Design and Retrofit of Sensor Networks in Process Plants

Miguel J. Bagajewicz
School of Chemical Engineering and Materials Science, University of Oklahoma, Norman, OK 73019

A novel procedure to find cost-optimal sensor networks is proposed. Cost is minimized subject to qualifying constraints that are related to certain requirements of data reconciliation. One basic qualifying constraint is a desired level of precision of reconciled values for a selected set of variables. Since precision requirements lead to multiple solutions, other qualifying constraints are proposed. These constraints are availability, resilience, and error detectability. Definitions for these terms are given and their impact on the results is presented.

Introduction

For a long time, the selection of sensors in chemical plants has been traditionally driven by the needs of basic control loop design. After sensors needed for control purposes were selected, additional sensors for process monitoring and fault detection were added. In recent years emerging data-reconciliation technologies have set up the scenario for a revision of the criteria used for sensor location. Data reconciliation provides results with improved precision for process economics (mainly accounting), for on-line modeling and optimization, and is an aid for instrument maintenance. Observability, redundancy, and reliability, as well as the ability to detect gross errors are features that are becoming increasingly important. The question is how to use these objectives in the context of a systematic procedure to design cost-optimal sensor networks.

Vaclavek and Loucka (1976) first explored this problem using graph theory to guarantee variable observability. Kretsovalis and Mah (1987) proposed a combinatorial search based on the effect of the variance of measurements on the precision of reconciled values. Madron and Veverka (1992) proposed classifying measured and unmeasured variables of linear systems according to preestablished criteria of "required" and "nonrequired." Unmeasured variables were later ordered from "hardly measured" to "easily measured." Madron proposed to use two objective functions: cost and overall precision of the system. Suboptimal structures are found by means of a matrix decomposition and an elaborate column permutation procedure. Madron (1992) also presents details of this procedure based on graph theory. The concept of cost-edged graph is introduced, and minimum spanning trees of these graphs are used to obtain minimum cost or optimal overall precision sensor networks. The method apparently cannot target the desired precision levels on individual variables. Ragot et al. (1992) presented a procedure that allows the set of sensors for which the system becomes observable to be identified. Luong et al. (1994) presented a method that provides optimal cost solutions that feature minimal observability of those variables required for control and a high degree of redundancy of variables. They use reliability as a means of screening alternatives with equal cost. Maquin et al. (1994) proposed to obtain the location of sensors by inverting the expression that provides the variance of reconciled variables as a function of the variance of measurements. Ali and Narasimhan (1993) proposed maximizing reliability, an idea that is based on sensor failure probability, observability of variables, as well as redundancy. While looking at all the networks containing the minimum number of sensors needed to achieve observability, they propose a max–min problem using reliability as the objective function. Another graph-oriented procedure was proposed by Meyer et al. (1994), who used cost as the objective function and provided solutions featuring networks containing the minimum number of sensors. Recently, Ali and Narasimhan (1995) extended their previous work to redundant networks. Their algorithm uses graph theory to build networks with a specified number of sensors and maximum reliability.

While all these methods provide fundamental insights, they fail to take into account the fact that sensor networks should be able to handle gross errors effectively, that is, detect them when they are too large and avoid extensive corruption of data when they are not detected. The answer to this problem was first introduced in a commercial software (DATACON,

1993). This software proposes using the concept of error-masking factors to measure the effectiveness of a sensor network in relation to error detectability. A sensor value function is proposed to measure the gross error-detection capabilities of a network. In turn, the concept of error masking relies on a test for gross error detection developed by Madron (1985).

This work picks up on the preceding important milestones in sensor network design and proposes a design strategy that incorporates some of these concepts as well as new ones. An MINLP problem is proposed to obtain cost-optimal structures subject to the desired level of precision in each variable. Additional constraints are proposed to guarantee network robustness of the sensor network. Robustness is defined as the ability of a sensor network to detect gross errors (error detectability), provide results at a certain level of precision in the presence of gross errors (availability), and minimize corruption of data by undetected gross errors (resilience).

The model for linear systems is presented first, and the solution procedure is discussed. Availability, error detectability, and resilience constraints are defined next, and their effect on the solutions is discussed. This article focuses on presenting these new basic concepts. Extensions of these concepts to nonlinear systems (component and energy sensor networks), as well as other important features to be addressed in future work, are briefly discussed throughout the text. Finally, the application of the methodology to an example from the literature is presented.

Model Formulation

Assume that z is the vector of all mass flows, and let q be a vector of binary variables defined by

$$q_i = \begin{cases} 1 & \text{if } z_i \text{ is measured} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Therefore if x denotes the measured flows and y denotes the unmeasured flows, we write the model of the plant as follows:

$$Dz = Ax + By = b, \tag{2}$$

where A is composed of the columns of D that correspond to variables and B is composed of the columns of D that correspond to unmeasured variables, that is, $A = A(q)$ and $B = B(q)$. This model corresponds to the usual case of material balance reconciliation. It is assumed that D is already in its canonical form (Madron, 1992), that is, all unobservable variables have been removed from the system of equations, and A is of full row rank and B is of full column rank. This procedure is equivalent to the matrix projection method developed by Crowe (1983) and is presented by Madron (1992): it consists of simple linear combination and rearrangement of rows as well as column reordering, which guarantees that A is of full row rank and B is of full column rank. The reconciliation problem is written as follows:

$$\begin{aligned} &\text{Min} \frac{1}{2}(x - x_m)^T S_{x,m}^{-1}(x - x_m) \\ &\text{s.t.} \\ &\quad Ax + By = b, \end{aligned} \tag{3}$$

where x_m and $S_{x,m}$ are the measured flow rates and their variance, respectively. The solution to this problem is

$$x = x_m - S_{x,m}A^T(I - G^{-1}BH^{-1}B^T)G^{-1}(-b + Ax_m) \tag{4}$$

$$y = H^{-1}B^TG^{-1}(-b + Ax_m), \tag{5}$$

where $G = AS_{x,m}A^T$ and $H = B^TG^{-1}B$. Since matrix B has been assumed to have a rank equal to the number of unmeasured quantities, H is nonsingular. If not, there are unobservable variables in y and a model reduction has to be performed to obtain a canonical form (Madron, 1992). The variance of the reconciled results is

$$S_y = H^{-1} \tag{6}$$

$$S_x = S_{x,m} - S_{x,m}A^T(I - G^{-1}BH^{-1}B^T)G^{-1}AS_{x,m}, \tag{7}$$

and the precision of each variable is

$$\sigma_i = \begin{cases} \sqrt{[S_x]_{ii}} & \text{if } z_i \text{ is measured} \\ \sqrt{[S_y]_{ii}} & \text{otherwise.} \end{cases} \tag{8}$$

Thus, when the variance of each instrument is known, the variance of the reconciled values depends on what the set of sensors selected is, that is, $\sigma_i = \sigma_i(q)$.

The literature on data reconciliation often uses the term accuracy when referring to the variance $S_{x,m}$. However, the literature on sensors (Liptak, 1995; Whitaker, 1980) refers to the accuracy of an instrument as the degree of conformity with a standard or true value, that is, the variance obtained through calibration. Moreover, the distinction between accuracy (as per this definition) and precision is made specifically (Liptak, 1995). Following these guidelines, in this article the term precision is used for S_x and S_y. Accuracy and precision are only equivalent in the absence of systematic errors or biases. In data reconciliation, one should also add model inaccuracy as another source, since leaks can affect the accuracy of reconciled data.

Sensor location goals

Now that a few background concepts on data reconciliation for linear systems have been introduced we are in a position to state the objective of a sensor location design/retrofit problem.

The objective of a sensor location design/retrofit problem is to determine which variables should be measured such that the cost of the instrumentation is minimum, providing a certain level of precision of the reconciled values in a robust, resilient, and reliable manner while satisfying the needs of control strategies.

This definition calls for

- An optimization procedure based on minimizing overall instrumentation cost
- A sensor network satisfying control goals
- A certain level of precision upon reconciliation
- Robustness
- Reliability.

Robustness includes three properties: availability, error detectability, and resilience:

• Availability is a property that makes values of desired variables available within a certain precision level, after gross errors are detected and the corresponding measurements are eliminated.

• Error detectability is the ability of the network to detect gross errors of a certain size or larger.

• Resilience is a property that limits the impact of undetected gross errors. Since undetected gross errors can corrupt the reconciliation results, the change of accuracy in key variables in the presence of undetected gross errors is limited.

Network reliability is a property that was introduced by Ali and Narasimhan (1993, 1995). These authors used reliability to design minimum sensor networks (Ali and Narasimhan, 1993) and later to design redundant networks (Ali and Narasimhan, 1995). As outlined earlier, in the context of minimum cost networks, reliability should be a constraint. This article concentrates in robustness, leaving the inclusion of reliability for future work.

Optimization model

If for each variable z_i there is only one potential measuring device with associated cost c_i, the total cost is given by

$$C(q) = \sum_{\forall i} c_i q_i. \tag{9}$$

Let I_s represent the set of variables for which a certain desired precision is required and let $\sigma_{k,*}$, be the corresponding maximum value of standard deviation for each variable in I_s. Then, the design of the sensor network is an optimization problem that can be written as

$$\begin{aligned} &\text{Min} \sum_{\forall i} c_i q_i \\ &\text{s.t.} \\ &\quad \sigma_k(q) \le \sigma_{k,*} \qquad \forall k \in I_s \\ &\quad q_i = 0,1 \qquad \forall i. \end{aligned} \tag{10}$$

If more than one device is being considered as a potential candidate to be used in each variable measurement the objective function requires the use of additional binary variables and additional constraints. Indeed, let n_i^m be the number of different alternative candidates of measurement devices and let the cost of each of these candidates be given by $c_{i,k}$ $(k = 1, \ldots, n_i^m)$. Finally for each variable z_i, introduce binary variables $p_{i,k}$ $(k = 1, \ldots, n_i^m)$ to determine which candidate will be used, that is,

$$p_{i,k} = \begin{cases} 1 & \text{if device } k \text{ is used to measure variable } i \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The total cost is now a function of p and is given by

$$C(p) = \sum_{\forall i} \sum_{k=1}^{n_i^m} c_{i,k} p_{i,k}. \tag{12}$$

Therefore, the optimization problem is written as

$$\begin{aligned} &\text{Min} \sum_{\forall i} \sum_{k=1}^{n_i^m} c_{i,k} p_{i,k} \\ &\text{s.t.} \\ &\quad \sigma_k(p) < \sigma_{k,*} \qquad \forall k \in I_s \\ &\quad \sum_{k=1}^{n_i^m} p_{i,k} \le 1 \qquad \forall i \\ &\quad p_{i,k} = 0,1 \qquad \forall i, \forall k = 1, n_i^m, \end{aligned} \tag{13}$$

where the constraint

$$\sum_{k=1}^{n_i^m} p_{i,k} < m_i, \tag{14}$$

guarantees that at the most, m_i devices are assigned to each variable.

When $m_i = 1$, constraint (Eq. 14) restricts the model by not allowing more than one measurement per variable, that is, only systems with spatial redundancy and no hardware redundancy are considered. The standard deviation $\sigma_k(p)$ can be easily obtained for $m_i > 1$ through slight modifications of the derivation leading to Eqs. 6 and 7. One way of doing this is to add a ghost unit with one input and one output in the place of each duplicate measurement. This results in a new equation in D representing just the equality of both variables.

Solution Procedure

This is an MINLP problem with special characteristics. Matrices A and B as well as S_x, have dimensions that are a function of q (or p). For this reason, a relaxation of the integer constraints is not possible, and consequently, as lower bounds cannot be generated, branch-and-bound procedures cannot be implemented. Although implicit tree-type enumeration is impractical for fairly large systems, in the case of this problem its special characteristics allow a fairly efficient search. Consider the tree depicted in Figure 1. This tree has some important properties:

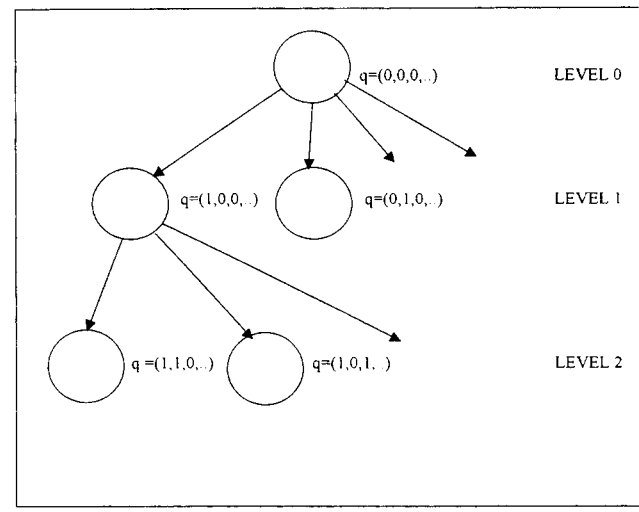

Figure 1. Tree of solutions.

1. The node $q = 0$ is trivially infeasible.

2. As measurements are added to form a branch, the nodes are infeasible. This is because key variables are unobservable or not enough precision has been achieved (low redundancy). Eventually, as measured variables are added, a node will become feasible. The node level at which feasibility is attained varies from branch to branch.

3. Cost increases from one level to the next if only one type of instrument is used ($n_i^m = 1$). In that case, the first feasible node in each branch is the one with the lowest cost in the tree it spans.

4. If more than one type of instrument is used ($n_i^m \geq 1$), then the cost of each node has a lower bound of $\ddot{C}(q) = \sum_{\forall i} \bar{c}_i q_i$, where $\bar{c}_i = \underset{\forall k}{\text{Min}}\{c_{i,k}\}$. This property is used below in the branching stopping criterion.

5. Finally, retrofit can be handled easily by freezing corresponding values of q.

Consider now the following tree enumeration scheme:

- Start with a root node with no variables being measured ($q = 0$).
- Use the branch first rule, that is, construct the tree by developing each branch (making one element of q active) until a stopping criterion is met. Then back up one level and develop the next branch. At each node, solve the optimization problem (Eq. 13).

Stopping criteria

Consider the case where only one type of instrument is used ($n_i^m = 1$). It is obvious that in such a case one should stop at the first feasible node and not develop the tree below that node, as any node below will be more expensive.

In the case where more than one type of instrument is used ($n_i^m \geq 1$), it is possible that some nodes belonging to the tree spanned by a feasible node can be cheaper. If the solution of the corresponding feasible node costs more than its lower bound, a cheaper solution can only exist in a few levels below this node. This number of levels is given by the nearest lower integer of $[C_b - \ddot{C}(q)]/\tilde{c}$, where $\tilde{c}$ is the cost of the cheapest instrument and C_b is the current best solution. Thus after a feasible node is obtained the maximum number of additional levels to explore is determined. The stopping criterion therefore becomes to stop after the maximum level of each feasible branch is reached and the node is evaluated.

Efficiency of the procedure

In the case where $n_i^m > 1$, the enumeration of the tree can also be based on developing a tree for p, rather than for q. In that case the simple criterion of stopping the tree at the first feasible node should be used. However, the tree will contain more nodes at each level and therefore it will be larger.

It is obvious that this stopping criterion scheme prevents the enumeration procedure from being exhaustive and impractical. However, it is still not an efficient procedure for large systems. Since the focus of this article is the conceptual development of the problem, investigations of numerically efficient schemes to solve this problem will be addressed in the future. In particular, a scheme where the changing dimensions of matrices A and B should allow a more efficient branch-and-bound procedure, or even facilitate the solution of the problem through some other methods.

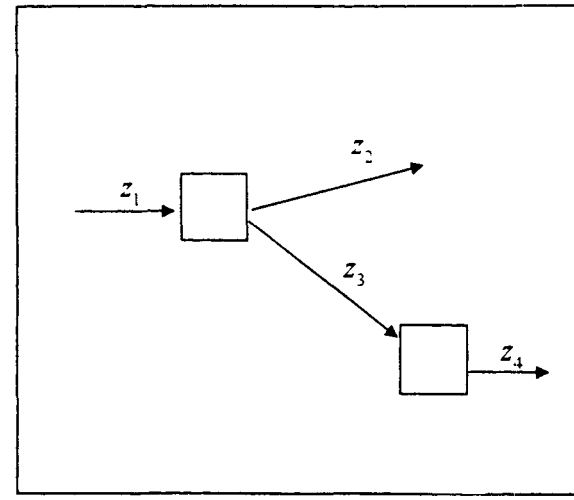

Figure 2. Example 1.

Example 1

Consider the process flow diagram of Figure 2. Flow rates are given by $z = (150.1, 52.3, 97.8, 97.8)$. Assume that for each rate, flowmeters of precision 3%, 2%, and 1% are available at costs 800, 1,500, and 2,500, respectively, regardless of size. Precision is only required for variables z_1 and z_4 ($I_s = \{1,4\}$), with $\sigma_{I_*} = 1.5\%$ and $\sigma_{4_*} = 2.0\%$. Two solutions are obtained featuring a cost of $C = 3{,}000$. The corresponding meters are shown in Table 1. Even though these solutions are low in cost, they feature nonredundant sensor networks. Though precision is achieved, gross errors are impossible to detect. Therefore the feasible region should contain redundant networks. If at least one degree of redundancy is requested, then there are two solutions with a cost of $C = 3{,}100$ (Table 2).

Availability

To avoid solutions consisting of nonredundant networks, one may ask each key variable to be redundant and/or the network to have a certain degree of redundancy. This can be achieved easily by requesting that the variance of estimates be lower than the variance of the measurement. In addition, redundancy is required to be able to perform data reconciliation after gross errors are deleted. But, how much redundancy is really required? This question is better answered by introducing the concept of *availability*.

Once a redundant measurement is found to have a gross error and is eliminated, the overall redundancy is lowered and the precision of all variables decreases. It is desirable

Table 1. Solutions of the Precision Constrained Problem

z_1	z_2	z_3	z_4
—	2%	2%	—
—	2%	—	2%

Table 2. Solutions of the Precision Constrained Problem (Redundancy Required)

z_1	z_2	z_3	z_4
3%	3%	2%	—
3%	3%	—	2%

that the precision of certain key variables remain above the specified threshold upon any deletion of gross errors.

Availability of a sensor network is defined as the ability of the network to guarantee a certain level of precision in key selected variables when gross errors are detected and the measurements are eliminated.

Since many gross errors can be found, the availability order is defined next. Availability of order one is achieved when precision is guaranteed upon deletion of one measurement, regardless of position. Availability of order two guarantees the desired level of precision upon the deletion of two measurements regardless of position, and so on.

Let A and B be the matrices describing the system for a particular choice of q, and let $A_{a,s}$ and $B_{a,s}$ be the corresponding matrices once x_s is eliminated. Also, let $\sigma_{a,i,s}$ be the variance of z_i once x_s has been eliminated, and $\sigma_{a*,i}$ the threshold value. Therefore, availability of variable i is guaranteed if the following inequality is satisfied for all measurements:

$$\sigma_{a,i,s} \leq \sigma_{a*,i} \qquad \forall s \in q. \tag{15}$$

If $\sigma_{k*} = \sigma_{a*,k}$, the constraint on precision can be dropped, as the former contains the latter. To require availability automatically means to require redundancy. In this sense availability is a more general concept than redundancy. Indeed, to achieve redundancy of a particular measurement it is sufficient to require a sufficiently low threshold value $\sigma_{a*,i}$.

Example 1 (continued)

Consider requiring that $\sigma_{a*,1} = 1.5\%$, and $\sigma_{a*,4} = 3\%$. Then there is one solution: $z = (2\%, 3\%, 3\%, 3\%)$ with cost $C = 3{,}900$. Now assume that availability is requested to the same level as precision. Then two alternative solutions with cost $C = 5{,}500$ are obtained (Table 3). Not only is the cost higher, but there is also one more degree of redundancy. For larger problems, the number of alternatives increases, requiring new criteria to further screen alternatives. Two more qualifying constraints are introduced next. They are related to the capacity of the synthesized network to handle gross errors, that is, to be able to detect them and limit their data-corrupting impact.

Table 3. Solutions of the Availability-constrained Problem ($\sigma_{a*,k} = \sigma_{k*}$)

z_1	z_2	z_3	z_4
1%	2%	2%	—
1%	2%	—	2%

Error Detectability

Another property that makes a sensor network robust is its capacity for detecting gross errors. In the presence of a gross error of size d_i in variable x_i, the objective function $Q = (x - x_m)^T S_{x,m}^{-1} (x - x_m)$ has a noncentral $\chi^2_{1-\alpha}$ distribution with a noncentrality parameter δ_i, that is, the mean of Q is $\delta_i + \nu$, where ν is the degree of freedom. This noncentrality parameter is related to d_j as follows (Madron, 1985):

$$\delta_i = d_i \frac{(\sigma_{i,m}^2 - \sigma_i^2)^{1/2}}{\sigma_{i,m}^2}. \tag{16}$$

Assume that the power of the gross error detection test is γ. Then, a gross error of size d_i larger than the threshold value

$$d_{i,\gamma} = \delta_{i,\gamma} \frac{\sigma_{i,m}^2}{(\sigma_{i,m}^2 - \sigma_i^2)^{1/2}} \tag{17}$$

will be detected with probability γ. Values of $\delta_{i,\gamma}$ are functions of the degrees of freedom ν and the significance level α used to detect the errors. These values have been tabulated and are available (Madron, 1985).

Thus, although the global test in itself cannot pinpoint the location of the gross error, one can quantify the effect of a gross error in the noncentrality of the global best, as discussed earlier. Having discussed this effect, we are in a position to work out this concept backwards. One can choose a threshold size of gross error for each variable d_{i*} and request that those errors larger than this threshold be detected by the network. If we define d_{i*} as a multiple of the standard deviation of the measured flow ($d_{i*} = n_i^D \sigma_{i,m}$), the error detectability criterion becomes

$$n_i^D > \delta_{i,\gamma} \frac{\sigma_{i,m}}{(\sigma_{i,m}^2 - \sigma_i^2)^{1/2}}. \tag{18}$$

Resilience

A sensor network is robust if it is capable of preventing undetected gross errors from corrupting the reconciled values. In the event that a gross error of a certain magnitude in any variable occurs and is not detected, a certain corruption of data will take place after the reconciliation is performed. Let $\Delta z_{i,s}$ be the change in the reconciled value of variable z_i when a gross error of size h is present in variable z_s. Thus, by using $x_m + e_j h$ instead of x_m in Eqs. 4 and 5, one obtains

$$\Delta z_{i,s} = \begin{cases} [-S_{x,m} A^T (I - G^{-1} B H^{-1} B^T) G^{-1} A e_s] h & \text{if } z_i \text{ is measured} \\ [H^{-1} B^T G^{-1} A e_s] h & \text{otherwise.} \end{cases} \tag{19}$$

Then, a desired level of resiliency is fixed for variable z_s, requesting that $\Delta z_{i,s}$ be lower than a certain threshold, Δz_{i*}. For each variable z_s the test is performed to find the gross

Table 4. Effect of Error Detectability Constraints ($n_i^D = 3.9$)

z_1	z_2	z_3	z_4
1%	3%	—	2%
1%	3%	2%	—

errors in every other measurement. Sometimes this threshold can be quantified in terms of cost (e.g., loss of product). Let the threshold be expressed in terms of the measured standard deviation through $\Delta z_{i*} = n_i^R \sigma_i$, and assume that h_s is chosen the same way, using the error detectability threshold, that is, $h_s = n_s^D \sigma_{s,m}$, the resilience criterion becomes:

$$n_i^R > \beta_{i,s} \left(\frac{\sigma_{s,m}}{\sigma_i} \right) n_s^D, \tag{20}$$

where

$$\beta_{i,s} = | e_i^T S_{x,m} A^T (I - G^{-1} B H^{-1} B^T) G^{-1} A e_s | \tag{21}$$

if variable z_i is measured, and

$$\beta_{i,s} = | e_i^T H^{-1} B^T G^{-1} A e_s | \tag{22}$$

otherwise. Usually n_i^R should exceed n_i^D.

Example 1 (Continued)

Error detectability has a large impact on feasibility, increasing the precision of the sensor network. Consider adding an error detectability of $n_i^D = 3.9$ (with $\gamma = 50\%$) to the availability-constrained problem ($\sigma_{a*,1} = 1.5\%$, $\sigma_{a*,4} = 3\%$). Two solutions from a set of only four feasible solutions are found with cost $C = 4{,}800$ (Table 4). If an error detectability of $n_i^D = 3.4$ for all measurements is requested, the problem has only one solution, namely $z = (1\%, 3\%, 1\%, 1\%)$, with cost $C = 8{,}300$. Now consider the addition of resilience: if error detectability is requested at a level of $n_i^D = 3.9$ for all measurements, and resilience is requested at a level of $n_i^R = 3$ for all flow rates, then the solution is again $z = (1\%, 3\%, 1\%, 1\%)$, with cost $C = 8{,}300$. Relaxing (increasing) the resilience levels maintaining the error detectability at the same level may actually lead to solutions of higher cost, even to infeasibility. The identification of sub-optimal networks and the addition of other constraints related to control and fault analysis is left for future work.

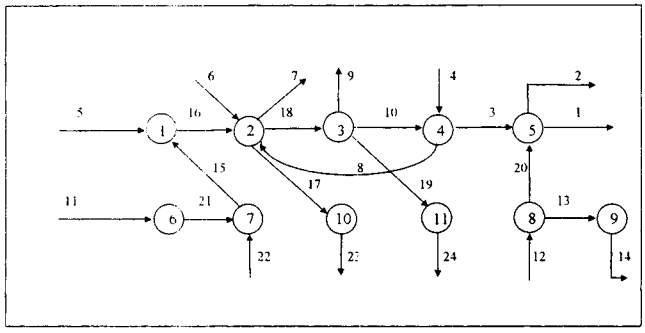

Figure 3. Example 2.

Table 5. Example 2. Cost of Flowmeters

Stream	1	2	3	4	5	6	7	8	9
Cost	19	17	13	12	25	10	7	6	5

From Meyer et al. (1994).

Example 2

The following example is of larger size. Consider a network proposed by Madron and Veverka (1992) (Figure 3). It contains several measured variables, and the problem consists of selecting what measurements should be added to make streams 1 and 5 observable. Madron and Veverka report the solution $S = \{2,4,8\}$, with a cost $C = 35$. Meyer et al. (1994) report the same solution using their own cost data (Table 5). Neither Madron and Veverka (1992) nor Meyer et al. (1994) reported flow rates for the measured variables. Thus the values shown in Table 6 were chosen to illustrate the method presented in this article.

If the requirement of precision is sufficiently relaxed, the same solution is found, that is, $S = (2,4,8)$. When flowmeters of 2.5% precision are used, this network has a precision of 27.8% in stream 5. Since this value is too high, resulting in poor monitoring performance, a 2.5 precision on stream 1 and 5 was requested. The solution found was $S = \{2,4,5,8\}$, with a cost $C = 60$. However, the network is nonredundant so error detectability cannot be requested. If error detectability is required at a level of $n_i^D = 3.9$ (with $\gamma = 50\%$) for flows 1 and 3, which are the larger flows, the solution obtained is $S = \{1,2,4,5,6,7\}$, with a cost of $C = 90$.

Conclusions

A novel method for designing and retrofitting cost-optimal sensor networks has been introduced. In addition to precision requirements, three new constraints were introduced. Networks that can guarantee the detection of gross errors (error detectability), prevent their data-corrupting effect (resilience), and maintain precision when they are detected (availability), can be identified by using these new constraints. Future work will include improved solving procedures, nonlinear networks (component and energy data reconciliation based), as well as network reliability issues. Because component, and especially energy data reconciliation, are used nowadays to add redundancy, and thus enhance data reconciliation, addressing this issue will allow the design and retrofit of more realistic networks. Sensor network reliability

Table 6. Flow Rates for Example 2

Stream	Flow	Stream	Flow
1	140	13	10
2	20	14	10
3	130	15	90
4	40	16	100
5	10	17	5
6	45	18	135
7	15	19	45
8	10	20	30
9	10	21	80
10	100	22	10
11	80	23	5
12	40	24	45

has been addressed by other authors and should be added as another constraints, so that this feature is taken into account in the design. The interaction of the concepts of reliability and availability are also another issue of future development, since it might be possible to effectively combine those issues in a single measure. Finally, the procedure needs to be nurtured with practical criteria as to what levels of error detectability and resilience are appropriate for a given problem.

Notation

D = linear model matrix for all variables
e = unit vector
G = auxiliary matrix
H = auxiliary matrix
m = hardware redundancy maximum
σ = square rooted diagonal element of S
β = auxiliary variable in resilience

Subscripts and superscripts

a = availability
D = error detectability
R = resilience
$*$ = desired threshold

Literature Cited

Ali, Y., and S. Narasimhan, "Sensor Network Design for Maximizing Reliability of Linear Processes," *AIChE J.*, **39,** 5 (1993).

Ali, Y., and S. Narasimhan, "Redundant Sensor Network Design for Linear Processes," *AIChE J.* **41,** 10 (1995).

Crowe, C. M., Y. A. Garcia Campos, and A. Hrymak, "Reconciliation of Process Flow Rates by Matrix Projection: I. The Linear Case," *AIChE J.*, **29,** 818 (1983).

DATACON Manual, "Application: Optimal Location of Instrument," Simulation Sciences, (1993).

Kretsovalis, A., and R. S. H. Mah, "Effect of Redundancy on Estimation Accuracy in Process Data Reconciliation," *Chem. Eng. Sci.*, **42,** 2115 (1987).

Liptak, B. G. *Instrument Engineers' Handbook*, 3rd ed., Chilton, (1995).

Luong, M., D. Maquin, C. T. Huynh, and J. Ragot, "Observability, Redundancy, Reliability and Integrated Design of Measurement Systems," *IFAC Symp. on Intelligent Components and Instrument Control Applications*, Budapest (1994).

Madron, F., "A Note on Detection of Gross Errors in Process Data," *Coll. Czech. Commun.*, **50,** 758 (1985).

Madron, F., and V. Veverka, "Optimal Selection of Measuring Points in Complex Plants by Linear Models," *AIChE J.*, **38**(2), 227 (1992).

Maquin, D., M. Luong, and J. Ragot, "Observability Analysis and Sensor Placement," *Safe Process IFAC/IMACS Symp. on Fault Detection, Supervision and Safety for Technical Process*, Espoo, Finland (1994).

Meyer, M. J., M. Le Lann, B. Koehret, and M. Enjalbert, "Optimal Selection of Sensor Location on a Complex Plant Using a Graph Oriented Approach," *Comput. & Chem. Eng.*, **18** (Suppl.), S535 (1994).

Ragot, J., D. Maquin, and G. Bloch, "Sensor Positioning for Processes Described by Bilinear Equations," *Rev. Diag. Surete Fonct.*, **2**(2), 115 (1992).

Vaclavek, V., and M. Loucka, "Selection of Measurements Necessary to Achieve Multicomponent Mass Balances in Chemical Plants," *Chem. Eng. Sci.*, **31,** 1199 (1976).

Whitaker, N. R., *Process Instrumentation Primer*, Petroleum Publishing (1980).

Manuscript received Oct. 23, 1996, and revision received May 19, 1997.