Diagnosis and Dynamic Data Reconciliation in Complex Technological Systems

MINCHO B. HADJISKI, NIKOLINKA G. CHRISTOVA*, PETER P. GROUMPOS*
Department of Automation in Industry
University of Chemical Technology and Metallurgy
Bul. Kl. Ohridski, 8, 1756 Sofia
BULGARIA

* Laboratory for Automation and Robotics
Dept. of Electrical and Computer Engineering
University of Patras, GR-265 00 Rion,
GREECE

Abstract: In this paper, data verification and data reconciliation methods, based on the mass and energy balances in complex technological systems are considered. Diagnostic procedures for gross and systematic measurement errors detection and their isolation in dynamical mode are developed. The fuzzy logic methods are used for the problem solving of fault measurement diagnosis and data reconciliation. Some results of the proposed methods application on chemical plant are presented. CSCC'99 Proceedings, Pages:4831-4837

Key-Words: - dynamic data reconciliation, fault measurement diagnosis, mass and energy balances, fuzzy logic-based diagnosis and correction, soft computing method, complex technological systems

1 Introduction

Because of the increasing demands on reliability and safety of complex technological systems and their elements, methods for improving the supervision and monitoring as a part of the overall control of the processes are getting an increasing interest. The problem of diagnosis and correction of measurement errors in dynamic data provided from continuous process is important for the correct control and operation of the complex systems [1], [4], [6]. As a result a smooth management and dynamic operation of the plant can be achieved. Process measurements may contain (i) random errors, normally introduced by the instrumentation systems; (ii) gross errors, caused by different external forces interfering with the system; (iii) systematic errors, usually caused by improper operation or incorrect adjustment of the measuring instruments. These errors make raw measured data show discrepancies in fundamental physical laws.

The incorrect usage of the local balances may leed to serious difficulties in the integrated system operation. This is due to the impossibility to realise optimal process scheduling or adequate operational actions in the presence of considerable internal and external disturbances [4], [6]. The modern complex production systems consist several scores of technological units, connected by feedforward meterial streams and recycles. There are many material as well energy invertors of different types and construction. The technological units themselves (reactors, distilation couloms, absorbers etc.) are accumulators of mass end energy. Special buffer tanks are provided for to relieve the mass flows control. The mass and energy balances control is a considerable part of the integrated control of the plant-wide, which is formed at the higher hierarchical levels [4]. The range of each local balance may be quite different and it is defined of the general strategy of the coordinated control.

In this study the local balances are described on the base of generalized models, which include the direct and indirect measurements of inlet and outlet streams parameters (flow rate, temperature, pressure), as well the state parameters of the technological units and equipment, where mass and energy are accumulated.

The usual steady-state approach for estimation of the mass balances and the measurements correction [2], [3], [8], [9] in the most of the cases is insufficient precise for the operational control of the balances because of dynamical errors. Due to the various temporal scale of the subsystems, different requirements for reconciliation and control of the local balances are created. Some approaches for dynamical diagnosis and data reconcilation, which add and replace the statical ones are proposed below. A modification of Fuzzy Logic-Based diagnosis and data reconciliation method [8], created for steady-state mode to be applied for dynamic modes of complex technological systems is presented.

2 Statement of the Problems

Several modes of operation are possible in the complex integrated manufacturing systems (Fig. 1):

a) Steady-state or close to it regime with small random deviations of the thrue values of the technological parameters and lack of accumulation in the units.

b) Change of the operational point of a part or all of the units due to the supervision control signals of scheduling system or unforeseen operational actions.

c) Regimes of considerable changes in the technological mode – change of technological structure (switch-on, shut-down, malfunctions).

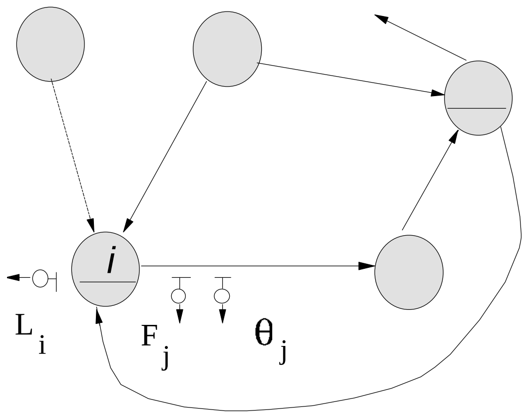


Fig. 1. Integrated production system

The existing methods for steady-state diagnosis and data reconciliation [2], [3], [8], [9] can not operate in the conditions of wide-spread considerable technological disturbances into the whole system.

Dynamic diagnosis and correction of the measurements must be performed in the case of unsteady-state behavior of the systems [1]. Dynamic data reconciliation is closely related to the field of nonlinear state estimation. The solution of the dynamic data reconciliation problem using a complex sequential-modular model generally is a difficult problem. An approach for solving this problem is to exploit the modular structure of the models and to use different data reconciliation algorithms for each type of module [1]. Several dynamic data reconciliation methods have been proposed for the equation oriented case [5], [7].

The next task is formulated: The problems of fault measurement diagnosis and data reconciliation in steady-state and dynamic modes of the process operation, using generalized approach must be solved.

Let us assume that the continuous industrial system under investigation consists of *m* units and *n* streams. The index set **I** = {1, 2, ..., *m*} represents the units numbers of the technological system. Similarly **J** = {1, 2, ..., *n*} is the index set containing the streams numbers.

- ***Material balance***

In dynamic operation of the system the following mass balance equations hold:

$$\rho_i \frac{d\Delta V_i(t)}{dt} = \sum_{j \in J} a_{ij} F_j(t), \quad i \in I \qquad (1)$$

where:

ρ_i – density of the fluid [kg/m^3];

$\Delta V_i(t)$ [m^3] is the increment of the accumulated volume (tank, vessel, part of unit) and is defined as $\Delta V_i(t) = V(t) - V(t_0)$ (t_0, t – initial and current time [s]);

a_{ij} are the coefficients of the (*m* x *n*) - dimensioned *Incidence Matrix* **A** with elements as follows:

a_{ij} = +1, if j is an inlet stream for unit i;

a_{ij} = -1, if j is an outlet stream for unit i;

a_{ij} = 0, if the j-th stream is not connected to the i-th unit;

$F_j(t)$ [kg/s] is the flow rate of the j-th stream at the moment t.

Usually the volume increment ΔV could be presented as follows:

$$\Delta V_i(t) = \Omega_i \Delta L_i(t), \qquad (2)$$

where:

Ω_i – cross-section of the i-th accumulator, [m^2];

$\Delta L_i(t)$ is the level increment of the i-th unit at moment t and is calculated as:

$$\Delta L_i(t) = L_i(t) - L_i(t_0) \qquad (3)$$

The liquid level into the i-th accumulator is measurable, [m].

In more complicated cases nonlinear equation should be used instead of the linear one (2):

$$\Delta V_i(t) = f_i(\Delta L_i(t)). \qquad (4)$$

To eliminate the stochastic disturbances, integration of the mass flow rate measurements $F_j(t)$ is suitable during the interval t_f. Equation (1) could be presented in the following form:

$$\rho_i \Delta V_i(t_f) = \sum_{j \in J} a_{ij} \int_0^{t_f} F_j(t)dt, \quad i \in I \qquad (5)$$

We denote the j-th integral flow rate $G_j(t_f)$, [kg] as:

$$G_j(t_f) = \int_0^{t_f} F_j(t)dt \qquad (6)$$

Equation (5) could be presented in the form:

$$\sum_{j\in J} a_{ij}G_j(t_f) + \rho_i \Delta V_i(t_f) = 0, \quad i \in I \qquad (7)$$

If we assume that only *r* units are accumulators and taking account (2), equation (7) could be presented as a function of measurable (directly or indirectly) variables:

$$\sum_{j\in J} a_{ij}G_j(t_f) + p_i d_i \Delta L_i(t_f) = 0, \quad i \in I \qquad (8)$$

where:

$$p_i = \begin{cases} 0, & \text{if } i \le r; \\ 1, & \text{if } r < i < m; \end{cases} \qquad (9)$$

$d_i = \rho_i\Omega_I$ - known parameter for the i-th accumulator.

- ***Heat balance***

The dynamic heat balance equation of every unit may be described by:

$$\rho_i c_i \frac{d\Delta(V(t)\theta(t))_i}{dt} = \sum_{j\in J} a_{ij} c_j \theta_j(t) F_j(t) + q_i(t), \qquad (10)$$

where:

c_i – specific heat of the fluid in the i-th unit, [kJ/kgK];

$\theta(t)$ – temperature into the i-th unit at moment t, [K];

c_j – specific heat of j-th stream, [kJ/kgK];

$\theta_j(t)$ – temperature of j-th stream at moment t, [K];

$q_i(t)$ – intensity of heat generation in i-th unit (accumulator) due to external heating chemical reaction at moment t, [kJ/s].

Following the same way as above we have:

$$\rho_i c_i \Delta(V\theta)_i\Big|_{t_f} = \sum_{j\in J} a_{ij} Q_j(t_f) + u_i(t_f), \qquad (11)$$

where:

$$Q_j(t_f) = c_j \int_0^{t_f} \theta_j(t) F_j(t) dt$$

$$u_i(t_f) = \int_0^{t_f} q_i(t)dt \qquad (12)$$

$$\Delta(V\theta)_i\Big|_{t_f} = V(t_f)\theta(t_f) - V(0)\theta(0)$$

Equation (11) has another representation, substituting (2) and (9):

$$\sum_{j\in J} a_{ij} Q_j(t_f) + p_i g_i \Delta S_i(t_f) + u_i(t_f) = 0, \qquad (13)$$

where:

$g_i = d_i c_i$; (14)

$\Delta S_i(t) = \Delta(V(t)\theta(t))_i$. (15)

3 Error Correction

Statement of the problem of the measurement error correction is described below. Further, we will deal with the solving of the data reconciliation problem, based on the mass balance.

As it was assumed above the continuous industrial system consists of *m* units and *n* streams. The index set **I** = {1, 2, ..., *m*} represents the units numbers (each of the unit volume or level is measured) and the numbers of the mass balance equations φ_i, ($i \in$ **I**) for the system. Similarly **J** = {1, 2, ..., *n*} is the index set containing the streams numbers, each of them being measured as a total or component mass flow rate F_j, $j \in$ **J**. The measuring instruments for F_j and L_i have the same numbers in **J** and **I**. The index set **R** is defined as **R** = **J** ∪ **I**.

If all the flow rates F_j, $j \in$ **J** and all the levels L_i, $i \in$ **I** are measured by respective measuring instruments or in some other way, then from a balance point of view all F_j and L_i are called *over determined* balance parameters. It is assumed that the measurements are taken in discrete samples with a sampling time T_f. However, instead of their exact values, the values $F_j^*(t_k)$ and $L_I^*(t_k)$ (k is an integer sampling index) including measurement errors $e_j(t_k)$ and $\varepsilon_i(t_k)$ are measured (observed) as follows :

$$F_j^*(t_k) = F_j(t_k) + e_j(t_k),\ j \in \mathbf{J}; \qquad (16)$$

$$L_i^*(t_k) = L_i(t_k) + \varepsilon_i(t_k),\ i \in \mathbf{I}. \qquad (17)$$

The errors $e_j(t_k)$, ($j \in$ **J**) and $\varepsilon_i(t_k)$, ($i \in$ **I**) can be either *gross* or *systematic errors*. Further we assume long time acting errors refereed to as *systematic errors*. From statistical viewpoint they can be regarded as random variables having a normal distribution with a standard deviation σ_r, ($r \in$ **R**) and unknown but *non zero mean* e_j^o and ε_i^o.

Taking into account the definitions in the previous section, in dynamic operation of the system the material balance equations (8) may be obtained. The integral flow rates $G_i(t_f)$, $j \in$ **J**, for the interval t_f will be used in the next consideration instead the measurement flow rates $F_j^*(t_k)$, $j \in$ **J**, at the moment t_k. The relationships of the integral flow rates $G_i(t_f)$, $j \in$ **J**, and the measurement flow rates $F_j^*(t)$, $j \in$ **J**, could be presented as follows:

$$G_j(t_f) = \int_0^{t_f} F_j^*(t)dt = G_{j1}(t_f) + G_{j2}(t_f),\ j \in J, \qquad (18)$$

where:

$G_{j1}(t_f)$ is the true value of the integral flow rate of the j-th stream;

$G_{j2}(t_f)$ is the error estimation of the integral flow rate of the j-th stream and will be obtained from the solving of the data reconciliation problem, that is given below.

In the proposed Soft Computing Method in [8] the assumption for the faulty status of each instrument is made in a *fuzzy manner* as an *n+m*-dimensional vector called *Hypothesis Pattern* **H** = { h_1, h_2, ... h_r, ... , h_{n+m}}, where $0 \le h_r \le 1$, $r \in \mathbf{R}$ represents the preliminary assumed ***faulty degree*** of each measuring instrument.

By introducing the Lagrange multipliers λ_i and μ_i, $i \in \mathbf{I}$, the problem of dynamic data reconciliation is transformed into solving the following (2*m*+*n*) x (2*m*+*n*) system of linear equations :

$$\sum_{j \in J} c_{ij} x_j + \eta_i x_{n+i} = b_i^F + b_i^L, \quad i \in \mathbf{I} \qquad (19)$$

$$h_j x_j + \sum_{i \in \mathbf{I}} c_{ij} \lambda_i = 0, \quad j \in \mathbf{J} \qquad (20)$$

$$h_{n+i} x_{n+i} + \eta_i \mu_i = 0, \quad i \in I \qquad (21)$$

where:

x_r, $r \in \mathbf{R}$ are the relative errors;

c_{ij}, b_i^F, b_i^L and η_i are normalized coefficients of the balance equation system (8) calculated as follows:

$$b_i^F = \frac{\sum_{j \in J} a_{ij} G_j(t_f)}{\sqrt{\sum_{j \in J} a_{ij}^2 \sigma_j^2}}, \quad i \in I \qquad (22)$$

$$b_i^L = f_i(\rho_i, L_i^*, \sigma_{n+i}), \quad i \in I \qquad (23)$$

$$c_{ij} = \frac{a_{ij} \sigma_j}{\sqrt{\sum_{k \in J} a_{ik}^2 \sigma_j^2}}, \quad i \in I; \quad j \in J \qquad (24)$$

$$\eta_i = \phi_i(\rho_i, \sigma_{n+i}), \quad i \in I \qquad (25)$$

The diagnosis of measurement errors can be performed as *n+m*-dimensional non-linear optimization task where the vector **H** represents *one feasible point* in the searching procedure. The following performance index is proposed to be used:

$$Q = K \sum_{r \in R} h_r \sigma_r + \sum_{r \in R} (1 - h_r) |x_r| \qquad (26)$$

The coefficient K is introduced as a kind of *penalty* for wrong guesses of the numbers of the faulty instruments. Its value is selected in a subjective way, but is advised to be more than the maximal expected systematic error. Thus the criterion (26) favors the diagnosing solutions having smaller possible number of faulty instruments.

Due to existence of many equivalent solutions [2], [8], the above criterion (26) represents a multiextremal function. Then a respective method of finding a *limited number* of equivalent solutions with a given *threshold* Δ would be very useful from a practical point of view.

4 Data Preparation for the Diagnosis Procedure and Data Reconciliation

Due to the interconnections in complex systems, the different temporal behavior of separate units and the measurement errors it is necessary to be provided with:

- a synchronization of the data, which are used in equations system (8);
- a data filtration for decreasing of the random errors in the measurable or inferable variables;
- a robust estimation of the mass accumulation.

In this connection, it is reasonable the technological units to be conditionaly divided in «fast» and «slow» units. It is assumed all of the units, which have own frequency at least three times greater than that one of the «slowest» unit, to be called «fast» ones [6]. As the «slow» units generate input signals with low frequency spectrum, it is advisable to be used different algorithms for data preparation of «fast» and «slow» variables and «fast» and «slow» technological nodes. The «slow» technological units and variables are setting for the data synchronization.

A) «Slow» units

A steady-state test verification of the input and output variables is performed. If a steady-state situation is exist, the values of the input and output variables are mean, using moving average method:

$$\tilde{v}(t) = \frac{1}{T} \int_{t-T}^{t} v(t) dt \qquad (27)$$

The derivatives in the left part of equations (1) and (10) for the accumulation estimation are defined on this basis.

If an unsteady-state situation is exist, the next sequence is applied:

- filtering of the input and output data by means of a low order filter;
- estimation of a consequence of the accumulated mass.

B) «Fast» units

A moving average method is used for the data filtration by obtaining massives of average data at each moment of time, defined by the «slow» units.

It should be noted, the steady-state of the input and output signals doesn't guarantee a steady-state of the accumulated mass in them, due to the integrating properties of the tanks, receivers and gas-holders. Because of that the usage of the dynamic balance equations is obligatory.

The accumulating components in (1) and (10) should be presented in the terms of measurable variables – level, pressure and temperature. Their values calculating can be accomplished as follows:

- through direct approximately setting of the derivatives in the left part of equations (1) and (10) on the basis of the filtrated input and output variables;
- on the basis of the mathematical models, if such ones exist, of the correspondent units;
- combined usage of both of the above approaches.

5 Diagnostic Procedures

In the followings a simple *heuristic method* of finding *equivalent solutions* for the fault diagnosis procedure is proposed. It uses random search technique where a preliminary given number of P random Hypothesis Patterns (vectors) is generated. The main steps of this optimization method are as follows:

1. Generate randomly the *current* Hypothesis Pattern (vector) $\mathbf{H}_p = \{ h_1, h_2, ... h_r, ... h_{n+m}\}, (1 \le p \le P)$;

2. Perform the modified measurement correction procedure (19), (20) and (21) for the generated vector $\mathbf{H}_p$. Calculate the respective relative errors x_r, $r \in \mathbf{R}$.

3. *Fuzzify* the relative errors x_r obtained, according to the maximum expected value x_{rmax} as a deviation from the normal (true) case [8]. As a result the vector $\mathbf{S}_q = \{s_1, s_2, ..., s_r, ..., s_{n+m}\}$ is obtained.

4. By using the *C-means* algorithm for *cluster analysis* [8] divide the elements in $\mathbf{H}_p$ as well as the elements in $\mathbf{S}_p$ into 2 *crisp* clusters representing *true* and *faulty* measurements. Then the measurements defined as clearly *faulty* in both vectors $\mathbf{H}_p$ and $\mathbf{S}_p$ form the respective clusters $\mathbf{C}_H$ and $\mathbf{C}_S$. Here $\mathbf{C}_H$ represents the hypothesis for faulty instruments while $\mathbf{C}_S$ is the calculated faulty status of the instruments.

5. Check the *feasibility* of the current hypothesis $\mathbf{H}_q$. If $\mathbf{C}_S \subset \mathbf{C}_H$ then a conclusion that $\mathbf{C}_S$ is a *candidate solution* is made. Otherwise *Go to Step* **7**.

6. Calculate performance index Q for the current candidate solution $\mathbf{C}_S$ according to (26). If $Q \le Q_{min} + \Delta$, (Q_{min} is the current minimal obtained value) then this candidate solution is called *equivalent candidate solution* and it is saved, otherwise it is discarded.

7. If $p \le P$ *Go to Step* **1**; otherwise *Stop*.

The next specific characteristics of the dynamic fault measurement diagnosis and data reconciliation can be noted:

1. At the difference of the steady-state approach in this case additional variables are included, which are connected to the state behavior of the technological nodes – such as a level of the tanks and units, a pressure, a temperature. In such way the number of the parameters, which have to be diagnosed, increases.
2. Due to the estimation inaccuracy of the mass and energy accumulation it is necessary a frequent repetition of the diagnostic procedures performance to be carry out at all the same hypotheses of the fault measurement (Fig. 2). This puts on the most computationally efficient algorithms to be used in the procedures of dynamic diagnosis and data reconciliation.

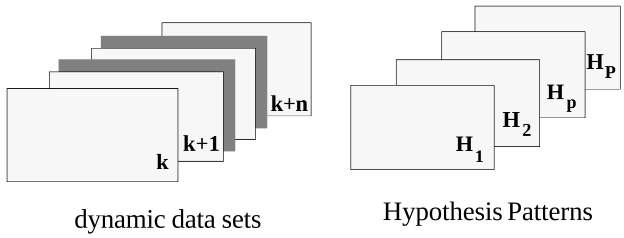


Fig. 2. Diagnostic procedure

6 Application and Analysis of the Results

This section shows the principal possibility of real application of the above proposed Soft Computing Method for diagnosis and dynamic correction of the measurements of mass flow rates and levels in real plants with complex structure. In Fig. 3 the structure of one subsystem of a plant for production of ammonia (NH_3) is presented. The main goal of this subsystem is to remove the contained CO gas in the so called process gas. There are 10 technological units, as follows: No. 1 and 2 are converters where CO is converted into CO_2; No. 3 is reactor for oxidation of CO into CO_2; No. 5 is a separator; No. 6 is a column for absorption of CO_2; No. 8 and 9 are Desorption columns where the process solution is regenerated. All of the mass flow rates F_j, j = 1,…, 22 as well as the levels of the units No. No. 5, 6, 8, 9 are measured.

The number of assumed faulty measurements in this example is 4. They are denoted as bold streams in Fig. 3 as follows: F_5, F_6, F_{12} and F_{15}.

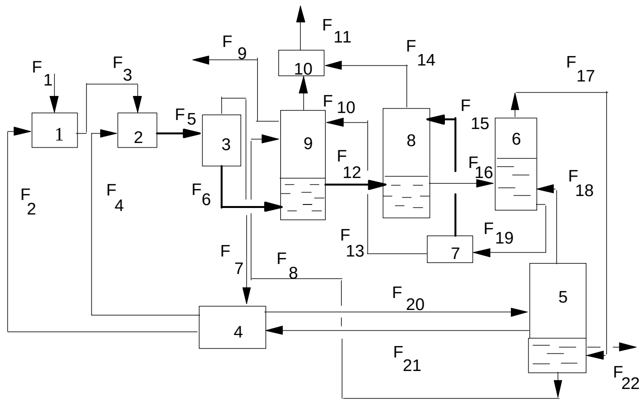
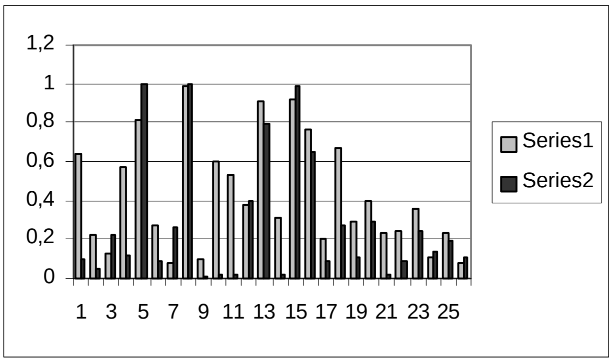


Fig. 3. Structure of a subsystem of a plant for production of ammonia (NH_3)

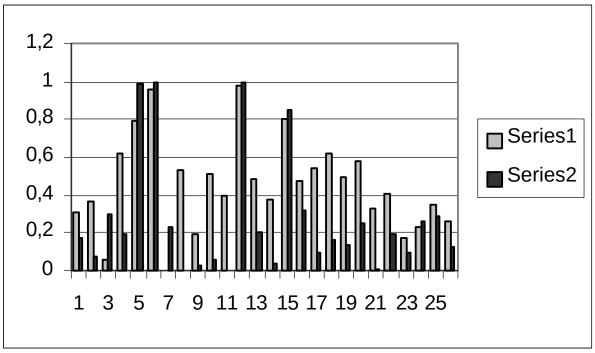
Applying the above described Soft Computing Method for diagnosis of faulty instruments, two equivalent candidate solutions have been found and shown in Fig. 4. They are obtained within a relatively small number of random generated Hypothesis Vectors (P = 150) for a threshold Δ = 200. The real solution F_5, F_6, F_{12} and F_{15} is also among them. In Fig. 4 Series 1 denotes the hypothesis vector $\mathbf{H}_p$ and Series 2 denotes the candidate solution $\mathbf{S}_p$. The cluster $\mathbf{C}_H$ of faulty instruments for the hypothesis vector and the cluster $\mathbf{C}_S$ of the faulty instruments for the candidate solution are presented too. Performance index $\mathbf{Q}$ is also depicted in Fig. 4 together with the number of iteration *It* for each candidate solution obtained.

It = 67; $\mathbf{Q}$ = 958.7;
$\mathbf{C}_H$ = {1, 4, 5, 8, 10, 11, 13, 15, 16, 18};
$\mathbf{C}_S$ = {5, 8, 13, 15, 16}



a)

It = 106; $\mathbf{Q}$ = 793.4;
$\mathbf{C}_H$ = {4, 5, 6, 8, 10, 12, 15, 17, 18, 20};
$\mathbf{C}_S$ = {5, 6, 12, 15}



b)

Fig. 4. Equivalent candidate solutions in diagnosing measurement errors

By increasing the number P of random generated hypothesis vectors $\mathbf{H}_p$ more equivalent candidate solutions can be found within the given threshold Δ. By decreasing the threshold candidate solutions with smaller number of faulty instruments is obtained. Finally an additional analysis of the plausibility of the solution obtained is needed taking into account some other data or knowledge for the system under investigation.

7 Conclusions

In this paper diagnosis procedure and dynamic data reconciliation method, based on the mass and energy balances in complex technological systems have been presented. At the distinction of the steady-state approach in this case additional variables are included, which are connected to the state behavior of the technological nodes. An optimization method of finding equivalent solutions for the fault diagnosis procedure has been developed. In the proposed Soft Computing Method the faulty status of each instrument is made in a fuzzy manner as a Hypothesis Pattern represents the preliminary assumed faulty degree of each measuring instrument.

An extensive analysis and evaluation of the methods proposed have shown their ability to be implemented into distributed industrial control systems for real time measurement correction and diagnosis.

References:

[1] Becerra V., P. Roberts and G. Griffiths, Dynamic Data Reconciliation for Sequential Modular Simulators, *Conference Publication No. 455, IEE*, 1998, pp. 1230-1235.

[2] Christova N., *Ph.D. Thesis on Diagnosis and Correction of Complex Measurement Errors in Technological Systems*, Sofia, 1999.

[3] Crowe C.M., Data Reconciliation - Progress and Challenges, *Proceedings of PSE'94*, Seoul, 1994, 111.

[4] Jamshidi M., *Large-Scale Systems: Modelling and Control*, Elsevier Science Publishing, New York, 1983.

[5] Liebman M., T. Edgar, L. Lasdon, Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques, *Computers in Chem. Eng.*, 16, 1992, pp. 963-986.

[6] Ng C. and G. Stephanopoulos, Synthesis of Control Structures for Chemical Plants, *Proceedings of the IFAC LSS'98*, July 15-17 1998, Rio Patras, Greece, pp. P16-P28.

[7] Ramamurthi Y., P. Sistu and B. Bequette, Control Relevant Dynamic Data Reconciliation and Parameter Estimation, *Computers in Chem. Eng.*, 17, 1993, pp. 41-59.

[8] Vachkov G. and N. Christova, Soft Computing Method for Solving the Problem of Diagnosis and Correction of Measurement Errors, Computers Chem. Engng., v. 20, Suppl., 1996, S659-S664.

[9] Yamamura K., M. Nakajima, H. Matsuyama, Detection of Gross Errors in Process Data Using Mass and Energy Balances, *Int. Chem. Engineering*, v. 28, No. 1, 1988, 91.