Evidential Evolving Gustafson-Kessel Algorithm (E2GK) and its application to PRONOSTIA's Data Streams Partitioning

Lisa Serir Emmanuel Ramasso Patrick Nectoux Olivier Bauer Noureddine Zerhouni

***Abstract*— Condition-based maintenance (CBM) appears to be a key element in modern maintenance practice. Research in diagnosis and prognosis, two important aspects of a CBM program, is growing rapidly and many studies are conducted in research laboratories to develop models, algorithms and technologies for data processing. In this context, we present a new evolving clustering algorithm developed for prognostics perspectives. E2GK (Evidential Evolving Gustafson-Kessel) is an online clustering method in the theoretical framework of belief functions. The algorithm enables an online partitioning of data streams based on two existing and efficient algorithms: Evidantial *c*-Means (ECM) and Evolving Gustafson-Kessel (EGK). To validate and illustrate the results of E2GK, we use a dataset provided by an original platform called PRONOSTIA dedicated to prognostics applications.**

I. Introduction

Cluster analysis, as a multivariate statistical analysis method, is a classification approach that aims at grouping a set of N data into c clusters $\omega_1, \ldots, \omega_c$ whose members are similar in some way. Clustering refers to a wide variety of algorithms that aim at discovering groups (also called clusters or partition) in the data in order to summarize the data or to emphasize an existing structure. In most cases, a cluster is defined as a subset of data for which the similarity is larger than the similarity with other data in other subsets. A natural way of data grouping is based on similarity measures such as the Euclidean distance, the Mahalanobis distance and the Kullback-Leibler divergence. In the context of prognostics and diagnostics applications, clustering makes it possible to partition data streams into different fault categories.

A wide variety of clustering methods have been developed. Application of cluster analysis in machinery fault diagnosis was discussed in [1]. The most commonly used methods for clustering are divided into two main categories: hierarchical and non-hierarchical methods. We also distinguish hard and fuzzy partitioning. The former aims at grouping data in an exclusive way, making a given data unable to belong to several clusters, whereas the former allows each data to belong to more than one cluster with different membership degrees. The most popular fuzzy partitioning method is Bezdek's Fuzzy C-means (FCM) algorithm [2]. One can also mention the Gustafson-Kessel fuzzy clustering algorithm [3] that is capable of detecting hyperellipsoidal clusters of different sizes and orientations by adjusting the covariance matrices.

A relatively new concept of partition, introduced in [4], is the *credal* partition (also called evidential partition) based on belief functions theory. It extends the existing concepts of hard, fuzzy, probabilistic, and possibilistic partitions by allocating, for each data, a *mass of belief*, not only to single clusters, but also to any subset of $\Omega = \{\omega_1, \ldots, \omega_c\}$. This particular representation makes it possible to code all the situations, from certainty to total ignorance of membership to clusters. For instance, doubt is generally encountered in data transition and can be useful to limit the number of clusters in the final partition. Moreover, outliers (atypical data) are well managed using the conflict degree explicitly emphasized in the belief function framework. These concepts were exploited in the Evidencial c-Means (ECM) algorithm [5] for static data.

Numerous techniques have been developed for clustering data in a static environment [6]. However, in many real-life applications, non-stationary data (i.e. with time-varying parameters) are commonly encountered. *Online clustering*, also called *incremental clustering* in machine learning [7], aims at sequentially grouping data into clusters. It is generally unsupervised and has to manage recursive training in order to incorporate new information gradually and to take into account model evolutions over time.

E2GK (Evidential Evolving Gustafson-Kessel Algorithm), presented in this paper, is an online clustering technique that relies on two existing algorithms: the Evolving (fuzzy) Gustafson-Kessel algorithm (EGK) proposed in [8] used for the determination of the centers of the clusters, and the Evidential c-means algorithm (ECM) proposed in [5] used for obtaining the credal partition. In E2GK, the credal partition is obtained online by allocating belief masses to the different subsets of clusters as data arrive. The size of the partition can evolve according to the novelty of the data.

To illustrate our results, a dataset provided by an experimental platform called PRONOSTIA is used. This platform is dedicated to bearing prognosis. PRONOSTIA is developed within the Department of Automatic Control and Micro-Mechatronic Systems (AS2M) of FEMTO-ST institute[1] for the test and validation of bearing prognostics approaches. The originality of this experimental platform lies in the characterization of both the bearing functioning and its degradation and also in the possibility to make the operating conditions of the bearing vary during its useful life.

FEMTO-ST Institute, UMR CNRS 6174 - UFC / ENSMM / UTBM, Automatic Control and Micro-Mechatronic Systems Dep., 25000, Besançon, France (lisa.serir, emmanuel.ramasso, noureddine.zerhouni, patrick.nectoux, olivier.bauer)@femto-st.fr

[1] FEMTO-ST stands for "Franche-Comté Electronics, Mechanics, Thermal Processing, Optics - Science and Technology". The plateform was developed in AS2M department (Automatic control and Micro-Mechatronic Systems).

Section II describes PRONOSTIA platform. In Section III, we present the GK algorithm and belief functions giving the necessary background for Section IV in which we introduce E2GK. Results are finally presented in Section V.

II. PRONOSTIA Platform

PRONOSTIA is an experimentation platform (Figure 1) dedicated to the test and validation of the machinery prognosis approaches, focusing on bearing prognostics. The main objective of PRONOSTIA is to provide real experimental data that characterise the degradation of a ball bearing along its whole operational life (until fault/failure). The collected data are vibration and temperature measurements of the rolling bearing during its functioning mode. As prognosis algorithms need statistical data, it is necessary to conduct an experiment in a few hours, and to collect a large amount of data in a few weeks. To do so, we developed a device that is able to maintain the bearing into hard operating conditions.

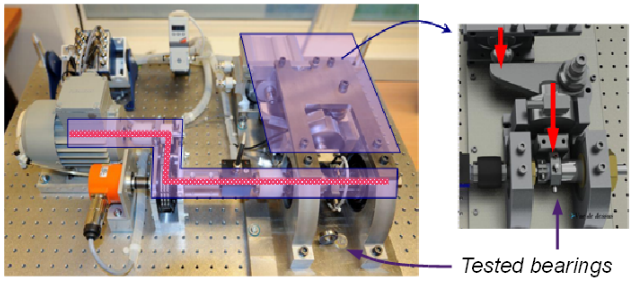


Fig. 1. PRONOSTIA platform.

A. Operating Conditions of the Bearing

The internal bearing ring is put in rotation, while the external bearing ring is maintained fixed. A radial load is applied on the external bearing ring in order to simulate its functioning. To speed up the degradation, the load exceeds the maximal load recommended by the supplier. The originality of this experimental platform lies not only in the conjunction of the characterization of both the bearing functioning (speed, torque and radial force) and its degradation (vibrations and temperature), but also in the possibilities, offered by the platform, to make the operating conditions of the bearing vary during its useful life. Figure 2 depicts a bearing before and after the experiment.

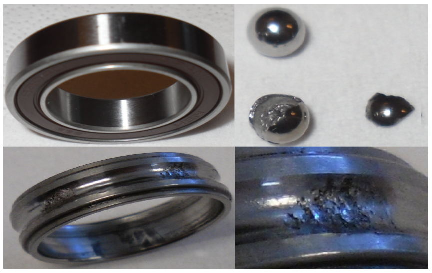


Fig. 2. Bearing degradation.

B. Measurement of the Bearing's Behavior

1) Measurement of the parameters imposed to the bearing: The bearing operating conditions are determined by instantaneous measures of the radial force applied on the bearing, the rotation speed of the shaft handling the bearing, and of the torque inflicted on the bearing. Thus, three sensors are used: a load cell and its transducer amplifier, an incremental encoder and its analog signal converter, and a torque transducer with its converter. Each of these analogue measurements are sampled at 100Hz.

2) Measurement of the degradation's characteristics: During a test, the rolling bearing starts from its nominal mode until the fault state. The bearing behavior is measured using different types of sensors (Figure 3). Resulting from the study of heat and vibrations, the characterization of the degradation lies on data provided by two types of sensors:

- Two miniaturized acceleration sensors for vibrations study (frequency) are positionned at 90 degrees one from the other, on the bearing clamping ring.
- A temperature probe of ceramic type PT100 in 1/3 DIN permits the visualization of the evolution of temperature.

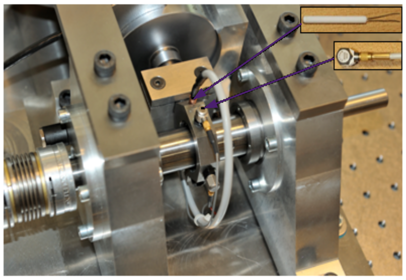


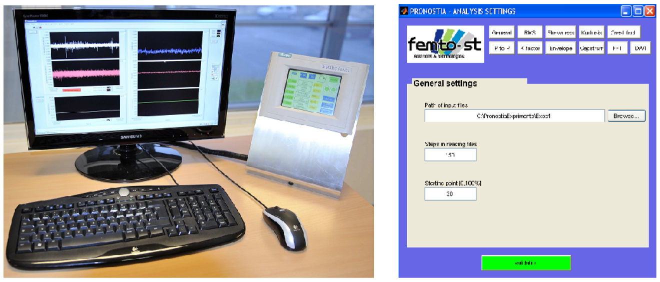
Fig. 3. Sensors for degradation measurement.

C. Data acquisition system

This apparatus permits the integration of three modules: a first module that is specific to acceleration sensors of type IEPE, a second specific to PT100 probes, and a last one of standard tension references. The first module includes the two acceleration sensors, the second, the PT100 temperature probe, and the third, the load cell and the torque transducer. Synchronous measures can also be made on the three modules. The resulting data are transmitted via a USB 2.0 link to the computer in charge of the data management. The dedicated application - developed using *Labview* - ensures the visualization of the signals provided by the different sensors and sampled in a specific manner. Thus, all data can be monitored in real time on scrolling graphs. They are also formated, timestamped, and recorded locally in different files in order to be exploited for bearing fault prognosis.

D. Signal processing toolbox

The raw signals provided by the sensors are processed in order to extract relevant information concerning bearings states. Several techniques have been implemented and gathered in a signal processing toolbox with Matlab: time-domain methods (RMS, skewness and kurtosis, crest factor, K-factor, Peak-to-Peak), frequency-domain methods (spectral and cepstrum analysis, envelope detection), time-frequency domain (short-time fourier transform) and discrete wavelets.



General settings

Path of input files
C:\PronostiaExprimernts\Exp1

Steps in reading files
150

Starting point [0,100%]
30

Fig. 4. (left) Labview VI for raw signal visualization and (right) the graphical user interface to set the optional parameters (if required) of the signal processing algorithms.

For example, Figure 5 is the power spectral density of the vertical acceleration sensor computed during the last half of the test period. One can visualize the amplitude growing at the end of the experiment when the bearing is gradually degrading. Various other data processings are possible to provide the necessary tools for bearing prognostics.

III. Background

A. Belief Functions

Dempster-Shafer theory of evidence, also called belief functions theory, is a theoretical framework for reasoning with partial and unreliable information. Ph. Smets proposed the *Transferable Belief Model* (TBM) [9] as a general framework for uncertainty representation and combination of various pieces of information without additional priors. In particular, TBM offers the possibility to explicitly emphasize doubt and conflict. Typically, the former represents ignorance and the latter emphasizes the contradiction within a fusion process. We present the basic notions of the theory and refer the reader to [9] for a thorough description.

Considering a finite set called the frame of discernment Ω of variables ω, the *belief* of an agent in subsets of Ω can be represented by a *basic belief assignment* (BBA):

$$\begin{array}{rccc} m: & 2^{\Omega} & \to & [0,1] \\ & A & \mapsto & m(A)\ , \end{array} \tag{1}$$

with $\sum_{A\subseteq\Omega} m(A) = 1$. A belief mass can not only be assigned to a singleton ($|A| = 1$), but also to a *subset* ($|A| > 1$) of variables *without assumption concerning additivity*. This property permits the explicit modelling of doubt and conflict, and constitutes a fundamental difference with probability theory. The subsets A of Ω such that $m(A) > 0$, are called the *focal elements* of m. Each focal element A is a set of possible values of ω. The quantity $m(A)$ represents a fraction of a unit mass of belief allocated to A. Complete ignorance corresponds to $m(\Omega) = 1$, whereas perfect knowledge of ω is represented by a categorical BBA, i.e. the whole mass is assigned to a ω. In the case where all focal elements are singletons, m boils down to a probability distribution.

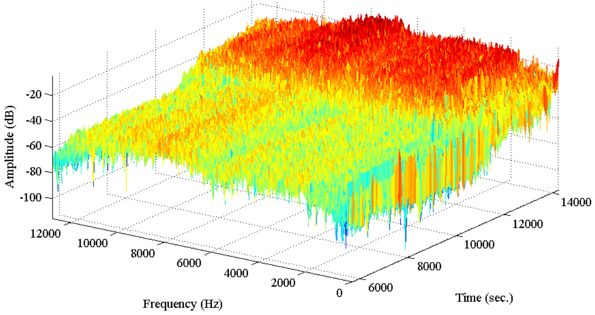


Fig. 5. Power spectral density of the vertical acceleration sensor.

A positive value of $m(\emptyset)$ is considered if one accepts the *open-world assumption* [10] stating that the set Ω might not be complete, and thus ω might take its value outside Ω. The conflict is then interpreted as a mass of belief given to the hypothesis that ω might not lie in Ω. This interpretation is useful in clustering for outliers detection [5].

B. Evidential C-Means

Belief functions theory is largely used in clustering and classification problems [11], [12]. Recently (2003), the concept of *credal partition* was introduced as a generalization of fuzzy partitioning. It particularly permits a better interpretation of the data structure and makes it possible to code all situations, from certainty to total ignorance. A credal partition is constructed by assigning a BBA to each possible subset of clusters. Partial knowledge regarding the membership of a data i to a class j is represented by a BBA m_{ij} on the set $\Omega = \{\omega_1, \ldots, \omega_c\}$. We introduce the concept of credal partition as described in (ECM, [5]).

In ECM, the objective function was defined as:

$$J_{ECM}(M,V) = \sum_{i=1}^{N} \sum_{\{j/A_j\neq\emptyset, A_j\subseteq\Omega\}} |A_j|^{\alpha} m_{ij}^{\beta} d_{ij}^{2} + \sum_{i=1}^{N} \delta^2 m_i(\emptyset)^{\beta}\ , \tag{2}$$

subject to

$$\sum_{\{j/A_j\neq\emptyset, A_j\subseteq\Omega\}} m_{ij} + m_i(\emptyset) = 1 \quad \forall i = 1,\ldots,N\ , \tag{3}$$

where:

- α is used to penalize the subsets of Ω with high cardinality,
- $\beta > 1$ is a weighting exponent that controls the fuzziness of the partition,
- d_{ij} denotes the Euclidean distance between data i and prototype v_j,
- δ controls the amount of data considered as outliers.

The $N \times 2^c$ partition matrix M is derived by determining, for each data i, the BBAs $m_{ij} = m_i(A_j)$, $A_j \subseteq \Omega$ such that m_{ij} is low (resp. high) when the distance d_{ij} between data i and focal element A_j is high (resp. low). The matrix M is computed by the minimization of criterion (2) and was shown to be [5], $\forall i = 1 \dots n$, $\forall j / A_j \subseteq \Omega$, $A_j \neq \emptyset$:

$$m_{ij} = \frac{\left|A_j\right|^{-\alpha/(\beta-1)} d_{ij}^{-2/(\beta-1)}}{\sum\limits_{A_k \neq \emptyset} |A_k|^{-\alpha/(\beta-1)} d_{ik}^{-2/(\beta-1)} + \delta^{-2/(\beta-1)}} \quad , \tag{4}$$

and $m_i(\emptyset) = 1 - \sum_{A_j \neq \emptyset} m_{ij}$. In ECM, centers of clusters are optimized by minimizing criterion (2). The distance between a data and any non empty subset $A_j \subseteq \Omega$ is then defined by computing the center of each subset A_j. This latter is the barycenter $\overline{v_j}$ of the centers of clusters composing A_j.

From the credal partition, the classical clustering structures (possibilistic, fuzzy and hard partitions) can be recovered [5]. One can also summarize the data by assigning each data to the set of clusters with the highest mass. One then obtains a partition of the points in at most 2^c groups, where each group corresponds to a set of clusters. This makes it possible to find the points that unambiguously belong to one cluster, and the points that lie at the boundary of two or more clusters. Moreover, points with high mass on the empty set may optionally be rejected as outliers.

Example 1: Considering $N = 4$ data and $c = 3$ classes, Tab. I gives an example of a credal partition. BBAs for each data given in Tab. I illustrate various situations: data 2 certainly belongs to class 1, whereas the class of data 4 is completely unknown. Partial knowledge is represented for data 1. As $m_3(\emptyset) = 1$, data 3 is considered as an outlier, i.e., the class of data 3 does not lie in Ω.

IV. E2GK: Evidential Evolving Gustafson-Kessel Algorithm

The GK algorithm [3] has the advantage to adapt the clusters according to their real shape. The obtained clusters are hyperellipsoids with arbitrary orientation and well suited for a variety of practical problems. In [8], an online version of GK clustering algorithm (EGK) that enables online partitioning of data streams was developed. The partition matrix was updated as in GK but rules were added to decide whether a new cluster has to be created or existing prototypes adapted.

In this section we propose an adaptation of this algorithm to the context of belief functions (Alg. 1). The main idea is to derive, online, a credal partition matrix from the data.

A. Step 1 – Initialization

At least one cluster center should be provided. Otherwise, the first point is chosen as the first prototype. If more than one prototype is assumed in the initial data, the GK or ECM algorithm can be applied to identify an initial partition matrix. The result of the initialization phase is a set of c prototypes v_i and covariance matrices F_i.

B. Step 2 – Decision making

The boundary of each cluster is defined by the cluster radius r_i, defined as the *median* distance between the cluster center v_i and the points belonging to this cluster with membership degree larger or equal to a given threshold u_h:

$$r_i = \underset{\forall x_j \in \text{ } i\text{-th cluster and } P_{ij} > u_h}{\text{median}} \left\| v_i - x_j \right\|_{A_i} \quad . \tag{5}$$

where P_{ij} is the confidence degree that point j belongs to $\omega_i \in \Omega$ and can be obtained as follows: either by using the belief mass $m_j(\omega_i)$, or the pignistic transformation [9] (converting a BBA into a probability distribution), or by using the plausibility transform [13]. We propose to choose the belief mass for which the computation is faster.

Compared to EGK, where the *maximum* rule is used, we here apply the *median* value which is less sensitive to extreme values. Moreover, the minimum membership degree u_h - initially introduced in [8] and required to decide whether a data point belongs or not to a cluster - can be difficult to assess. It may depend on the density of the data as well as on the level of cluster overlapping. We rather set u_h automatically to $1/c$ in order to reduce the number of parameters while ensuring a natural choice for its value.

C. Step 3 – Computing the partition matrix:

Starting from the resulting set of clusters at a given iteration, we need to build the partition matrix M. We here consider the Mahalanobis-like distance d_{ik} assuming that each cluster volume ρ_i is one as in standard GK algorithm:

$$d_{ik}^2 = \|x_k - v_i\|_{A_i}^2 = (x_k - v_i) A_i (x_k - v_i)^T \quad , \tag{6a}$$

$$A_i = \left[\rho_i \cdot \det(F_i)\right]^{1/n} F_i^{-1} \quad , \tag{6b}$$

$$F_i = \frac{\sum_{k=1}^{N} (m_{ik})^\beta (x_k - v_i)^T (x_k - v_i)}{\sum_{k=1}^{N} (m_{ik})^\beta} \quad . \tag{6c}$$

where F_i is the fuzzy covariance matrix. We then compute the credal partition using Eq. 4.

Remark 1: Storing the whole partition is however not efficient. Indeed, we simply need to store the belief masses on singletons in order to make the decision concerning the radius. As shown in Eq. 4, values on singletons are easy to compute but the problem is to estimate the normalization factor. For that, all values of masses have to be computed but not stored. This little trick exponentially decreases memory comsumption.

D. Step 4 – Adapting the structure

Given a new data point x_k, two cases are considered.

Case 1: x_k belongs to an existing cluster, thus a clusters' update has to be performed. Data point x_k is assigned to the closest cluster p if the distance d_{pk} is less or equal to the radius r_p. Then, an update of the p-th cluster has to be performed as follows:

$$v_{p,new} = v_{p,old} + \theta \cdot \Delta \quad , \tag{7}$$

where $\Delta = x_k - v_{p,old}$ and

$$F_{p,new} = F_{p,old} + \theta \cdot \left(\Delta^{\mathrm{T}} \Delta - F_{p,old}\right) \quad , \tag{8}$$

where θ is a learning rate, $v_{p,new}$ (resp. $v_{p,old}$) and $F_{p,new}$ (resp. $F_{p,old}$) denote the new (resp. old) values of the centers and covariance matrix.

Case 2: x_k is not within the boundary of any existing cluster (i.e. $d_{pk} > r_p$), thus a new cluster may be defined and a clusters' update has to be performed. The number of clusters is thus incremented. Then, the incoming data x_k is

TABLE I
EXAMPLE OF A CREDAL PARTITION.

A	$\emptyset$	ω_1	ω_2	$\{\omega_1,\omega_2\}$	ω_3	$\{\omega_1,\omega_3\}$	$\{\omega_2,\omega_3\}$	$\{\omega_1,\omega_2,\omega_3\}$
$m_1(A)$	0	0	0	0	0.2	0.5	0	0.3
$m_2(A)$	0	1	0	0	0	0	0	0
$m_3(A)$	1	0	0	0	0	0	0	0
$m_4(A)$	0	0	0	0	0	0	0	1

accepted as a center v_{new} of the new cluster and its covariance matrix F_{new} is initialized with the covariance matrix of the closest cluster $F_{p,old}$. In order to quantify the credibility of the estimated clusters, the number of points belonging to the i-th cluster is estimated as in EGK [8] where the authors suggested a threshold parameter P_{tol} to guarantee the validity of the covariance matrices and to improve the robustness. This parameter corresponds to the desired minimal amount of points falling within the boundary of each cluster (its value is context-determined). The new created cluster is then rejected if it contains less than P_{tol} data points.

We also propose an additional step in E2GK. After creating a new cluster, the data structure evolves. However, the new cluster may contain data points previously assigned to another cluster. Thus, the number of data points in previous clusters could change. We propose to verify, after the creation of a new cluster, that all clusters have at least the required minimum amount of data points (P_{tol} or more). If it is not the case, then remove the cluster with the minimum number of points. Compared to the initial EGK algorithm, in which the number of clusters only increases, E2GK is more flexible because the structure can change by *decrementing* or *incrementing* the number of clusters.

The overall algorithm is presented in Alg. 1 where the proposed adaptations appear in bold.

Algorithm 1 E2GK algorithm

```
**Initialization:** Take the first point as a center or apply the off-line GK algorithm to get the initial number of clusters $c$ and the corresponding matrices $V$ and $F_i$, $i = 1 \cdots c$
**Calculate** $\overline{v_j}$, the barycenter of the cluster centers composing $A_j \subseteq \Omega$
**Calculate the credal partition** $M$ using Eq. 4 (store only singletons and normalize)
for all new data point $x_k$ do
  Find the closest cluster $p$
  Calculate the radius $r_p$ of the closest cluster (Eq. 5)
  if $d_{pk} \leq r_p$ then
    Update the center $v_p$ (Eq. 7)
    Update the covariance matrix $F_p$ (Eq. 8)
  else
    Create a new cluster: $v_{c+1} := x_k$ and $F_{c+1} := F_p$
    Keep it if the number of points is $\geq P_{\text{tol}}$
  end if
  **Recalculate** the credal partition $M$
  **Check the new structure:** estimate the number of points in each cluster and remove the cluster with the minimum number of points if $\leq P_{tol}$
end for
```

V. APPLICATION OF E2GK

A. A benchmark 1-D problem

As a first example of E2GK, we propose to consider the Mackey-Glass chaotic time series:

$$x(t) = \frac{0.2 \cdot x(t-\tau)}{1 + x^{10}(t-\tau)} - 0.1 \cdot x(t) \ , \tag{9}$$

with $\tau = 17$ and $x_0 = 1.2$. A total of 270 samples were generated. The obtained series is depicted in Figure 6 as well as the resulting segmentation by E2GK (using $[t \ x(t)]$ as inputs). E2GK parameters were set to $\delta = 20$, $\alpha = 1$, $\beta = 2$, $\theta = 0.01$ and $P_{\text{tol}} = 15$. Figure 7 shows the number of clusters evolving along time. The online segmentation generated 10 segments well located on the curve.

B. A multidimensional real problem on PRONOSTIA

As a second application, we considered the PRONOSTIA plateform. We here applied E2GK in order to automatically find a partioning (online) of the data streams. We consider here the power spectral density made of 512 points at each time slice. This huge dataset is then post-processed by a principal components analysis in order to automatically select the 6 most representative frequencies. These 6 features are used as inputs of E2GK (with 250 points).

E2GK algorithm is initialized ramdomly using 2 centers and 20 data. Parameters were set *as in the previous application*. The first 20 data correspond roughly to the first 0.5 hour of the experiment where the bearing does not present any default. Then, data arrives sequentially and make the clustering structure possible to evolve. E2GK adapted its structure until obtaining 7 clusters as pictorially described in Fig. 8. First of all, a third cluster is obtained into the cloud around the initialisation points. This shows that the bearing only degrades from about the third hour. Then 4 clusters are gradually added according to the degradation. Cluster ω_4 represents a *transition* between the *normal modes* (ω_1, ω_2 and ω_3) towards the *degrading modes* (ω_5 and ω_6). Finally the *fault mode* is detected with ω_7. Figure 9 shows the assignments (cluster chosen for each data point w.r.t. the maximum of belief) along with the first three dimensions.

Some similar segmentation can be obtained by EGK by precisely tuning its parameters. However, E2GK is able to provide the real location of transition between clusters. Moreover, conflict in E2GK can be exploited for outlier

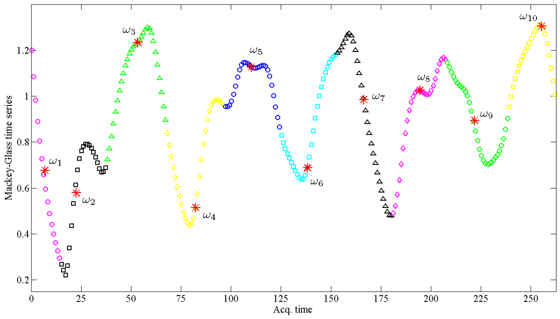


Fig. 6. The Mackey-Glass time series and its online segmentation. Prototypes appear in red stars.

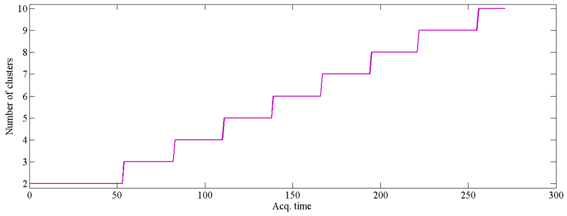


Fig. 7. Number of clusters along time for the first application.

detection (this point has not been demonstrated here but in the static case one can refer to [5]). Finally, one can notice that the *same* parameter values were used in both applications showing that E2GK may be easy to tune. In comparison, we tested EGK for which different parameters values must be chosen to obtain a satisfactory result in particular in the second application with noisy data.

VI. Conclusion

E2GK algorithm is an evolving clustering algorithm using belief function theory, which relies on the credal partition concept. This type of partition permits a finer representation of datasets by emphasizing doubt between clusters as well as outliers. Doubt is important for data streams analysis from real systems because it offers a suitable representation of gradual changes in the stream. E2GK also relies on some parts of EGK algorithm [8], initially based on a fuzzy partition, to which we bring some modifications. We treated implementation issues inherent to belief masses computation. We finally tested E2GK on a real plateform called PRONOSTIA designed to provide data concerning the degradation of a ball bearing during its operational life.

A thorough analysis of parameters sensitivity (P_{tol} and θ) is now required to properly and automatically set them. The interpretation of clusters provided by E2GK is also an important problem for industrial applications [14]. Besides, the automation of the variation of bearing operating conditions on PRONOSTIA is under study in order to simulate different scenarios such as real industrial cycles.

Finally, we are currently developing a prognostic approach fully based on belief functions and plan to compare it with the fuzzy evolving systems proposed by Angelov [15].

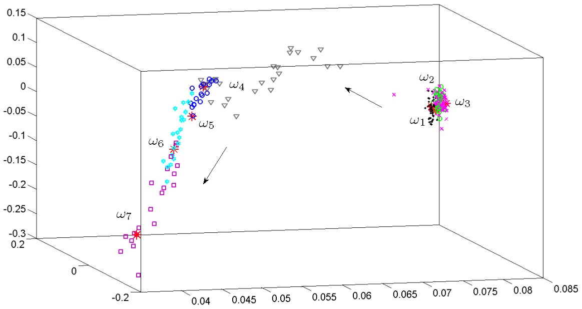


Fig. 8. Online segmentation into clusters for PRONOSTIA's data (centers appear in red cross, arrows represent order of data arrival).

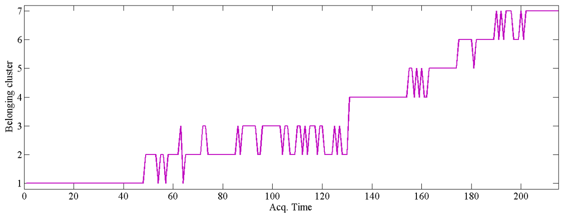


Fig. 9. Belonging cluster for the second application.

References

[1] V. Skormin, L. Popyack, V. Gorodetski, M. Araiza, and J. Michel, "Applications of cluster analysis in diagnostics-related problems," in *IEEE Aerospace Conference*, vol. 3, Snowmass at Aspen, CO, USA, 1999, pp. 161–168.

[2] J. C. Bezdek, *Pattern Recognition with fuzzy objective function algorithms*. New York: Plenum Press, 1981.

[3] E. Gustafson and W. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in *IEEE Conference on Decision and Control*, 1978.

[4] T. Denoeux and M.-H. Masson, "Evclus: Evidential clustering of proximity data," *IEEE Transactions on Systems, Man and Cybernetics Part B*, vol. 34(1), pp. 95 – 109, 2004.

[5] M.-H. Masson and T. Denoeux, "Ecm: An evidential version of the fuzzy c-means algorithm," *Pattern Recognition*, vol. 41(4), pp. 1384 – 1397, 2008.

[6] C. M. Bishop, *Neural networks for pattern recognition*, Oxford, Ed. Oxford University Press, 1995.

[7] S. Janichen and P. Perner, "Acquisition of concept description by conceptual clustering," *Perner, P., Imiya, A. (Eds.), Machine Learning and Data Mining in Pattern Recognition. Lecture Notes in Artificial Intelligence*, vol. 3587, pp. 153 – 163, 2005.

[8] O. Georgieva and D. Filev, "Gustafson-kessel algorithm for evolving data stream clustering," in *International Conference on Computer Systems and Technologies - CompSysTech 09*, 2009.

[9] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[10] P. Smets, "Analyzing the combination of conflicting belief functions," *Information Fusion*, vol. 8, pp. 387–412, 2004.

[11] T. Denoeux, "A k-nearest neighbor classification rule based on dempster-shafer theory," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 25 (5), pp. 804 – 813, 1995.

[12] H. Kim and P. H. Swain., "Evidential reasoning approach to multisource-data classification in remote sensing," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25(8), pp. 1257 – 1265, 1995.

[13] B. R. Cobb and P. P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models," *Int. Journal of Approximate Reasoning*, vol. 41, pp. 314–330, 2006.

[14] D. Filev, R. Chinnam, F. Tseng, and P. Baruah, "An industrial strength novelty detection framework for autonomous equipment monitoring and diagnostics," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 4, pp. 767–779, 2010.

[15] P. Angelov, D. Filev, and N. Kasabov, *Evolving Intelligent Systems: Methodology and Applications*, J. Willey and Sons, Eds. IEEE Press Series on Computational Intelligence, 2010.