1. Introduction

Assessment of a high risk situation in plant operation represents the identification of all process conditions that may lead to a hazardous process behaviour and implementation of a cause-effect strategy able to handle it automatically. This strategy comes as an if-then set of rules and provides solutions from actuator interlock conditions to emergency shut-down strategies.

In order to evaluate the risk of an operating plant, main approaches address offline methods that use specific operating strategies and suggestions to make industrial plants safe-by-design, thus minimizing the consequences and risks associated with human intervention.

Process control reconfiguration strategies come to provide alternatives considering an abnormal behaviour of the process in a nominal state, control reconfiguration determines how the process should react under to faulty operation. Efficient strategy reconfiguration to assess possible faults that may occur in the process operation usually requires a priori knowledge of the process model. When this information is not available alternative methods that use statistical process analysis or intelligent – based reasoning can be applied.

An integrated system that implements automatic unexpected situation detection and management strategies is referred to as a fault-tolerant system [15]. There are four main stages in the fault handling process [5]:

- **Fault detection** that identifies an abnormal behaviour in the process output;
- **Fault identification** that determines the most relevant variables in diagnosing the fault;
- **Fault diagnosis** that determines the type of fault, as well as its location, magnitude and time when it occurred;
- **Process recovery or control reconfiguration** which tries to eliminate the fault effect by replacing the missing signal by an observed value.

1.1 Fault detection and diagnosis

First three stages presented above are commonly addressed as the fault detection and diagnosis (FDD) process. Different methods have been proposed in literature to find solutions for this problem: model-based [11] (including MPC - model predictive control [2]), based on structural analysis [7, 14], advanced computing methods (like artificial neural networks or fuzzy logic) [16, 17] and statistical analysis methods [4, 6, 18].

Considering real-time processes for which an analytical plant model is not available, statistical methods provide a handy approach for addressing fault detection and diagnosis, as they rely on historical process data.
1.2 Process reconfiguration

Process reconfiguration or system recovery after occurrence of a fault has the main purpose of ensuring the fault will not affect the stability of the process. A secondary objective is represented by achieving the best under-optimal solution when operating in a faulty state. As detailed in [9, 19], numerous researches using different methods have been conducted in this domain: model-matching, structural analysis, fuzzy-based methods or trajectory tracking approaches.

1.3 System configurations

There are two possible directions for implementing a fault tolerant control structure: having different controller models for each type of possible fault, or actively reconfiguring the control strategy each time a fault occurs [15]. The first approach proved to provide a better approach in handling possible performance degradation caused by the fault occurrence [20]. It also has an advantage from an increased speed in identifying a new control solution but needs more resources for storing and processing all available options. The second approach of an adaptive controller has an increased complexity as the controller is modified “on the way” but provides the possibility of identifying increased performance solutions.

The main purpose of this work is to identify most suitable fault detection and diagnosis methods for a standard water management application and to provide an automatic reconfiguration strategy that would use the quantification of the cross-correlation between process variables for identifying the best control alternative in presence of faults.

2. Methods and Algorithms

2.1 Fault-tolerant control method

Our fault analysis and control reconfiguration approach is based on the idea presented in [13] of “hiding” the faults by placing a reconfiguration block between the nominal controller and faulty plant at the reconfiguration time. This way, we can find a minimum invasive solution so that the reconfigured plant has the same behaviour as the nominal plant. The control scheme is illustrated in Figure 1. $u$ and $y$ represent the plant input and output in absence of faults, considering a standard control scheme. The fault-tolerant control strategy introduces the $uf$ parameter as the alternative control signal in case of a faulty actuator, $u^*$ as being the reset on the actuator control signal in case of fault detection and $yf$ as the process output through a faulty sensor.

![Figure 1. Fault-tolerant control based on fault-hiding technique](http://www.sic.ici.ro)
2.2 Statistical methods for FDD

Statistical methods for fault detection and diagnosis have two main approaches: univariate analysis, when the variation of each process variable is taken into consideration, and multivariate analysis, which takes into consideration also the relationship between different parameters. Each of them provides minimum (LCL – lower control limit) and maximum (UCL – upper control limit) values under normal plant operation. Exceeding one of these limits is signalled as a fault. First approach is easier to implement and provides easier interpretation for the process control engineer but may miss fault signals because it doesn’t take into consideration how simultaneous faults may affect the process dynamics. Second approach is characterised by increased accuracy and space reduction but its complexity makes it hard to implement in process controllers.

To deal with this limitations we can use a hybrid approach that implements univariate methods at the controller level, ensuring a fast reaction to individual device-specific faults, and multivariate methods at the supervisory level where advanced processing tools can be used to respond to inconsistencies in the dynamic behaviour of the plant as a whole.

2.3 Univariate parameter analysis

Statistical analysis techniques provide tools to define the expected variation of a parameter based on its prior behaviour. There are several methods for choosing the upper and lower control limits of a process variable, out of which the most common are Shewhart chart, exponentially weighted moving average (EWMA), X and R or S chart or individuals control chart [6]. Differences between these methods come from the detection speed of the process shifts and the computational effort.

Another important aspect regarding their applicability comes from the fact that most statistical process control methods assume several observations are available for each sample in the data set. This can be easily obtained in cases where several identical experiments can be conducted, or if we can read several values almost simultaneously from process devices, but in a real-time processes specific approaches that use a single data set should be used. In this case two methods are available [10]: Western Electric rules and individuals control chart. These methods illustrate the variation of the process measurements relative to the average value (represented as the center line CL), and to the upper and lower control limit (UCL and LCLC, respectively).

Considering a process data set

\[ X = [X_1, X_2, \ldots, X_n], X \in \mathbb{R}^n \]

of a considerable number of samples n, we can define the average value with the following equation:

\[ \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \]  

(1)

The standard deviation describes the variability of the data values. It is computed based on the variables sample average with the following formula:

\[ \sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2} \]  

(2)

The Western Electric rules compute a parameter’s plot center line as being the average value and considers a parameter as being out of control if its value is outside the following upper and lower limits defined by the following equations:

\[ UCL = \bar{X} - 3 \sigma \]  

(3)

\[ LCL = \bar{X} - 3 \sigma \]  

(4)

This is also called the 3σ rule. In addition to this, the occurrence of any of the following conditions signals a parameter fault:

- Two of three consecutive samples go beyond the ±2σ control limits;
- Four of five consecutive samples go beyond the ±σ control limits;
- Eight consecutive points plot on one side of the CL.

The individuals control chart method computes the moving range between two consecutive samples to provide a quantitative measure of the process variability. The moving range \( MR_i \) is defined by:

\[ MR_i = |X_i - X_{i-1}| \]  

(5)

The average moving range on all the n samples is given by:
An estimator for the standard deviation from this method is:

\[
\hat{\sigma} = \frac{MR}{d_2}
\]  

where \(d_2\) represents a tabulated constant dependent on the sample size, equal to 1.128 in this case because of the computing of the difference between two consecutive samples [10]. Using this method \(UCL\) and \(LCL\) are defined by:

\[
UCL = \bar{X} + 3 \times \frac{MR}{1.128}
\]  

\[
LCL = \bar{X} - 3 \times \frac{MR}{1.128}
\]

As the method uses the average of all process values at each step, the correct identification of a process fault depends on the magnitude of the fault measured value relative to the \(UCL\) and \(LCL\), respectively, as well as on the number of fault values \(m\) from the total number of samples \(n\). Under these conditions, it is necessary to determine the magnitude of the fault value that can be identified using this method and how the ratio between the fault values and normal operation values can affect its correct identification.

### 2.4 Multivariate parameter analysis

The most commonly used multivariate method for fault diagnosis are Principal Component Analysis (PCA) and PLS (Partial Least Squares) [5]. While other methods like Fisher’s Discriminant analysis (FDA) or Discriminant partial least squares (DPLS) have shown better accuracy in fault diagnosis [5] their increased complexity make them less suitable for practical implementations.

PCA is an efficient method for determining the variance of the data that reduces the initial data set to one of “principal components” or loading vectors that represent a linear combination of the original data set.

Let us consider a data matrix \(X \in \mathbb{R}^{n \times m}\), where \(n\) represents the number of observations and \(m\) is the number of variables, and \(x\) the last observation vector, \(x \in \mathbb{R}^n\). As detailed in [5], the loading vectors can be computed using SDV (Singular Value Decomposition):

\[
\frac{1}{\sqrt{n-1}}X = U \Sigma V^T
\]

where \(U \in \mathbb{R}^{n \times a}\) and \(V \in \mathbb{R}^{m \times a}\) are unitary matrices and \(\Sigma \in \mathbb{R}^{a \times a}\) is a diagonal matrix.

Critical limits can be derived from this method using the \(T^2\) statistic. The Hotelling \(T^2\) statistic provides an elliptical confidence area based on the covariance matrix between the observed parameters that accurately identifies potential outliers.

Considering \(P\) the loading vectors matrix associated with \(a\) largest singular values and \(\Sigma_a\) containing first a rows and columns of the singular values matrix, the Hotelling \(T^2\) statistic is given by:

\[
T^2 = x^T P \sum_a^{-2} P^T x
\]

The threshold for the \(T^2\) statistic is:

\[
UCL = T^2_a = F_{\alpha}(a,n-a) - \frac{(n+1)(n-1)\alpha}{n(n-a)}
\]

where \(F_{\alpha}(a,n-a)\) represent the upper 100\(\alpha\)% critical point of the \(F\)-distribution with \(a\) and \(n-a\) degrees of freedom. If the value for the \(T^2\) statistic is greater than \(UCL\) then a fault occurred.

As the result of the multivariate analysis is affected by variations of multiple parameters, additional actions must be performed to diagnose the cause. An overview of existing methods is presented in [1]. The main approaches are the investigation of the principal components that lead to exceeding the control limits or the graphical interpretation using Andrew curves or the cross-correlation ratios between process variables and principal components.

### 2.5 Virtual sensor estimation

The main objectives of control strategy reconfiguration are to ensure system stability and to track as much as possible the initial control loop. Considering the FDD step implemented using statistical analysis methods, increased execution performance can be achieved if we follow the same approach for the system reconfiguration step.
Faults of measuring devices can be overcome by taking into consideration the cross-correlation between process variables. In order to quantify the strength of the correlation we can compute the sample correlation coefficient \( r_{xy} \) of two variables sample vector \( X \) and \( Y \) as:

\[
r_{xy} = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2 \sum_{i=1}^{n} (X_i - \bar{X})^2}}
\]

(13)

The closer its value is to \( \pm 1 \), the stronger is the correlation between the variables and the sign of this coefficient indicates if we have a positive or a negative relationship [10].

Based on the assumptions that, under normal operation, the variation of a parameter \( X \) should follow a dynamic behaviour similar to the parameter \( Y \), and that each parameter has a normal distribution, we can estimate the value of \( X \) in the presence of a sensor fault based on its behaviour prior to the fault occurrence and on the sample correlation coefficient using the following approximation:

\[
X_n \approx \frac{(n-1)\sigma_{(n-1)} + \sigma_{xy}r_{xy} - cov_{xy}}{Y_n - \bar{Y}} + \bar{X}_{n-1}
\]

(14)

where \( \sigma_x \) and \( \sigma_y \) represent the standard deviations of the cross-correlated variables, \( cov_{xy} \) in the covariance between \( X \) and \( Y \), and \( \bar{X}_{n-1} \) is the average value of \( X \) prior to the occurrence of the fault.

In this case, the average value of \( X \) at step \( n \) is computed with the following recursive formula:

\[
\bar{X}_n = \frac{(n-1)\bar{X}_{n-1} + X_n}{n}
\]

(15)

In the same way, we obtained a recursive formula for computing the time series variance, given by:

\[
\sigma_{x_n}^2 = \frac{n-1}{n} \sigma_{(n-1)}^2 + \frac{1}{n-1} (X_n - \bar{X}_n)^2
\]

(16)

By considering the entire data set of \( X \) when computing the average values and sample variation, the accuracy of the approximated value may be affected by process behaviour in different operating modes. For this, a limited set of last measured variables should be used.

### 2.6 Virtual actuator implementation

Design of virtual actuators can be implemented only if at least one of the following conditions is true:

- the faulty actuator was designed in a redundant structure (for example in multiple pumps configuration where at least one is considered reserve);
- the process design allows fault isolation and activation of an alternative route (for example, in power or water management systems).

The virtual actuator will implement specific reconfiguration strategies considering actions both in case of single faults and for multiple faults. An implementation solution is by use of fault decision trees with different costs associated to multiple control solutions in applications with functional redundancy.

### 4. Experimental Results and Discussions

#### 4.1 Application

Using real-time process data from a water management system application, we investigated how we can improve the process reaction to faults by implementing additional state evaluation and control reconfiguration modules that would not interfere with the functionality of the controller for the steady-state operation of the plant. Until now, fault detection and diagnosis for water distribution systems have been proposed in literature using either fuzzy-based methods [8, 12] or statistical analysis using PCA [3]. Our approach uses the univariate method, parameter estimation and strategy reconfiguration for implementing a fault-tolerant control system.

System configuration is illustrated in Figure 2. The process consists in three wells, each equipped with a pump that will extract water and push it through the network pipes until it reaches the accumulation tank. The pumps are functioning in a 2 active and 1 stand-by configuration. We measure the water level in each well and in the tank, and also the water pressure and output flow on the main pipe, immediately after the link of each secondary pipe to the main pipe. The control objective is to maintain the tank water level between two predefined limits.
In the design phase we identified several technological interlocks and alarms that can be considered in the process operation:

1) A pump failure can be identified if the start or stop command is not executed.
2) A sensor failure is signaled if the input signal is less than 0 (under 4 mA).
3) If flow in any of the measuring points exceeds the value of 7 m/s than the main collecting pump is broken.
4) If flow is 0 and its corresponding pump is on than its secondary pipe is broken.
5) Pump operation is 2 active and 1 reserve, and the switching is performed manually by the plant operator considering total functioning hours.

These conditions were implemented in the standard operation of the process control strategy. Limitations of this standard approach come from the reaction to the fault occurrence, rather than predicting a possible abnormal operation and also from the limitation in providing solutions for addressing all possible faults. For example, fault 1 can be caused not only by a pump failure, but also by a controller input or output module failure or a power relay failure. Each of this causes may have different solutions but a correct fault diagnosis is needed.

In order to improve process reaction to unexpected situations, we designed a fault detection and diagnosis module and interconnected it to a process reconfiguration module.

4.2 Implementation and results

The implemented fault tolerant structure is the one presented in Figure 1. We analysed the benefits and limitations of using either PCA multivariate or X-chart univariate fault diagnosis and identification techniques. For each case we considered a pressure sensor fault resulting in a step variation with a magnitude of 6% over the average value. We chose this value based on the computed value of the univariate UCL limit. The system reconfiguration module was designed following the fault-hiding technique with statistical estimation in case of a sensor fault and predefined reconfiguration model in case of an actuator fault.

In the univariate analysis this results in a variation slightly above the computed upper control limit. As illustrated in Figure 3, even this slight variation was correctly identified as a process fault. Also, fault identification accuracy is not affected by adding more faulty signals to the data set but its performance rises as the time period for capturing the process behaviour is higher.

We applied the PCA multivariate method on the same data set, including also the flow and level parameters. The $T^2$ squared distribution is illustrated in Figure 4. As can be seen, the fault is again correctly detected but for the identification of the faulty parameter we need additional operations. For this, the $T^2$ threshold can be computed on 2 by 2 data sets where occurrence of a fault can be identified as one of the variables being the faulty one.

Fault identification steps were illustrated in Figure 5. Figure 5a) indicates a fault in the simultaneous analysis on pressure and flow and Figure 5b) indicates no fault for the flow and level analysis. Thus we can identify the pressure as being the faulty parameter.

For the process reconfiguration we analysed possible cross-correlations between process variables in order to design a virtual sensor that would replace the faulty signal of the pressure value.
We computed the sample correlation coefficient between the three variables associated with a pump functioning. The resulted scatter plot is illustrated in Figure 6.

Because this step is performed offline, on a known process behaviour, fault data was not used here. As can be seen, these three parameters $C1$, corresponding to pressure, $C2$, corresponding to flow, and $C3$, corresponding to level, are strongly related and have the correlation coefficients equal to:

- $r(C1, C2) = -0.8130$
- $r(C1, C3) = 0.8526$
- $r(C2, C3) = -0.9735$

As the maximum correlation value for pressure is obtained from the level value, the virtual pressure sensor value was computed starting from the after fault variation of the level parameter.
Figure 7 shows the evolution of the computed value for the pressure parameter in the presence of faults, based on its correlation to the level parameter. The estimated signal is represented by the dashed line. As can be seen the moment the signal fault is detected, the virtual sensor is activated and the estimated value is transmitted to the controller.

![Figure 7. Pressure estimation in presence of faults](image)

A second experiment was conducted to evaluate the estimation performance based on the proposed method. The evaluation was conducted on a set consisting in the last 20 data samples. Figure 8 presents a comparison between the measured value (represented by the straight line), its estimation considering a constant value of the sample correlation coefficient (represented by the dashed line) and its estimation after recalculating the sample correlation coefficient at each step (represented by the dotted line).

![Figure 8. Comparison between real pressure variation (straight line), estimated value with r constant (dashed line) and estimated value with r recalculated (dotted line)](image)

As can be seen, a better estimation is achieved by maintaining the initial high sample correlation coefficient. This is because the error between the estimated value and the real value results in a weaker correlation to the associated parameter and also a less accurate approximation of the virtual sensor.

As this application has three pumps with similar characteristics and functionality, we implemented an actuator fault handling that processes the commands issued by the controller and analysis it in the context of existing faults received from the FDD module.

A simplified strategy for actuator control reconfiguration considering a 3 pump system is illustrated in Figure 9. If the controller issues a start command for an unfaulty pump, the command is sent to the actuator. In case the request was issued for a faulty pump, the strategy will investigate if other pumps are available. If so, it will send the command to the corresponding pump. To ensure process stability, the FDD module must also implement the necessary adjustments at the pump feedback to simulate a normal operation at the controller level.

![Figure 9. Actuator control reconfiguration](image)

4. Conclusions

In this paper it is identified a control methodology suitable for implementing fault tolerant control based on process historical data and physical control structure. The methodology was validated on a case study for a simple water management application. We analysed the balance between increased performance and reliability of FDD multivariate methods, compared to the ease of implementation on univariate solutions. The conclusion was that univariate solutions provide satisfactory results for this type of...
applications and can be implemented efficiently at the controller level using recursive functions, thus being preferred to multivariate solutions. Even more, the computed mathematical parameters can be used to estimate the value of a virtual sensor. Also, a methodology for implementing a virtual actuator was proposed.

Our main objective was the identification of practical limitations concerning both the complexity of the algorithms and the adaptability to existing control applications. Future work will be focused on identifying better statistical methods for estimating the variation of a sensor which entered a fault state and evaluating the proposed method on a multivariate control application.

Acknowledgements

This work was supported by the Romanian National Research Programme PNII, project: Cloud Architecture for an open Library of Complex re-Usable Logical function blocks for Optimized Systems – CALCULOS.

REFERENCES


