Information visualisation for industrial process monitoring

Elodie Toufaili, Youssef Miloudi, Christophe Bortolaso, Jean-Marc Petit, Vasile-Marian Scuturici

Information visualisation for industrial process monitoring

Elodie Toufaili[1,2], Youssef Miloudi[2], Christophe Bortolaso[2], Jean-Marc Petit[1], and Vasile-Marian Scuturici[1]

[1]INSA Lyon, Villeurbanne, France
[2]Berger-Levrault, France

Abstract

In the context of process monitoring and predictive maintenance, an adapted visualisation of sensor data is essential in order to help the domain experts to make the right maintenance decision. The large volume and diversity of data leads us to aggregate the data to obtain semantically rich information useful to the domain expert. We study the case of industrial machinery equipped with several sensors producing time series, and we consider that this machinery has different operating states in its operation. We propose a method to identify an optimal representation of the data in 2 dimensions, understandable by the domain expert. This representation allows to easily identify the operating modes of the equipment and the possible deviation from a "normal" behavior. We use co-occurrence matrices to synthesise the time series data, and the features of interest and discretization are selected using two proposed criteria to measure the separation of working modes.

Keywords— Time series, Data visualisation, Co-occurrence matrices

1 Introduction

The use of data in the management of a company's technical assets is becoming increasingly common in industry, particularly with a view to optimising maintenance of equipement. In the case of equipment monitoring and predictive maintenance, the data, often in the form of time series, can be used to anticipate the future failures of the equipment being studied. Behrisch [Carvalho et al.(2019)] provides a literature review of machine learning methods applied to predictive maintenance. This literature review shows the importance of this topic in the Industry 4.0 era, but that there is still work to be done, especially because the existing approaches depend on specific equipement, and it is difficult to generalise these results. It can also be noted that most methods are based on equipment degradation mechanisms. This is notably the case of Cocheteux [Cocheteux et al.(2010)], who uses RULs (Remaining Useful Life) indicators to build his models, or of Vinson [Vinson et al.(2012)], who uses physical and statistical aging laws for the prognosis.

We used datasets corresponding to the monitoring of luggage conveyors in an airport. We have equipped these conveyors with IoT sensors to monitor the evolution of the equipment. The airport needs the conveyors in high availability. This constraint imposes the implementation of preventive maintenance and the use of redundant systems that will take over immediately in case of failure. This solution is expensive, especially as failures are actually quite rare on these systems. In order to reduce the number of preventive maintenance actions, the objective is to propose a predictive solution for possible failures, so that maintenance actions are only carried out when necessary, but still before breakdown.

Usually a dashboard is used in a command and control room displaying data to a domain expert. Our approach is to graphically represent these data in 2D permitting domain experts to observe the equipment behaviour on time. To do this, we propose to explore the use of co-occurrence matrices to represent the behaviour of a piece of equipment. This representation intends to be a tool for domain experts, to easily detect deviations from the usual behaviour, which could indicate a breakdown or a malfunction. We consider it is impossible to graphically represent all the raw data available while obtaining a representation that is readable and interpretable by domain experts. Dimension reduction approaches are not useful in the context where the meaning of the represented data is incomprehensible to the domain expert. The problem lies in the choice of the attributes to be used in the graphical representation, as well as the granularity of the discretization for continuous data. To address this problem, we propose two criteria to guide this choice in order to obtain a representation that helps the expert to visualise the different states of the equipment.

We will first describe the context of this study, including the elements available to us, then we will define what these co-occurrence matrices are, and finally explain our approach based on two criteria to measure the visualisation quality of co-occurrence matrices. These criteria are used to select the most interesting representations for a real world usecase.

2 Motivation and context

We are working on baggage conveyors, for which we want to implement a predictive maintenance strategy. Indeed, these baggage conveyors are subject to strong constraints of availability (100% for certain equipment) and reactivity: the maintenance team is thus reinforced to ensure that the downtime of the conveyors is reduced to a minimum. In addition, the system is oversized, so that if some of the conveyors are stopped, it is still possible to move luggage onto other conveyors. There is also the public dimension of some of the equipment, which is directly visible to passengers. A breakdown in these systems gives a bad image of the airport. Finally, some equipment is difficult to access for certain repairs, and therefore requires more time for the interventions to be carried out. The aim is to follow the evolution of the machines, in order to have a log allowing us to understand their normal functioning, and then to anticipate failures before they put the equipment out of use. This way maintenance can be scheduled, for example at times when the conveyor is not in use. We want to provide a decision support solution for the various maintenance actors, and one of the criteria for the success of this solution will be the confidence of the domain experts. It should also be kept in mind that the system is not intended to replace human expertise.

The conveyors are equipped with IoT sensors that provide us with data on the status of the equipment. We measure the speed of the conveyor belt, the intensity of

the motor and its temperature, and the temperature of the motor oil. From these measurements, we add an attribute corresponding to the difference of the motor and the oil temperature compared to the outside temperature at the time of the measurement, and the temperature difference between the motor and the motor oil.

This data arrives in the form of time series. We have chosen to represent these data by means of co-occurrence matrices in 2 dimensions, the formalisation of which is given in the following paragraph. We assume that there is at least one co-occurrence matrix representation of the discretised time series that allows the phenomena observed on the equipment to be translated by distinct states/groupings of data.

3 Formalisation

3.1 General definitions

Let $\mathcal{U} = \{A_1, ..., A_n\}$ be a set of attributes, and $\mathcal{D}$ an infinite set of real values. We define for all $A \in \mathcal{U}$ the **domain** of A, noted $dom(A) \subseteq \mathcal{D}$, as the set of possible values of the attribute A. For example, for an attribute "Temperature", the domain $dom(Temp)$ can be the interval $[-20; 100]$.

Let $X \subseteq \mathcal{U}, X = \{B_1, ..., B_p\}$. A **tuple** v on X is an element of $dom(B_1) \times ... \times dom(B_p)$. We note $v[B_i]$ the value of v corresponding to the i^{th} component of $dom(B_1) \times ... \times dom(B_p)$, with $i \in [\![1, p]\!]$. By extension, for $Y \subseteq X$, we note $v[Y]$ the projection of v on Y.

Let $X \subseteq \mathcal{U}$. A **time series** TS on X is defined by a sequence of pairs of the form (i, v), where $i \in \mathbb{Z}$ is an index representing the temporal order of the elements of TS and v a tuple on X. It is denoted

$$TS = \langle (1, v_1), ..., (m, v_m) \rangle$$

Let TS be a time series defined on $X \subseteq \mathcal{U}$, and $A \in X$. We denote $TS[A]$ the **projection** of TS on A, defined by $TS[A] = \{(i, v[A]) \mid (i, v) \in TS, i \in \mathbb{Z}\}$. Let $Y \subset X$, we denote $TS[Y]$ the projection of TS on the attributes of Y, i.e.

$$TS[Y] = \{(i, v[Y]) \mid (i, v) \in TS, i \in \mathbb{Z}\}$$

An **univariate time series** corresponds to a time series composed of tuples defined on a single attribute, i.e. $|X| = 1$. When $|X| > 1$, we speak then of **multivariate time series**.

Let $A \in X$ and TS a time series on X. The **active domain** of A in TS, denoted $Adom(A, TS)$, is the set of values taken by A in TS. It is defined by $Adom(A, TS) = \{v[A] \mid (i, v) \in TS, i \in \mathbb{Z}\}$.

We define the **minimum** of A on TS as the minimum of the active domain of A in TS. We denote it $min(A, TS)$. The **maximum** $max(A, TS)$ is defined in a similar way.

We call **discretizer** on an attribute A of a time series TS, an increasing function $D : \mathbb{R} \to \mathbb{Z}$.

An example of a discretizer is the function $D_1 : x \mapsto \lfloor x \rfloor$ that associates its integer part with x.

The discretization of the attribute A of TS by the discretizer D consists in applying D to all the values $v[A]$ of the attribute A in TS. We note $\hat{A}$ the discretized attribute A: $\hat{A} = \{D(v[A]) \mid v[A] \in Adom(A)\}$

Let TS be a time series defined on X such as the attribute $A \in X$ of TS is discretized by D. We define the **data** range of the attribute $\hat{A}$ in TS as the ordered sequence of the integers between $min(\hat{A})$ and $max(\hat{A})$. It is denoted $range(\hat{A}, TS) = \langle min(\hat{A}, TS), min(\hat{A}, TS) + 1, ..., max(\hat{A}, TS) - 1, max(\hat{A}, TS) \rangle$

3.2 Two-dimensional co-occurrence matrices

Let $\mathcal{U}$ be a set of attributes, $X = \{A_1, A_2\}$, $X \subseteq \mathcal{U}$, and TS a time series on X. Let D_1 and D_2 two discretizers applied respectively to the attributes A_1 and A_2 of TS. We denote $\hat{TS}$ the time series from TS whose two attributes have been discretised. We denote $m = \mid range(\hat{A}_1, TS) \mid$ and $n = \mid range(\hat{A}_2, TS) \mid$.

Given $A1$ and $A2$ two attributes and $\hat{TS}$ a discretized time series, the co-occurrence matrix M is a matrix of size $m \times n$, defined by : for all $i \in [\![1, n]\!]$ and $j \in [\![1, m]\!]$, $M(i, j)$ is the number of elements (a, v) of $\hat{TS}$ for which $v[A_1] = range(\hat{A}_1, TS)[i]$ and $v[A_2] = range(\hat{A}_2, TS)[j]$.

3.3 Example

We wish to calculate the co-occurrence matrix of the two-attribute time series of the table 1 :

TS	t1	t2	t3	t4	t5	t6	t7	t8	t9
A_1	0.4	1.2	1.9	1.8	1.3	1.4	2.3	3.9	3.1
A_2	10.1	11	11.7	11.2	12.6	12.8	10.1	10.6	12.1

Table 1: Initial timeserie

We choose as discretizer the function D which associates to each value of the attribute the nearest integer which is lower than it. This gives the table 2.

$\hat{TS}$	t1	t2	t3	t4	t5	t6	t7	t8	t9
$\hat{A}_1$	0	1	1	1	1	1	2	3	3
$\hat{A}_2$	10	11	11	11	12	12	10	10	12

Table 2: Discretized timeserie

The data ranges for $\hat{A}_1$ et $\hat{A}_2$ are therefore :

$$range(\hat{A}_1, TS) = \langle 0, 1, 2, 3 \rangle$$

$$range(\hat{A}_2, TS) = \langle 10, 11, 12 \rangle$$

Then we can calculate the co-occurrence matrix (table 3), by counting the observations with the same discretised values.

The co-occurrence matrices are similar to the 2D histograms and M-histograms presented by Plaud [Plaud et al.(2019)].

		$\hat{A}_2[t]$		
		10	11	12
$\hat{A}_1[t]$	0	1	0	0
	1	0	3	2
	2	1	0	0
	3	1	0	1

Table 3: Co-occurrence matrix.

Each element (i, j) of the matrix thus gives us the number of records in the dataset for which the 2 attributes considered have the same two discretised values. Then, when plotting the matrices, we choose to display the frequency of occurrence of pairs of attributes, instead of the number of occurrences, in order to more easily compare the matrices between them, in case some of them have been calculated on much less data.

The construction of a co-occurrence matrix depends on the choice of certain parameters: the two attributes of the time series and the discretizer used for each attribute. It is possible to build co-occurrence matrices with more than two dimensions, but the graphical representations may involve difficulties of readability. The eventual use of a dimensionality reduction algorithm, like PCA or t-SNE, impacts the understability of the visualisation from the domain expert. We consider here only two-dimensional matrices. The selection of parameters cannot be left to chance, as the representation must allow domain experts to better visualise and understand the normal behaviour of a piece of equipment on the co-occurrence matrices, and thus facilitate the detection of deviations from this behaviour.

4 Quality of the representation

Our objective is to detect how equipment anomalies are reflected in the data aggregated in a co-occurrence matrix. The approach considered is to detect deviations from a behaviour identified as normal. It is therefore necessary to determine how to represent this "normal" behaviour in order to clearly distinguish deviations from it, and thus allow the domain expert to easily identify anomalies.

For a set of n available attributes available, we have a multitude of possibilities for the choice of a two dimensional co-occurrence matrix to present to the user. The objective is to obtain a representation on which it will be easy to distinguish regions representative of the same operating mode of the studied equipment, in order to visualize deviations from normal behaviour. How can we measure the adequacy of the representation of a co-occurrence matrix with this objective?

4.1 Graphical visualisation parameters

4.1.1 Attributes

In a multi-attribute time series, it is necessary to choose carefully which attributes to represent. It can be assumed that not all attributes are useful to the domain expert.

Trying to represent all possible pairs of attributes will be costly and difficult to follow by the human user. It is therefore necessary to determine which attributes work

better together, and which actually provide useful information to the domain expert when represented in a matrix.

4.1.2 Discretization

The choice of discretization will determine the shape of the matrix. If the discretisation is too fine, the level of detail will be too high and there will be a risk of differentiating groups of points when they correspond to the same state. Conversely, a discretisation that is too large will bring together points that do not correspond to the same state.

4.2 Proposed solution

Let $\mathcal{C}$ be a set of classes. A labelled time series corresponds to a time series as described in the previous paragraph, to which for each pair (i, v), we also associate a class $c \in \mathcal{C}$.

$$TS = \langle (1, v_1, c_1), ..., (m, v_m, c_k) \rangle$$

This notion of class, in the field of maintenance, corresponds to the operating mode of the studied equipment, for each element of the time series. It can be, for example, a state of running, of stopping, or an operating mode in general. We are interested here in the distribution of classes in a co-occurrence matrix representation. The idea is to find a best configuration, selection of two attributes and discretization granularity, which optimize the intra-class and interclass distances.

For the choice of parameters, i.e. the attributes to be represented and the discretisation, we take into account some criteria based on the visual separation of the states. The objective is to check that the states are well delimited, to ensure that there are distinct groupings of data representative of the same state of the equipment. We also want to ensure that overlapping points mostly corresponds to the same class.

Let TS be the labelled time series defined on $X \subseteq \mathcal{U}$, whose classes belong to $\mathcal{C}$. The co-occurrence matrix of the class $c \in \mathcal{C}$ corresponds to the co-occurrence matrix calculated on the elements of $\hat{TS}$ whose class is c. It is noted M^c.

We also denote I the set of pairs of indices for which the elements of the co-occurrence matrix are non-zero.

$$I = \{(i, j) \mid i \in [\![1, n]\!], j \in [\![1, m]\!] \text{ and } M(i, j) \neq 0\}$$

4.2.1 Superposition criterion

A first criterion is proposed to determine the purity of the points in the matrix. Indeed, on each element of the co-occurrence matrix, we indicate the number of occurrences of a pair of values in the time series. But among these occurrences, not all are of the same class. Our goal is to find representations where these class collisions are minimal, because the representations with minimal collisions are more understandable by domain expert. To mesure these collisions, for each element of the matrix, we determine the dominant class and compare the number of occurrences of this class in this element to the total number of occurrences.

$$sup(M) = \frac{1}{|I|} \sum_{(i,j) \in I} \frac{\max\limits_{c \in \mathcal{C}}(M^c(i, j))}{M(i, j)} \tag{1}$$

4.2.2 Separation criterion

The second criterion measures the separation of the classes from each other. For this, we are interested, for each point of the matrix, in the classes represented on its 8 direct neighbours.

Let be F the following matrix $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Let $M * F$ be the matrix resulting from the convolution product of M with the matrix F. This convolution gives as result the number of neighbours for each element of M.

For each class represented on a point of the matrix, its number of neighbours of the same class is compared to the total number of its neighbours (Equation 2). The convolution product by the F matrix allows us to explore the 8 direct neighbours of each point of a co-occurrence matrix. For the points located at the edge of this matrix, we will perform the same convolution operation, even if these points have less than 8 neighbours.

$$sep(M) = \frac{1}{|I|} \sum_{\substack{(i,j)\in I \\ c_0 \ s.t. \\ \max_{c\in\mathcal{C}}(M^c(i,j)) = \\ M^{c_0}(i,j)}} \frac{M^{c_0}(i,j)}{\sum_{(i,j)\in I} M(i,j)} \times \frac{(M^{c_0} * F)(i,j)}{(M * F)(i,j)} \tag{2}$$

Each term of the equation corresponds to the ratio between the number of neighbours of the class the most represented on the term (i,j) and the total number of elements in the same neighbourhood. this ratio is weighted by the frequency of the majority class on the considered element of the matrix compared to all the points of the matrix, in order to take into account the importance of the representation of the class on this point: if there are few points of the majority class, the ratio will be minimized compared to an element of the matrix on which the majority class counts more points. If it happens that $(M * F)(i,j) = 0$, we replace $\frac{(M^c * F)(i,j)}{(M*F)(i,j)}$ by 0. This criterion allows us to determine whether the points of different classes are well separated and points are surrounded by points of the same class. If a point has no neighbours, i.e. $(M * F)(i,j) = 0$, then it is considered to be too far from other points of his class.

The values obtained with these two criteria are between 0 and 1, and we are looking for values close to 1 for these two criteria, corresponding to a good separation of the classes and a low overlapping.

5 Experiments

We apply these criteria to a maintenance dataset from luggage conveyors, called here Conveyor dataset, which contains 12 860 records. The time series in this dataset represents measurements such as temperatures or motor intensity. This dataset is composed of 7 numerical attributes and one class attribute, describing the state of the conveyor.

We are interested in the class of each measure, and how these classes are represented. We first use the criteria stated above to determine the two most interesting attributes to represent, among the possible 21 combinations of attributes in this data set, allowing us to visualise groups of distinct points relating to the same class.

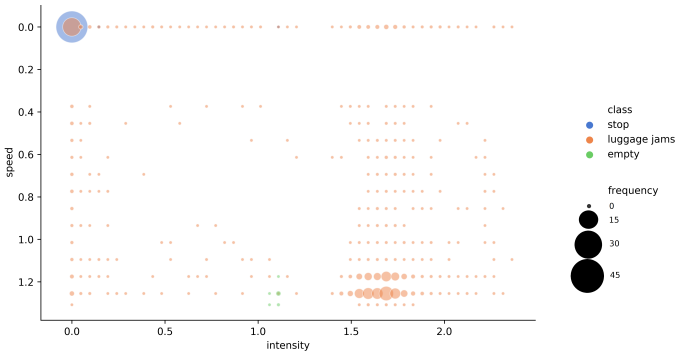

Figure 1: Representation of the co-occurrence matrix of the pair of attributes with the best superposition criteria of the Conveyor dataset (sep = 0.496; sup = 0.998; avg = 0.747).

5.1 Choice of attributes

In the presence of a dataset containing a large number of attributes, it can be very difficult to choose which ones to represent, and which will allow to separate the states of the equipment, and one cannot use dimension reduction methods, such as PCA or t-SNE, which produce visualizations that are difficult for the domain expert to interpret. The criteria detailed in the previous paragraph were therefore first applied to choose the pairs of attributes to be used for the co-occurrence matrices. For each possible pair of attributes, the same discretization was applied, and then its score for the different criteria was calculated.

In fact, the two attributes with the best separation criterion are not necessarily the ones with the best superposition criterion. This is the case for the conveyor data on which we have worked: figure 2 shows the representation with the best separation criterion, whereas figure 1 has the best superposition criterion. We can see that the shape of the data differs a lot on these 2 figures, on the figure 2 the criteria reflect that the classes are well separated from each other, and on the figure 1 the superposed points are mostly of the same class. In order to help us choose which attributes to select, we decide to compute the average of the two criteria for all pairs of attributes. We therefore choose the two attributes shown in figure 3. This choice allows us to maximise both the superposition and separation criteria on the same figure. However, in this case, the two attributes with the worst superposition criterion are also those with the worst separation criterion. They are shown in Figure 4. Indeed, as these two attributes are proportional, all the data are located on the same line, and it is impossible to distinguish the classes.

5.2 Choice of discretization

The discretisation method used here consists of dividing the interval comprising the values of the time series into a number of intervals of equal size. The number of intervals for each attribute is determined empirically, to get an idea of how the criteria evolve with the applied discretisation. We applied the scoring criteria for the pair of attributes with the best scores for the average of the two criteria, over several different

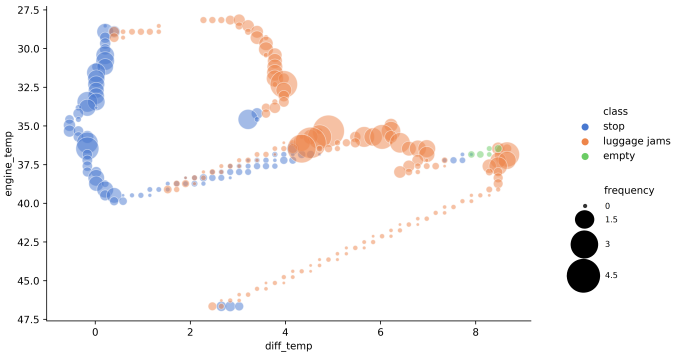

Figure 2: Representation of the co-occurrence matrix, discretized in 50 intervals, of the pair of attributes with the best separation criteria for of Conveyor dataset (sep = 0.927; sup = 0.976; avg = 0.951).

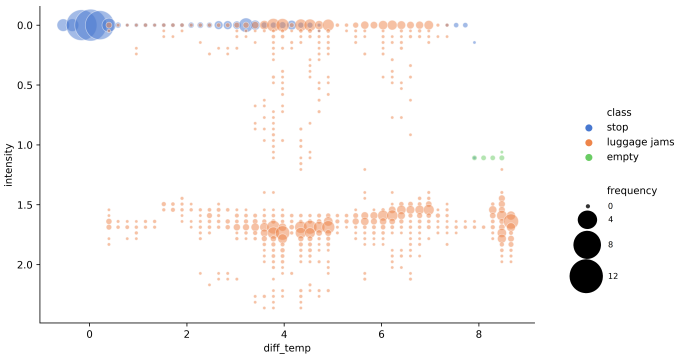

Figure 3: Representation of the co-occurrence matrix, discretized in 50 intervals, of the pair of attributes with the best average for the two criteria of the Conveyor dataset (sep = 0.924; sup = 0.987; avg = 0.956).

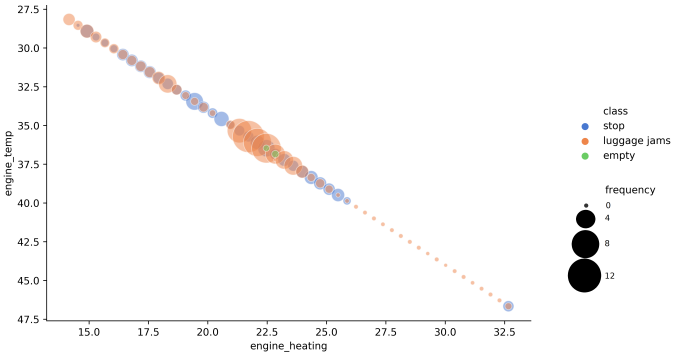

Figure 4: Representation of the co-occurrence matrix, discretized in 50 intervals, of the pair of attributes with the worst criteria of the Conveyor dataset (sep = 0.533; sup = 0.82; avg = 0.677).

discretisations. While the separation criterion allows us to choose a discretization that is neither too fine nor too large (the one shown in the figure 7), this is not the case for the superposition criterion. Indeed, as we can see in figure 5, this criterion varies very little with the discretization, and tends to privilege the finest discretizations. However, this is not what we want.

Inspired by the decision tree construction algorithms, we decide to use the gain of this criterion to make a discretization choice. Table 4 shows the results of the gain for each discretization intervalls we tested.

It can be seen that the largest discretization for which the absolute term of the gain reaches zero is the 50-interval discretization, so the 40-interval discretization (figure 6) can be considered sufficient for this criterion. In order to take into account the result of the separation criterion as well, in order to determine the optimal discretization for this dataset, the results of the discretizations in 40, 50 and 100 intervals are compared. For 100, which obtains the best score for this criterion, the result is 0.9245 (figure 7), for 50, the result is 0.9238, and for 40 it is 0.9126. Taking into account that the 100 and 50 discretisations have a very close separation criterion, and that the gain of the superposition criterion between 40 and 50 is null, we can consider that the optimal discretisation is the discretisation in 50 intervals (figure 3). For comparision, figure 8 shows a very large discretization : we can see that it is very difficult to distinguish the classes on this representation.

5.3 Comparison with Gini Impurity and Entropy

The Gini index and entropy are used in the process of building a decision tree to measure the purity of a set. We applied them on each elements of the co-occurrence matrices, to compare with the results of the superposition criterion. But as for the Gini indexes and the entropy we try to get closer to 0, while for the superposition we want values closer to 1, we take into account the value of $1 - sup(M)$. Thus, the Gini index is quite similar to the superposition criterion, due to the fact that we use the probability of class in the calculation, so the results for selecting the pair of attributes are the same (see Figure 9). As before, we used the gain for the Gini index and the entropy to determine the smallest discretization for which the gain becomes zero. We

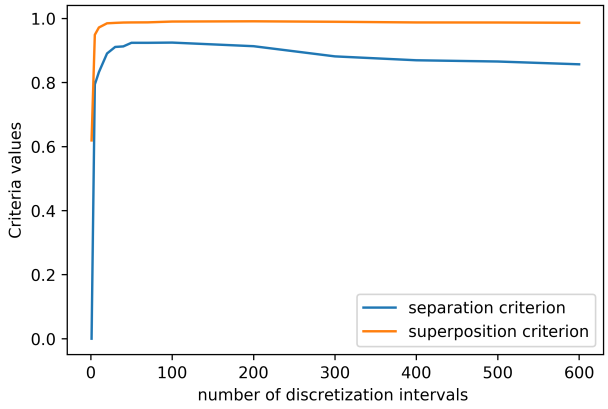

Figure 5: Evolution of the two criteria as a function of the discretization for the Conveyor dataset, for attributes *intensity* and *diff temp*.

Discretization	Gain (absolute term)
1	x
5	0.33
10	0.023
20	0.013
30	0.001
40	0.001
50	0
70	0
100	0.003
200	0
300	0.001
400	0.002
500	0
600	0.001

Table 4: Gain in absolute terms of the superposition criterion as function of the discretization, rounded to the thousandth, for attributes *intensity* and *diff temp* of the Conveyor dataset.

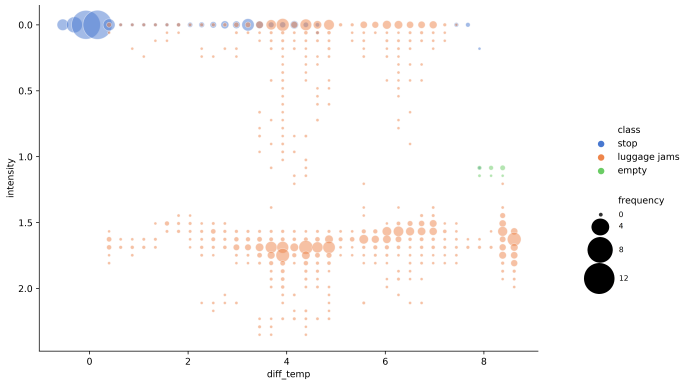

Figure 6: Representation of the co-occurrence matrix discretised into 40 intervals of the highest scoring attributes of the Conveyor dataset (sep = 0.913; sup = 0.987).

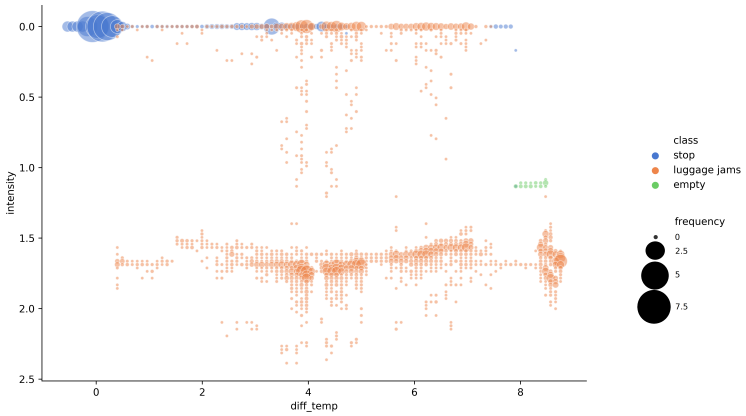

Figure 7: Representation of the co-occurrence matrix discretised into 100 intervals of the highest scoring attributes of the Conveyor dataset (sep = 0.924; sup = 0.99).

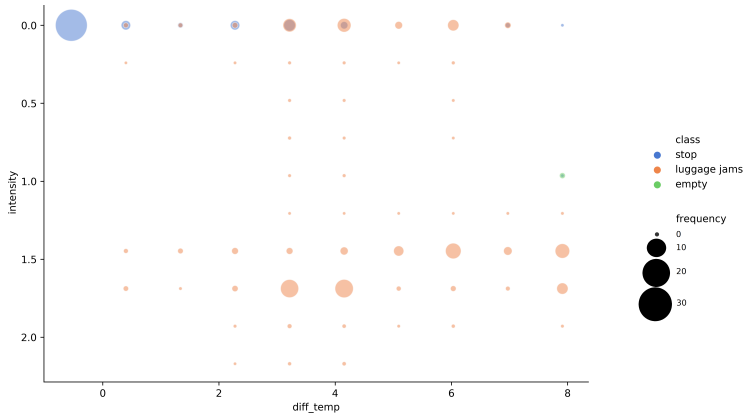

Figure 8: Representation of the co-occurrence matrix discretised into 10 intervals of the highest scoring attributes of the Conveyor dataset (sep = 0.832; sup = 0.972).

obtain the same result as with the superposition criterion, i.e. the discretisation into 50 intervals.

5.4 Validation of the results

It is difficult to automatically evaluate the results obtained above precisely. Indeed, the main objective is to make it easier for domain experts to read their data on these co-occurrence matrices and to understand their equipment. We carried out a study in the form of a questionnaire to evaluate whether the representation in co-occurrence matrices meets the requirements we have set. The questionnaire include a number of questions to validate their understanding of these representations, as this is one of the initial objectives, and then confront several representations in multiple choice questions in order to gather the users' opinions. We first ask a series of questions in which they have to choose between two representations of co-occurrence matrices with different couple of attributes. Then the questions focus on the variation of discretization.

We have so far had 11 responses from domain experts. It can already be noted that in the case where only one of the two representations has both the maximum overlap and separation criteria, the majority of respondents chose this representation as the best. For discretization variations, the results are less easy to highlight, we will need more data.

6 State of the art

We have chosen to answer the problem by means of a representation of multivariate time series from the equipment studied. Indeed, as the end users of our work will be technicians working on the equipment, we wish to offer them a solution that is simple and quickly interpretable. One of the challenges is to represent large volumes of data in a very short period of time. Tufte [Edward R. Tufte(2001)] outlines a set of good practices for achieving a representation of data that is easily understandable, without

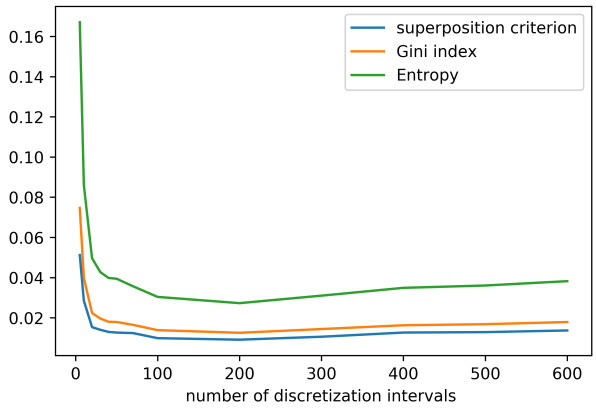

Figure 9: Evolution of Gini index, Entropy and superposition criterion.

distorting reality. For example, he introduces the "data-to-ink ratio", which measures the proportion of a graph that truly and uniquely represents the data, without redundancy, compared to the totality of what is represented on the graph. Peng [Peng(2008)] proposes a method for representing large multivariate time series in two dimensions in a single graph, by discretising the values into 3 categories: "low", "medium" and "high". These representations are very complete, but lack visibility and are therefore difficult to interpret. More recently, Plaud [Plaud et al.(2019)] uses M-histograms which make it possible to visualise the frequency of appearance of the elements of the temporal series in the same interval. These M-histograms are used for the purpose of time series classification, not for subsequent representation. In continuation of this work, co-occurrence matrices, computed from discretized data, are used to represent data from sensor on helicopters [Chazelle et al.(2021)]. This method requires a reduction in dimension, especially in the choice of the elements to be represented. We provided a method to help with the dimension reduction.

In order to visualise the different states of a piece of equipment, it is necessary to find a representation that presents groupings of points that are representative of the same state of the equipment. We have chosen to assume that such a representation exists. In order to determine which attributes to select and which discretization to apply, we have determined criteria that allow us to guide the choices among all the possibilities. According to Behrisch et al. [Behrisch et al.(2018)], the scale of the axes influences the visibility of clusters. In our case, it is the discretization that will influence this parameter. Sedlmair and al. [Sedlmair et al.(2012)] presents a classification of factors that affect the visual separation of clusters, for point cloud representations. The paper also provides a rather critical evaluation of the performance of the [Tatu et al.(2009)] and [Sips et al.(2009)] methods in comparison to human judgement. The first is based on grids, and the second proposes to use the distances of the points from the centroids of each class to evaluate the representations. Indeed, for both methods, the observed results are very far from the judgement made by a human on the given examples.

7 Conclusions

We have established tools for choosing the parameters that will allow us to propose a "best" 2D representation of a multi-variate time series so that each class is well distinguished from the others in the representation. Once the parameters have been selected from labelled data, we can distinguish areas of the figure in which the points relating to the same state are grouped together. Then, we can apply these same parameters to co-occurrence matrices calculated from unlabelled data, in order to analyse the behaviour of the equipment with respect to the zones delimited in the previous step. If points deviate too much from these zones, they can be considered as potentially related to an equipment malfunction, and generate an intervention, or at least a verification by the technicians working with these tools. We carried out a study comparing the results obtained with the criteria with the assessments of a group of people. Indeed, the end-users of the co-occurrence matrix representations will be the technicians who work on the maintenance of the equipment, so it is important to ensure that these criteria visually match their perception.

These tools aim to help maintenance technicians and engineers to monitor their equipments by providing a reference of the normal behavior of the equipments, in the shape of co-occurrence matrices with labels, like the examples in this paper, which can then be compared to the shape of the co-occurrence matrices of data with no labels, coming as it happens. This will allow to detect deviations from the normal behavior and lead to vigilance over the equipment. In a future work, we can add an automated detection of failures, based on these cooccurrence matrices, to create a system that helps technicians to detect the changes in the equipment behavior.

The authors would like to thank Berger-Levrault for the funding of the study project conducted by Elodie Toufaili.

References

[Behrisch et al.(2018)] M Behrisch et al. 2018. Quality Metrics for Information Visualization. *Computer Graphics Forum* (2018), 38.

[Carvalho et al.(2019)] Thyago P. Carvalho et al. 2019. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering* 137 (Nov. 2019).

[Chazelle et al.(2021)] Benjamin Chazelle, Pierre-Loic Maisonneuve, Ammar Mechouche, Jean-Marc Petit, and Vasile-Marian Scuturici. 2021. From Large Time Series to Patterns Movies: Application to Airbus Helicopters Flight Data. In *Advances in Databases and Information Systems*. Vol. 12843. Cham, 213–226.

[Cocheteux et al.(2010)] Alexandre Cocheteux, Pierre Voisin, Eric Levrat, and Benoît Iung. 2010. System performance prognostic: context, issues and requirements. *1st IFAC Workshop on Advanced Maintenance Engineering, Services and Technology* (2010).

[Edward R. Tufte(2001)] Edward R. Tufte. 2001. *The visual display of quantitative information* (second edition ed.). Graphics Press USA.

[Peng(2008)] Roger D. Peng. 2008. A Method for Visualizing Multivariate Time Series Data. *Journal of Statistical Software* 25, Code Snippet 1 (2008).

[Plaud et al.(2019)] Angéline Plaud, Mephu Nguifo Engelbert, and Jacques Charreyron. 2019. Classification des séries temporelles multivariées par l'usage de Mgrams. *CAP 2019* (2019).

[Sedlmair et al.(2012)] M. Sedlmair, A. Tatu, T. Munzner, and M. Tory. 2012. A Taxonomy of Visual Cluster Separation Factors. *Computer Graphics Forum* 31, 3pt4 (June 2012), 1335–1344.

[Sips et al.(2009)] Mike Sips, Boris Neubert, John P. Lewis, and Pat Hanrahan. 2009. Selecting good views of high-dimensional data using class consistency. *Computer Graphics Forum* 28, 3 (June 2009), 831–838.

[Tatu et al.(2009)] Andrada Tatu et al. 2009. Combining automated analysis and visualization techniques for effective exploration of high-dimensional data. In *2009 IEEE Symposium on Visual Analytics Science and Technology*. IEEE, Atlantic City, NJ, USA, 59–66.

[Vinson et al.(2012)] Garance Vinson, Michel Combacau, and Thomas Prado. 2012. Synchronous machine faults detection and diagnosis for electromechanical actuators in aeronotics. *38th Annual Conference of IEEE Industrial Electronics (IECON)* (2012).