Isolation of plant-wide faults using causality detection methods

Guilherme M. Garcia∗ Celso J. Munaro∗

∗Department of Electrical Engineering, Universidade Federal do Espírito Santo, Vitória, ES, 29075-910 Brazil (e-mail: guiﬁles@gmail.com; cjmunaro@gmail.com).

Abstract: Isolation of plant-wide faults in large-scale complex systems is particularly challenging. A methodology to detect and isolate faults is proposed, detecting the faulty variables using univariate control charts and the causality information between them to indicate the source. The clustering of faulty variables uses univariate analysis to avoid the smearing effect brought by multivariate analysis. The variable where the fault took place is indicated, handling fault novelties in a very natural manner. The proposed method is discussed and illustrated through its application to the Tennessee Eastman Process and to routine operating data from a thermoelectric power plant.

Keywords: Causality Detection, Fault isolation, Transfer Entropy, PCA, Process Control.

1. INTRODUCTION

Detection and isolation of plant-wide faults are major problems in the process industry. To isolate a fault in large-scale complex systems is particularly challenging because of the high degree of interconnections among different parts. A simple failure may propagate along information and affect other parts of the system.

Fault detection techniques range from fault trees, digraphs and analytical frameworks to knowledge-based systems and neural networks. From the perspective of modeling Venkatasubramanian et al. (2003) divides into 3 categories: quantitative models, qualitative models, and process history based models or simply data-driven. In contrast to methods based on prior knowledge (both quantitative and qualitative) data-driven methods only require access to a large amount of historical process data. There are different ways to transform and present this data as a priori knowledge to a diagnostic system. This process, known as feature extraction, can be qualitative, such as specialist systems and trend modeling, or quantitative, such as statistical models and neural networks. Statistical process monitoring (SPM) is successful in feature extraction, Qin (2003) reviewed the use of fault detection indices such as $T^2$ and $Q$, which were calculated directly from statistical projection methods using normal operating data.

Unlike the fault detection problem, fault isolation has received less attention in the SPM research community. The detection of cause-effect relationships in signals from industrial processes is useful to discover signal flow paths. The estimation of time delays has been used to find the source of disturbances caused by faults (Bauer and Thornhill, 2008; Stockmann et al., 2012). However, the estimation of time delay is very prone to errors. Additionally, this method fails in estimating the propagation in non-linear systems. One of the methods to find the causal relationships is the transfer entropy (Schreiber, 2000), which relies on conditional probabilities. This measure can be used to identify the propagation path of disturbances (Bauer et al., 2007), to overcome the problem of time delay estimation.

Any method used to find the causal relationships to indicate the source of the fault may fail if variables that are only correlated with faulty variables are included. This paper presents a methodology that can circumvent this issue, highlighting the results that one can obtain for the usual situations.

The paper is organized as follows. In Section 2, the concepts of fault detection and isolation and causality are reviewed. Section 3 describes the proposed methodology. In Section 4 Tennessee Eastman Process and a Thermoelectric Power Plant are used to demonstrate the effectiveness of the proposed approach, followed by concluding remarks in Section 5.

2. FAULT DETECTION AND ISOLATION

Due to the wide scope of the fault detection area and the difficulty of real-time solutions, various techniques have been developed over the years. The most well-known statistical methods for feature extraction in data-driven fault detection are the Principal Component Analysis (PCA), the Partial Least Square (PLS), and the Independent Component Analysis (ICA).

PCA is the most widely used method in industrial systems (Chiang et al., 2001). Given a set of $n$ variables and $m$ observations stacked into a matrix $X \in \mathbb{R}^{mxn}$, the loading vectors are calculated by solving an eigenvalue decomposition of the sample covariance matrix,

$$S = \frac{1}{(m-1)} \sum_{i=1}^{m} X_i X_i^T = \Lambda \Lambda^T$$ (1)

where the diagonal matrix $\Lambda$ contains the non-negative real eigenvalues in decreasing magnitude ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$)
... ≥ λ_n ≥ 0). The loading vectors are the orthonormal column vectors in V. The variance of the training set projected on the \( i^{th} \) column of V is \( \lambda_i \). The columns of the loading matrix \( P \in \mathbb{R}^{nxa} \) are the columns of matrix V associated to the \( a \) largest eigenvalues, while \( \tilde{P} \in \mathbb{R}^{nxn-a} \) holds the remaining loading vectors. The projections of an observation \( x \in \mathbb{R}^n \) into the lower-dimensional principal and residual subspaces are \( \tilde{x} = PP^T x \) and \( \tilde{x} = \tilde{P} \tilde{P}^T x \).

With the normal condition modeled (feature extraction), the next step in statistical data-driven methods is to calculate multivariate statistics to detect abnormal situations. Hotelling’s \( T^2 \) statistic (eq. 2) and the squared prediction error known as Q statistic (eq. 3) can be used to detect faults for multivariate process data.

\[
T^2 = x^T P A^{-1} P^T x \tag{2}
\]
\[
Q = x^T \tilde{P} \tilde{P}^T x \tag{3}
\]
Thresholds for these statistics for a given significance level \( \alpha \) (\( T^2_\alpha \) and \( Q_\alpha \)) can be calculated (Chiang et al., 2001) to detect faults.

2.1 Fault Isolation

When an out-of-control situation is detected, a search for its cause is accomplished. This task has been traditionally performed with pattern classification methods, and requires data collected during out-of-control operations that are categorized into separate classes according to the faults. Assuming that the detected fault is present in the database, the fault can be properly diagnosed (Chiang et al., 2001).

In the context of SPM several methods were introduced in the literature for fault isolation. Contribution plots are based on the idea that variables with the largest contributions to the statistical indices are most likely to be associated with the fault. Alcala and Qiu (2011) classified these methods as complete decomposition contributions (CDC), partial decomposition contributions (PDC), diagonal contribution (DC) and reconstruction-based contributions (RBC). The merit of the contribution plots is to narrow down the search for the variables where the fault took place. A drawback of these methods is the possibility of including in the group of candidate variables those that are correlated to variables affected by the fault, but are not affected by the fault. This problem, called smearing effect, was tackled by many authors. Recently Van den Kerkhof et al. (2013) showed that, after a multivariate fault detection, univariate fault isolation is superior to contribution plots in most of the situations to isolate the variables that were affected by the fault. These results are fundamental to the fault isolation method proposed here.

With the assumption that the faulty variables were clustered, information about the causality between these variables may allow improving fault isolation. Bauer and Thornhill (2008) proposes the use of causality to find the propagation paths of oscillatory disturbances in industrial plants based on time delay estimation. More recently, Stockmann et al. (2012) proposed the use of k nearest neighbor to have a better estimate of time delays in fault isolation. Although these methods require less computational effort, estimation of time delays is prone to errors when slow dynamics are present. This fact motivated the use of tests to check the consistency of the time delay estimates. These methods rely on properly clustering only the faulty variables. An incorrect source may be indicated if a variable is included due only to its correlation to faulty variables. Therefore, the combination of statistical methods to detect faults and methods to detect causality to isolate the faults is carefully examined in this paper.

Several methods are available to detect causality as discussed in Marques et al. (2015). The method used here is the transfer entropy, a nonparametric method that can be used even in the presence of nonlinearities. Schreiber (2000) proposes the transfer entropy as the measure that shares some of the desired properties of mutual information but takes the dynamics of information transport into account. With minimal assumptions about the dynamics of the system and the nature of their couplings, one is able to quantify the exchange of information between two systems, separately for both directions, and, if desired, conditional to common input signals.

2.2 Transfer Entropy

Let us consider two time series \( x = \{x_1, x_2, ..., x_n\} \) and \( y = \{y_1, y_2, ..., y_n\} \). The transfer entropy from \( x \) to \( y \) (\( T_{x \rightarrow y} \)) is different to the transfer entropy from \( y \) to \( x \) (\( T_{y \rightarrow x} \)), because of its inherent asymmetry:

\[
T_{y \rightarrow x} = \sum_{x_{n+h}, y_n} p(x_{n+h}, x_n, y_n) \log \left( \frac{p(x_{n+h}, x_n, y_n) \cdot p(x_n)}{p(x_n) \cdot p(y_n)} \right) \tag{4}
\]
\[
T_{x \rightarrow y} = \sum_{y_{n+h}, x_n} p(y_{n+h}, y_n, x_n) \log \left( \frac{p(y_{n+h}, y_n, x_n) \cdot p(y_n)}{p(x_{n+h}, y_n) \cdot p(y_n)} \right) \tag{5}
\]
Joint PDFs (Probability Density Function) for two stationary signals sequential in time are denoted by \( p(x_{n+1}, y_n) \) with the same PDF for \( x_n, x_{n+1} \), because of stationarity, that is, \( p(x_n) = p(x_{n+1}) \), where \( l \) is the prediction horizon of \( x_n \) and will be substituted by parameter \( h \). The generalization of this joint PDF is the joint PDF for \( k+l \) variables giving \( p(x_n, y_n) \), where \( x_n = \{x_n, x_{n-1}, ..., x_{n-(k-1)}\} \) and \( y_n = \{y_n, y_{n-1}, ..., y_{n-(l-1)}\} \) are embedded vectors. The parameters \( k \) and \( l \) are referred to as the embedding dimension of \( x_n \) and \( y_n \), respectively (Bauer et al., 2007). The conditions for the successful application of this method are discussed in Naghoosi et al. (2013).

3. PROPOSED METHODOLOGY

The direct application of causal maps in faulty variables clustered by faulty isolation methods can be misleading. This problem arises mainly because of the possible inclusion of variables that are only correlated to faulty variables, and also due to limitations of methods to detect causality. These situations are discussed in this Section yielding the proposed methodology.
3.1 Training and Monitoring

Data from Normal Operation Condition (NOC) are used to build the PCA model, to calculate thresholds for fault detection in both $T^2$ and $Q$ statistics, and to estimate causality relations as well. In order to avoid false alarms in fault detection, an Exponentially Weighted Moving Average (EWMA) filter is applied to calculated statistics, and the weighting parameter ($\rho$) is tuned using NOC data.

The reason for detecting causality relations using NOC data is twofold: the causality measures require in general a certain amount of data. If the calculation is performed after the fault is detected, using data related to the fault, some delay would follow waiting for enough data. The second reason is the possibility of structural changes during the fault that could result in causality relations different from those that caused the fault. For example, if the fault comes from a sensor that stops working, keeping the last measurement constant, no causality from this signal to the others will be found after the fault. Thus, the variable where the fault happened would be excluded from the search.

Since the parameters $k$, $l$ and $h$ greatly affect the calculation of the transfer entropy, a systematic method can be used: two parameters can be fixed, $k = 0$, $l = 1$ and $h$ is chosen to maximize the transfer entropy, as proposed in Naghoosi et al. (2013). To establish a threshold for causality significance, Monte Carlo methods using surrogate data can be used. For surrogate time series construction, the iterative amplitude adjusted Fourier transform (iAAFT) method is used in all the following computations. The significance level is then defined as $s_{z \rightarrow y} = \frac{T_{z \rightarrow y} - \mu_k}{\sigma_k} > 6$ where $\mu_k$ and $\sigma_k$ are mean and standard deviation of causality calculated on surrogate data. A 6-sigma threshold for the significance level is chosen as in Bauer et al. (2007).

The time consumed for computing the causality detection depends on the number of variables. It is certainly smaller if only clustered variables are evaluated, using NOC data also. However, since these causality computations are performed during training, time restrictions are rarely violated.

Although a single operating point for the plant is considered in this paper, multiple steady-state operation modes can be considered as well. The results from Maestri et al. (2010) can be readily applied to this approach.

3.2 Fault isolation

Fault isolation is carried out using univariate control charts for all variables. The use of the absolute value of the mean centered and scaled to unit variance ($\frac{x - \mu}{\sigma}$) with a threshold of 2.58 (99% confidence) is proposed here. In order to avoid false alarms, an EWMA filter is used.

As previously discussed, the smearing effect is avoided using this technique. The control charts must be checked for all variables in a data window. Since a delay may be introduced by the filter, the analysis should start with measurements prior to the fault detection instant. The number of samples after the fault detection must be also chosen to warrant that all faulty variables are included.

4. APPLICATIONS

4.1 Tennessee Eastman process

The proposed method is applied to simulation data from the Tennessee Eastman process. Details on the process description are further explained in Chiang et al. (2001). The process has 12 manipulated variables, 22 continuous process measurements sampled every 36 s and 19 composition measurements sampled less frequently (360 s). A total of 31 variables are used for monitoring, we excluded all composition measurements because they are hard to measure on-line and 3 of the manipulated variables (viz., compressor recycle valve, stripper steam valve, and agitator Speed)
Fig. 2. $T^2$ and Q statistics for Fault 1
because they are static during all the observations. The
faults are introduced after 3000 observations (30 hours).

The first 2000 observations (20 hours) of the simulated
data were used as NOC data. The number of principal
components was selected in order to explain at least 80%
of the total variance given a total of 17 components.
The calculated thresholds for 99% confidence limit are
$T^2_a = 33.85$ and $Q_a = 12.10$.

After generating a PCA-based monitoring model the next
step is to utilize this model for detecting an abnormal event
(i.e. fault). In Fault 1 scenario a step change is induced in
de A/C feed ratio. The composition of A thus changes from
48.5 mol% to 45.5 mol% meanwhile, the composition of C
changes from 51 mol% to 54 mol%. In order to maintain
the composition of A, the valve opening of the A feed
flow was increased by controller. This demonstrates that
the process fault was propagated to the controller output
variables in order to maintain its set point. The
$T^2$ and Q statistics detected this fault as shown in Fig. 2, where
the statistics are in blue and their filtered values with
$\rho = 0.2$ are in green. The same EWMA filter was used
on univariate control chart.

With a fault detected, the univariate isolation is performed to
all variables in a window of 10 samples before and after
the detection fault instant (red vertical line in Fig 3). Nine
faulty variables were identified as shown in Fig. 3. The
isolation of faulty variables using diagonal contribution
plot (Alcala and Qin, 2011) in the same window would
produce 14 variables.

The next step in the proposed methodology is to investiga-
te the causal relations between the isolated faulty vari-
able. Only the transfer entropy calculated using NOC for
these variables is presented in Table 1 where the variables
in rows affect the ones in the columns ($x_{row} \rightarrow x_{col}$).
The bold values indicate the causality measures that are
significant.

The causality measures from Table 1 are illustrated via
the directed graph of Fig. 4. The results show that the A
feed flow ($x_{25}$) causes all other variables, similar to the
example shown in Fig. 1. Therefore, the A feed flow was
successfully isolated as the variable that caused the fault.
From process knowledge the A feed is increased in order
to compensate the loss in A composition. The variations
in all other variables are a result of this variation in A feed
flow.

This example demonstrates the case where it is possible to
isolate a single variable and find the source. The reliable
detection of causality plays an important role to ensure
correct source indication.

4.2 Thermoelectric power plant

The proposed method is now applied to routine operating
data from a thermoelectric power plant (Fig. 5). The
control loops shown are those responsible for producing
steam at a given temperature (TIC400) and pressure
(PIC400). This steam generates energy in the turbine and

<table>
<thead>
<tr>
<th>TE</th>
<th>$x_{1}$</th>
<th>$x_{7}$</th>
<th>$x_{13}$</th>
<th>$x_{16}$</th>
<th>$x_{20}$</th>
<th>$x_{21}$</th>
<th>$x_{23}$</th>
<th>$x_{25}$</th>
<th>$x_{29}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{1}$</td>
<td>-</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>$x_{7}$</td>
<td>0.04</td>
<td>-</td>
<td>0.22</td>
<td>0.45</td>
<td>0.06</td>
<td>0.12</td>
<td>0.05</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>$x_{13}$</td>
<td>0.05</td>
<td>0.30</td>
<td>-</td>
<td>0.47</td>
<td>0.10</td>
<td>0.13</td>
<td>0.03</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>$x_{16}$</td>
<td>0.06</td>
<td>0.16</td>
<td>0.15</td>
<td>-</td>
<td>0.07</td>
<td>0.14</td>
<td>0.04</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>$x_{20}$</td>
<td>0.04</td>
<td>0.08</td>
<td>0.10</td>
<td>0.12</td>
<td>-</td>
<td>0.05</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>$x_{21}$</td>
<td>0.03</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.04</td>
<td>-</td>
<td>0.04</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>$x_{23}$</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>-</td>
<td>0.04</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>$x_{25}$</td>
<td>0.12</td>
<td>0.09</td>
<td>0.07</td>
<td>0.10</td>
<td>0.06</td>
<td>0.05</td>
<td>0.04</td>
<td>-</td>
<td>0.04</td>
</tr>
<tr>
<td>$x_{29}$</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>
is distributed to all other consuming processes. Changes in steam consumption or in fuel consumption cause out-of-control status in these variables, which may propagate to other variables. When the pressure in the super-heater decreases, a new set point is generated for the water level in the drum to produce more steam. This loop adjusts the set point for the water flow loop (FIC408). This flow comes from a reservoir (deaerator) whose level is controlled by LIC430. The temperature is also reduced according to the ideal gas law; this reduction causes the fuel flow to increase. The condensed steam leaving the turbine feeds the tank whose level is controlled by LIC434, and this tank feeds the deaerator.

The data used came from 39 loops selected to be sampled every 5 s, yielding 8640 samples. Figure 6 shows the signals from the main variables related to the fuel consumption. The variables FIC401 (LDG) and FIC43A (Main COG) remain stable during a long time (manual mode), when their flow changes they affect the flow of BFG (FIC402) and the flow of stabilizing COG (FIC43B), which are automatic and has set point given by FIC400 (Calorie master). Therefore, when there is an exchange of these fuels near sample 5300, a failure occurs, spreading through other control loops. The proposed methodology is applied to detect and to isolate this event.

![Fig. 5. Instrument diagram of the thermoelectric power plant](image)

![Fig. 6. Variables from fuel consumption](image)

The samples between 2000-5000 of the observed data are used as NOC data to construct the PCA model. The number of principal components was selected in order to keep at least 80% of the total variance, resulting in 16 principal components. Both statistical thresholds are calculated for the 99% confidence limit, $T_2^{\alpha} = 32.26$ and $Q_\alpha = 18.04$. As shown in Fig. 7, they are able to detect the abnormal behavior, where the red line corresponds to filtered values of the statistics with an EWMA filter ($\rho = 0.3$).

The faulty variables isolated using univariate control charts are shown in Fig. 8. The directed graph obtained is shown in Fig. 9 where variables PIC408, PIC400, PICC40L PIC440 and TIC400 do not have causal relation with the other variables and can be eliminated from the analysis as proposed. FIC490 is a sink variable and is removed too. Without FIC490, FIC406 becomes a sink (causes only FIC490) and is also removed.

Different to the first example, it is not possible to isolate a single variable as a source. The remaining variables are the combustion air pressures (PIC405 and PIC406), flow rates of each fuel (FIC402, FIC43B and FIC400) and air...
burning (FIC405), and the amount of O$_2$ (AIC400), which are strongly connected. Those variables can all be grouped as a single subsystem, isolating the burning fuel subsystem as the origin of the fault. In fact, we can infer that the failure is due to a change in fuel composition as shown in Fig. 6, where the flows of fuels FIC43A (main COG) and FIC401 (LDG), change their values dramatically in the instant between samples 5000-6000. Thus, the proposed method was able to isolate a subsystem where the fault happened, simplifying tremendously the task of finding the source of the fault by plant engineers.

5. CONCLUSION

A method to isolate faults combining SPM and causality detection methods was presented. NOC data are used to obtain PCA models, to calculate thresholds for the statistics and to detect relations of causality between the variables. In order to avoid the well known smearing effect, once a fault is detected, the faulty variables are clustered using univariate control charts. The causality shows the topology behind the variables and allows the isolation of the variable where the fault happened. When feedback is present a sub-process may be isolated. In all cases, the approach allows us to indicate the fault or to substantially narrow down the search for the root cause. In the Tennessee Eastman example, fault propagation due to the actions of the process controllers was analyzed through the proposed approach demonstrating that the isolated variable was identified as the source. In the industrial application, the proposed method was capable of identifying the subsystem that was responsible for the abnormal event.

The great advantage of the proposed approach is its ability to deal with novelties, isolating the faults through the indication of the faulty variables instead of using predefined fault classes.

ACKNOWLEDGEMENTS

The authors would like to thank ArcelorMittal for providing data for the third example.

REFERENCES


