A New Comprehensive Sensor Network Design Methodology for Complex Nonlinear Process Plants

Mohammadnia, Vahid; Salahshoor, Karim*+

Department of Automation and Instrumentation, Petroleum University of Technology, Tehran, I.R. IRAN

ABSTRACT: *This paper presents an optimal integrated instrumentation sensor network design methodology for complex nonlinear chemical process plants using a Combinatorial Particle Swarm Optimiazation (CPSO) engine. No comprehensive sensor network design approach has been addressed yet in the literature to simultaneously incorporate cost, precision and reliability requirements for nonlinear plants. The presented approach attempts to accomplish this objective via enhancement of the estimation accuracy of the aimed instrumentation sensor network subject to desired cost, reliability and redundancy constraints. An Unscented Kalman Filter (UKF)-based data reconciliation algorithm has been developed to present evaluating measures through comparisions of the estimated and real variables in terms of Modified Root Mean Squared of Error (MRMSE), while CPSO maintains the provisions of the Network Fault Tolerence (NFT) including sensor netowrk reilability (R), strong and weak redundancy degrees (i.e., SRD and WRD). The developed CPSO engine searches in a diverse variety of possible sensor networks to adopt the most fitted one based on the imposed NFT and cost design constraints. The effective capabilities of the proposed design methodology has been illustrated in a simulated nonlinear Continuous Stirred Tank Reactor (CSTR) as a complex process plant benchmark.*

KEY WORDS: *Sensor network design, Unscneted kalman filter, CPSO, MRMSE, Reliability, Redundancy degree.*

INTRODUCTION

Measurements of all process variables are not practically cost-effective and yet operationally feasible in complex industrial plants. Accordingly, only a limited number of process variables are decided to be measured directly and hence reconciliation techniques could be beneficial for estimating the non-measured variables using the process model dynamics. Generally, a sensor network design methodology mainly deal with location or/and precision of sensors in large-scale plants so that some desired criteria, viz: observability [1], precision [2, 3], reliability of estimation of variables [4,5] and gross and error detectability [6,7].

Bagajewicz [8] used a tree type enumeration procedure to design a minimal cost network subject to constraints on precision, availability, resilience and error detectability. He proposed a design strategy that incorporates these criteria simultaneously for linear systems and suggested a MINLP to solve the problem. Further, *Bagajewicz & Sanchez* [9] showed that problem of minimizing the variance subject to cost constraint can be converted to the problem of minimizing the cost subject to the variance constraints via determining measurement locations in linear networks. *Bagajewicz & Cabrera* [3] presented a new MILP

* *To whom correspondence should be addressed.*
+ *E-mail: salahshoor@put.ac.ir*
1021-9986/12/3/ 7/$/2.70

formulation, replacing the previous tree search solution procedures for minimizing cost subject to explicit constraints of precision, error detectability, resilience and availability. Their method has been developed upon non-redundant linear systems and hence the resulting network evaluation was performed in steady state mode. *Sen et al.* [10] integrated graph theory and genetic algorithm concepts to develop a generalized sensor network design algorithm for non-redundant linear mass flow processes. In comparison with graph-theoretic algorithms [4], GA-based method provides near optimal solutions. Furthermore, in contrast with the previous methodologies which were able to design networks subject to one measure; general objective designs such as optimizing cost, estimation accuracy and network reliability ones could be addressed with this method. *Bagajewicz et al.* [11] developed an instrumentation network design scheme that could reflect the potential benefit of adding sensors in networks and used value and cost concepts separately for the integrated design, enabling to satisfy fault detection, material accounting and control criteria simultaneously. *Kotecha et al.* [12] proposed a duality between the precision and reliability problems for non-redundant sensor network design in linear processes. This method enables one to convert any reliability design measure to precision framework and use explicit optimization algorithm, which was already developed for precision-based design [3], to design sensor network in the precision domain, satisfying initial reliability requirements. As seen almost all the research performed in this field are limited to the non-redundant linear plants.

Staroswieck et al. [13] addressed the problem of fault tolerant estimation and the design of fault tolerant sensor networks. They defined fault tolerance regarding to a principle that a given functional of the system state should remain observable when sensor failures occur. All sensor sets were shown in an automaton which contains all the subsets of sensors such that the estimation objective can be achieved. They introduced three criteria evaluating the system fault tolerance with respect to sensor failures when a reconfiguration strategy is used: (strong and weak) Redundancy Degrees (RD), sensor network Reliability (R), and Mean Time To Non-Observability (MTTNO). Sensor networks are designed by finding redundant sensor sets whose RD and/or R and/or MTTNO are larger than some specified values. However, their regressive method works well on small designs, but in large-scale plants that have many sensor sets to be examined their method does not work well. Moreover, their optimized algorithm, which searches for a solution that satisfies design constraints and uses a minimum number of sensors in its topology, merely investigates existence of the sensors in the network and the variety of sensors, has not been considered. This method works well in one-sensor based designs. But, when it comes to designing of networks with multiple sensors available for measuring a variable, due to the drastic increase of number of possible solutions, the calculation effort highly increases and the presented algorithm fails. In addition, a main criterion in instrumentation design procedure, i.e. cost of instrumentation, has been neglected in this approach.

Only a few works have addressed both type and location of sensors simultaneously; *Muslin et al.* [2] discussed both location and type of sensors in precise linear sensor network designs. If type of sensors is not a consideration in design procedure, number of possible networks that can be constructed via given set of sensors decrease drastically. Moreover, most of the current design formulas have used static reconciliation technique to estimate the variables. Implementing this type of reconciliation does not take much time in design. But in contrast with steady state reconciliation, dynamic data reconciliation techniques should be performed during a period of time and this consumes much time to be implemented. Employing the two mentioned issues, i.e. static reconciliation technique and neglecting the variety of sensor type in the design procedure can lead to a substantial saving of design time. Subsequently, designer can take advantage of the saved time in order to implement designs based on the enumeration methods. These methods try to examine all possible candidates based on a logical algorithm, and suggest the optimal solution whose optimality is guaranteed because of analytical behavior of the methods [3,9].

In order to address the mentioned issues altogether, a new instrumentation design methodology for precise and fault-tolerant sensor networks which is more comprehensive, flexible and practical than other designs given in the literature has been presented in this article. In the proposed method, instead of following the analytical techniques which use a regular and determined approach to check all the possible nodes that fulfill the constraints, a modified and efficient search engine is used in order to maximize the accuracy of the estimator, i.e. UKF, subject to fault tolerance and cost constraints. Different from enumeration based methods, the search engine does not examine all solutions to look for the most optimal one, this enables us to implement larger designs with more flexibility of using different types of sensor. The results of this paper has the potential to be used in the previous design works to improve the presented methods [14,15].

In this article, first procedure of precision assessment is provided via UKF as the data reconciliation technique used in this method. A brief introduction to the terms related to the fault tolerance capability of the networks, i.e. redundancy degrees and reliability of networks, is provided in the next section which is followed by the proposed comprehensive model and design algorithm procedure. The case study used in this paper, i.e. a CSTR, is introduced and the method is implemented on it in the next section. Then, some verifying tests for the results will be presented to validate the method performance. At last, in the conclusion section, two suggestions to complement this approach will be provided for future works.

$$x_{k+1} = f\left(x_k, u_k\right) + w_k \qquad (1)$$
$$u_k = g\left(x_k\right) + v_k$$

THEORITICAL SECTION

Precision assessment

In this section, an optimized sensor network design is presented using UKF and CPSO and the presented method is tested on 15-state, nonlinear CSTR to illustrate the design procedure. The model of a non-linear system can be represented by following state and measurement equations:

Where x_k represents the unobserved state of the system, u_k is a known exogenous input and y_k is the observed output through measurement instruments. The process noise w_k adds on the model equation and the measurement noise is represented by v_k.

The UKF algorithm uses a "deterministic sampling" approach to calculate the mean and covariance estimates of Gaussian random state variables (i.e., x) with a minimal set of 2L+1 sample points (L is the state dimension), called as sigma points, through the actual nonlinear system dynamics without any linear approximations. The UKF algorithm used in this paper is the one developed by *Julier & Uhlmann* [16].

To have an optimized sensor network, it should be determined which variables be measured and what sensors be utilized from the point of accuracy. The performance of each sensor network is dependant on the measurement noise covariance matrix (R) and the resulting observation function ($g(x_k)$). R is related directly to the accuracy of sensors and $g(x_k)$ determines sensor locations in the network. Theses two functions represent the sensor network and hence, determine its topology. In most cases, abundant budget is not in hand to buy sensors; on the other hand, according to a conceptual understanding that says "the more expensive sensors you buy, the more efficient network you have", more precise and consequently more expensive sensors have been always desired for a satisfactory instrumentation. To resolve this problem, the maximum cost is considered limited in the presented algorithm and a search is performed to seek for the most optimal network for the specified cost. Some metric to present the performance of each candidate network is required (which represents the precision in this literature). Hence, Modified Root Mean Squared Error (MRMSE) is introduced to evaluate each network topology as follows:

$$\mathrm{MRMSE} = P_c^i = \frac{1}{[x_j]_i}\sqrt{\frac{\sum_{j=0}^{n}([\hat{x}_j]_i - [x_j]_i)^2}{n}} \qquad (2)$$

where x represents the estimated value of variable x and n denotes the number of samples of variable xi recorded in a specified time range.

$$x_{k+1} = Ax_k + Bu_k + w_k \qquad (3)$$
$$u_k = Cx_k + v_k$$

From the view of monitoring, all states and inputs/manipulates of the system should be measured or estimated; thus, it is wise to add the inputs/manipulates as new states to the original system states and design sensor network for the new system. Consider a typical linear time-invariant system like (3).

To identify the state transition matrix, it is assumed that input variables during a given time period are correlated with the input during the previous time period as (4).

$$u_{k+1} = u_k + w_k \qquad (4)$$

Inputs are combined with states to form new states of the system x*=[x,u]`. New system parameters should be modified as follows:

$$* = \begin{bmatrix} I^{n_u+n_u} & O^{n_u+n_x} \\ B^{n_u+n_u} & A^{n_x+n_x} \end{bmatrix} \quad (5)$$

giving

$$x_{k+1}^* = A^* x_k^* + w_k^* \quad (6)$$
$$y_k^* = C^* x_k^* + v_k^*$$

When steady state estimation is desired, this simple modification works well in linear systems. Likewise, this is true for non-linear systems providing that appropriate modifications are made. However; assumption (4) is not valid in dynamic data reconciliation case, since set points and consequently manipulates should be changed in order to observe dynamic behavior of the system. Therefore, in order to handle this in dynamic reconciliation, Eq. (4) should be replaced by appropriate control loop equation so that manipulated variables can be monitored and considered in instrumentation design. This procedure is demonstrated on a non-linear case study having three control loops (PI controllers) and subsequently three manipulated variables in following sections.

In each Sensor Network Topology (SNT), measurement matrix, $h_k(x_k,v_k)$, should be modified properly to show which variables are measured and which are not.

Fault tolerence assesment

Minimality and redundancy

Consider the continuous time deterministic system:

$$\dot{x} = f\left(x(t), u(t)\right) \quad (7)$$

$$y(t) = g\left(x(t)\right) \quad (8)$$

$$z(t) = h\left(x(t)\right) \quad (9)$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the control input, $y \in R^p$ is the measurement vector, and $z \in R^q$ is the functional of the state which is to be estimated. The inputs u(t) are assumed to be sufficiently differentiable and f, g, h are sufficiently smooth vector fields. Let $J \subseteq R$ be a subset of the system sensors, and introduce the notation obsv(z/J) where (for a given definition of observability):

$$\text{obsv}(z/J) = \begin{cases} 1 & \text{if } z \text{ is observable with J} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Let 2^R be the set of all subsets of R; then (10) induces a two- class partition:

$$2^{R+} = \left\{J \subseteq R; \text{obsv}(z/J) = 1\right\} \quad (11)$$
$$2^{R-} = \left\{J \subseteq R; \text{obsv}(z/J) = 0\right\}$$

The class 2^{R+} contains all the subsets of sensors by which z is observable, and it is assumed that $R \in 2^{R+}$; i.e. the system is observable by the whole set of sensors. Accordingly, minimal sensor set and redundant sensor sets are defined a following: A subset of sensors $J \in 2^{R+}$ is minimal, if

$$\forall k \subset J \qquad k \notin 2^{R+} \quad (12)$$

and a subset of sensors $J \in 2^{R+}$ is redundant, if it is not minimal.

Interpretation of fault tolerence

The interpretations of Minimum Sensor Set (MSS) and Redundant Sensor Set (RSS) are as follows: suppose that at a given time, the system is operating with a subset of sensors $J \in 2^{R+}$ such that the functional z is observable (therefore, J is a MSS or a RSS). Assume that one or several sensor failures occur at time t_f so that the set of sensors J can be decomposed into the normal and the faulty ones: $J = J_n \cup J_f$.

Therefore, the measurement equations can be written

$$y_n(t) = g_n\left(x(t)\right) \quad (13)$$

$$y_f(t) = g_f\left(x(t)\right) \quad (14)$$

where y_n (resp. y_f) represent the normal (resp. the faulty) outputs of the sensor network J and g_n (resp. g) are the normal (resp. the faulty) measurement equations. The fault tolerance problem used in this paper can be interpreted as follows: the faulty sensors J_f are switched off, and the problem is to assess the possibility of still estimating the functional z by using the remaining set of sensors J_n which is indeed true, provided system is still observable. This method that is named reconfiguration strategy only needs fault detection and isolation (fault estimation is not necessary), and that the fault tolerance property is a structural one, since it is associated with triple (7), (13) and (14), it does not depend on the type of fault which affects the sensors J_f, In this paper, we consider only the reconfiguration strategy.

System redundancy degrees

Let $J \subseteq I$ be any subset of the sensors, i.e. some state of the system instrumentation. Let $K \in MSS(J)$, then the quantity |J\K| represents the maximal number of sensors which can be lost while z can still be estimated by K. In the 'best' situation, as many sensor losses as can be accepted. The weak redundancy degree evaluates the size of this 'best' situation.

$$|J| - \min_{K \in MSS(J)} |K| \tag{15}$$

The weak redundancy degree associated with the pair (z, J) is From the interpretation of WRD(z,J) it follows that the following statement is true:

$$WRD(z, J) = |J| - \min_{K \in MSS(J)} |K| \tag{16}$$

$$\exists J' \subset J \text{ such that } |J'| = WRD(z, J) \text{ and } J \setminus J' \in MSS(J) \tag{17}$$

Of course, in many cases, z will no longer be observable after less than WRZ(z,J) sensors are lost.

The strong redundancy degree SRD(z,J) evaluates the maximal number of sensors which can be lost while keeping z observable for sure (i.e. considering the worst case situation). This means that the following statement is true

$$\forall J' \subset J : |J'| = SRD(z, J) \text{ and } J \setminus J' \in RSS(J) \tag{18}$$

The strong redundancy degree associated with the pair (z,J) is

$$SRD(z, J) = |J| - \max_{J^* \in RSS(J)} |J \setminus J^*| - 1 \tag{19}$$

Availabillity and the estimation service

Let t_0=0 be the time at which the system operation was started, and let J(t) be the subset of the non-faulty (available) sensors at time t. Let J_0 =J(0), assuming such data to be available, the fault tolerance of the z-estimation process can be evaluated by the probability for the estimation of z to be possible during the given time interval [0,*t*] assuming that it was possible using the set J_0 at time 0, R(z/J_0). Let $K \subseteq J_0$ be any subset of sensors. The probability for the estimation of z to be possible during the time interval [0,t] using K is given by (20):

$$R(z / K, t) = P(z / K).R(K, t) \tag{20}$$

where P(z/K) = 1 if K is a MSS or a RSS and P(z/K) = 0 otherwise, and R(K,t) is the reliability of the set of sensors K; which is defined as the probability that no sensor of K fails during the interval [0,t]. If sensor failures are independent, i.e. there is no common mode failure, one has

$$R(K, t) = \prod_{k \in K} R_k(t) \prod_{k \notin K} (1 - R_k(t)) \tag{21}$$

Where $R_k(t)$ is sensor k reliability. The reliability of such individual components is often modeled using the Poisson distribution:

$$R_k(t) = e^{-\lambda_k t} \tag{22}$$

Where λ_k is sensor k failure rate, which is supposed to be constant.

Now, considering the whole set J_0; it follows from the fact that all its subsets K are exclusive, that the probability for the estimation of z to be possible during the time interval [0,t] is given by

$$R(z / J_0, t) = \sum_{K \subseteq J_0} P(z / K) R(K, t) \tag{23}$$

in (23), P(z/K) is 1 if subset K is observable and 0 if not.

Comprehensive model and design procedure

The precise and fault tolerance requirements have been discussed sufficiently in previous sections. Now that the definitions of required terms are determined, we can go through the comprehensive model which utilize all mentioned criterion in order to suggest us a comprehensive instrumentation network. Precision is interpreted in this model as the accuracy of the UKF obtained estimations that will be used in control and monitoring applications. Reliability and redundancy degrees as the corresponding definitions presented in section 3. In our model, precision is treated as the optimization criterion while cost and fault tolerance criteria take the role of constraints in this optimization problem. Accordingly. the eligibility of networks can be judged by estimation accuracy in some or all variables of great importance for designer; in general, the optimization object suggested in this model can be represented by any function of MRMSE (P_c^i). This metric which we name it Instrumentation Criterion (I.C) is shown by (24).

$$\text{instrumentation Criterian}\left(\text{I.C}\right)=F\left(P_c^i\right),\ i=1,2,\ldots n \quad (24)$$

Definition of function $F(P_c^i)$ is highly independent on the requirements of the control and monitoring systems. For instance, in an application that precise estimation of all variables is important equally, $F(P_c^i)$ can be defined as sum of P_c^i and in an application that estimation of some variables are more important than others, the weighted sum of P_c^i values is the proper selection criterion. Such a model that satisfies all mentioned characteristics can be shown in (25).

$$\text{Min}\,(\text{Instrumentation Criterion}) \quad (25)$$

$$\text{S.t.}\quad \sum_j (C_j S_{ji}) \le \text{Cost}^*$$

$$R(\text{network}) \ge R^*$$

$$SRD \ge SRD^*$$

$$WRD \ge WRD^*$$

Where S_{ji} represents the integer number showing the placement of the variable of sensor type j at network location i. For any sensor placed on variable i, its Corresponding Variance δ_{i2} is entered in R. The Instrumentation Criterion (I.C) for any sensor network is highly dependant on accuracy of sensors that matrix R feeds to the model.

Now that the accurate model of the optimization problem is in hand, a component tool is necessary to solve this problem. Accordingly, a combinational particle swarm optimization algorithm will be utilized as a search engine to solve optimization problem in (25). The CPSO should be modified so that before selecting the particles in a repetition, the determined constraints are checked inside the engine to verify that whether cost and fault tolerance inequalities are satisfied or not. If any of these requirements are not fulfilled, the engine puts away the corresponding the particle and chooses another one according to the engine's defined regulations. This scenario goes on until CPSO gathers enough competent particles to form a swarm to survive. Then, most eligible particles are detected and other particles in the swarm is replaced by new particles driven from the eligible ones.

Any particle of the swarm used here is a string whose length is set equal to number of variables whose values are desired to be estimated. Any index of the string can be assigned to an integer representing the type of sensor used for that variable. CPSO algorithm is modified so that observability, cost, reliability and redundancy degree constraints are validated for any topology before assigning that topology to a particle.

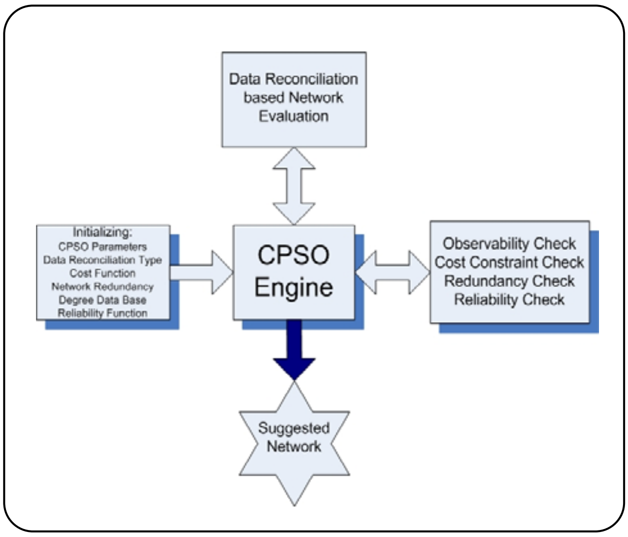

Fig. 1: Block diagram of the overall instrumentation design based on CPSO, two sided arrows represent two sided relationships.

Before CPSO starts searching, all the necessary information related to the instrumentation should be provided for the search engine. This includes all the CPSO parameters, constraints, cost function and data reconciliation algorithm (Fig. 1). The block diagram shown in Fig. 1 shows the relevant modules involved in the optimization problem along with the interactions between them. For more illustration, the suggested algorithm will be presented in next section.

RESULTS AND DISCUSSION

Implementing the design procedure in the case study and results

The case study used in this paper is an nonlinear CSTR that *Bhushan & Rengaswamy* [17] introduced in their article (Fig. 2). This process involves an exothermic liquid-phase reaction. The process involves an exothermic liquid-phase reaction: A(l) → B(l) + C(g). As shown, the temperature controller (T_C) controls the temperature of the reactor by manipulating in the reactor is controlled by the level controller (V_C) which manipulates the outlet flow rate from the reactor. The pressure in the reactor is controlled by changing

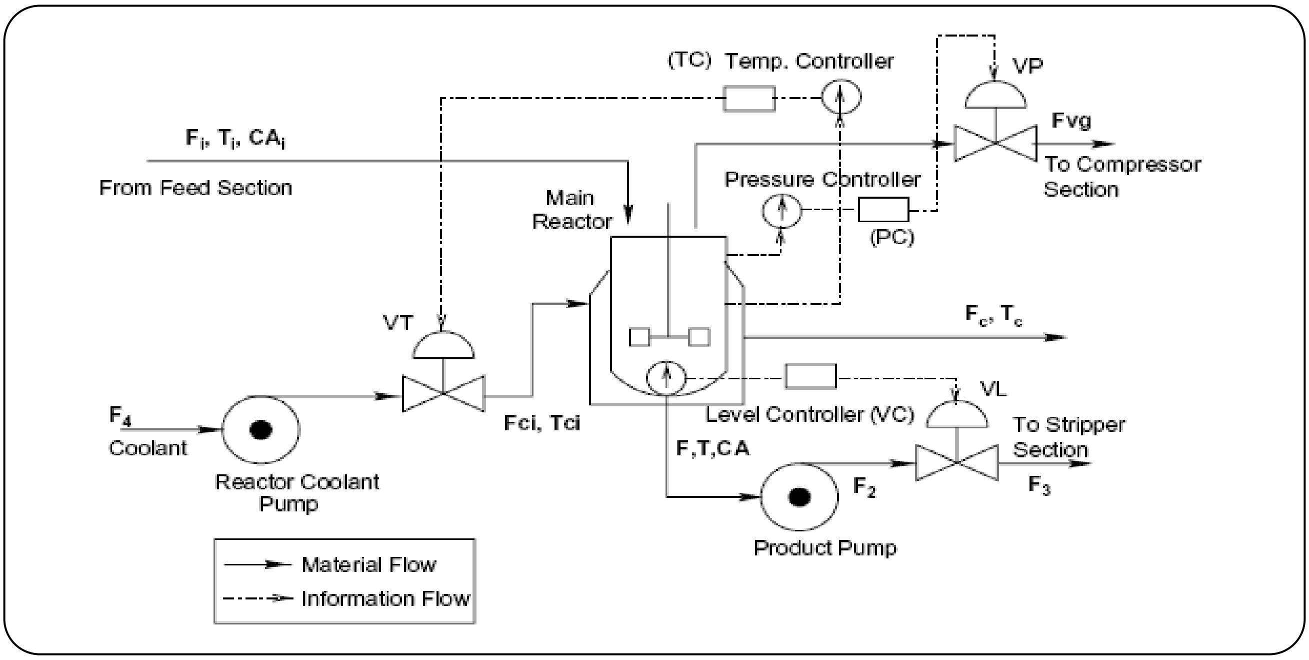

Fig. 2: The schematic diagram of CSTR process.

the vent gas flow rate. Both the reactor and the jacket are modelled with perfectly mixed tank dynamics.

The CSTR model equations are given as follows:

Global mass balance:

$$F_i - F = \frac{dV}{dt} \tag{26}$$

Component mass balance (C_A):

$$\frac{F_i}{V}(C_{Ai} - C_A) - r_A = \frac{dC_A}{dt} \tag{27}$$

Overall heat balance on the reactor (Result obtained assuming constant heat capacities and densities):

$$\frac{F_i}{V}(T_i - T) + \frac{r_A(-\Delta H)}{\rho C_p} - \frac{UA(T - T_c)}{V\rho C_p} = \frac{dT}{dt} \tag{28}$$

Overall heat balance on the jacket:

$$\frac{F_c}{V_j}(T_{ci} - T) + \frac{UA(T - T_c)}{V_j \rho_j C_{pj}} = \frac{dT_c}{dt} \tag{29}$$

Gas phase balance:

$$r_A V - F_{vg} = \frac{dn}{dt} \tag{30}$$

$$r_A = C_d C_A k_0 e^{-E/RT} \tag{31}$$

Elemental mass balances in valves and pumps, assuming no accumulation:

$$F_3 - F_2 = 0 \ , \ F_2 - F = 0 \ , \ F_4 - F_c = 0 \tag{32}$$

Pressure in the Reactor (where V_g is the vapor space and is assumed constant, assuming ideal behavior):

$$PV_g = nRT \tag{33}$$

As shown in Fig. 2, variables T, V, P are controlled via three PI control-loops, using variables F, F_{ci}, F_4 as manipulated variables. Writing dynamic equations for these three closed-loops and denoting proportional and integral constants by K_{pi} and K_{Ii} respectively, the following equations are obtained.

$$K_{p_1}\frac{dT}{dt} + K_{I_1}T - K_{p_1}\frac{dU_1}{dt} - k_{I_1}U_1 = \frac{dF_4}{dt} \tag{34}$$

$$K_{p_2}\frac{dV}{dt} + K_{I_2}V - K_{p_2}\frac{dU_2}{dt} - k_{I_2}U_2 = \frac{dF}{dt} \tag{35}$$

$$K_{p_3}\frac{dP}{dt} + K_{I3}P - K_{p_3}\frac{dU_3}{dt} - k_{I_3}U_3 = \frac{dF_{vg}}{dt} \tag{36}$$

Equations (34-36) should be inserted in the new state equation of the system. But, dynamics of other inputs are

Table 1: Available sensors for instrumentation design:V: Volume; C: Concentration; T: Temperature; P: Pressure; F: Flow.

	V_1	V_2	V_3	C_1	C_2	C_3	T_1	T_2	T_3	P_1	P_2	P_3	F_1	F_2	F_3
Failure Rate ($\times 10^{-2}$)	1/81	1/50	1/28	1/80	1/50	1/28	1/88	1/49	1/23	1/90	1/48	1/22	1/82	1/44	1/2
Cost ($)	3000	2200	1600	2500	1800	800	1500	1000	400	1400	1000	800	4500	3000	2200
Accuracy (%)	0.1	2	5	0.5	2	4	0.1	1	5	0.1	3	6	0.1	3	5

considered as before. The model parameters along with their nominal operating values are presented in Table 4.

In this case study, the measurable variables are V, C_A, T, T_C, P, F_4, F, F_{vg}, F_c, F_i, T_i, C_{Ai}, T_{ci}, F_2, and F_3. Thus, there should be five types of sensors to cover these 15 variables. For each type of sensor, three different sensor sets are considered. The failure rates of these sensors along with their corresponding cost and precision values have been tabulated in Table 1.

The CPSO algorithm used here is the one suggested by *Jaboui et al.* [18] and parameters of search engine are set as follows: ω=1.1, c_1=0.6, c_2=0.5, $\upsilon_{max} = 2$, $\upsilon_{min} = -2$ and α=1.2. Moreover, 20 particles and 100 iterations have been considered in this CPSO.

Now that our model of the instrumentation for the purpose of overall design has been created and CPSO is ready to use, implementation of the instrumentation procedure on the case study will be straightforward. We perform the design procedure for different costs with keeping reliability and weak redundancy degree constraints fixed at the values of 0.8 and 4 for all scans. Five types of sensors are available and each type consist of three sets of sensors so that in spite of their similar structures, they differ in their precision and failure rates. Increasing the bound cost in design sensor network enables us to take advantage of using more precise and expensive sensors in our design and thus, increases time consumption of search engine since it should examine more variety of networks constructed by combination of available sensors.

Weak redundancy degree and network reliability are set 4 and 0.8 in our all scans while cost constraint varied in designs and takes different values [$30000, $35000, $40000, $45000]. The search engine is forced to looking for the most deserved network that satisfies specified constraints. The selection criterion is the precision, i.e. if the search engine runs into two networks that both satisfying all constraints, it will select the one with less estimation error and put away the other one which produce less accurate estimations.

In all designs, the WRD constraint is set four but no limit is imposed on SRD. This is due to the fact that this value is zero for all topologies in this plant; hence, the topology of the plant enforces us to put sensors on variables F_2, F_3 and F_{14} in order to get the estimations of these variables. The search is carried out ten times so that it does not fall into local optimization points. The best network among the ten obtained solutions is chosen as the main solution for that design problem. The spectrum of cost versus MRMSE values for all designs are shown in Fig. 3. Note that lower bold line in Fig. 3 represents the candidates out of all ten solutions for cost constraints. The best and worst solutions obtained in each case are tabulated in Table 2. The effect of repeating scans is clearly observed by comparing the two solutions that they are considerably different from each other.

Verification Tests

One remarkable advantage of search engines is their inherit capability to deal with large-scale designs in which analytical methods fails to be successful. But no profit is free; the cost of achieving such a valuable ability should be paid. This payment includes lack of optimality guarantee in obtained results. The proposed CPSO algorithm which was utilized to offer the best possible networks enforced to fulfill the requirements of the design problem, but there is no appropriate tool that assures us the obtained solution is the optimal network among all the other practical networks that can be built by the given sets of sensors. On the other hand, it should be checked that whether the suggested solution satisfies all the design constraints or not. The mentioned issues necessitate applying verification tests on the obtained results. Thus, three tests were conducted to verify whether the obtained solution satisfies the defined metrics or not.

Our case study consists of 15 variables to measure, and too many feasible networks can be constructed by combination of variables; thus, performing an exhaustive search to verify the results is not practical. If we consider different possible types of sensors given in Table 1.

Table 2: The best and worst solutions obtained for different reliability constraints.

	V	CA	T	Tc	P	F4	F	Fvg	FC	Fi	Ti	CAi	Tci	F2	F3	Cost*	MRMSE	WRD	R	Cost($)
Best	×	1	3	3	3	1	×	1	1	×	1	×	2	1	1	30000	0.0313	4	0.8166	29300
Worst	×	1	1	3	×	1	3	1	3	×	1	×	2	1	1	30000	0.0441	4	0.8180	29800
Best	×	1	×	2	3	1	×	1	1	1	1	×	1	1	1	35000	0.0260	4	0.8655	34800
Worst	4	3	3	×	×	1	×	1	1	1	2	1	1	1	1	35000	0.0324	4	0.8303	33400
Best	1	×	×	3	3	1	1	1	1	1	1	×	1	1	1	40000	0.0231	5	0.8654	39300
Worst	×	2	×	3	3	1	1	1	1	1	1	1	2	1	1	40000	0.0296	6	0.8658	39700
Best	1	1	1	3	1	1	1	1	1	1	1	×	1	1	1	45000	0.0183	7	0.8713	44200
Worst	3	1	1	×	2	1	1	1	1	1	1	1	1	1	1	45000	0.0232	8	0.8702	44900

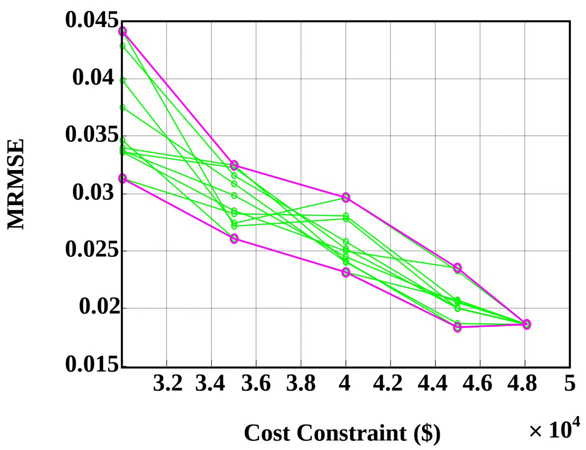

Fig. 3: The MRMSE values for different cost constraints obtained by CPSO-based design, the bold upper and lower lines shows the worst and best solutions.

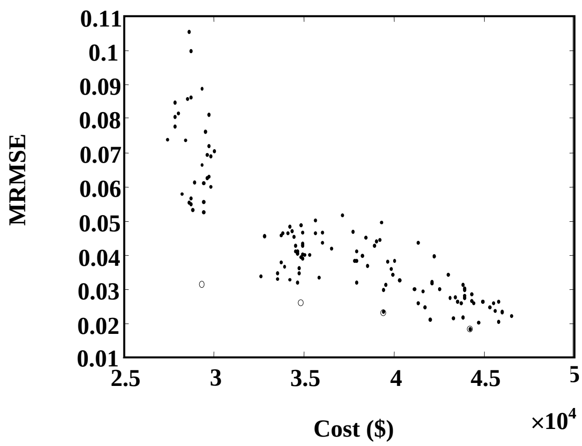

Fig. 4: 200 networks have been chosen randomly in order to verify the result of the design; the circles are the solutions of our design that all are below or as high as the dots.

to measure each variable, there will be 384,422,112 observable networks. Each scan on average took 570 seconds to complete, considering that there are ten scans for each design, the average time for a typical design will be 95 minutes. So, implementing four designs took almost 6 h and 20 min, whereas performing an exhaustive search, i.e. without any search engine, will take more than ten years for such a plant!

Obviously, it is not possible to perform a comprehensive verification for our design results. However, in order to assess the presented approach performance, 200 randomly chosen networks are shown with their corresponding costs and MRMSE values in Fig. 4. In this figure the circles represent the solutions suggested by the CPSO that all are below or as high as the dots. Although this type of verification can not completely approve the results, it can lend additional support to the performance of the presented method.

In all designs the reliability constraint has been considered to be 0.8. The reliability of the solutions in each scans have been tabulated in Table 2. In order to verify these values and make sure that solutions fulfil the reliability requirements of the problem, we took a number of randomly chosen networks and let their sensors fail according to their reliability to see that whether the new obtained network is observable or not, the statistical reliability is obtained by division of the observable network to the total number of networks. The diagram of the statistic reliability versus the number of networks undertaken in this test for all searches is depicted in Fig. 5. As seen as the number of networks applied in the test increases the reliability approaches the mentioned

Table 3: Weak redundancy degree verification for five design sets.

Cost*		V	C_A	T	Tc	P	F_4	F	Fvg	F_C	Fi	Ti	C_{Ai}	Tci	F_2	F_3
30000	Initial Net	×	1	3	3	3	1	×	1	1	×	1	×	2	1	1
	Last Net	×	F	F	3	F	1	×	1	1	×	1	×	F	1	1
35000	Initial Net	×	1	×	2	3	1	×	1	1	1	1	×	1	1	1
	Last Net	×	F	×	2	F	1	×	1	1	1	F	×	F	1	1
40000	Initial Net	1	×	×	3	3	1	1	1	1	1	1	×	1	1	1
	Last Net	F	×	×	3	F	1	1	1	1	F	F	×	F	1	1
45000	Initial Net	1	1	1	3	1	1	1	1	1	1	1	×	1	1	1
	Last Net	F	F	F	3	F	1	1	1	1	F	F	×	F	1	1

Table 4: Nomenclatures and Nominal Values of the CSTR.

Notation	Variable	
V	Volume of reactor	48 ft^3
C_A	Reactant Concentration in reactor	0.2345 lb.mol of A/ft^3
T	Reactor temperature	600° R
F	Outlet flow rate	40 ft^3/h
N	No. of moles of vapor	28.3657 lb. mol
P	Pressure in vapor space	2116.79 lb/ft^2
F_{vg}	Vent flow rate	10.6137 lb. mol/h
Fi	Inlet feed flow rate	40 ft^3/h
C_{Ai}	Inlet reactant concentration	0.5 lb. mol of A/ft^3
T_C	Jacket temperature	590.51° R
F_C	Coolant flow rate	56.626 ft^3/h
Ti	Inlet feed temperature	530° R
V_j	Volume of jacket	3.85 ft^3
K_0	Frequency factor	7.08×10^{10} h^{-1}
C_d	Catalyst activity	1
E	Activation energy	29,900 btu/lb. mol
R	Universal gas constant	1.99 btu/lb. mol°R
U	Heat-transfer coefficient	150 btu/h.ft
RD	Redundancy Degree	
WRD	Weak Redundancy Degree	
SRD	Strong Redundancy Degree	
NFT	Network Fault Tolerence	
MRMSE	Modified Root Mean Square Error	
Obsv	Observability	
MTTNO	Mean Time to Non-Observabiliy	
CSTR	Continuous Stirred Tank Reactor	
CPSO	Combinatorial Partical Swarm Optimization	

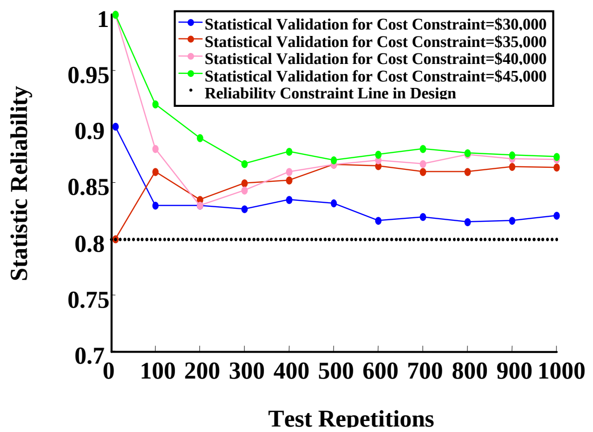

Fig. 5: The reliability validation through failure experiment; the dotted line represents the reliability constraint of our design and the colored dots at the end of each diagram are above it.

corresponding reliability values. Note that the statistic reliability obtained with low number of repetitions cannot be valid for verification of real network reliability, but as more experiments are conducted the reliability gets closer to the expected value. The lines drawn in dots indicate the design reliability constraint which is 0.8 for all scans. As seen in the plot, all diagrams converge to their expected value and all are above the limit line used in this design.

Reliability and cost validations have been investigated. Now only WRD values remain to be verified. Table 3 shows the verification results for weak redundancy degree. For each design, two networks have been shown. The first is the initial one, indicating the main solution suggested by the CPSO algorithm, while the second one represents the network that has been obtained after occurrence of some sensor failures in the initial network.

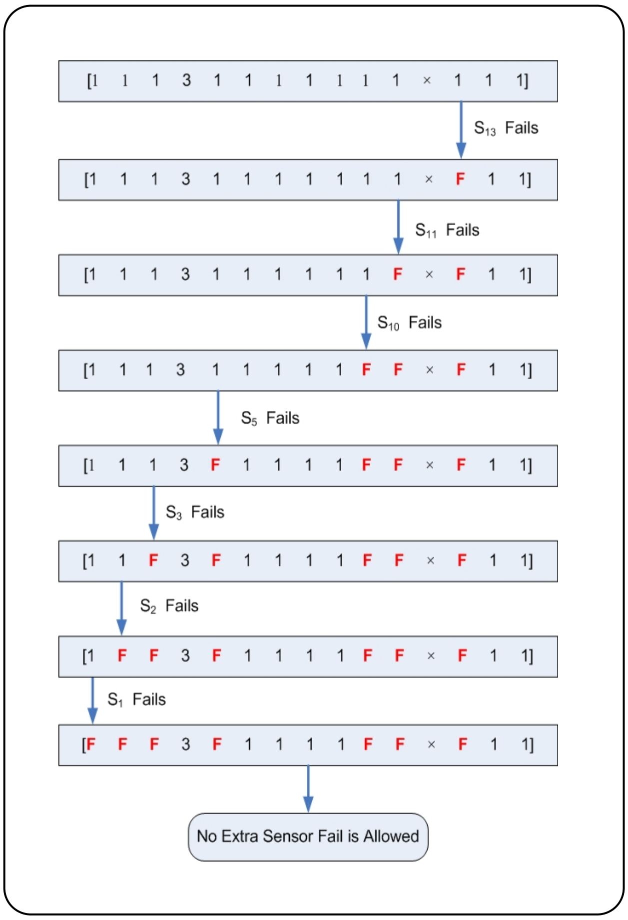

***Fig. 6:** The sensor failure diagram of the design undertaken for Cost Constraint=\$45,000, it approves that at most seven failures can occur in this network.*

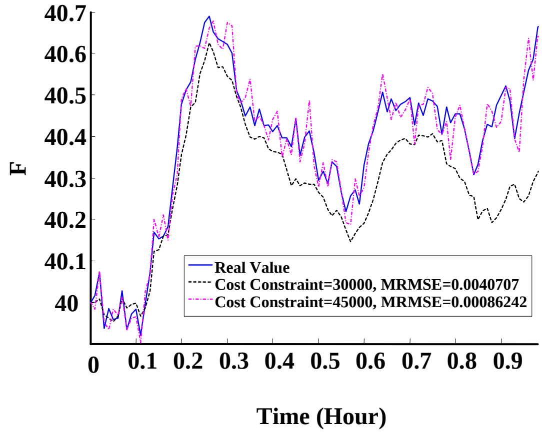

***Fig. 7:** The real and estimated plots for F.*

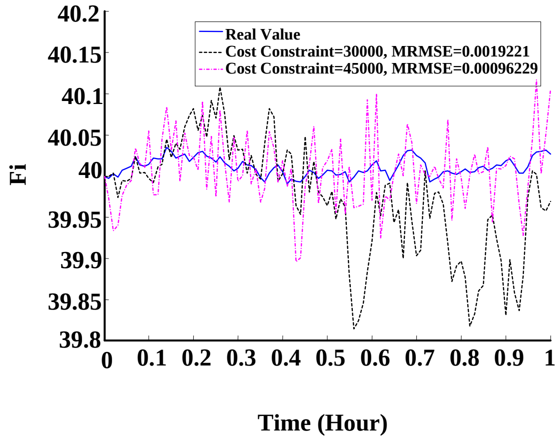

***Fig. 8:** The real and estimated plots for Fi.*

Examining Table 3 infers that the networks have been remained observable after specified failures appear in the sensor sets. Of course, the networks with fewer failures that are located between initial sensor set and its corresponding Minimum Sensor Set (MSS) are feasible too. For instance, in each sensor failure, a new observable network is obtained. This procedure continues until it reaches a node that has the minimal number of sensors and hence no extra sensor failure can occur, indicating that sequence of failures ends at this node. The number of failures in this sequence determines the weak redundancy degree of initial network which is more than four in all designs. For instance, the initial network shown in Fig. 6 that shows the solution suggested by CPSO in the design cost constraint = \$45000, has a WRD value of 7 and can tolerate a sequence of seven failures, S_{13}, S_{11}, S_{10}, S_5, S_3, S_2 and S_1 causing all networks laid between upper and lower networks to be feasible and observable. At last, to compare the effect of investing more money on the estimation accuracy of variables, the estimated and real values of variables F, F_i and C_{Ai} for cost=\$30,000 and \$45000 are shown in Figs. 7, 8 and 9. As seen, thanks to more investment in the network, the accuracy of states have been improved drastically.

CONCLUSIONS

A reliable and precise sensor network design has been proposed in this paper as a new comprehensive methodology. The efficiency and accuracy of this method has been approved by different test scenarios which were undertaken on the CSTR study in a large-scale design. However, the presented method can be established in an industrial software similar to the common data reconciliation packages available in the market. Our future work includes preparing such a software.

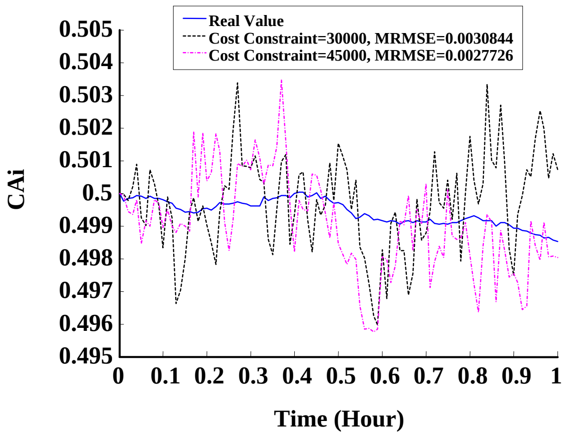

Fig. 9: The real and estimated plots for C_{Ai}.

In addition, fault detective and diagnosis considerations can be added to the suggested model as the final complement.

Received : Mar. 6, 2010 ; Accepted : Sep. 19, 2011

REFERENCES

[1] Vaclavek V., Loucka M., Selection of Measurements Necessary to Achieve Multicomponent Mass Balances in Chemical Plant, *Chem. Eng. Sci.*, **31**, p. 1199 (1976).

[2] Musulin E., Benqlilou C., Bagajewicz M.J., Puigjaner L., Instrumentation Design Based on Optimal Kalman Filtering, *Journal of Process Control*, **15**, p. 629 (2005).

[3] Bagajewicz M., Cabrera E., New MILP Formulation for Instrumentation Network Design and Upgrade. *AIChE J*, **48**, p. 2271 (2002).

[4] Ali Y., Narasimhan S., Sensor Network Design for Maximizing Reliability of Linear Processes, *AIChE J*, **39,** p. 820 (1993).

[5] Ali Y., Narasimhan S., Redundant Sensor Network Design for Linear Processes, *AIChE J*, **41**, p. 2237 (1995).

[6] Bhushan M., Rengaswamy R., Design of Sensor Network Based on the Signed Directed Graph of the Process for Efficient Fault Diagnosis, *Ind. Eng. Chem. Res*, p. 999 (2000a).

[7] Raghuraj R., Bhushan M., Rengaswamy R., Location of Sensors in Complex Chemical Plants Based on Fault Diagnostic Observability Criteria. *AIChE J*, **45** (2), p. 310 (1999).

[8] Bagajewicz M., Design and Retrofit of Sensor Networks in Process Plants, *AIChE J*, **3**, p. 2300 (1997).

[9] Bagajewicz M., Sanchez M., Duality of Sensor Network Design Models for Parameter Estimation. *AIChE J*, **45**, p. 661 (1999).

[10] Sen S., Narasimhan S., Deb K., Sensor Network Design of Linear Processes Using Genetic Algorithms, *Comput. Chem. Eng*, **22**, p. 385 (1998).

[11] Bagajewicz M., Chmielewski D., Rengaswamy R., Integrated Process Sensor Network Design. Presented at: *The Annual AIChE meeting*. Austin, Texas (2004).

[12] Kotecha P.R., Bhushan M., Gudi R.D., Keshari M.K., A Duality Based Framework for Integrating Reliability and Precision for Sensor Network Design, *Journal of Process Control*, **18**, p. 189 (2008).

[13] Staroswiecki M., Hoblosl G., Aitouche A., Sensor Network Design for Fault toLerant Estimation. *Int. J. Adapt. Control Signal Process*, **18**, p. 55 (2004).

[14] Panjeshahi M.H., Hasan M., Ataei A., Gharaie M., Comprehensive Approach to an Optimum Design and Simulation Model of a Mechanical Draft Wet Cooling Tower., *Iran. J. Chem. Chem. Eng. (IJCCE)*, **29**(1), p. 21 (2010).

[15] Razzaghi S., Kharrat R., Rashtchian D., Vossoughi Sh., Saraji S., Investigation of Auto Ignition Condition under Different Parameters, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **27**(2), p. 92 (2008).

[16] Julier S.J., Uhlmann J.K., A New Extension of the Kalman Filter to Non-Linear Systems, in: "Proc. of AeroSense, The 11th Int. Symp.", A.D.S.S.C. (1997).

[17] Bhushan M., Rengaswamy R., Design of Sensor Location Based on Various Fault Diagnosis Observability and Reliability Criteria, *Compt. Chem., Eng.*, , **24** (2000b).

[18] Jarboui B., Damak N., Siarry P., Rebai A., A Combinatorial Particle Swarm Optimization for Solving Multi-Mode Resource-Constrained Project Scheduling Problems, *Applied Mathematics and Computation*, **195**(1**)**, p. 299 (2008).