

Evaluation of fault detection techniques using receiver operating characteristic curve and adaptive thresholds

Thiago Cabral de Souza¹, Paulo Victor Padrao Lopes¹, Yago Pessanha Correa¹, Wladimir Pinheiro¹ and Alexandre Carvalho Leite¹

¹ Rua Dr. Siqueira, 273. Campos dos Goytacazes/RJ - 28030130 - Instituto Federal Fluminense

Abstract: Any engineering application is susceptible to fault occurrence which can cause a major damage on processes. To detect these faults, classic methods can be used, such as fixed thresholds or structural residues. This paper discusses the application of detection fault theory using the ATLMS, an adaptive threshold technique, with structured residual generation. The efficiency of these techniques will be evaluated by an receiver operating characteristic curve, which can be used as a performance metric to the true positive and the false positive rates, showing the better way to detect faults. Here, is also presented a case study with real data and fault detection using the ATLMS and a ROC analysis of the result.

Keywords: *Fault Detection Theory, ATLMS, ROC Curve, Structural Residues, Principal Component Analysis*

INTRODUCTION

Faults are unpermitted deviation of at least one characteristic property of the system from the acceptable, usual or standard condition (Isermann, 2005). With the growth and domain of the technological applications, system complexity has been playing an important role (Lopes,Souza,2018). Such systems require reliability, which make fault detection a fundamental task. Fault detection and diagnosis (FDD) theory came to create ways to increase the system reliability, lifetime expansion of monitored components and failure probability reduction of the overall system (Gertler, 2015), (Chiang, Russel, 2001).

BACKGROUND CONCEPTS

Receiver Operating Characteristic Curve

The Receiver Operating Characteristic Curve (ROC) is a performance metric that calculates the efficiency of a binary classifier that demonstrate the distribution behavior scanning all the possible thresholds, as can be seen in Fig. 1.

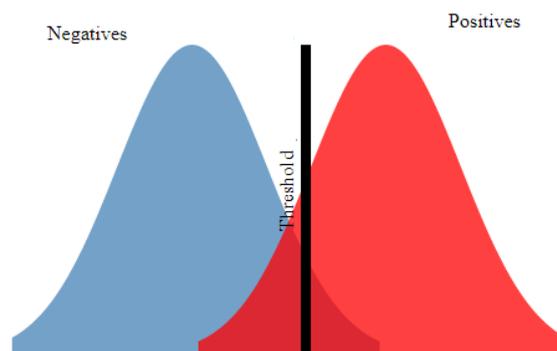


Figure 1 – Intersection Between Distributions (Navan, 2013)

To determine the classifier performance of the ROC curve, it is necessary to calculate the true positive rate (TPR) and the false positive rate (FPR) (Krzanowski,2009), as shown in Eq. 1 and Eq. 2:

$$TPR = \frac{\text{Number of Real Alarms Detected}}{\text{Number of Faulty Samples}} \quad (1)$$

and the FPR is calculated by:

$$FPR = \frac{\text{Number of False Alarms Detected}}{\text{Number of Normal Samples}} \quad (2)$$

Fixed Threshold

A simple method to detect faults is the use of fixed thresholds. If the variable exceeds a predefined limit, then the detection system indicates a fault by raising a flag. It is the most used in industry, since it is easy to implement and does not require a complex theoretical background (Isermann, 2005). However, this method has some disadvantages such as:

1. Sensitivity to noise and disturbances;
2. Sensitivity to control signal.

Adaptive Threshold Least Mean Squares

The Adaptive Threshold Least Mean squares technique was developed at the German Aerospace Center (DLR) in order to detect faults in a space rover (Leite, 2012). It allows the threshold tuning by changing well known parameters independently of the case study. This signal-based approach does not require a priori knowledge either of the fault models or plant, but relies on a suitable set of fault signatures delivered by residual generators.

The ATLMS technique is based on the Sequential Probability Ratio Test (SPRT) and an adaptive non-recursive filter. Alarm activation results from a comparison between a residual non-sensitive residual to the applied fault and an adaptive threshold (Leite, 2012), (Padrão,Souza, 2018).

The adaptive threshold is calculated by Eq.3, as follows:

$$f = \phi^T \theta \quad (3)$$

Resulting in Eq. 4:

$$f = [u_1 \quad \cdots \quad u_{N_c} \quad y_{0k} \quad \cdots \quad y_{0k-m} \quad 1] \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{N_c+m+2} \end{bmatrix} \quad (4)$$

Figure 2 presents the structure of the adaptive threshold,

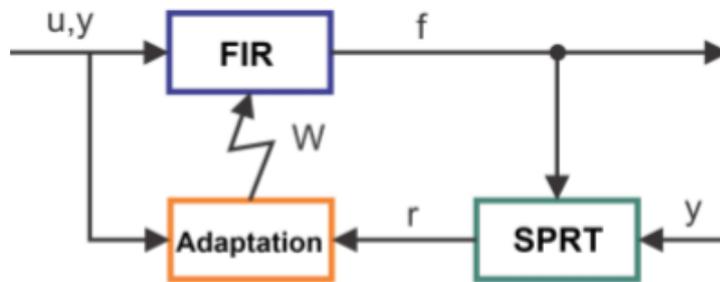


Figure 2 – ATLMS diagram (Leite, 2012)

where

- u : control signal vector;
- y_0 : insensitive residue to the fault to be detected (obtained as in [Isermann]);
- y_1 : sensitive residue to the fault to be detected;
- r : instantaneous *SPRT*;
- W : weight vector of the non-recursive filter.

In that way the ATLMS calculates a new threshold using the residuals and signal control which brings the following advantages:

1. Detectability of residual trend;
2. Insensitivity to noise and disturbances;
3. Insensitivity to control signal.

In the case of adaptive methods, the ROC is represented by a point insted of a curve. This happens because the ATLMS scans the useful thresholds and it returns the best case, as demonstrated in Fig. 3:

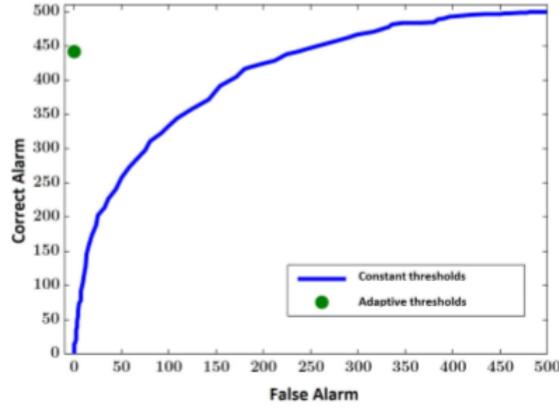


Figure 3 – Difference between a ROC used in fixed and adaptive threshold (Leite, 2012)

The main objective of the ATLMS is to decrease false alarms, which imply low detection sensitivity to noise and disturbances and also to increase correct alarms in order to have a faster detection (Leite, 2012).

Principal Component Analysis

Principal Component Analysis (PCA) is a dimensionality reduction technique without much loss of information (Rus- sel, Chiang, 2000). It is a statistical technique that is greatly used in face recognition, image compression, and also used for finding pattern in high dimension data (Kandula, 2011). It is assumed that x is a vector of M random variables (measurement signals) and that the variance of the random variables and the structure of the covariances or correlations between M variables of interest.

According to Isermann (2001), PCA algorithm follows the described methodology, which consists in six steps:

1. Calculation of the correlation matrix:

$$A = X^T X \quad (5)$$

2. Calculation of the eigenvalues λ_j of the matrix A and the eigenvectors P_j

$$(A - \lambda_j I) P_j = 0 \quad j = 1, \dots, m \quad (6)$$

3. Selection of the most significant eigenvalues

$$X' = t_1 p_1^T + t_2 p_2^T + \dots + t_r p_r^T \quad (7)$$

4. Determination of the transformation matrix P

$$P = [p_1 \ p_2 \ \dots \ p_r] \quad (8)$$

5. Calculation of the new data matrix

$$T = X' P [t_1 \ t_2 \ \dots \ t_r] \quad (9)$$

6. Back-transformation in the original data coordination system yields

$$X_{[N \times m]}^* = T_{[N \times r]} P_{[r \times m]}^T \quad (10)$$

To accomplish fault detection, a residual has to be calculated, according to Eq.9.

$$R = X^* - X \quad (11)$$

Monte Carlo Simulation

According to Palidase (2018), a Monte Carlo simulation consists in values which are sampled at random from the input probability distributions. Each simulation runs a set of samples, and the outcome is recorded. Monte Carlo simulation does this hundreds or thousands of times, and the result furnishes the decision-maker with a range of possible outcomes and the probabilities they will occur for any choice of action.

The signal used in the Monte Carlo simulation was a white Gaussian noise which has a fault at $t = 5s$. It is a signal which is commonly found in engineering applications and it describes accurately how the noisy measurements influences fault detection methods. The faulty signal is presented in Fig. 4:

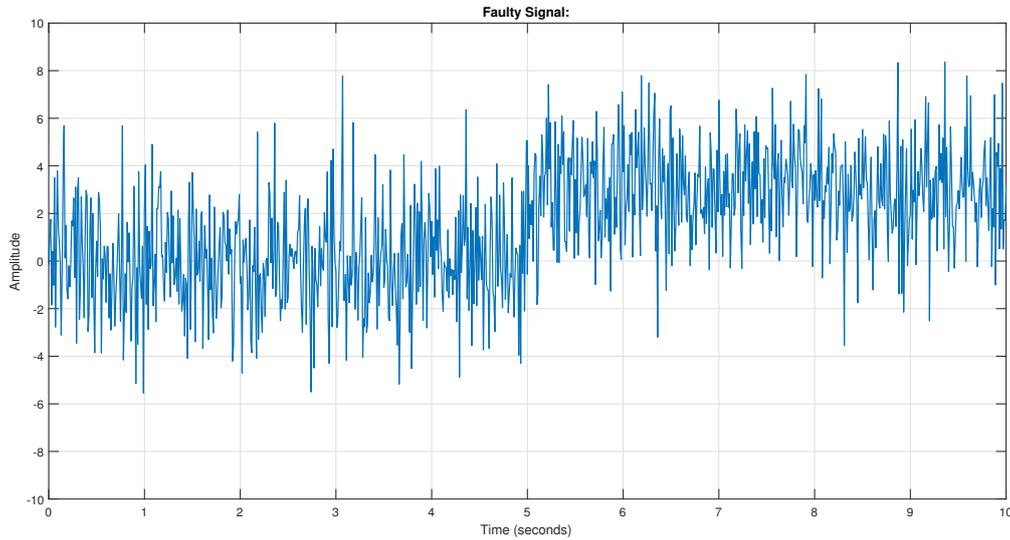


Figure 4 – Input Signal

RESULTS

Robustness Analysis

With the purpose of analysing the robustness, the simulation was performed using the fixed threshold technique. This signal has passed through 1000 simulations, each simulation using a fixed threshold. From one simulation to another, the step of the threshold was increased of 0.02. Thus, all data was processed, then the *FPR* and *TPR* were calculated, which made the ROC curve. This result can be demonstrated in Fig. 5:

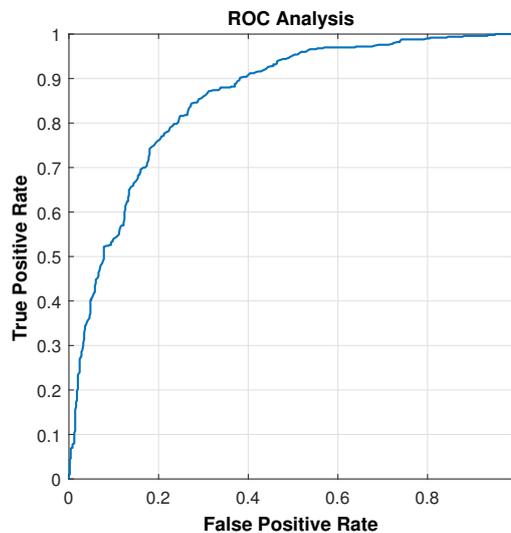


Figure 5 – ROC of fixed threshold response

Here, the ROC shows that at any point of the threshold it has never reached the point (0,1). That point means zero false alarms and all alarms detected correctly. It does not happen because there are points of intersection between this fault and normal distribution and this method does not follow the trend, only the actual value. The reason for that was a fixed threshold cannot treat differently points in intersection, as it was exposed beforehand in Fig. 1.

This result was not acceptable, then the ATLMS was used. To prove its efficiency and robustness, a Monte Carlo simulation has been made, which consists in 5 millions of simulations. Each simulation is a different signal with normal distribution. The result of the simulations can be seen in Fig. 6 and the Fig. 7 shows the 6th image zoomed on the region of interest. These pictures show how the fault detection is efficient using ATLMS.

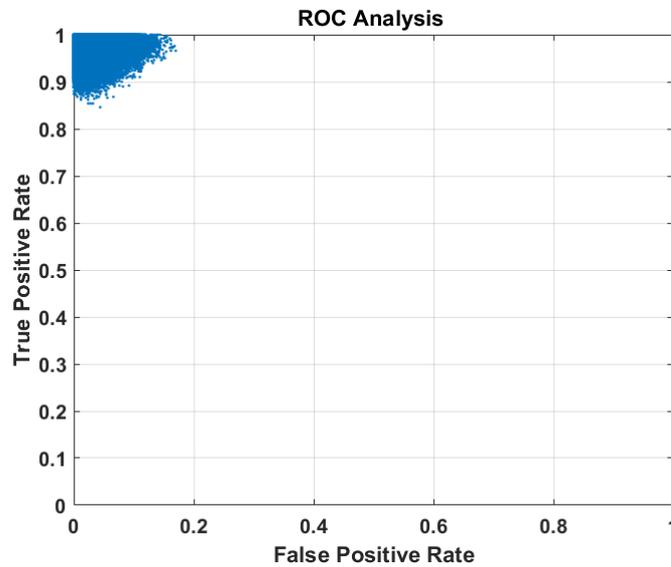


Figure 6 – ROC of adaptive threshold response

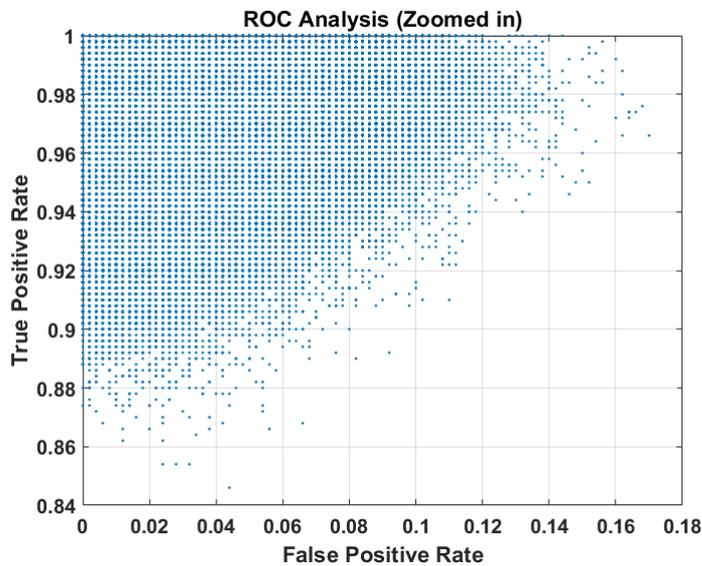


Figure 7 – ROC of adaptive threshold response zoomed

As presented in Fig. 6 and Fig. 7, the ATLMS has shown a better result, showing millions of signals which had the fault detected at the same moment that it happened. The insensitivity to noise combined with the threshold variability made the ATLMS a improved binary classifier; that can treat the points in the intersection differently, because it does not only classify by the absolute value but also by the trend. The robustness of the ATLMS to signals which contain gaussian noise is remarkable. It shows 100% of detection, demonstrating its robustness.

CASE STUDY

Trebuchet - Nasa Prognostics Center

This process consist in a trebuchet launching a 16.674 g ball and using a 505.52 g counterweight, which can be seen in Fig. 8:

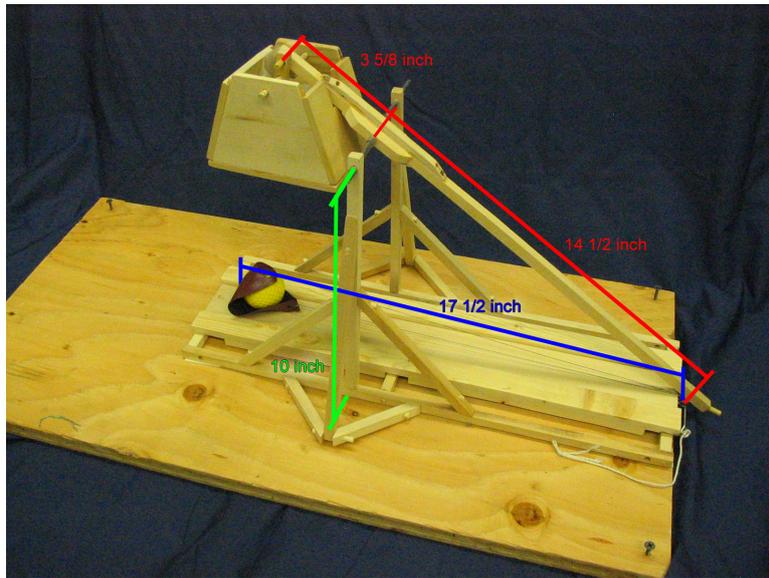


Figure 8 – Trebuchet dimensions of NASA (Morton, 2008)

The launched ball has its dynamics determined by drag force and gravity. To determine the position of the ball in each video frame, a simple change detection routine was used between adjacent frames. A Gaussian low pass filter was then applied to the resultant frame, to eliminate noise from the image, then the maximum value of the frame was recorded as the position of the ball.

To demonstrate the efficiency of the ATLMS in a real system, a case study was made, using real data of a trebuchet (Morton, 2008) and simulated a sensor delay fault. That process disposes two measures, the position in x and the position in y . As described in Fig. 9, the sensor shows a delay at the x position, and it kept in a failure state, showing the expected, or normal, value and the fault behavior. In order to extract the residuals, the method used was PCA, which results in two graphics, that can be seen in Fig. 10. The first residual is sensitive to the fault and the second insensitive to the fault applied. Fig. 11 shows the flag response for fault occurrence (flag = 1). With respect to delay fault, the ATLMS detected a deviation from normal behavior at only one sample after the occurrence. To evaluate the detection, it was also calculated a ROC analysis, and the result is 0% of false positive alarmes and 100% of true positive rate, as demonstrated in Fig. 12.

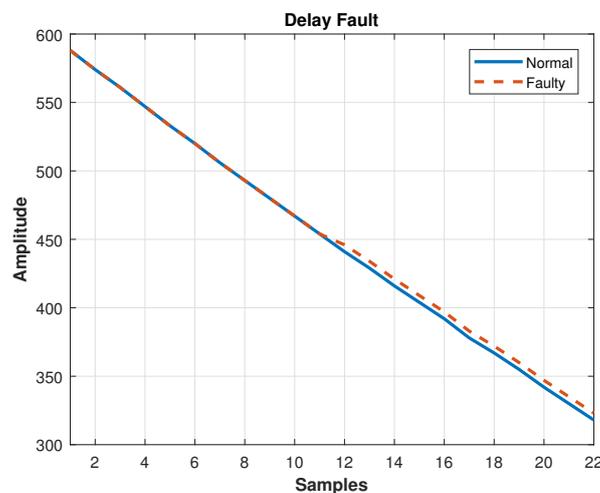


Figure 9 – Sensor delay fault in x

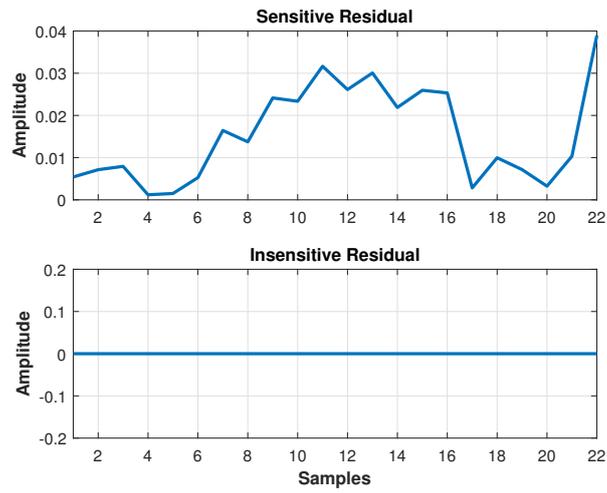


Figure 10 – Residuals calculated using PCA

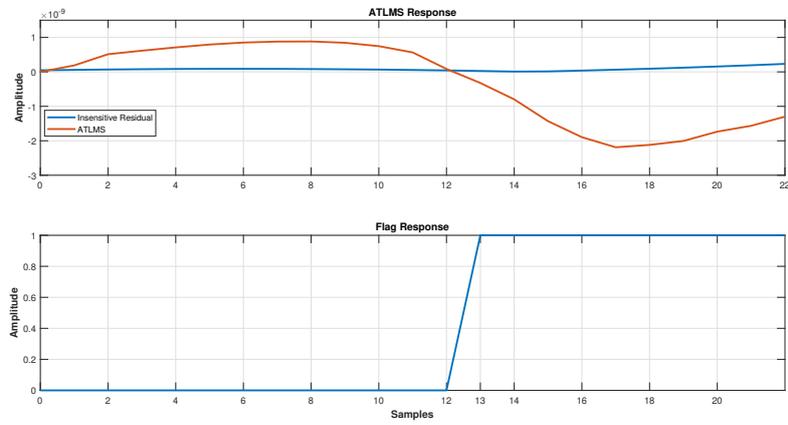


Figure 11 – ATLMS Response - Case study

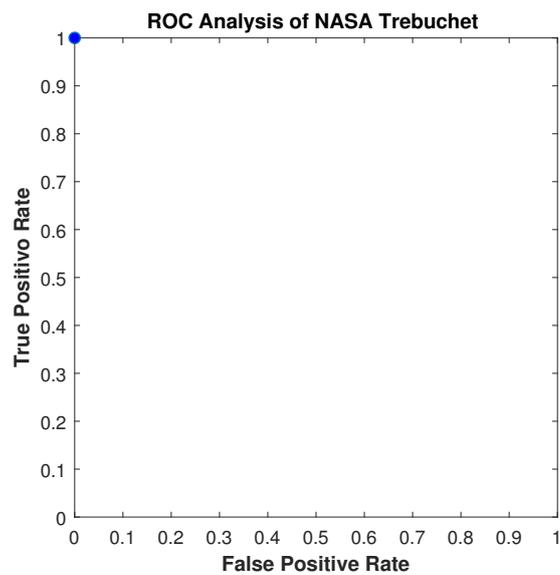


Figure 12 – ROC Analysis - Case study

CONCLUSION

The ATLMS technique was proven its robustness and a very useful tool for detecting faults. It shows a high number of true positive alarms and a low number of false positive alarms. Furthermore, the ATLMS also shows a high capability of dealing with noisy signals and accomplish fault detection regardless of it.

REFERENCES

- Chiang, E.L. Russel, 2001, Fault Detection and Diagnosis in Industrial Systems.
- Gertler, 2015, Fault Detection and Diagnosis in Engineering Systems.
- Isermann, 2016, Fault-Diagnosis Systems, 1stedn, Springer-Verlag.
- Kandula, 2011, "Fault detection in process control plants using principal component analysis", Louisiana State University, United States of America.
- Krzanowski, Hand, 2009, "ROC Curves for Continuous Data (Chapman Hall/CRC Monographs on Statistics Applied Probability)", 1st, Chapman an Hall/CRC.
- Leite, 2012, "Fault detection, diagnosis and re- configuration for optimized design of safe space vehicles applied to the PMM and planetary rovers", Deutsches Zentrum fr Luft- und Raumfahrt, Germany.
- Morton, B. Sentient Corporation. "Trebuchet Data Set",NASA Ames Prognostics Data Repository (<http://ti.arc.nasa.gov/project/prognostic-data-repository>), NASA Ames Research Center, Moffett Field, CA, retrieved on August 18, 2018.
- Navan.name/ROC, retrieved on July 15, 2018.
- Palisade.com/risk/monte_carlo_simulation.asp, retrieved on november 3, 2018.
- E. L. Russell, L. H. Chiang, R. D. Braatz, Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis, Chemometrics and Intelligent Laboratory Systems 51, 8193 (2000).

RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.