Fault diagnosis in dynamic processes: a data mining and SVM application

Addison Ríos-Bolívar, Francisco Hidrobo, Pablo Guillén, Francklin Rivas

***Abstract*—** The major inconvenient for the fault detection and isolation (FDI) in technical processes based on analytical redundancy is the requirement of a very accurate model of the system. By contrary, in the methods based on data handling it is not required of a precise model, since they are based on the manipulation of the information by means of the measured data. Thus, in this work a set of statistical indices allowing quantifying the amount of information contained in the collected data is presented, in order to realize the FDI. These indices are used for the reconstruction of the fault patterns, and next for its classification is used the machine learning technique, in particular the support vector machines (SVM). For verification of the results, generated data by two nonlinear models are used; one of discreet time that simulates the Logistic Application, which is used under different types of dynamic states behavior that represent the occurrence of faults. The other model is a continuous time that represents the control of a magnetic levitation system. For the first model, the results show that by means of the obtained statistical indices the reconstruction of fault patterns is obtained, which allow separating the different dynamic behaviors (faults) and using a SVM considering different kernel, a classification between the faults is obtained (classes). For the second model, the classification of the faults by means of a SVM is realized, obtaining one diagnosis index.

***Keywords*—**Fault diagnosis, Data mining, Operational classification, Machine learning, Support Vector Machines.

I. Introduction

In the processes of industrial production, the safety and operational reliability must be guaranteed by means of the correct operation of the processes, by the associate control systems and the coordination between them. In the context of a reliable and safe operation, some systems are developed in order to allow the events recognition, which must orient the decision making when the performance of the productive process is affected by the presence of any abnormality.

As the reliability is highly related to the security concept, then, it is fundamental to equip the industrial processes with demanding safety mechanisms, whose basic elements are the Monitoring, Detection, and Diagnosis (MDD) systems, which, by means of the indicators and the measured variables of the processes, maintain a continuous and constant supervision of the evolutionary behavior in the production time, reporting symptomatic conditions that are considered abnormal. The MDD systems are based on their capacity to respond under unexpected situations of the process behavior, so that its main task is the one of the FDI. A FDI system, such as is shown in Figure 1, uses the measurements of the process in order to produce residues, from which, by means of evaluation functions and logic decisions, it looks for the faults identification and the isolation, with perspective of prognosis and autonomic maintenance [11].

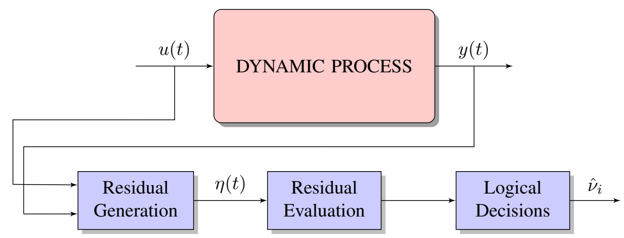

Fig. 1. FDI System

From the point of view of the residual generation by comparison, the FDI filters design techniques can be classified in

1) *Physical Redundancy*: in which physical duplicates of the devices and systems under study are used. The residues are obtained by comparison of the outputs of the different elements. These techniques have the main disadvantage of the costs involved for its implementation and maintenance.
2) *Model based Methods*: from which they produce estimated values of the process outputs for the generation of the residuals, by means of the comparison with measured real outputs. The main disadvantage of this technique is concerning the construction of very precise models.

When for the residual generation some models are used, it is necessary to consider aspects important of the technical processes [14], [15]:

- Noisy industrial environments (Disturbances).
- Approximate models (Uncertainties).
- Different structures for the production.
- Physical limits of production.

Those aspects entail to consider that the FDI filters must be able to respond satisfactorily under adverse conditions due to the reality of the productive processes [16], [17].

Financial supports for this research from CDCHTA Universidad de Los Andes, under the project No. I-1268-11-02-A, are gratefully acknowledged.
A. Ríos-Bolívar, P. Guillén, and F. Rivas are with Los Andes University, Sector La Hechicera, Facultad de Ingeniería, Escuela de Ingeniería de Sistemas, Mérida, Estado Mérida, Venezuela, Tlf. +58-274-2402811 / 3336, Fax: +58-274-2402811, E-mail: ilich@ula.ve, pguillen@ula.ve, rivas@ula.ve
F.Hidrobo is with Los Andes University, Sector La Hechicera, Facultad de Ciencias, Departamento de Física, Mérida, Estado Mérida, Venezuela, Tlf. +58-274-2401284, Fax: +58-274-2402811, E-mail: hidrobo@ula.ve

Thus, the models based on knowledge are based more on the information or data available, which entails to have mechanisms of handling of data: data reconciliation, conventional numerical methods, machine learning, among others. These methods based on historical data have the strength of which they are possible to be integrated of complementary way to object to extract qualitative and quantitative characteristics, that allow to the detection and the diagnosis of faults [1], [6], [8], [12], [19].

Following this context, the taxonomy for the residual generation can be extended taking into account methods based on the handling of data (see Figure 2).

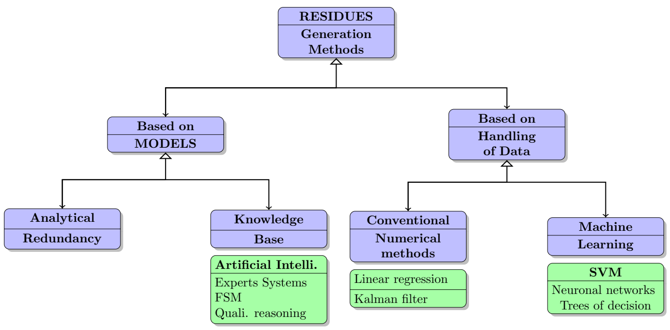

Fig. 2. A taxonomy for residual generation.

Thus, there is necessity to develop mechanisms for design and implantation of FDI systems in productive processes, from residual generation filters that consider environments of real industrial production. Then, in this work a set of statistics is presented corresponding to fault patterns, each one of as they measure different characteristics from the measured data, quantifying the dynamic activity inherent to the presence of faults, which allows his location for a greater exactitude, according to the MDD techniques. A SVM, considering different Kernel, is used to classify the reconstruction of the fault patterns, then, the potentiality of the use of the SVM in the recognition of anomalies is verified, and next a classification algorithm based on decision trees is used from the values of the obtained statistical indices and it verifies that the decision tree can identify with very good precision and exactitude the faults, determining the patterns of the operation conditions.

II. FDI filters

With reference to the analytical methods, the dynamic systems can be described by models that enter within two categories: the representative models and the diagnosis models. The representative models allow to describe the dynamic behavior of the systems in terms of a structure standardized that, of satisfactory way, comes near to the behavior input-output or the behavior of state of the systems, in them are the heuristic models. On the contrary, the diagnosis models allow to describe to the dynamic behavior through an retort of the structure or physical architecture of the systems. There, the basic functional units, or of interest, are modeled in explicit form: models of sensors, actuators, etc. To the aims to design the filters of FDI, the diagnosis models are most useful, whereas for control intentions, the representative models are used.

A. FDI filters based-on state observers

Since in this analytical technique of design of FDI filters a mathematical model is fundamental, we consider the dynamics of a linear, continuous and invariant system in the time, LTI,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t) \end{aligned} \tag{1}$$

where $x \in \mathfrak{X} \subset \mathbb{R}^n$ are the states, $u \in \mathfrak{U} \subset \mathbb{R}^m$ are the control signals, $y \in \mathfrak{Y} \subset \mathbb{R}^q$ are the outputs ; and the matrix $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{q \times n}$ are known. There, the pair (A, C) is observable.

In the description of the system (1) is easy to distinguish three subsystems: the structure of the process represented by A; the actuators that are described by B; and the sensors represented by C. In anyone of those subsystems inadequate situations of behavior with respect to the established characteristics of performance in the design can appear. So that the monitoring schemes are due to construct that generate the residuals based on the possible faults in each one of the functional units. The design of FDI filters can be divided is two stages: first stage is the generation of the residuals, (the detection problem). The second stage is the evaluation of the residuals to object to determine the origin of the faults, (the problem of separation of the faults). The residuals take place when comparing the exit considered with the measured output of the physical plant. Thus, for the system (1) exists a gain matrix $D \in \mathbb{R}^{n \times q}$ so that estimated state vector $\hat{x}(t)$ of the state vector $x(t)$ will be the solution for the equation of the observer of complete order:

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + D(y(t) - C\hat{x}(t)), \\ \hat{y}(t) &= C\hat{x}(t). \end{aligned} \tag{2}$$

There, $\hat{y}(t)$ are the estimated outputs and D the feedback gain of the observer, which should be select suitably. That is, D is selected such that $(A - DC)$ is stable, and the estimation error will be null, $e(t) = 0$. Since for $t < t_0$, the process has a normal operation, that is, faults do not exist, then the residual signal is approximately equal zero. When any fault becomes present, in $t \geq t_0$, the residual is different from zero and it is propitious for the fault detection. We can observe, from the different representations from the faults, that the same can be described like additional entrances in the dynamics of the process, in addition, of the susceptibility to the existence of different faults in a same subsystem, in future we will adopt the following model of diagnosis for the systems put under faults:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + \sum_{i=1}^{k} L_i \nu_i(t), \quad x(0) = x_0 \\ y(t) &= Cx(t) + \sum_{i=1}^{k} M_i \nu_i(t), \end{aligned} \tag{3}$$

where: L_i, M_i are the fault directions in the sensors and actuators, respectively. One assumes that these directions are known. In the future, one assumes that the directions of faults are linearly independent.

$\nu_i \in \mathfrak{V}_i$ corresponds to the fault mode. It is an independent, arbitrary, and unknown function. In addition, $\nu_i = 0$ if $t < t_0$; and $\nu_i \neq 0$ for $t \geq t_0$. k is the number of faults that functional and statistically are significant. Under this representation, where multiple possible faults exist, with different directions, the fault isolation problem becomes evident, we must construct a generator of residuals that allows us to distinguish to what direction of the state space of (3) the fault belongs that has become present. In the case of filters based on observers, we must construct a gain of the observer so that the vector of residuals, that is, the output of the estimation error, has characteristics in a single direction associated with some direction of well-known fault, the result is the isolation of the faults in the output space. As the important information is in the fault direction, it is not required of the knowledge of the fault mode [13], [17].

We consider an observer of states like in (2), for the system (3), then, the dynamics of the estimation error will be governed by

$$\begin{aligned} \dot{e}(t) &= (A - DC)e(t) + \sum_{i=1}^{k} (L_i - DM_i)\, \nu_i(t) \\ \eta(t) &= Ce(t) + \sum_{i=1}^{k} M_i \nu_i(t) \end{aligned} \tag{4}$$

with $e(0) = x_0 - \hat{x}_0$.

If D is selected so that $(A - DC)$ is stable, and if $\nu_i(t) \neq 0$, then $e(t) \neq 0$, therefore take place the residuals, since $\eta(t) \neq 0$. Any change in the process, as a result of faults, will be accentuated in the innovation in the observer output, in this way completes the phase of residual generation. On the other hand, due to the properties of the observers, the uncoupling in the initial conditions does not have major effect. At this moment, if the unique condition that prevails for the selection of the matrix of gain of the observer, D, is that $(A - DC)$ is stable, is not in capacity to establish a clear distinction between the effects of the different faults. In principle, it would be possible to be thought about the design of a set of observers, each one of as it is made correspond with a direction of specific fault, which suggests the design of different gains which is not nor practical, nor elegant feasible. The idea is to construct a unique filter of FDI. Of previously exposed two important questions stand out: when a fault is detectable?, the detection problem, and when the faults are separable?, the diagnosis problem. If $M_i = 0$, the conditions ensuring that faults are detectable and separable are given by, [11]:

$$\ker(CL_i) = 0, \qquad i = 1, \ldots, k \tag{5}$$

$$Im(CL_i) \bigcap \left(\textstyle\sum_{i,j=1}^{k}{}_{j \neq i} Im(CL_j) \right) = 0, \tag{6}$$

which, in the technical processes, are difficult to satisfy, so that the generation of residues by means of the handling of measured data is necessary [18].

In order to generate the residues, some knowledge based models can be used, that are based on the information or historical data, which entails to have mechanisms of data handling: data mining, data reconciliation, conventional numerical methods, machine learning, among others [6], [12], [19]. These methods based on historical data have the strength of being possible to be integrated in a complementary way to objects in order to extract qualitative and quantitative characteristics, that allow the faults detection and diagnosis [1], [8], [18].

In that context, in this work a set of statistics is presented corresponding to fault patterns, each one of them measures different characteristics from the measured data, quantifying the dynamic activity inherent to the presence of faults, which allows its location for a greater exactitude, according to the MDD techniques. A SVM, considering different Kernel, is used for classifying the reconstruction of the fault patterns, then, the potentiality of the use of the SVM in the recognition of anomalies is verified, and next a classification algorithm based on decision trees is used from the obtained statistical indices values [10], and it verifies that the decision tree can identify with very good precision and exactitude the faults, determining the patterns of the operational conditions.

III. Methodology based on data mining

A. Statistical indices

In machine learning applications, for the information or characteristics extraction, some statistical indices are required, such that certain information zones can be classified. In our case, two indices are used.

Curve size: This characteristic is useful for knowing the stability of the signal values. If in an interval, the value of this characteristic is "low" indicates that the signal is stable, in the other case, the signal is unstable. The equation (7) defines how index can be measured:

$$L = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{7}$$

where each x_i corresponds to a sample of the data set $X = (x_1, x_2, \cdots, x_N)$.

Threshold: the determination of the threshold γ is based on the calculation of the data deviation, in order to know how dispersed they are inside of a data window of size . The threshold is determined by means of:

$$\gamma = \frac{3}{N-1}\sqrt{\sum_{i=1}^{N}(x-\bar{X})} \quad (8)$$

where $\bar{X}$ is the average of the data set. In the context of fault detection, this threshold allows minimizing false alarms.

B. Support Vector Machines

The theory of the SVM was developed by Vapnik [1], [2]. It is based on the idea of minimization of the structural risk (inductive principle SRM). First, the SVM maps the input points to a space of characteristics, which will have a greater dimension (for example: if the input points are in R^2 then these points are mapped to R^3 by the SVM), immediately an hyperplane is found in order to separate the characteristic points by maximization of the m margin between the classes, see Figure 3

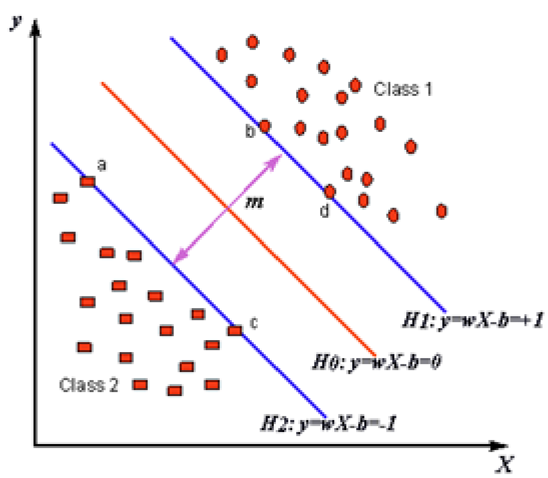

Fig. 3. Maximization of margin m

As a result of the maximization of m margin, the class separation problem is transformed into a quadratic programming problem, which can be solved by its dual problem introducing Lagrangian multipliers [2], [3]. The SVM can find the optimal hyperplane using the dot product with functions (kernels) in the space of characteristics. The support vectors are combinations of a few input points that allow defining the solution of the hyperplane in a simpler way. Thus, let us consider the case with:

- a set of N points of training data $\{(X_1, y_1), \cdots, (X_1, y_1)\}$.
- a hyperplane

$$H_0 : y = wX - b = 0 \quad (9)$$

  where w is normal with respect to the hyperplane, $b/||w||$ is the perpendicular distance to the origin and $||w||$ is the Euclidean norm of w.
- two parallel hyperplanes to H_0:

$$H_1 : y = wX - b = 1 \quad (10)$$

$$H_2 : y = wX - b = -1 \quad (11)$$

  with the condition that does not exist data points between H_1 and H_2.

This situation is illustrated in the Figure 3. If the distance $d_+(d_-)$ corresponds to the smallest separation from the separation of the hyperplane H_0 to the point nearest positive (negative), where the hyperplane $H_1(H_2)$ is located, then the distance between the H_1 and H_2 planes is $d_+ + d_-$. Thus, $d_+ = d_= 1/||w||$, then the margin is equal to $2/||w||$. The problem is to find the hyperplane that gives the maximum margin. The parameters w and b are called weight and slant vectors, respectively. Thus, the optimization problem is defined by the following equation:

$$\min_{w,b} \frac{1}{2} w^t w \ restricted\ by: \ \gamma_i(wX - b) \geq 1 \quad (12)$$

The optimization problem presented in the previous equation can be declared as convex and quadratic optimization problem in (w, b), into a convex set. Using the Lagrangian formulation, the limitations can be replaced by limitations of Lagrangian multipliers in themselves. Additionally, introducing the Lagrangian multipliers in this reformulation, like a consequence of the training data, it could only appear as the dot product between the data vectors, $\alpha_1, \cdots, \alpha_N \geq 0$. Therefore, a Lagrange function for the optimization problem can be defined as:

$$L_p(w, b, \alpha) = \frac{1}{2} w^t w - \sum_i (\alpha_i \gamma_i (wX - b) - \alpha_i) \quad (13)$$

Using the dual formulation and the limitations of the Lagrange optimization problem, the parameters α_i can be calculated and the parameters w and b, which specify the separation of the hyperplane, can be calculated using the following equations:

$$w = \sum \alpha_i \gamma_i X_i \quad (14)$$

$$\alpha_i(\gamma_i(wX + b) - 1) = 0; \forall i \quad (15)$$

According to (14), the parameters α_i that are not equal to zero correspond to X_i, Y_i data, which are the support vectors for b, c and d (see Figure 3). If the separation surface of two classes is nonlinear, the data can be transformed into another characteristic space with higher dimension, then the problem is linearly separable.

If the transformation to a greater dimensional space is $\phi()$, then the Lagrangian function can be expressed as:

$$L_D = \sum a_i - \frac{1}{2} \sum_{i,j} a_i a_j \gamma_i \gamma_j \phi(X_i)\phi(X_j) \quad (16)$$

The scalar product $\phi(X_i)\phi(X_j)$, in the space of greater dimension, defines a kernel function $k(X_i, X_j)$ and therefore this is not necessarily explicit from the transformation $\phi()$, as long as the kernel function, corresponding to the dot product, is known in the same characteristic space of higher dimension. From an appropriate kernel, the SVM can separate, in the characteristic space, the data that in the original inputs space are not separable. There are some kernel functions that can be used, for example:

Gaussian kernel function of radial base

$$k(X_i, X_j) = e^{-\gamma\|X_i - X_j\|^2} \quad (17)$$

Polynomial kernel

$$k(X_i, X_j) = (X_i X_j + m)^p \quad (18)$$

A kernel function has a good efficiency if the support vectors that were calculated using the corresponding transformation are few and if the classification of the testing data is sufficient [9].

In summary, with the purpose of separating a data set, a training data set (X, Y) is selected, the optimization problem is solved and the α_i, w and b parameters are calculated. Then, a X data vector given of the initial data set is classified according to the value from $(wX^* + b)$.

The efficiency of the calculated support vectors is proven using the testing data set derived from the initial data set.

In principle, the SVM have been proposed for binary classification. Then, this method has been extended to multiclass classifications, in which binary combinations of SVM are used in order to consider all the classes simultaneously, also there are SVM with multiclass classification. One of the methods for approaching the multiclass problem is the denominated "one against one", in which each class is compared with the rest of separated way, this method entails a total of $K(K-1)/2$ classifiers, and it is applied in this work.

IV. Fault Diagnosis using SVM

First of all, it is possible to be combined the potentialities of the filters based on state observers and the SMV. The technique consists of the generation of residues, for the fault detection problem, from filter based on state observer, which entails to that the condition given by (5) is satisfied. Then, for the fault isolation problem, the residues are process for recognition and classification of fault patterns using SVM. This has the advantage that is possible to recognize, immediately, the presence of faults. In addition, it is not necessary to satisfy the hard condition of fault separability given by (6). In a schematic way, the Figure 4 shows the different elements to construct to obtain this FDI system.

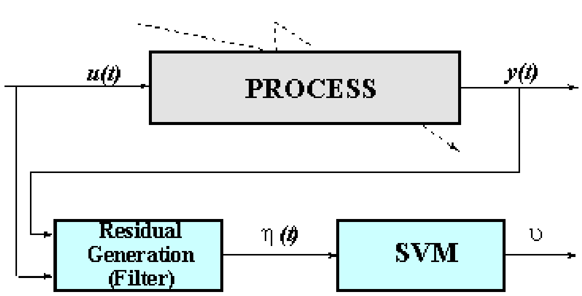

Fig. 4. A model for FDI based on SVM.

If the fault detectability condition given by (5) is not satisfied, as it happens in many technical processes, mainly of nature nonlinear, becomes the generation of the residues by means of observers complicated. Then, it is possible to resort to the SVM to reconstruct certain fault patterns from the inputs and outputs of the process, to see the Figure 5.

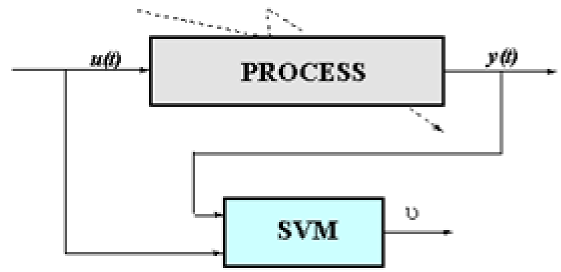

Fig. 5. Fault diagnosis with SVM

Thus, the fault diagnosis by means of the data handling consists, first of all, in the information extraction of the measured data, from which the fault detection problem can be solved. For this, statistical indices are used. Secondly, in order to solve the fault separation problem, some classification techniques are used, that in this work correspond to SVM. Thus, in the technical processes the operative conditions according to faults presence or not, can be distinguished. Consequently, through distinguishing the status or the operational condition, the faults presence can be associated.

This is the case that will be presented next, where the characterization of operative conditions, represented as faults, for a dynamic process denominated logistic application, is obtained, in order to evaluate the previously shown techniques as tools for fault diagnosis.

A. Logistic application

The logistic application or logistic map is mainly applied to problems of population growth and has the following expression:

$$x_{n+1} = r * x_n * (1 - x_n) \quad (19)$$

Where x_{n+1} (dependent variable) represents the population in the $n+1$, stage or generation, of the related species, x_n (independent variable) symbolizes the population in n stage, and r is the rate of growth that will depend on the environmental conditions, climatic, nourishing, etc.

One of the properties that the logistic maps present is the arrival to a stationary state, fixed or attractor, as usually it is called, which is characterized for being equal for the dependant variable as for the independent variable, this means, that the population stays invariable. If r is increased, a state with two attractors appears, and if r continues being increased, the number of attractors or attractor cycles begins to be duplicated. For example, when r is 3.5, the cycle will be of four attractors because the previous one had two attractors. As a result of the duplication of the cycle, the separation or distances between the points (values of r) becomes decreasing, but the relation between the separation distance, for two consecutive cycles, and the previous analogous distance stays constant. This is other of the interesting properties that presents iterative maps like the logistics. The graph that turn out to take in the X-axis the r values in each bifurcation, and in the ordinates one (Y-axis), the values of the attractors is known as Feigenbaum Diagram. This way, for the abscissa $r = 2$, corresponds the ordinate 0.5. From the point $(2, 0.5)$ a parallel line to the X-axis draws up until arriving to the beginning point of the cycle of two attractors. For each of these two attractors, a parallel line to the X-axis that will arrive until the point of the following bifurcation is made. It will be completed satisfied when in the graph, the figure of a parallel bracket or the X-axis is found. When arriving at the r value of the following duplication, each branch of the bracket will be branched off, and so on they will be formed small and more and more nearer brackets to each other.

As it is possible to be observed in Figure 6, when r goes from 1 to 3, there is a single attractor (simple attractor), from $r = 3$ until $r = 3.5$, there are two attractors (double attractor), and finally, when r has a value very near 4, no longer appear repetitions or attractor cycles, that is, the process has lost regularity and also appears the fact that very small variations of the initial value of the independent variable, generates remarkable variations of the values that are being obtained. This situation, where the regularity does not exist and the great sensitivity to the variations of the initial conditions is present, is well-known as chaos.

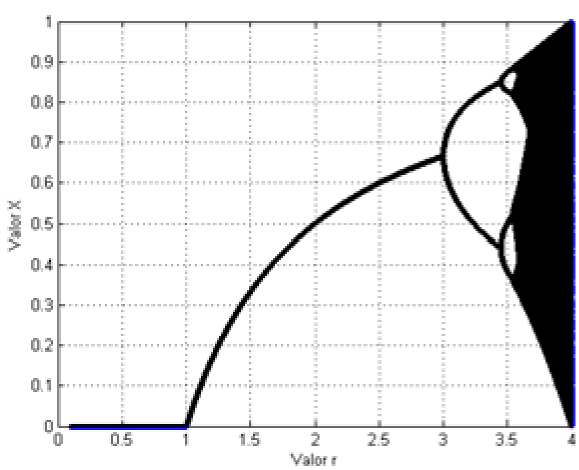

Fig. 6. Feigenbaum Diagram for $x_0 = 0.3$

According to this, in a logistic map we can distinguish three zones (operational conditions), depending on the attractors number (states). These zones are presented in the Table I.

The logistic application or Feigenbaum diagram was generated with $N = 1000$ samples. Following, this series was divided into consecutive windows of 5 samples and for each of these windows; we have determined both statistical indices or characteristics, yielding a total of 200 instances (patterns) for each characteristic. Then, we have labeled the first 200 instances as Class 1 (C1) and the second 200 instances as Class 2 (C2).

TABLE I
ZONES IN A LOGISTIC MAP BY ATTRACTORS NUMBER

Zone	Initial value r	Final value r
Simple Attractor	1	3
Double Attractor	3	3.5
Chaos	3.5	4

Figure 7 shows the patterns extracted at logistics application, the most irregular pattern (upper pattern) corresponds to the curve size, and the other one corresponds to statistical threshold.

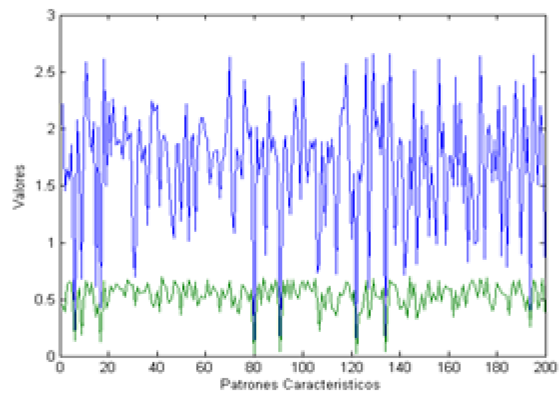

Fig. 7. Reconstructed patterns of the logistic application

The training and validation of SVM was made considering the linear, polynomial and RBF kernels; with all them, the prediction for the known patterns (C1 and C2) has been done. Table II shows the best results, and the parameters associated with each kernel, with respect to the most efficient percentage in the classification model. These results are obtained using a cross-validation of 10 layers.

TABLE II
CLASSIFICATION RESULTS WITH SVM

Kernel	Parameters	Succes %
Linear	C=100	94.7%
Polynomial	C=10, p = 2	94.2%
RBF	C=10000, $\gamma = 0.01$	95.2 %

Table II shows a better efficiency of the classification with the RBF Kernel.

In order to obtain a better efficient of classification with SVM, we have considered the instances that were correctly predicted using regression analysis, such as are immediately shown:

In practical sense, as it is shown in the values for classification percentage and relative statistics for each class,

TABLE III
STRATIFIED CROSS-VALIDATION (SUMMARY)

Parameter	Value
Time taken to build model	0.86 seconds
Correctly Classified Instances	387 (96.9849%)
Incorrectly Classified Instances	13 (3.0151%)
Kappa statistic	0.9397
Mean absolute error	0.063
Root mean squared error	0.1688
Relative absolute error	12.5959 %
Root relative squared error	33.7512%
Coverage of cases (0.95 level)	98.4925%
Mean rel. region size (0.95 level)	63.6935%
Total Number of Instances	400

TABLE IV
DETAILED ACCURACY BY CLASS

Precision	Recall	F-Measure	ROC	Class
0.94	0.969	0.975	0.97	C1
1	0.971	0.975	0.91	C2

both classes have high precision and recall. ROC areas as well as F-Measure, which is the harmonic mean of the precision and recall values, are above 0.9. All these values can be considered excellent, in this way, it is possible determining the operating condition patterns and occurrence of each of the classes (faults).

Once analyzed that is possible to distinguish operative conditions, regarding the faults presence, by means of statistical indices and SVM from the measured data, these techniques will be applied for the fault diagnosis in a technical process.

B. Magnetic Levitation

A diagnostic model for the magnetic levitation system is given by:

$$\begin{aligned} \dot{x_1} &= x_2 \\ \dot{x_2} &= -\frac{k}{2m}\left(\frac{x_3}{x_1}\right)^2 + g + v_1(t) \\ \dot{x_3} &= \frac{x_2 x_3}{x_1} - \frac{R}{k} x_1 x_3 + \frac{1}{k} u(t) + v_2(t) \\ y_1 = x_1, &\quad y_2 = x_3 \end{aligned} \tag{20}$$

where $x_1 = z$ (vertical separation distance), $x_2 = \dot{z}$ (vertical relative speed), $x_3 = i$ (electrical flow in the magneto), g (gravitational force), m (suspended object mass), $u(t)$ (voltage in the magneto) and $k = \frac{u_0 N^2 A}{2}$ (force factor). The physical parameters are: $A = 0.04m^2$, $N = 660$, $\mu_0 = 4\pi 10^{-7}$, $R = 1\Omega$, $g = 9.8\frac{m}{s^2}$, $m = 0.300Kg$. The signals $\mu(t)$, y_1, y_2 $v_1(t)$ and $v_2(t)$ are control, outputs and faults, respectively. The fault classification data is obtained according to Figure 5.

We have simulated 300 seconds with four possible states:

- Normal: Non faults ($t < 100$ and $t >= 250$).
- Fault 1: The first fault appears. ($100 \leq t < 150$). This fault has a sinusoidal behavior with amplitude 0.5 and frequency 1.
- Fault 2: The second fault appears. ($200 \leq t < 250$). This fault has a sinusoidal behavior with amplitude 15, and frequency 1, and its absolute value is considered.
- Multiple Faults: Both faults occur. ($150 \leq t < 200$).

Using the generated data, a SVM with a Gaussian kernel has been trained. In this case, the control and output of the system are used as input data values. Taking a time window of 0.5 second, with simulation step of 0.05 second, 10 values in each window are obtained. In order to reproduce and smooth the trend of the time series in each window, a third degree polynomial interpolation is applied, then each set of 10 values is represented by a polynomial of the form: $ax^3 + bx^2 + cx + d$. For training data the SVM achieves the classification in each state without any error.

Subsequently, in order to test the effectiveness of the SVM with a different setting, the same parameters were used for the faults, but the system was modified to produce them at different time (This is shown in the Figure 8):

- Normal: Non faults ($t < 100$ and $t \geq 250$).
- Fault 1: The first fault appears. ($200 \leq t < 250$).
- Fault 2: The second fault appears. ($100 \leq t < 150$).
- Multiple Faults: Both faults occur. ($150 \leq t < 200$).

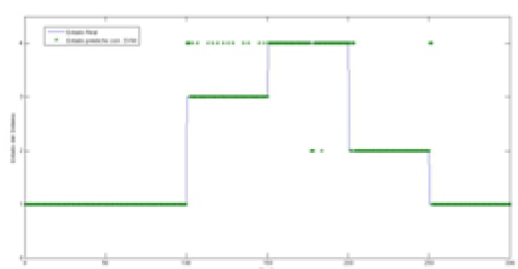

Fig. 8. Classification of the operation condition

In this case, an operating condition prediction of the system (classification) with a precision greater than 97% is achieved. Then, the amplitude of fault 2, from 15 to 20 is changed, obtaining a prediction error of 3.62%, so that the SVM is able to make a good classification even when the operating condition have changed. Subsequently, the frequency of fault 1 is changed, from 1 to 1.1, simultaneously with the above change. With these new values, the prediction of operating conditions is correct by 95%. For this evaluation, Figure 8 shows the comparison of the prediction found using the SVM with the real state of the system. The X-axis represents the time and the axis of the ordinates the state of the system: 1 means Normal condition, 2 means Fault 1, 3 means Fault 2, and 4 means multiple Faults.

V. CONCLUSIONS

In order to provide industrial processes with demanding security mechanisms, it is possible to use statistical indices, in conjunction with machine learning algorithms, such as those presented in this paper. Then, we can build reliable systems for monitoring, diagnosis and detection (MDD).

The obtained results by considering different Kernel in SVM methodology, on mathematical model for the logistic application, confirm the potential of SVM for the patterns recognition and classification, which could be used in MDD for responding to unexpected behavior of the process.

For the logistic application, the results show that the decision tree generator algorithm is able to select the variables and theirs values for achieving the best classification of two analyzed classes. The decision tree, and decision rules that can be easily extracted from it, can help us in order to identify with good precision and accuracy the operating condition patterns and occurrence of each of the classes (faults).

Finally, the use of data mining techniques, as presented in this study, can be used for real time monitoring of industrial processes, reporting the behavior of the production as well as symptomatic conditions that are considered as abnormal situations.

REFERENCES

[1] C. Batur, Z. Ling and C. Chien-Chung C., *Support vector machines for fault detection*, Proc. 41st IEEE Conf. on Decision and Control, vol. 2, 2002, pp. 135-156.

[2] C. Burges, *A tutorial on support vector machines for pattern recognition*, Data Mining and Knowledge Discovery, 1998, pp. 2:121-167.

[3] X. Carreras, L. Márquez and E. Romero, *Máquinas de Vectores de Soporte, en Hernández, J., Ramírez, M. y Ferri, C., Introducción a la Minería de Datos*, Editorial Pearson, España, 2004, pp. 353-382.

[4] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and H. Witten, *The WEKA Data Mining Software: An Update*, SIGKDD Explorations., 2009, pp. 11: 10-18.

[5] J. Hernández-Orallo, M. Ramirez-Quintana and C. Ferri-Ramirez, *Introducción a la Minería de Datos*, Pearson Educación S.A. Prentice Hall. Madrid, 2004.

[6] C. Isaza, J. Aguilar-Martin, M. LeLann, J. Aguilar and A. Ríos-Bolívar, *An Optimization Approach for the Data Space Partition Obtained by Training Methods for the Monitoring of Dynamic Processes*, IOS Press, 2006, pp. 80-87.

[7] L.B. Jack and A.K. Nandi, *Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms*, Mechanical Systems and Signal Processing, 16(2-3), 2002, pp. 373-390.

[8] S. Mahadevan and S.L. Shah, *Fault detection and diagnosis in process data using one-class support vector machines*, Journal of Process Control, 19(10), 2009, pp. 1627-1639.

[9] T. Okba, T. Ilyes, G. Tarek and M. Hassani, *Supervised learning with kernel methods*, Proc. 10th WSEAS Int. Conf. on Wavelet analysis and multirate systems, WAMUS'10, 2010, pp. 73-77, Stevens Point, Wisconsin, USA.

[10] J. Quinlan, *Induction of Decision Trees. Machine Learning*, 1(1), 1986, pp. 81-106.

[11] A. Ríos-Bolívar, *Sur la synthèse de filtres de dètection de dèfaillances*, Ph.D. dissertation, (2001), Universitè Paul Sabatier, Toulouse.

[12] A. Ríos-Bolívar, F. Hidrobo and P. Guillén, *Diagnóstico de fallas en procesos dinámicos: Un enfoque basado en SVM*, Proc. V Congreso Iberoamericano de Estudiantes de Ingeniería Eléctrica, CIBELEC, 2012, pp. M1-M8, Mérida, Venezuela.

[13] A. Ríos-Bolívar and F. Szigeti, *A FDI Filter Based-on Inversion for Nonlinear Systems*, Advances in Systems Science: Measurements, Circuits and Control, WSES Press, 3, 2001, pp. 35-40.

[14] A. Ríos, L. Parraguez, F. Hidrobo, F. Rivas, M. Heraoui, J. Anato : *An Imprecise Computation Framework for Fault Tolerant Control Design*. WSEAS Trans. on COMPUTERS. Issue 7, 8, 2009, pp. 1093-1102.

[15] A. Ríos-Bolívar, R. Márquez: *Fault Tolerance for Robust Anti-Windup Compensation Implementation*. WSEAS Trans. on Systems. Issue 9, 5, 2006, pp. 2197-2203.

[16] A. Ríos-Bolívar and G. Garcia. *An FDI Robust Filter Based-On LMI Control*. Proc. 7th WSEAS Int. Conf. on Automatic Control, Modeling and Simulation (ACMOS '05), Prague, Czech Republic, 2005, pp. 389-394.

[17] A. Ríos-Bolívar, W. Acuña: *Robust FDI in Uncertain LTI Systems: A Multiobjective H_2/H_∞ Setting*. Int. J. of Electronics, Electrical and Communication Engineering (IJEECE). 2, 1, 2010, pp. 25-45.

[18] A. Ríos-Bolívar, P. Guillén, F. Hidrobo, and F. Rivas-Echeverria. *Data Mining for Fault Diagnosis in Dynamic Processes: An Approach based on SVM*. Proc. 15th Int. Conf. on Mathematical and Computational Methods in Science and Engineering (MACMESE '13). 2013.

[19] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri and K. Yin, *A review of process fault detection and diagnosis Part III: Process history based methods*, Computers and Chemical Engineering, 27, 2003, pp. 327-346.

A. Ríos-Bolívar was born in Upata, Venezuela 1961. He is Electrical Engineer (1987), Master in Control Systems (1994) from the Universidad de Los Andes, Mérida, Venezuela. He has obtained his Doctor Degree in Automatic Systems from the Université Paul Sabatier, Toulouse, France (2001), and Doctor in Applied Sciences from the Universidad de Los Andes, Venezuela (2003).

He is full time professor at the Department of Control Systems in the Universidad de Los Andes. Dr. Ríos-Bolívar has published more than 300 scientific articles in journals, books, and international conference proceedings. He is coauthor of the books, "Sistemas MultiAgentes y sus Aplicaciones en Automatización Industrial" and "Implementando Técnicas de Control Acotado: Un Enfoque Basado en Tolerancia a Fallos". He is part of various editorial committees for journals and conferences. He has been a member of the jury of several scientific awards, masters and doctoral works at the national and international level. He has developed fault detection tools for Venezuela Oil Company (PDVSA). He has been recognized by the Research Promotion Program and Direct Support to Groups at the Universidad de Los Andes. In addition, he is a maximum level researcher for the Research Promotion Program of the Venezuelan Foundation for the Promotion of Research and the National Science, Technology and Innovation Observatory in the Program Promotion of Innovation and Research.

F. Hidrobo is Systems Engineer from the University of Los Andes (ULA); he obtained a Master degree in Computer Science in the same university. In 2004, he obtained the DEA and the PHD degree, both at the Polytechnic University of Catalonia (Spain). He has published several papers in journals and conference in the areas of Automation, Intelligent Computing and High Performance Computing. He has worked as consultant in various research and development projects led by the government oil company of Venezuela.(PDVSA). He has been recognized by the Research Promotion Program at the University of Andes and by the National Program to Promotion Research. Currently, he is full time professor at the Science Faculty of the Universidad de Los Andes.

P. Guillén is Bachelor of Science (B.Sc.) in Mathematics, College of Science, University of Los Andes, Mérida, Venezuela (1992); Magister Scientiae (M.Sc.) in Applied Mathematics to Engineering, College of Engineering, University of Los Andes, Mérida, Venezuela (1997); Advanced Studies and Research for the Doctor in Engineering (DEA), Polytechnic University of Catalonia (UPC), Barcelona, Spain (2000); Doctor in Engineering (PhD), Polytechnic University of Catalonia (UPC), Barcelona, Spain (2002); Postdoctoral Research in Computational Science, Program in Computational Science, University of Texas at El Paso, El Paso, TX, USA, (2013). Currently, He is associate professor at the center of simulation and modeling (CESIMO),University of Los Andes, and Research Assistant Professor at the Department of Computer Science, University of Houston, Houston, TX, USA. During the last 15 years he has been working on Research and Development projects related to Oil, Gas, and Biomedical Sciences. These projects have been in different areas such as: Reservoir Simulation, Geophysics, Geothermal, Gas, Artificial Intelligence, Data Mining, Processing of Signals and Visualization and High Performance Computing. These projects were funded by Universities, Ministry of Science and Innovation and Companies. He has published several papers in journals and conference.

F. Rivas (M01) was born in Mérida, Venezuela 1969. He is Systems Engineer, Master in Control Systems and Doctor in Applied Sciences from the Universidad de Los Andes, Venezuela.

He is full time professor at the Universidad de Los Andes. Also he is the Director of Intelligent Systems Laboratory and the University General Rector Coordinator. Dr. Rivas-Echeverría has published more than 200 scientific articles in journals, books, and international conference proceedings. He is coauthor of the books, Introduccin a las Tcnicas de Computacin Inteligente and Control de Sistemas No lineales, the latter published by Pearson Education of Spain. He is part of various editorial committees for journals and conferences. He has been a member of the jury of several scientific awards, masters and doctoral works at the national and international level. He has presided over several symposiums, workshops and has been invited to give master conferences and tutorials in various parts of the world. He has created, taught, and participated in national and international training courses and has received several international and national awards. He has been recognized by the Research Promotion Program and Direct Support to Groups at the Universidad de Los Andes. In addition, he is a level II researcher for the Research Promotion Program of the Venezuelan Foundation for the Promotion of Research and the National Science, Technology and Innovation Observatory in the Program Promotion of Innovation and Research. Level B. Halliburton gave him a recognition for contributions and dedication to the development of petroleum technology. Recognition awarded by Magazine Revista Gerente as one of the 100 most successful Managers in Venezuela. September 2012.