dint_d, Dext, Dext_n, Dext_a, Dext_d, Dextint_n, Dextint_a, Dextint_d. Les détails des attributs et des classes désignés sont résumés dans les tableaux 9 et 10.

Tableau 9 Attributs de l'arbre de décision

	<i>attributs</i>	<i>Valeurs</i>
1	Kurtosis	Numérique
2	Facteur de crête	Numérique
3	<i>RMS</i>	Numérique
4	$A_{Fbe} > A_{2Fbe}$	Oui : A_{Fbe} et A_{2Fbe} existent et $A_{Fbe} > A_{2Fbe}$ Non : A_{Fbe} et A_{2Fbe} existent et $A_{Fbe} < A_{2Fbe}$ Pas : A_{Fbe} et A_{2Fbe} n'existent pas.
5	$A_{Fbi} > A_{2Fbi}$	Oui : A_{Fbi} et A_{2Fbi} existent et $A_{Fbi} > A_{2Fbi}$ Non : A_{Fbi} et A_{2Fbi} existent et $A_{Fbi} < A_{2Fbi}$ Pas : A_{Fbi} et A_{2Fbi} n'existent pas.

Avec A_{Fbe} l'amplitude à la fréquence d'un défaut localisé sur la bague extérieure
 A_{Fbi} l'amplitude à la fréquence d'un défaut localisé sur la bague intérieure

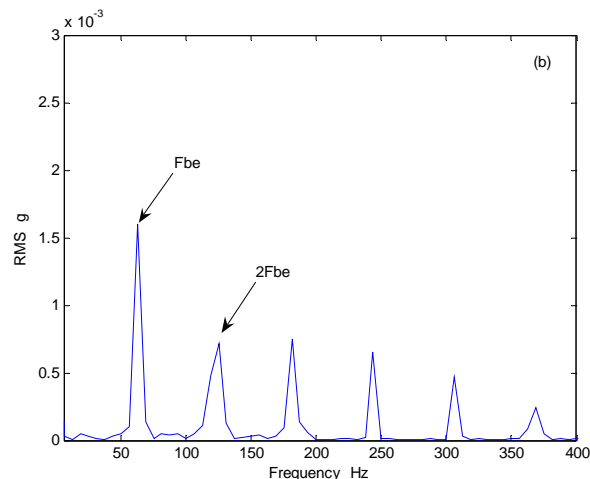


Figure 9 Spectre d'enveloppe d'un signal d'un défaut sur la bague externe, Pour une charge de 5000N et une vitesse de 1000tr/min.

L'Application de l'algorithme de classification J48 sur l'ensemble de données a permis de construire l'arbre de décision illustré par la figure 10. D'après cette figure on constate que tous les attributs choisis de l'ensemble de données étudié apparaissent dans l'arbre de décision sauf le facteur de crête. Sachant que cet indicateur est destiné à la détection des défauts et non à la localisation des défauts. La présence des indicateurs choisis confirme l'importance de ces attributs pour prendre une décision dans le processus d'identification des anomalies.

Tableau 10 Classes de l'arbre décision

	<i>Classe (ou décision)</i>	<i>symbole</i>
1	Roulement sans défaut	sans
2	Roulement avec un défaut sur la bague interne et état normal	Dint_n
3	Roulement avec un défaut sur la bague interne et état alerte	Dint_a
4	Roulement avec un défaut sur la bague interne et état danger	Dint_d
5	Roulement avec un défaut sur la bague externe et état normal	Dext_n
6	Roulement avec un défaut sur la bague externe et état alerte	Dext_a
7	Roulement avec un défaut sur la bague externe et état danger	Dext_d
8	Roulement avec deux défauts : Un sur la bague externe et l'autre sur la bague interne et état normal.	Dextint_n
9	Roulement avec deux défauts : Un sur la bague externe et l'autre sur la bague interne et état alerte.	Dextint_a
10	Roulement avec deux défauts : Un sur la bague externe et l'autre sur la bague interne et état danger.	Dextint_d

En outre, on constate qu'il est facile d'extraire des règles de décision. A titre d'exemple, suivant le chemin à gauche de l'arbre, on en déduit la règle suivante :

Règle 1 : **Si** A_{Fbi} et A_{2Fbi} existent **et** $A_{Fbi} > A_{2Fbi}$ **et** $RMS \leq 0.02$ g

Alors Roulement avec un défaut sur la bague interne et état de sévérité normale.

Règle 2 : **Si** A_{Fbi} et A_{2Fbi} existent **et** $A_{Fbi} > A_{2Fbi}$ **et** $0.02 < RMS \leq 0.148$ g

Alors Roulement avec un défaut sur la bague interne et état de sévérité alerte.

De la même manière, on peut traduire l'ensemble des chemins sous forme des règles de décision.

D'après les expressions des règles obtenues, on constate que ces règles semblent plus proches que celles obtenues par l'expérience [13].

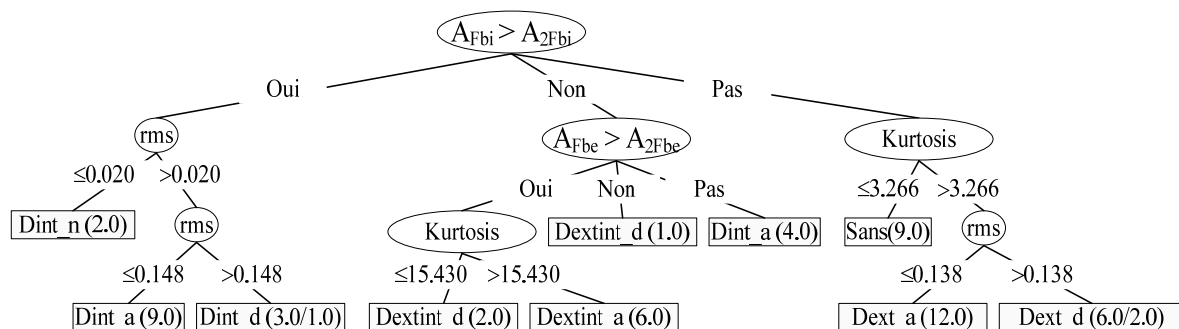
**Figure 10 Arbre de décision pour l'identification des défauts de roulement.**

Tableau 11 Matrice de confusion pour l'identification des défauts

<i>Classes</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>J</i>
A=Sans	9	0	0	0	0	0	0	0	0	0
B=Dext_n	0	0	0	0	0	0	0	0	0	0
C=Dext_a	0	0	12	2	0	0	0	0	0	0
D=Dext_d	0	0	0	4	0	0	0	0	0	0
E=Dint_n	0	0	0	0	2	0	0	0	0	0
F=Dint_a	0	0	0	0	0	13	1	0	0	0
G=Dint_d	0	0	0	0	0	0	2	0	0	0
H=Dextint_n	0	0	0	0	0	0	0	0	0	0
I=Dextint_a	0	0	0	0	0	0	0	0	6	0
J=Dextint_d	0	0	0	0	0	0	0	0	0	3

Le tableau 12, présente quelques paramètres de performances pour l'évaluation de l'arbre de décision du processus d'identification des défauts de roulement. Il montre que ce processus est parfaitement modélisé.

Tableau 12 Paramètres de performances pour l'identification des défauts

	<i>Taux de classification</i>	<i>Kappa statistique</i>	<i>Erreur absolue moyenne</i>	<i>Nombre de feuilles</i>	<i>Nombre d'exemples</i>
Identification	94.44 %	0.93	0.0145	10	54

4. CONCLUSION

L'objectif de cette étude réside dans l'utilisation de l'arbre de décision dans la prise de décision dans un processus de détection et d'identification de défaut de roulement. Elle vise à extraire des connaissances à partir de données vibratoires issues d'historiques d'une machine. L'algorithme C4.5 décrit précédemment a permis la construction des arbres de décision et la modélisation graphique permettant de caractériser l'état d'un roulement. Cette approche est validée à partir des valeurs des paramètres de performances qui indiquent que les arbres obtenus donnent des informations pertinentes quant à l'état d'un roulement. La lisibilité et la simplicité de ces arbres de décision offrent une aide dans l'extraction des règles de décision. Cette approche a un avantage certain sur les systèmes experts qui n'arrivent pas toujours à dégager les règles expertes. Pour que les arbres et les règles soient significatifs, il faut utiliser un ensemble de données plus important en considérant d'autres types de défauts de roulement tels que les défauts de cage ou de billes. Il faut également porter une correction sur l'imprécision des seuils limites en terme d'amplitude et sur l'imprécision sur les fréquences (mesurées et théoriques) en intégrant des facteurs de certitude dans les règles expertes.

5. REFERENCES BIBLIOGRAPHIQUES

1. Morel J., 1992, Vibrations des machines et diagnostic de leur état mécanique, Eyrolles.
2. Harris C. M., Piersol A. G., 2002, shock and vibration Handbook, McGraw-Hill.
3. Dron J. P., Bolaers F., Rasolofondraibe L., 2004, Improvement of the sensibility of scalars indicators thanks to de-noising method by spectacle subtraction. Application to the detection ball bearing defects, Journal of sound and vibration, vol. 270, 61-73.
4. Ericsson E., Grip N., Johansson E., Persson L. E., Sjoberg R., Stronberg J. O., 2005, Towards automatic detection of local bearing defects in rotating machines, Mechanical systems and signal processing, 19 :509-535.
5. Boulenger A., Pachaud C., 2003, Diagnostic vibratoire en maintenance préventive, Dunod.
6. Yang B. S., Park C. H., Kim H. J., 2000, An efficient method of vibration diagnosis for rotating machinery using a decision tree, International journal of rotating machinery, 6(1): 19-27.
7. Sugumaran V., Ramachandran K. I., 2007, Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing, Mechanical systems and signal processing, 21 :22372247.
8. Augeix D., 2002, Analyse vibratoire des machines tournantes, Techniques de l'ingénieur, bm5145.
9. Quinlan J. R., 1986, Induction of decision trees. Machine Learning, 1, 81-106.
10. Quinlan J. R., 1993, C4.5: Programs for machine learning, Morgan Kaufmann Publisher.
11. Lavoie B., 2006, Arbres de décision: Synthèse de lectures, Séminaire sur l'apprentissage automatique, Université du Québec à Montréal.
12. Djebala A., Ouelaa N., Hamzaoui N., 2007, Optimisation de l'analyse multi-résolution en ondelettes des signaux de choc. Application aux signaux engendrés par des roulements défectueux, Mécanique et industries, 8:379-389.
13. Ebersbach, S., Zhongxiao, P., (2008). Expert system development for vibration analysis in machine condition monitoring. Expert systems with applications, 34,291–299.