Experiences on innovative trends in the field of rotating machinery condition monitoring and diagnostics

Expériences sur des tendances innovatrices dans le domaine de la surveillance d'état et du diagnostic des machines tournantes

A. Lucifredi, P. Silvestri
University of Genova
Laboratorio di Meccanica Generale e Meccanica delle Vibrazioni (MGMV)
Via Opera Pia, 15 A
16145 Genova
Italy
E-mail : mgmv@unige.it

Résumé

Le laboratoire MGMV consacre traditionnellement une grande activité à la surveillance et au diagnostic des machines. Dans l'effort continu pour améliorer les outils disponibles, une recherche soutenue a été menée pendant ces dernières années, principalement concentrée sur le traitement des signaux et sur la surveillance, par exemple en développant des outils basés sur la théorie de chaos pour une détection précoce des opérations erronées, et en général en améliorant considérablement le logiciel disponible ; l'évolution de la révision de la boîte à outils informatique spécifique a atteint un stade avancé d'accomplissement. Pour la surveillance, nous développons maintenant, entre autres, des techniques de détection de nouveauté, et des techniques de séparation aveugle des sources. Au sujet du premier point, on peut préciser que beaucoup de problèmes de surveillance d'état sont mieux abordés avec une approche de détection de nouveauté. La tâche de surveillance est de décider quand une machine ne fonctionne pas correctement. En général, il y aura une très grande quantité de données relevées lorsque la machine fonctionne correctement, et très peu ou aucune donnée ne seront disponibles lorsqu'elle fonctionne dans un état défectueux. En outre, souvent, bien des types différents de défaut peuvent se produire, ainsi il est impossible de recueillir assez de données pour chaque défaut et combinaison de défauts qui pourraient être présents. Une approche alternative est de perfectionner un modèle afin qu'il puisse identifier quand la machine fonctionne correctement, et traiter n'importe quel écart significatif de ce modèle comme un défaut potentiel qui exige davantage d'investigations.

Actuellement, notre nouvelle tendance de développement est davantage consacrée au diagnostic de machines, en essayant de développer, essayer, et introduire, des modules de diagnostic automatiques dans notre code informatique. Un grand effort de recherche que nous avons commencée et que nous projetons de développer dans un proche avenir, est lié à l'introduction dans les machines tournantes et de puissance, de techniques de diagnostic d'installations industrielles prises (avec de grandes modifications) dans d'autres champs, non seulement de la technologie, mais également de la médecine, de la biologie, des sciences économiques… Pour le diagnostic, nous testons l'utilisation des techniques d'identification de la voix, d'identification d'image et de forme, d'extraction de données, etc. Les systèmes experts codifient sous forme de règles la connaissance de l'expert, rendant possible, aux opérateurs de niveau moindre, l'interprétation des données quand l'expert n'est pas disponible ou plus disponible, ce qui permet des économies significatives. Si certaines règles ne sont pas connues, il est possible d'essayer de les extraire automatiquement, à partir des données acquises par des techniques d'extraction de données, par exemple en associant la présence de certains éléments (les antécédents) à la présence d'autres éléments (les conséquents). Comme alternative, nous pouvons employer les réseaux de neurones, capables de s'auto-instruire sur la base d'un nombre suffisant d'exemples. Puisque les réseaux de neurones émulent le comportement du cerveau humain, ils ont l'avantage d'une rigidité plus faible en ce qui concerne les méthodes basées sur des règles. En fait, ils sont capables d'identifier des cas similaires mais pas exactement égaux, d'extraire les caractéristiques essentielles à partir des données affectées par du bruit, et de s'adapter aux changements du processus au cours du temps ; en outre, leur ajustement est généralement plus rapide.
Le présent article décrit quelques aspects de l'expérimentation en cours. Les premiers consistent spécifiquement en essais et comparaisons d'applications, travaillées avec des modifications à

des cas spécifiques, d'algorithmes de classement. Les données utilisées ont été produites par un modèle expérimental de rotor, ce qui a rendu possible de créer différents états anormaux typiques de fonctionnement des machines tournantes (frottement, rotor fissuré, défaut de palier) : la grande quantité de données expérimentales ainsi obtenue a été traitée dans une phase ultérieure par le logiciel de diagnostic considéré.
Les méthodes de classement partitionnelles et hiérarchiques ont été prises en compte. Les premières sont basées sur des partitions constituées selon des critères particuliers, de façon à répartir les objets dans des classes sans empiétement, de telle manière que chaque objet appartienne à seulement une classe. Parmi celles-ci, on a appliqué et comparé des techniques où chaque classe se compose d'un ensemble de points choisis pour être plus proche du centre de leur classe que du centre des autres. Ainsi nous avons examiné la méthode de la k-moyenne (chaque faisceau est représenté par son vecteur moyen, et chaque point de données est affecté à la classe avec le vecteur le plus proche, en minimisant le carré de la distance), et la méthode de la k-medoid (qui est moins sensible aux points extérieurs, minimisant la distance moyenne ou dissimilitude de l'objet représentatif, appelé le medoid, à tous les autres objets dans la même classe). La figure 1 montre, comme exemple, une application possible de la méthode de la k-medoid au diagnostic de machines tournantes, les résultats du classement de 50 échantillons de données pour chaque défaut à classifier (machine saine, frottement, rotor fissuré, défaut de palier). Les paramètres considérés sont la dimension de l'information du signal de déplacement d'une section de l'arbre, l'amplitude du premier harmonique 1Ω, l'énergie du spectre dans la gamme (0-50 % de la fréquence du premier harmonique), l'énergie du spectre dans la gamme (1Ω, 2Ω).

Chaque état classifié est représenté par une valeur numérique indiquant le degré de concordance à la classe. De cette manière, il est possible de comprendre quels éléments ont un degré élevé d'appartenance à la classe (valeurs proches de 1), et quelles autres sont situées dans une zone de frontière, à proximité à d'une autre classe (valeurs proches de zéro). Les résultats sont assez satisfaisants, du fait que la quantité de données correctement classifiées est tout à fait significative.
Le second type de méthodes de classement correspond aux méthodes hiérarchiques, qui peuvent être agglomératives (état initial avec tous les objets séparés ; à chaque étape, deux des classes sont fusionnées) ou séparatives (état initial avec tous les objets ensemble, et à chaque étape suivante une classe est fractionnée). La rigidité des méthodes hiérarchiques est à la fois la clef de leur succès (temps court de calcul), et leur inconvénient principal (l'incapacité de corriger des décisions incorrectes). L'article rapporte un exemple de méthodologie agglomérative ; la procédure est basée sur le concept de proximité de classe. Un autre exemple se rapporte à une méthode séparative.
D'autres aspects de l'expérimentation par les auteurs, décrits dans l'article, se rapportent à la reconnaissance de formes, aux systèmes experts, à l'analyse en composantes principales (PCA).

Abstract

The paper considers methodologies for rotating machinery diagnostics based on the application of clustering techniques, i.e. on the automatic classification of data into groups. Different algorithms of clustering have been used on data taken from an experimental model of rotor in different anomalous operation conditions (rubbing, cracked shaft, bearing defect) and in healthy conditions. A vibrational diagnostics has been performed based on data of traditional techniques (e.g. the values of particular harmonics of the signal, of the energy of the spectrum in different frequency bands) and on some chaos quantifiers [1].
The paper reports also an example of rotating machinery automatic diagnostics through expert systems.

Introduction

The MGMV Lab traditionally devotes a large activity to monitoring and diagnostics of machinery. In the continuous effort of upgrading the available tools, a strong research has been performed during the past years, mainly focused on signal processing and monitoring, e.g. developing tools based on chaos theory for an early detection of misoperations and in general greatly improving the available software ; the evolution of the specific computer toolbox revision is in an advanced stage of completion [2, 3, 4].
Presently our new trend of development is more devoted to machinery diagnostics, trying to develop, to test and to introduce automatic diagnostic modules into our computer code. A large effort of research we started and we plan to develop in the near future is related to the introduction in the rotating machinery and power or industrial plants diagnostics of techniques taken (with large modifications) from other fields, not only of engineering but even also of medicine, biology, economics…
For diagnostics we are testing the use of techniques for voice recognition, image and pattern recognition, data mining, etc. Expert systems codify in the form of rules the knowledge of the expert making possible to lower level operators to interpret the data when the expert is not available or is no more available, enabling relevant economic savings. If rules are not known, it's possible to try to extract them automatically from the acquired data through data mining techniques, e.g. associating the presence of some elements (antecedents) to the presence of others (consequents). Alternatively we can use neural networks, able to self-instruct themselves on the basis of a sufficient number of examples. Since neural networks emulate the behavior of the human brain, they have the advantage of a lower rigidity with respect to the methods based on rules. In fact they are able to recognize cases similar but not exactly equal, to extract the essential characteristics from data affected by noise and to adapt to changes of the process during the time ; in addition their tuning is generally faster [5].
The present paper describes some aspects of the ongoing experimentation. The first specifically consist in tests and comparisons of applications, tailored with modifications to specific cases, of clustering algorithms. The data used have been generated by an experimental rotor model, which made possible to create different anomalous operation conditions typical of rotating machines (rubbing, cracked rotor, bearing defect) ; the great amount of experimental data so obtained has been processed in a later time by the diagnostic software considered.
Both partitional and hierarchical clustering methods have been taken into consideration. The former are based on partitions performed according to particular criteria to distribute the objects in non-overlapped clusters and in such a way that each object belongs to only one cluster. Among them have been applied and compared techniques where each cluster consists of a set of points such to be nearest to the center of its cluster than to the center of the others.
The second type of clustering methods corresponds to the hierarchical methods, which can be agglomerative (start with all objects apart ; at each step two clusters

are merged) or divisive (start with all objects together and in each following step a cluster is split up). The rigidity of hierarchical methods is both the key to their success (short computation time) and their main drawback (the inability to correct erroneous decisions). The paper reports an example of agglomerative methodology ; the procedure is based on the concept of cluster proximity. Another example refers to a divisive method.
Other aspects of Authors experimentation reported in the paper refer to expert systems.

Diagnostic applications through clustering techniques

Clustering modules have been used, for diagnostics purposes, for an automatic classification into groups.
Data have been taken from an experimental model of rotor Bently Nevada RT4 Rotor Kit, in different conditions of anomalous operation (rubbing, cracked shaft, bearing defect) and in healthy conditions. This made possible to test in different real cases different algorithms, to evaluate which of them are best fitted for each diagnostic goal. The data used for clustering consist of sets of variables or feature evaluated for each single operation condition of the machine ; initially the variable adopted have been the information dimension of a displacement signal of a section of the shaft, the magnitude of the first harmonic, the energy of the spectrum in the interval 0÷50 % of the first harmonic 1X, the energy of the spectrum in the interval 1X-2X. The choice has been motivated by the fact to have only numeric variables, more apt to be treated by the clustering software ; in addition the chosen variables show a good sensitivity to the changes from an operation condition to another, as requested to make each set of variables as more characteristic as possible of the particular condition of the rotor. Nevertheless in the present study also other sets of data have been considered, to evaluate the influence of the choice of the features on the result of clustering. Before performing the clustering, data have been properly standardized, in such away to avoid the influence of the units adopted : the original values have been converted into variables without units. The adopted procedure is the following. First the mean value m and the standard deviation s of the variable f are computed :

$$z_{if} = \frac{x_{if} - m_f}{s_f}$$

A first classification of the data coming from the Rotor Kit has been performed through the concept of similarity and dissimilarity of data. Similarity is expressed in terms of a distance function. The algorithms used adopt the Minkowski distance, given by the following expression :

$$d(i,j) = \sqrt[q]{\left|x_{i1} - x_{j1}\right|^q + \left|x_{i2} - x_{j2}\right|^q + \ldots + \left|x_{ip} - x_{jp}\right|^q}$$

where $i = (x_{i1}, x_{i2}\ldots, x_{ip})$ and $j = (x_{j1}, x_{j2}\ldots, x_{jp})$ are two p-dimensional object in the same variables x, and q is a positive integer. If q = 1, d is the Manhattan distance (also called city block distance), while if q = 2 d is the Euclidean distance [6].
In the present experimentations both the distances have provided comparable results.

Program "DAISY"

The comparison of the dissimilarities has been performed on the data coming from the rotor model through the module "DAISY".

Fig. 1 reports, as an example, the values of the dissimilarity of two operation conditions (the 6 corresponding to an healthy rotor and the 10 to a rotor rubbing) evaluated on a series of 16 data : the first 8 refer to the healthy rotor, the remaining ones to the rubbing case. The features adopted are those previously mentioned. Low values of dissimilarity have been obtained from the comparison of a set of variables with those in the same operation conditions, while higher values were obtained when data of different conditions were compared. The estimation of the dissimilarity has been reported both in the standardized data case and in the non-standardized case and for Euclidean and for Manhattan distance : the results obtained are coherent and comparable.

The module "DAISY" recognizes as similar, operation conditions belonging to the same category of defect ; vice-versa the software classified as dissimilar an operation condition corresponding to another category of failure. The dissimilarity estimation is a process of classification and therefore may be considered as an elementary methodology of diagnostics.

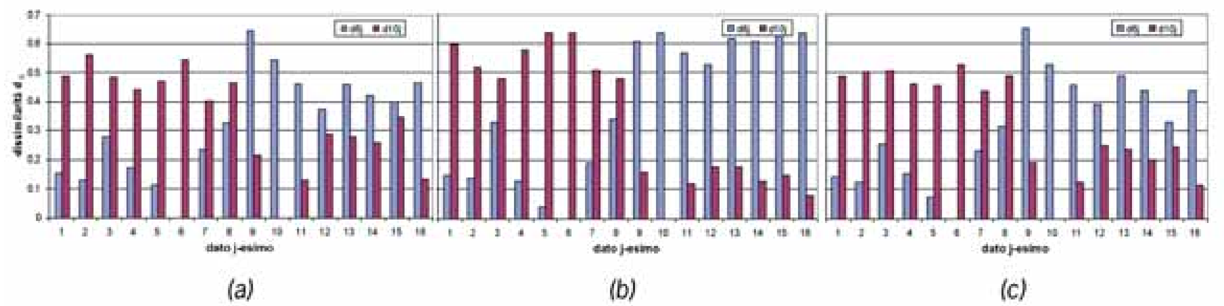


Fig. 1 : Dissimilarity of elements 6 (healthy rotor) and 10 (rubbing) evaluated with respect to 8 samples of healthy rotor (1-8) and to 8 samples of rotor rubbing : (a) standardized data and Euclidean distance ; (b) non-standardized data and Euclidean distance ; (c) standardized data and Manhattan distance

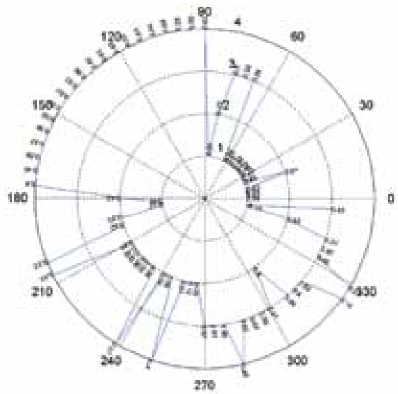
Partitioning algorithms

Such algorithms create various partitions of the data set on the basis of some criterion. They are iterative methods and therefore they may be expensive in terms of computation time.

Program "PAM" (Partitioning Around Medoids)

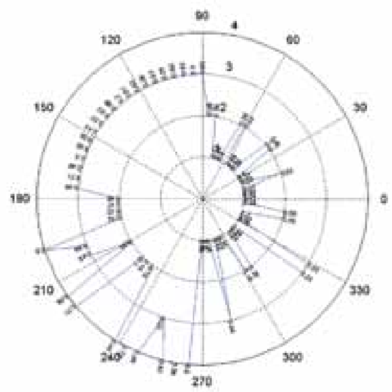
We tested in a previous work the k-means method [5] (each cluster is represented by its mean vector and each data point is assigned to the cluster having the closest vector, by minimizing the square of the distance) and now the k-medoids method (which is less sensitive to outliers and more robust in the case of cluster of non-globular shape and of different density). In the case of k-medoids each cluster is represented by one of the objects in the cluster, named medoid. The corresponding cluster are then obtained by assigning the objects to the one having the nearest significant element. To identify this element the concept of distance is adopted. "PAM" initially defines a set of medoid and iteratively attempts to replace one of the medoids with another non-medoids element. The modification is maintained if the replacement improves the total distance of the clustering.

Fig. 2 shows, as an example of a possible application of the k-medoids method to rotating machines diagnostics, the results of clustering of 20 data samples for each defect to be classified (healthy machine, rubbing, cracked rotor, bearing defect). The features considered are the information dimension of the displacement signal of a shaft section, the amplitude of the first harmonic, the energy of the spectrum in the range (0,50 % of



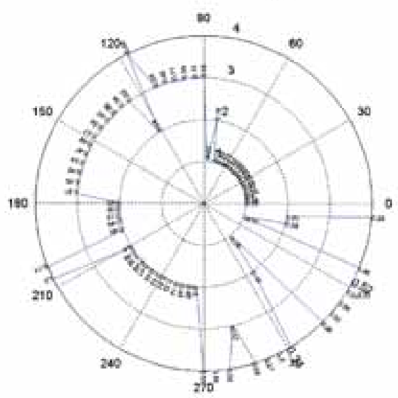
	healthy (1)	rubbing (4)	cracked (2)	bearing def. (3)
healthy	15	0	2	1
rubbing	0	20	4	3
cracked	2	0	14	2
bearing defect	3	0	0	14
overall classification accuracy: 78.7%				

(a)



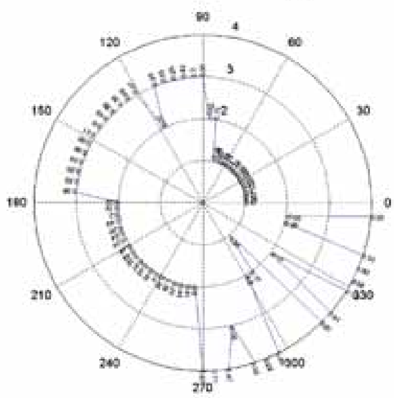
	healthy (1)	rubbing (3)	cracked (4)	bearing def. (2)
healthy	13	0	0	13
rubbing	0	20	3	0
cracked	7	0	9	4
bearing defect	0	0	8	3
overall classification accuracy: 68.7 %				

(b)



	healthy (1)	rubbing (3)	cracked (2)	bearing def. (4)
healthy	19	0	0	2
rubbing	0	18	0	1
cracked	1	1	18	3
bearing defect	0	1	2	14
overall classification accuracy: 86.5%				

(c)



	healthy (1)	rubbing (3)	cracked (2)	bearing def. (4)
healthy	18	0	0	1
rubbing	0	19	0	1
cracked	2	1	20	5
bearing defect	0	0	0	13
overall classification accuracy: 87.5 %				

(d)

Fig. 2 : Application of the k-medoids method to the diagnostics of rotating machines no-standardized data (a) Euclidean, (b) Manhattan distance - standardized data, (c) Euclidean, (d) Manhattan distance

the frequency of the first harmonic), the energy of the spectrum in the range (1X, 2X).

In the present paper a polar notation is adopted to visualize the results. The data used in the present study refer to measurements in 4 different operation conditions of the rotor and in an equal number for each category of failure : adopting a constant angular spacing to display all the measurements, the set of data referring to each operation condition corresponds to a circular sector (healthy machine 1st sector, rubbing 2nd sector, cracked 3rd sector, bearing defect 4th sector). In the interpretation of the display of the results the width of the peaks may be misleading : it must be remembered that the polar representation has a constant pitch between the radii from the circle center, while the angle between the segments outgoing from the points of the circumferences depends on the length of the corresponding segments. Each classified condition is represented by a numeric value indicating the degree of matching to the cluster. In such a way it's possible to understand which elements have a high degree of membership to the cluster (values near to 1) and which other are located in a border zone, in proximity to another cluster (values near to zero). The results are enough satisfactory, since the amount of correctly classified data is quite relevant.

Fig. 2 reports the results in absence (fig.2a, 2b) and in presence (fig.2c, 2d) of standardization and in addition, compares the algorithms outputs for the case of Euclidean and Manhattan distance.

In order to quantify the cluster accuracy it is possible to evaluate the confusion matrix, whose rows represent the true classes and the columns the predicted classes ; Cij is the number of samples of class i, which were classified in class j. The ideal matrix is diagonal.

From the analysis of the confusion matrix it's apparent that the data classification is much better if standardized measurements are used (the increase in accuracy is about 10 %) while in the other case more dispersions in more clusters appear in data coming from the same defect. In absence of standardization a worse recognition of the defect is apparent particularly for cracked shaft and bearing damage : the two conditions are often confused or classified as healthy condition. This could be explained by the fact that the absence of standardization makes more critical the not great difference in the values of the features ; in addition the dispersion of the values of the features causes a higher weight of some variable in the computation of the distance and therefore on the process of clustering. The classification is worse in the case of Manhattan distance, where only the 68 % of classifications are correct (fig.2b).

In the case of standardized data the efficiency of classification is 86.5 % in the case of Euclidean distance and 87.5 % in the case of Manhattan distance. In all cases the 4 objects representing the 4 clusters (medoids) refer to different conditions of the rotor and the majority of the elements of each cluster are in the same operation condition of the medoids. The accuracy of the classification remains higher then 65 % also when evaluated within a single misoperation (in absence of standardization the case of bearing defects the value was 15 %).

If we compare, for the same data set, the accuracy with the accuracy with the k-means algorithm, for the case of standardized measurements and Euclidean distance, the results are equivalent (86.5 % in the case of "PAM" and 86 % in the case of k-means), while a k-nearest neighbor clustering on the same data has provided somewhat better results with an efficiency of 92 % [5].

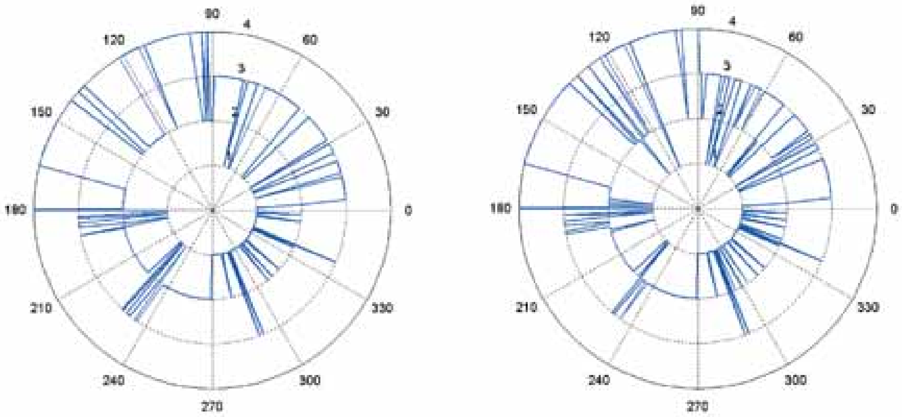
Program "CLARA" (Clustering Large Applications)

"PAM" is not suited to perform clustering of an elevated number of objects (<100 as reported in [6]) since the algorithm could be slow and not convergent to an acceptable solution. For application on elevated numbers another module ("CLARA") has been developed, performing the clustering in two step : first a sample of data is extracted randomly from the set of objects and is clusterized in a certain number of subsets using a procedure analogous to "PAM", providing the representative objects, i.e. the medoids. Then each object not belonging to the initial sample of data is assigned to the nearest of the k representative objects ; in such a way the clustering of all the data is performed. A measure of the quality of the clustering is provided by the average distance of each object of the dataset from the representing objects. The overall procedure is iterated and the accepted solution is the one minimizing the global value of the average distance [6]. The advantage is the aptitude of the algorithm to deal with an elevated number of samples without excessively loading the computer resources : the limit is that the process is highly dependent from the choice and the dimension of the initial sample of data. Hereafter we report the results of the classification referring again to the previous 4 different operation conditions but to a data set of 400 elements.

Fig. 3 compares the classification obtained by "PAM" (fig.3a) and by "CLARA" (fig.3b). By analyzing the data reported in the confusion matrix, the two programs provide comparable results ; the classification efficiency is 76.7 % for the case of "PAM" (modified to allow a dimension 400 of the variables instead of the original 100) and 77.5 % for the case of "CLARA".

Program "FUNNY" (Fuzzy analysis)

The software "FUNNY" is a generalization of the previous algorithms of partitioning. This methodology defines a degree of belonging of each object to the clusters fixed during the partitioning process, through a parameter called membership coefficient, ranging from 0 to 1. This implies that the different objects do not belong fully to only one cluster, but are near to a given cluster. In this way it's possible to get, with respect to the previous two algorithms, better information on the structure of the objects to be classified ; in some cases this could be a disadvantage, since the output data grow very much when the number of samples is high. In addition this software doesn't define any object representative of the cluster. Fig. 4 reports an application to the same data set of 80 objects already used before with "PAM". In the case of rubbing and cracked shaft, the objects corrected classified are characterized by an elevated membership coefficient, in the case of the rotor having bearing defects the misoperation is recognized by the algorithm in a less satisfactory extent and some elements have a low value, to indicate a bad clustering.

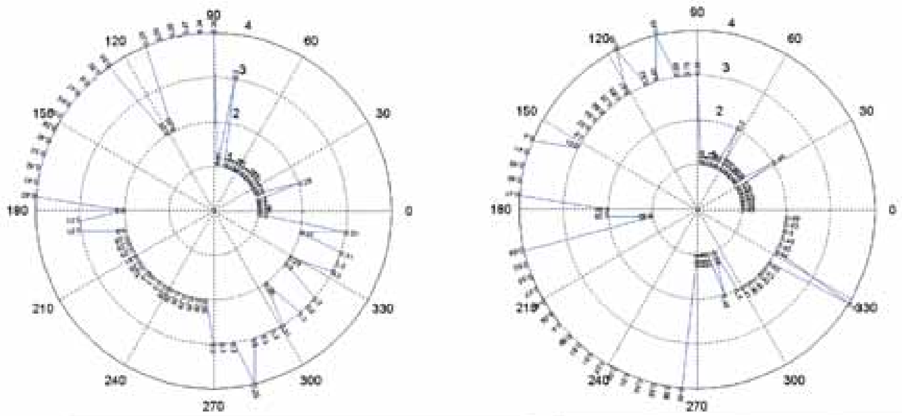


	healthy (3)	rubbing (4)	cracked (2)	bearing def. (1)
healthy	81	0	13	3
rubbing	0	60	1	0
cracked	1	40	76	7
bearing defect	18	0	10	90
overall classification accuracy: 76.7%				

	healthy (3)	rubbing (4)	cracked (2)	bearing def. (1)
healthy	72	0	9	3
rubbing	0	71	1	0
cracked	10	25	81	11
bearing defect	18	4	9	86
overall classification accuracy: 77.5 %				

(a) (b)

Fig. 3 : Application of the program (a) PAM, (b) CLARA to the diagnostics of rotating machines (standardized data, Euclidean distance)



	healthy (1)	rubbing (4)	cracked (2)	bearing def. (3)
healthy	18	0	0	1
rubbing	0	18	0	1
cracked	1	2	18	4
bearing defect	1	0	2	14
overall classification accuracy: 85.0 %				

	healthy (1)	rubbing (4)	cracked (3)	bearing def. (2)
healthy	18	0	1	5
rubbing	0	13	0	1
cracked	0	7	17	0
bearing defect	2	0	2	14
overall classification accuracy: 77.5 %				

(a) (b)

Fig. 4 : Application of the program FANNY standardized data, Euclidean distance ; (a) same features as in previous examples, (b) the energy in the frequency band 1X-2X is substituted by the third harmonic 3X

A test has been performed substituting the energy value of the signal in the frequency band 1X-2X with the one of the third harmonic 3X ; as reported in fig.4b, the classification process seems to become worse and in particular in the case of rubbing the accuracy falls from 90 % to 65 % (in the case of rubbing the energy contents in the range 1X-2X is elevated with respect with the other operation conditions).

Hierarchical algorithms

The programs of hierarchical type create a hierarchical decomposition or a hierarchical agglomeration of the objects on the basis of some criterion. The clusters are organized within a hierarchical tree generated by the software. The programs of hierarchical type use the concept of distance as a criterion for joining or separating the objects [6].

Program "TWIN" (AGNES and DIANA)

The program "TWIN" consists of an agglomerative module ("AGNES") and a divisive module ("DIANA"). "AGNES" starts considering all the points as individual cluster. At each step, it joins the couple of nearest clusters (joins the nodes having minimum dissimilarity) to form a new cluster ; this is repeated and at the end only one cluster remains (or an integer number of clusters if the process is interrupted before the end is reached). "DIANA" proceeds in the opposite order with respect to "AGNES" : it starts with a single cluster including all points and, at each step, breaks a cluster till at the end each cluster contains only one point. In general the agglomerative methods are more used, in particular in the genetic field.

Fig. 5 reports an application of "DIANA" to the previous 80 measurements. Initially (step 0) there is only one cluster containing all the 80 objects. At the first step the software divides the data into 2 clusters : in fig.6a we can see the presence of a cluster mainly composed by the measurements in rubbing conditions (circular sector II) and a cluster composed by the remaining measures (sectors I, III, IV). In the step 2, 3 clusters are present : in addition to the previous related to the rubbing conditions, a cluster corresponds to healthy machine and a third cluster contains the conditions of cracked shaft and of defective bearing. In the step 3 the program acts exclusively on the measurements in rubbing conditions : these are divided into two different cluster, one mainly formed by measurements at low speed, under the critical condition, the other in super critical conditions.

In the step 4 (fig.6d) a new cluster is originated, consisting in only one point ; this condition corresponds to a cracked shaft condition. The fact that this point is classified as an isolated cluster could suggest that the measurement is very different from the other and should be considered as an anomalous measurement.

In the step 5 we have 6 clusters : one corresponds mainly to healthy conditions, two correspond to rubbing, one refers to only one measurement (the anomalous one), one corresponds to cracked rotor and the last one corresponds to a defect of bearing.

Fig. 6 reports the results obtained with the agglomerative software on the same data. Analogous considerations hold for the results of this module. Also in this case the same measurement in cracked shaft condition seems to present anomalous characteristics being considered as an isolated cluster till the last steps of the agglomerative algorithm.

Program "HCE" (Hierarchical Clustering Explorer)

In parallel to the modules "DIANA" and "AGNES", we used a software named "Hierarchical Clustering Explorer" developed for genetic applications, making possible to identify interesting patterns in very large datasets having more than 500 000 features. The result of such a process of clustering is represented through a "dendrogram". The software provides also a graphical representation of the levels of the features for each object to be classified through a mosaic display, which better puts into evidence the agglomeration process of the objects and the presence of similar patterns. Additional information about this software package is available in [7].

Fig. 7 reports in a dendrogram form the results provided by the software for the same series of 80 data used previously. To estimate the proximity between the objects the euclidean distance has been used, (as for the algorithms "DIANA" and "AGNES") and also in this case the measurements have been standardized. Fig.8a

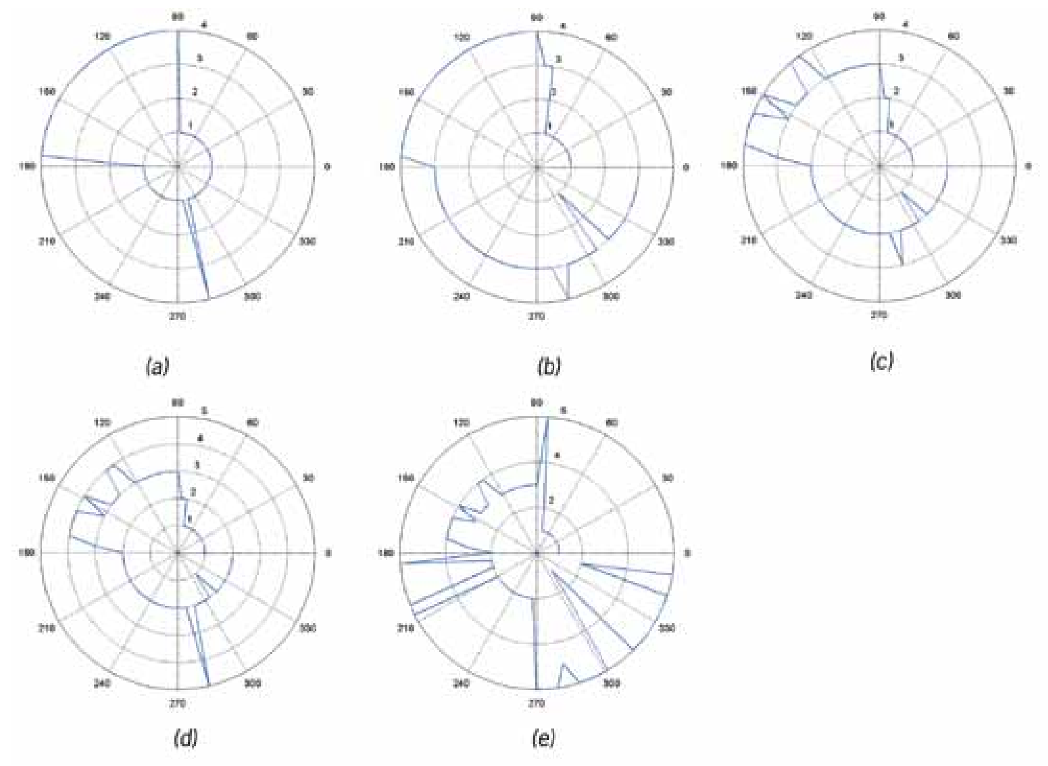


Fig. 5 : Application of the program DIANA (a) step1, (b) step2, (c) step3, (d) step4, (e) step5

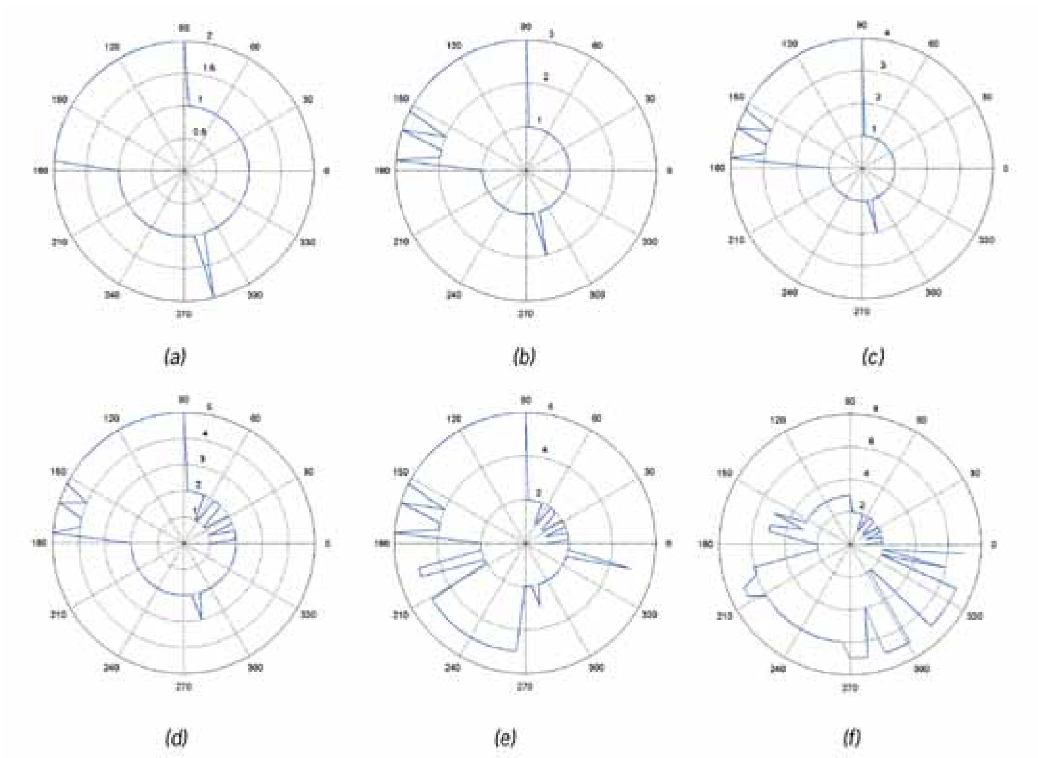


Fig. 6 : Application of the program AGNES (a) step79, (b) step78, (c) step77, (d) step76, (e) step75, (f) step74

shows, according to a polar notation, the result of the classification for the case where the agglomeration process was interrupted at a value 0.043 of the minimum similarity, corresponding to 5 clusters. Also in this case the element 64 (bearing defect) is recognized as an anomalous element, i.e. as an outlier, and is classified into a cluster consisting only of itself. Fig. 8b and fig.8c show a classification resulting from the use of the correlation coefficient of Pearson instead of the euclidean distance to evaluate the similarity. The Pearson correlation coefficient [7] is a measure of the degree of correlation of two variables ; its results for the present application were worse with respect to the case of the Euclidean distance. This could be due to the fact that the Pearson correlation coefficient should give interesting results in the case where the number of features is elevated, like in the genetic applications.

Program "MONA" (Monothetic Analysis)

Finally, in the area of the hierarchical algorithms, we used a clustering software ("MONA") accepting as inputs the values of the variables in binary format. The only two values 0 and 1 indicate the presence or the absence of an attribute. An advantage of this methodology could be to limit the influence of the noise by the process of variables discretization into two levels.

The basic idea of the software is to select one of the two variables and, based on it, to divide the set of objects into those having this attribute (value equal to 1) and those without it (value equal to 0). In each subset, one of the remaining features is selected and used in the same way to split the subset into two smaller groups. The process is iterated and ends when the subsets are reduced to a

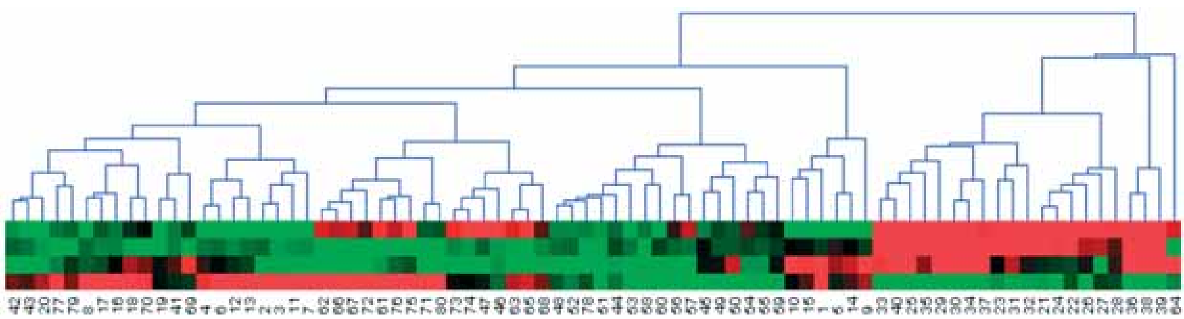


Fig. 7 : Application of the program Hierarchical Clustering Explorer – dendrogram

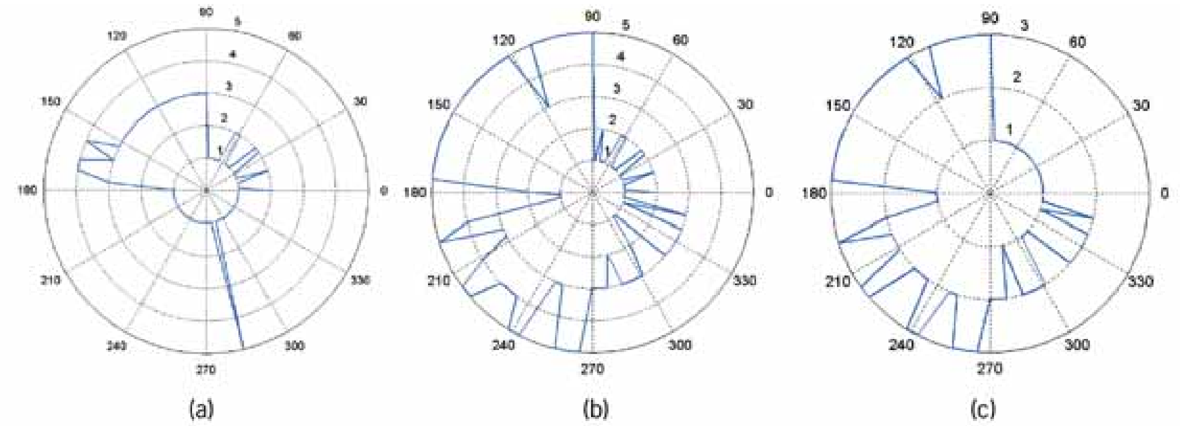
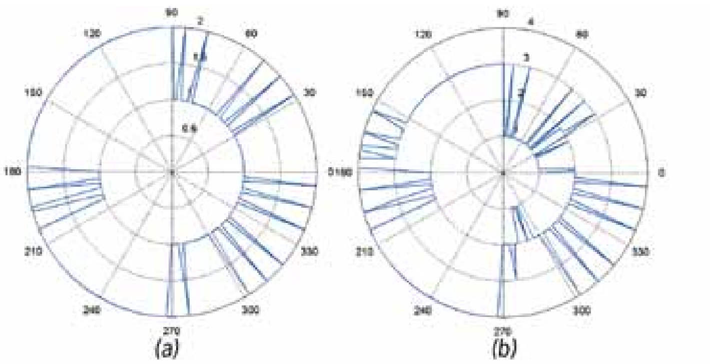


Fig. 8 : Application of Hierarchical Clustering Explorer – classification with a) minimum similarity = 0.043, 5 clusters, euclidean distance ; b) minimum similarity = 0.682, 5 clusters, Pearson correlation coefficient ; c) minimum similarity = 0.481, 3 clusters, Pearson correlation coefficient

single item. Clearly the method belongs to the category of hierarchical divisive algorithms. One of the most difficult aspects is the choice of the feature to split a set ; one criterion used is to choice the feature, which exhibits the maximum similarity to the other [6].

The diagnostic application tested was consisting in 200 measurements on a rotor, in the 4 operation conditions considered in the previous cases. For each measurement, 7 variables were examined : the 4 used by the previous algorithms and, in addition, the 2nd harmonic 2X, the 3rd harmonic 3X, and the energy of the signal in the range 2X-3X. The values have been discretized in binary form using a transformation of level assuming as the reference value the average value of the feature evaluated on all the samples under examination. The results are reported in fig. 9. The accuracy of the classification remains high (79 %), analogous to the ones obtained with the previous modules ; this could be justified by the fact that the loss of information due to discretization is compensated by the increase in the number of variables, passing from 4 to 7. In fact a clusterization using 4 variables only provided worse results.



	healthy	rubbing	cracked	bearing defect
healthy	38	0	0	5
rubbing	5	43	1	1
cracked	0	7	39	6
bearing defect	7	0	10	38
overall classification accuracy: 79 %				

Fig. 9 : Application of the program MONA – (a) step1 : the cluster is divided into 98 and 102 objects, using variable 1X, (b) step2, the first cluster is divided into 43 and 55 objects, using variable information dimension and the second cluster is divided into 50 and 52 objects using the same variable (1healthy, 3 rubbing, 4 cracked, 2 bearing defect)

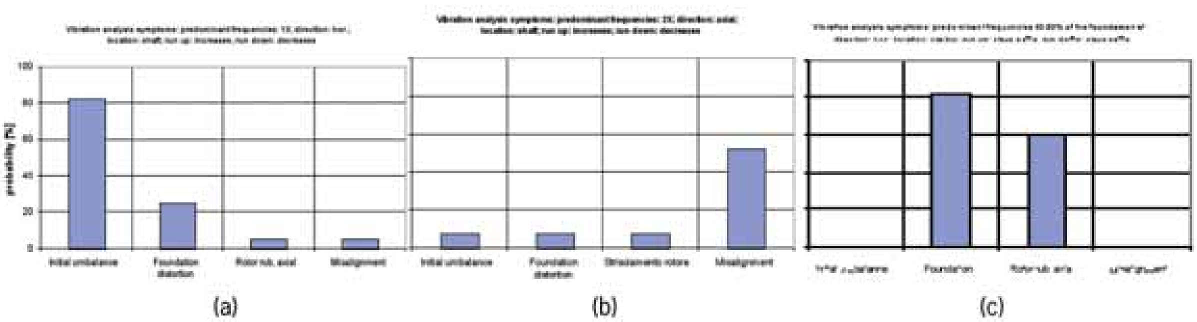


Fig. 10 : Results of an expert system

Diagnostic applications with expert systems

An automatic diagnostics may be also performed using expert systems, performing the analysis through rules given by experts. Expert systems codify trough rules the experts knowledge, making possible the interpretation of data by lower level operators, when the expert isn't or isn't any more available, and relevant economic savings. Fig. 10 shows an example of a possible application to rotating machinery diagnostics, using Bayes theory, based on a priori and conditioned probabilities. The reference for the conditioned probabilities was the Sohre tables for rotating machines [8, 9]. Sohre tables don't provide the a priori probabilities ; in the example developed as a test, lacking real precise historical data, they were assumed with modest variations, according to common sense judgments, with respect to an equiprobabile distribution. The plots of Fig. 10 report for three distinct cases the probabilities of the possible misoperations, evaluated by the expert system based on vibration data referring to the frequency of the highest amplitude harmonics, to its direction, to the measurement section of the machine corresponding to this component, to the dynamic behavior of the machine during run-up and run-down.

Conclusion

The results provided by the different clustering algorithms tested seem to be encouraging, since the amount of data correctly classified is high enough, and significant in the perspective of an advanced condition monitoring of rotating machines. The hierarchical methods in some cases showed some limitations, since a set of data corresponding to one category of defect has been split into two clusters. In addition, an aspect to be evaluated is the fact that, when a decision has been taken with respect to agglomeration or division of a cluster, it isn't possible to modify it. A consequence could be a sensitivity of these methods to the presence of disturbs and noise in the variables used [6].
Also the example developed for testing the feasibility of an expert system based on Sohre tables seems to indicate a satisfactory applicability of the method.

Bibliography

[1] M.L.Adams, I.A.Abu-Mahfouz : "Exploratory research on chaos concepts as diagnostic tools for assessing rotating machinery vibration signatures" Case Western University, Cleveland, 1994

[2] A. Lucifredi, A. Magnetto, P. Silvestri "Experimental validation of an original software package, based on chaos theory, for monitoring and diagnostics of rotating machinery" ; 6th International Conference on Rotor Dynamics, September 30 - October 3, 2002, Sydney, Australia

[3] M. Fontana, A. Lucifredi, P. Silvestri "A monitoring and diagnostic tool for machinery and power plants, based on chaos theory" IV International Conference Condition Monitoring and Diagnostic Engineering Management, University of Manchester, UK, September 2001

[4] A. Lucifredi, A. Magnetto, P. Silvestri "Condition monitoring and diagnostics based on chaos theory"; ISMA2002 Int. Conf. Noise and Vibration Engineering, Leuven, B, 16-18 September 2002

[5] A. Lucifredi, P. Silvestri "An overview of fundamental requirements for a condition monitoring and fault diagnosis system for machinery and power plants", Tenth International Congress on Sound and Vibration - ICSV10, Stockholm 7-10 July 2003

[6] L. Kaufman, P.J. Rousseeuw, "Finding Groups in Data : An Introduction to Cluster Analysis" Wiley-Interscience - New York 1987

[7] Internet web page : http://www.cs.umd.edu/hcil/multi-cluster

[8] J.S.Sohre "Foundation for High-Speed Machinery", ASME 62-WA-250, fifth revised edition, July 1976

[9] J.S.Sohre "Operating problems with High-Speed Turbomachinery, Causes and Correction" ASME petroleum conference, Dallas, Texas, 1968