PROCESS DESIGN AND CONTROL

An Evolutionary Approach for the Design of Nonredundant Sensor Networks

Mercedes Carnero,[†] José Hernández,[†] Mabel Sánchez,[‡] and Alberto Bandoni*[,‡]

Departamento Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de Río Cuarto, Campus Universitario, (5800) Río Cuarto, Argentina, and Planta Piloto de Ingeniería Química, Camino La Carrindanga Km 7, C.C. 717, (8000) Bahía Blanca, Argentina

In this work, solution strategies for the optimal design of nonredundant observable linear sensor networks are discussed. The Greedy algorithm allows the problem only to be tackled for a subset of optimization criteria. Particular deterministic techniques or general evolutionary strategies are necessary to solve the problem for more complex objective functions. In this context, a procedure based on the application of genetic algorithms (GAs) and linear algebra is presented. Ad hoc operators are designed for the crossover and mutation operations because the classic genetic operators perform poorly. In contrast to ad hoc deterministic codes, which find the design solution for each specific criteria, this strategy allows the problem to be solved with different objective functions using the same implementation. Furthermore, this code is extended to handle multiobjective problems through a modification of only the selection operator. An industrial example is provided to show the efficiency of the algorithm.

1. Introduction

Currently, the availability of accurate and complete process knowledge is crucial for on-line plant optimization, control, safety, and environmental issues. For this reason, new sensors are located in the process, and measurements are adjusted through the application of data reconciliation procedures to enhance precision.

The structure of sensor networks is defined by the type, amount, precision, reliability, and location of their instruments. This structure has a great effect on the quality and availability of process knowledge.

Different objectives are fulfilled in the optimal design of sensor structures, for example, minimization of the instrumentation costs, maximization of precision, estimation of all or a required set of variables, etc. In any case, a combinatorial optimization problem subject to constraints arises that can be solved using either deterministic or stochastic approaches.

Vaclavek and Loucka[1] first addressed the problem of sensor location using the concepts of graph theory. They selected sensor structures for observing a set of required variables for bilinear systems. Later, Madron[2] used the concept of spanning trees to obtain the minimum-cost linear sensor network that allows for the observability of all unmeasured variables. For the case of maximizing the precision of the measurement system as a whole, this author found suboptimal solutions by inspecting the spanning trees of distance 1 from an initial structure.

The concept of graph cutsets was applied by Ali and Narasimhan[3,4] to design a nonredundant structure of instruments for linear and bilinear systems that maximizes the least reliability of estimation among all variables. These authors extended the first work to select instruments for a redundant sensor network with the same objective function.[5]

A strategy based on linear algebra was applied by Kretsovalis and Mah.[6] They developed a combinatorial search to incorporate measurements into an observable linear system. An objective function made up of a weighted average of the cost of the measurements and the precision of the estimates was minimized. Later, Madron and Veverka[7] proposed a methodology for sensor placement that applies the Gauss−Jordan decomposition of a matrix. They considered as objective functions the cost and a measure of the overall inaccuracy. The procedure gives suboptimal solutions.

Departing from graph theory and linear algebra approaches, Bagajewicz[8] formulated an optimization problem to obtain cost-optimal network structures for linear systems subject to constraints on precision and robustness. In Bagajewicz and Sánchez,[9,10] the duality between the minimum-cost model and a proposed maximum-precision model was stated for the design and upgrade of instrumentation. A tree search procedure with stopping criteria was applied to solve academic problems in the aforementioned works, which aim at developing the conceptual aspects of these problems.

Stochastic approaches were introduced recently for solving sensor network design problems. Sen et al.[11] developed an algorithm based on concepts of graph theory and genetic algorithms (GAs). It allows for the selection of flowmeters for nonredundant sensor net-

* To whom correspondence should be addressed. Fax: 54-291-4861600, Phone: 54-291-4861700. E-mail: msanchez@plapiqui.edu.ar, abandoni@plapiqui.edu.ar.

[†] Universidad Nacional de Río Cuarto.

[‡] Planta Piloto de Ingeniería Química.

10.1021/ie000941k CCC: $20.00
Published on Web 10/23/2001

works that optimize criteria of cost, reliability, or estimation accuracy.

In this work, we first discuss the application of the Greedy algorithm[12] to tackle a particular type of sensor network design problem. Then, an evolutionary strategy is presented to design minimum-number sensor structures for linear processes that ensure the observability of all unmeasured variables while optimizing single or multiple criteria. The solution procedure is developed using a genetic algorithm whose operators are modified using linear algebra concepts. The strategy is applied to locate flowmeters for an industrial steam metering network.

2. Problem Formulation

Let us represent the operation of a process under steady-state conditions by the following set of linear equations

$$\mathbf{Dz} = \mathbf{Ax} + \mathbf{Bu} = \mathbf{0} \tag{1}$$

where **D** is the incidence matrix of dimension $m \times n$, **z** is the $n \times 1$ vector of mass flowrates, **x** represents the vector of measured variables, and **u** represents the vector of unmeasured variables; **A** and **B** are compatible submatrices.

The optimal design of nonredundant sensor networks that ensures the observability of all unmeasured variables aims at selecting the set of $g = n - m$ flowrates that should be measured to estimate the unmeasured ones while optimizing a certain objective. Accordingly, matrix **D** is divided into submatrix **A** ($m \times g$) and submatrix **B** ($m \times m$), and vector **z** is partitioned into vector **x** ($g \times 1$) and vector **u** ($n \times 1$).

The observability of all unmeasured variables for a given set of measurements is guaranteed if the rank of **B** is m.[13] Thus, the aforementioned sensor network design problem can be formulated as a general discrete optimization problem

$$\begin{aligned} &\min/\max \quad f(\mathbf{q}) \\ &\text{s.t.} \quad \operatorname{rank}(\mathbf{B}(\mathbf{q})) = r_B = m \end{aligned} \tag{2}$$

where f stands for a single or multiobjective optimization criterion and **q** represents a binary vector such that $q_i = 1$ if the process variable i is measured and $q_i = 0$ otherwise ($i = 1, ..., n$).

Depending on the selected optimization criterion, different solution strategies of the combinatorial optimization problem are considered in this work. First, the Greedy algorithm is applied to solve optimal designs that fulfill special mathematical conditions. Then, evolutionary techniques based on linear algebra and GA are developed for more complex problems that involve single and multiple optimization criteria.

3. Greedy Algorithm

Let us consider that a subset system $S = (E, \theta)$ is a finite set E together with a collection θ of subsets of E that is closed under inclusion (that is, if $N \in \theta$ and $N' \subseteq N$, then $N' \in \theta$). The elements of θ are called independent. A combinatorial optimization problem associated with a subset system (E, θ) is the following: Given a fixed weight $w(e) \geq 0$ for each $e \in E$, find an independent subset θ that has the largest possible total weight.

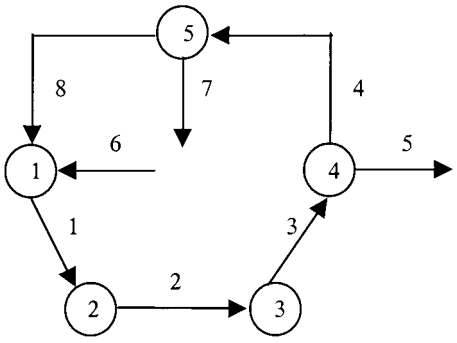

Figure 1. Simplified ammonia network.

If S is a subset system in which, for any subset $T \in E$, every maximal independent subset in T has the same cardinality, then S is a matroid. Consequently, the Greedy algorithm included in Appendix A correctly solves any instance of the combinatorial optimization problem associated with S.[14]

For our purposes, let us assume that (1) E is the set of columns of matrix **D**, that is, the set of columns of an $m \times |E|$ matrix; (2) θ is the collection of linearly independent sets of columns of **D**, that is, the collection of matrices **B** of full rank; and (3) w_i is the cost of each flowmeter, c_i , that can be located to measure the ith mass flowrate.

The subset system $S = (E, \theta)$ verifies the condition of a matroid because all maximal linearly independent subsets of a set of vectors E' have the same cardinality. In linear algebra, this cardinality is, in fact, called the rank of the submatrix **B**′ defined by E'. Consequently, the following combinatorial optimization problem can be efficiently solved using the Greedy algorithm

$$\begin{aligned} &\min_{\mathbf{q}} C = \sum_{i}^{n} c_i q_i \\ &\text{s.t.} \quad r_B = m \end{aligned} \tag{3}$$

Problem 3 arises as a particular case of the sensor network design problem 2 when capital cost C is considered as the objective function.

The Greedy algorithm is applied to determine the minimum-cost flowmeter location for the simplified ammonia network[3] depicted in Figure 1, which involves five units and eight streams. Its incidence matrix is

$$\mathbf{D} = \begin{bmatrix} -1 & & & & & & 1 & 1 \\ 1 & -1 & & & & & & \\ & 1 & -1 & & & & & \\ & & & 1 & -1 & -1 & & \\ & & & & 1 & & -1 & -1 \end{bmatrix}$$

If the flowmeter cost vector is $\mathbf{c}^T = [300\ 350\ 400\ 330\ 270\ 310\ 280\ 370]$, the procedure builds the matrix of unmeasured variables, **B**, by adding to the null matrix the columns of **D** corresponding to variables [3 8 2 4 6] one by one. These variables are associated with the most expensive flowmeters and form an independent set of columns of **D**. This condition is verified by matrix rank calculations before each column is incorporated to **B**. The optimal sensor network cost for this example is \$850.

In previous works,[2] it was shown that, to solve problem 3, sensors should be located on the chords of the maximum-cost spanning tree of the process graph,

and a deterministic procedure is given to obtain the solution. As matroid theory[14] unifies graph theory and linear algebra, we consider it valuable to pose the problem in terms of matroids. The matroids can be viewed as prototypes of independence systems and 0−1 integer programs with nice properties that allow efficient algorithms to be developed for the corresponding optimization problems.

If the design purpose is now to maximize the sensor system reliability R_s, defined by Maquin et al.[15] as the probability of estimating all variables, the following optimization problem arises

$$\max_{\mathbf{q}} R_s = \prod_{i=1}^{g} r_i(t) q_i \quad \text{s.t.} \quad r_B = m \tag{4}$$

where $r_i(t) = [1 - p_i(t)]$ is the reliability of a sensor at time t and p_i is its failure probability. Defining E as the set of columns of matrix $\mathbf{D}$, θ as the collection of linearly independent sets of columns of $\mathbf{D}$, and w_i as the reliability of each measurement, that is, $(1 - p_i)$, the above design problem can also be solved using the Greedy algorithm.

Let us consider again the process network illustrated in Figure 1. The Greedy algorithm is applied to maximize the system reliability for a nonredundant sensor structure. The vector of failure probability is $\mathbf{f}^T = [0.141\ 0.104\ 0.07\ 0.165\ 0.052\ 0.17\ 0.128\ 0.154]$. In this case, the optimal set of measurements corresponds to streams [3 5 7], and the maximum system reliability is $R_s = 0.769$.

Limitations of the Greedy Algorithm. This procedure allows the optimal solution of the aforementioned design problems to be obtained in a straightforward manner, because the objective functions are such that a constant weight can be associated with each member of E for the same set of problem constraints. However, several single and multiobjective criteria exist that do not verify this condition.

Single Criterion. If among the objective functions we consider, as an example, the measure of the overall precision defined by Kretsovalis and Mah,[6] the following optimization problem results

$$\min_{\mathbf{q}} \mathrm{GI} = \sum_{i=1}^{n} [\sigma_i^2 q_i + \hat{\sigma}_i^2(\mathbf{q})(1 - q_i)] \quad \text{s.t.} \quad r_B = m \tag{5}$$

where $\hat{\sigma}_i^2$ is the variance of the ith variable estimation. If i is a measured variable ($q_i = 1$), then $\hat{\sigma}_i^2 = \sigma_i^2$, where σ_i^2 is the standard deviation of the ith measurement error, but if i is an unmeasured variable ($q_i = 0$), then $\hat{\sigma}_i^2$ depends on the σ_i^2 values of the measured variables used to calculate it. In this case, the Greedy algorithm can not be applied because $\hat{\sigma}_i^2$ does not represent a set of constant weights.

In the same way, the maximization of the least reliability among all variables,[3] R_v, cannot be formulated in terms of constant weights. This problem is stated as follows

$$\max_{\mathbf{q}} \min_{\mathbf{q}} R_v(\mathbf{q}) \quad \text{s.t.} \quad r_B = m \tag{6}$$

To optimize both aforementioned criteria, specific deterministic strategies[16,3] and a GA approach[11] have been developed based on graph theory.

Multiobjective Criteria. Previously, the optimal selection of sensor structures that satisfy only one criterion was formulated. However, as different objective functions can be chosen, the designer should opt for one. To take into account different optimization criteria simultaneously, the problem can be written as follows

$$\min_{\mathbf{q}} f_a(\mathbf{q}) \quad a = 1, 2, ..., \mathrm{nof} \quad \text{s.t.} \quad r_B = m \tag{7}$$

where $f_a(\mathbf{q})$ ($a = 1, 2, ..., \mathrm{nof}$) represents each objective function to be optimized. No solution vector $\mathbf{q}$ exists that simultaneously optimizes all nof objective functions, but there exists a set of solution vectors $\mathbf{q}$ that is superior to the rest of the solutions in the search space when all objectives are considered and is inferior to other solutions in one or more objectives. Therefore, it is clear that the concept of optimality in multicriteria optimization deals with a set of solutions rather than a single solution.

Dominated and Nondominated Solutions. A solution $\mathbf{q}$ is said to dominate the other solution $\mathbf{p}$ if both of the following conditions are true: (1) The solution $\mathbf{q}$ is no worse (the operator $<$ denotes worse and $>$ denotes better) than $\mathbf{p}$ in all objectives, or $f_a(q) \not< f_a(p)$ for all $a = 1, 2, ...,$ nof. (2) The solution $\mathbf{q}$ is strictly better than $\mathbf{p}$ in at least one objective, or $f_a(q) > f_a(p)$ for at least one $a \in \{1, 2, ..., \mathrm{nof}\}$.

Considering a set P including nondominated solutions, if there exists no solution in the search space that dominates any member in the set P, then the solutions belonging to the set P constitute a global Pareto-optimal set.

4. An Evolutionary Technique

4.1. Motivation. If Greedy algorithms can not tackle the design problem, two alternatives emerge to solve it: (1) applying a deterministic strategy for each specific criterion and (2) applying a stochastic procedure that allows optimization of any single or multiple criteria.

As the second alternative is more flexible, in this work, a new evolutionary technique based on GA[18] is presented that applies well-known concepts from linear algebra. It is successfully applied to solve the nonredundant sensor network design problem for any optimization criteria.

4.2. Description of the Strategy. An evolutionary technique is a probabilistic algorithm that maintains a population of individuals $P(t) = \{p_1(t), p_2(t), ...\}$ for iteration t. Each individual represents a potential solution to the problem. Each solution p_i^t is evaluated to give some measure of its fitness. Then, a new population (iteration $t + 1$) is formed by selecting the individuals that fit better. Some members of the new population undergo transformations by means of genetic operators to form new solutions. There are unary transformations (mutations) that produce new individuals by a small change on a single individual, and higher-order transformations (crossovers) that form new individuals by combining parts from several individuals. The program converges after some number of generations. The best individual is considered a near-optimum solution.

For sensor network design, the solution is represented by a fixed-length binary string, **q**, as in classical GAs. Nevertheless, the basic genetic operators cannot be applied to solve the design problems. The main drawback is that the feasibility of new population members generated using these operators decreases significantly. To overcome this problem, a new technique based on linear algebra is proposed. The relevant features of the new procedure are considered next.

4.2.1. Initial Population. In contrast to classical GAs, we start with a feasible initial population that is obtained by the following steps: (a) Any set of unmeasured variables is selected that satisfies the condition that matrix **B** has full row rank. Thus, one feasible solution of the problem is generated ($\mathbf{q_0}$). (b) Matrix **D** is divided accordingly into two submatrices, **A** and **B**, the first of which corresponds to measured variables. (c) As matrix **B** has full rank, matrix $\mathbf{B}^{-1}$ exists; thus, the vector **u** of unmeasured variables is formulated as

$$\mathbf{u} = -\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \mathbf{R}\mathbf{x} \tag{8}$$

and the ith element of vector **u** can be evaluated as

$$u_i = \sum_{j=1}^{j=g} r_{ij}x_j \quad \forall i = 1, ..., m; \forall j = 1, ..., g \tag{9}$$

(d) Matrix **R** contains the necessary information to determine which measured and unmeasured variables in $\mathbf{q_0}$ can be interchanged while preserving system observability. The members of a feasible population are defined by applying the following statement: A pair of variables (u_i, x_j) can be interchanged if the coefficient $r_{ij} \neq 0$.

As an example, let us consider a matrix **D** of dimension (3 × 6)

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

If $\mathbf{q_0}$ = [1 1 1 0 0 0], then **R** is the following (3 × 3) matrix

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

from which vector **u** is estimated as

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + 0.x_3 \\ x_1 - x_2 + 0.x_3 \\ 0.x_1 - x_2 + x_3 \end{bmatrix}$$

where u_1 corresponds to flow z_4, u_2 to z_5, and u_3 to z_6. It can be seen that feasible interchanges are (x_1, u_1), (x_2, u_1), (x_1, u_2), (x_2, u_2), (x_2, u_3), and (x_3, u_3). Each interchange originates a matrix **B** of dimension 3 × 3 of full rank.

4.2.2. Selection. Selection is based on the value of the fitness of each member of the population. The roulette-wheel selection strategy[18] is used (proportional selection).

4.2.3. Crossover. Given two strings of feasible solutions $\mathbf{q_1}$ and $\mathbf{q_2}$, a crossover operation is performed as follows: (a) Compute two vectors, $\mathbf{px}_i$ and $\mathbf{pu}_i$, to store the ordinal position of ones and zeros for each chromosome $\mathbf{q}_i$. (b) Define $K = \{k \mid \mathbf{px_1}(k) \neq 0\}$, $L = \{l \mid \mathbf{px_2}(l) \neq 0\}$, and $H = K \ll L$. If at least one element k in K and one element l in L do not both belong to H, then

$$\text{set } i = \mathbf{pu_1}(l),\ j = \mathbf{px_1}(k),\ s = \mathbf{pu_2}(k),\ b = \mathbf{px_2}(l)$$

(c) If $r_{1,i,j} \neq 0$ (for $\mathbf{R_1}$ associated with chromosome $\mathbf{q_1}$) and $r_{2sb} \neq 0$ (for $\mathbf{R_2}$ associated with chromosome $\mathbf{q_2}$), then the first offspring adopts the elements of $\mathbf{q_2}$ in positions (k, l); accordingly, the second offspring adopts the elements of $\mathbf{q_1}$ in the same positions.

Let us suppose that two feasible parents are obtained for the example presented in section 4.1. They are represented by the strings $\mathbf{q_1}$ = (1 1 1 0 0 0) and $\mathbf{q_2}$ = (1 0 0 0 1 1). The corresponding $\mathbf{R}_i$ matrices are

$$\mathbf{R_1} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \qquad \mathbf{R_2} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

The ordinal positions of the ones and zeros for each chromosome are contained in the vectors

$$\mathbf{px_1} = [1\ 2\ 3\ 0\ 0\ 0] \quad \mathbf{pu_1} = [0\ 0\ 0\ 1\ 2\ 3]$$

$$\mathbf{px_2} = [1\ 0\ 0\ 0\ 2\ 3] \quad \mathbf{pu_2} = [0\ 1\ 2\ 3\ 0\ 0]$$

from which sets K, L, and H are obtained as K = {1,2,3}, L = {1,5,6}, and H = {1}. The intersection set H contains a variable only in position 1; thus, variables in position 2 and 3 from $\mathbf{q_1}$ and in position 5 and 6 from $\mathbf{q_2}$ can be considered for interchange. If the pair (2, 5) is selected, then $i = \mathbf{pu_1}(5) = 2$, $j = \mathbf{px_1}(2) = 2$, $s = \mathbf{pu_2}(2) = 1$, and $b = \mathbf{px_2}(5) = 2$.

The coefficients r_{22} of $\mathbf{R_1}$ and r_{12} of $\mathbf{R_2}$ are inspected. As they are not zero, the following offspring result: $\mathbf{o_1}$ = (1 0 1 0 1 1) and $\mathbf{o_2}$ = (1 1 0 0 0 1).

4.2.4. Mutation. The mutation operation is performed on a population member if its mutation probability exceeds the value of the mutation probability parameter. In this case, a measurement selected randomly is replaced by the unmeasured variable of lower cost that preserves the observability of the system, as explained in section 4.1.

For the example in the aforementioned section, let us consider the chromosome $\mathbf{q_0}$ = (1 1 1 0 0 0); the measured variable x_1 is selected for the mutation operation. It can be swapped with u_1 or u_2. If u_1 is the variable of lower cost, then the mutated chromosome is $\mathbf{q}_0^*$ = (0 1 1 1 0 0).

4.3. Design of Redundant Networks. The preceding algorithm can be extended to the design of redundant sensor networks, when a number s of sensors, greater than the minimum number (g), is used to ensure the observability of all unmeasured variables.

In this case, submatrices **A** and **B** corresponding to measured and unmeasured variables are of sizes $m \times s$ and $m \times (n - s)$, respectively. The constraint on observability is verified if the rectangular matrix **B** has exactly $n - s$ linearly independent columns.

This fact guarantees that **B** has a left inverse **B**** given by

$$\mathbf{B}^* = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T \tag{10}$$

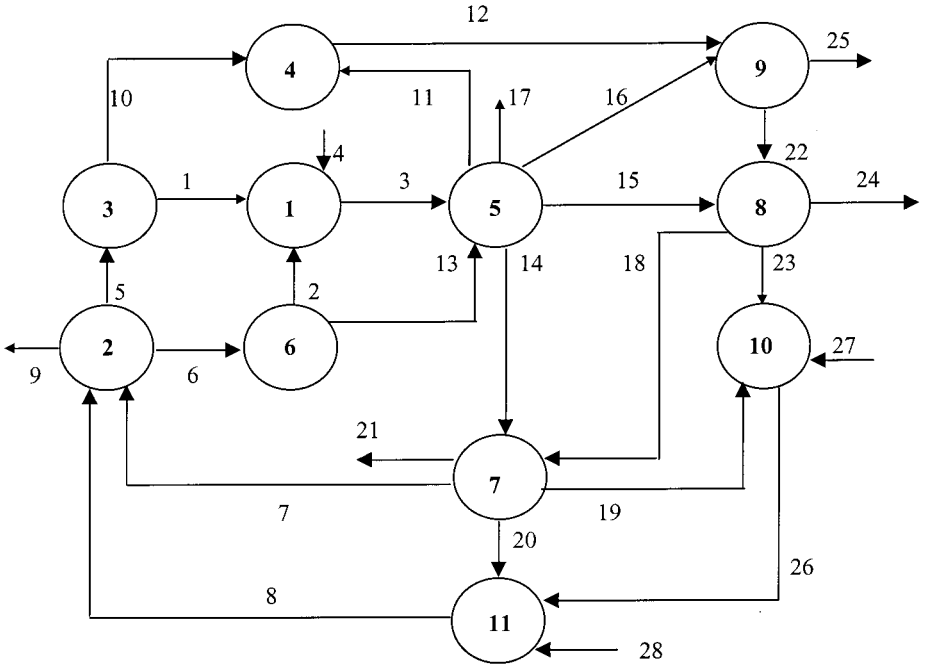

Figure 2. Steam metering network.

which allows us to rewrite eq 8 as

$$\mathbf{u} = -\mathbf{B}^*\mathbf{A}\mathbf{x} = \mathbf{R_1}\mathbf{x} \tag{11}$$

If **B** is a square matrix (nonredundant case), then the full row range also implies a full column range, and the left inverse coincides with $\mathbf{B}^{-1}$.

In this way, $\mathbf{R_1}$ again contains the necessary information to determine the measured and unmeasured variables that can be exchanged to preserve the feasibility of the solutions.

4.4. Multiobjective Optimization. Several methods based on a stochastic approach are available for solving multiobjective optimization problems. Srinivas[19] proposed the NSGA (nondominated sorting genetic algorithm). In this procedure, the solutions are ranked into nondominated order, and a niche method is used to maintain stable subpopulations of good points.

The selection method of the aforementioned evolutionary algorithm is replaced by a nondominated sorting procedure in conjunction with a sharing technique that avoids the bias toward some Pareto-optimal solutions. Before the selection is performed, the population is ranked on the basis on an individual's nondomination level, which is found by mean of successive classifications in fronts. A front is a set of nondominated solutions.

Once the first front is obtained, the solutions belonging to it are ignored, and the next front is formed with the solutions dominated by the first one. All solutions in each front are assigned the same dummy fitness value. Then, to maintain diversity in the population, these nondominated solutions are shared with their dummy fitness.

Sharing is achieved by dividing the dummy fitness value of an individual by a quantity proportional to the number of individuals around it. The population is then reproduced with the shared fitness values. Crossover and mutation remain without changes.

In this work, the sharing method is carried out by means of the following sharing function[20]

$$Sh(i,j) = \begin{cases} 1 - (d_{ij}/\sigma_{\text{share}})^2 & \text{if } d_{ij} \le \sigma_{\text{share}} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where d_{ij} is the Hamming distance between a solution i and another solution j.

The sharing function is then used to calculate the degradation factor $m_i = \sum_{j=1}^{n_k} Sh_{ij}$ (where n_k is the number of individuals in the kth front).

Finally, the shared fitness is and this new shared

$$f_{i-sh} = \frac{f_k}{m_i} \tag{13}$$

fitness is used to perform the selection procedure.

4.5. Examples and Results. The evolutionary technique is first applied to solve problem 5 for the case of a steam metering network of a methanol synthesis unit.[21] The process flowsheet is illustrated in Figure 2. It involves 11 units and 28 streams. From Sen et al.,[11] data for costs, failure probabilities, and standard deviations for measurement error are obtained and reproduced in Table 1. The evolutionary technique is run using the parameters indicated in Table 2 for the GA. The optimal set of measured variables and the objective function value are presented in Table 3; they are in agreement with those obtained using a graph-oriented approach.[11]

Then, the optimization criterion is changed. The same code and parameters are applied to maximize the least variable reliability among all variables, R_v. Five different optimal solutions arise for this design problem, which are verified using complete inspection.

Furthermore, the procedure is applied to minimize the instrumentation costs and to maximize system reliability. Although these problems can be satisfactorily solved by the Greedy algorithm, they are addressed using the new approach to show its usefulness with different objective functions. All results are included in Table 3; they correspond to the global optimum, as was verified by complete inspection and deterministic methods.

Table 1. Data for Steam Metering Network

variable	p_i	σ_i	c_i
1	0.141	0.0215	3.7
2	0.174	0.025	4.5
3	0.104	2.8	132.2
4	0.07	2.749	129.2
5	0.174	1.332	65.3
6	0.096	2.807	132.4
7	0.164	0.058	5.0
8	0.165	4.101	193.9
9	0.055	0.0215	2.06
10	0.099	1.31	62.8
11	0.153	0.3715	20.2
12	0.092	1.682	80.0
13	0.052	2.782	130.4
14	0.154	2.296	109.8
15	0.095	1.5	71.6
16	0.116	0.591	29.7
17	0.077	0.8182	39.5
18	0.08	0.4057	20.4
19	0.099	0.1987	11.1
20	0.074	0.2625	13.6
21	0.071	2.182	102.9
22	0.094	0.1362	8.1
23	0.170	0.0648	6.3
24	0.066	1.166	55.5
25	0.088	2.136	101.0
26	0.143	2.033	93.7
27	0.128	1.769	84.7
28	0.075	1.806	85.4

Table 2. Genetic Algorithm Parameters for the Single Criterion Optimization

parameter	value
population size	50
crossover probability	0.99
mutation probability	0.50
number of generations	20

Table 3. Single Criterion Optimization Results

objective criterion	objective function value	measured variables
GI	85.00	1, 2, 7, 9, 10, 11, 16−24, 27, 28
R_v	0.5282	3, 4, 5, 8, 9, 10, 13, 15, 16, 17, 22−28
		3, 4, 6, 8, 9, 10, 11, 13, 15, 17, 22−28
		3, 4, 6, 8, 9, 10, 12, 13, 15, 17, 22−28
		3, 4, 6, 8, 9, 10, 13, 15, 16, 17, 22−28
		1, 3, 4, 8, 9, 10, 11, 13, 15, 17, 22−28
C	555.5	1, 2, 7, 9, 10, 11, 16−24, 27, 28
R_s	0.235	3, 4, 6, 9, 10, 12, 13, 15, 17−22, 24, 25, 28

Table 4. Objectives Values for Each Optimization Criterion

	GI	R_v	C	R_s
GI	85	0.265	555.5	0.141
R_v	245.9	0.528	1291.3	0.151
	278.8		1349.0	0.116
	285.2		1408.8	0.124
	282.3		1358.5	0.137
	231.8		1220.3	0.110
C	85	0.265	555.5	0.141
R_s	268	0.376	1178.2	0.235

In Table 4, all objective functions values for each optimization criterion are shown. For this particular set of instrumentation data, it can be seen that a conflict exists between the different objectives, except for the cost and overall precision case.

For example, for a nonredundant network, the optimum design corresponds to a cost of 555.5, with a reliability R_s of 0.141, whereas the system reliability value is 0.235, almost double, for the most reliable network, which has a cost of 1178.2. The same analysis can be done for cost versus R_v.

Table 5. Pareto Optimal Set

reliability	cost	measure variables
0.1672	741.96	1, 4, 7, 9, 10, 11, 13, 15, 16−23, 28
0.1681	773.66	1, 2, 4, 7, 9, 10, 13, 15−24, 28
0.1457	616.06	1, 2, 4, 7, 9, 10, 11, 15−23, 28
0.1780	790.16	1, 4, 7, 9, 11, 12, 13, 15−20, 22, 24, 27, 28
0.1671	666.46	1, 2, 7, 9, 10, 11, 13, 15−22, 24, 28
0.1539	647.06	1, 2, 4, 7, 9, 10, 11, 15−20, 22, 24, 27, 28
0.1386	564.46	1, 2, 7, 9, 10, 11, 16−24, 26, 28
0.1551	664.26	1, 2, 4, 7, 9, 11, 12, 15−20, 22, 24, 27, 28

Table 6. NSGA Parameters

parameter	value
population size	50
crossover probability	0.7
mutation probability	0.10
number of generations	300
σ_{sh}	0.4

If combinations of cost (C) and system reliability (R_s) criteria are considered simultaneously, then a multiobjective optimization is performed with the NSGA code. The set of Pareto-optimal solutions is shown in Table 5. As can be observed, the cost is higher than that corresponding to a single objective, whereas the reliability of the network decreases as compared with that of the most reliable network.

The genetic algorithm parameters are shown in Table 6.

5. Conclusions

In this paper, strategies for the design of nonredundant sensor networks are discussed. The Greedy algorithm allows specific types of problems to be solved, for instance, when cost and system reliability are considered as the objective functions. To perform instrumentation design efficiently for any optimization criteria, an evolutionary technique based on GA is presented. The classical genetic operators are modified to avoid finding many infeasible solutions. New crossover and mutation operators are designed using concepts from linear algebra. A nondominated sorting procedure in conjunction with a niche formation technique are implemented to solve multicriteria sensor design problems in which a set of multiple Pareto-optimal solutions must be found. Application results are provided for an industrial steam metering network.

The main advantage of the evolutionary approach is that optimal or near-optimal solutions can be found in polynomial time, independently of the computational complexity of the problem. The methodology is very flexible, and alternative objectives or multiojectives functions can be easily used.

Appendix A: Greedy Algorithm[12]

The solution procedure of this algorithm is as follows:

; **I** set with final solution

; E set of columns of matrix **D** (m x $|E|$)

; **e** any element of E

; θ collection of linearly independent sets of columns of **D**

begin

$\mathbf{I} := \varnothing$

while $E \neq \varnothing$ do

begin

let e be the element of E that has the largest/smallest weight

remove e from E

if $\mathbf{I} + e \in \theta$, then $\mathbf{I} := \mathbf{I} + e$

end

end

Literature Cited

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

(2) Madron, V. *Process Plant Performance*; Ellis Horwood: Chichester, U.K., 1992.

(3) Ali, Y.; Narasimhan, S. Sensor Network Design for Maximizing Reliability of Linear Processes. *AIChE J.* **1993**, *39*, 820.

(4) Ali, Y.; Narasimhan, S. Sensor Network Design for Maximizing Reliability of Bilinear Processes. *AIChE J.* **1996**, *42*, 2563.

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

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

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

(8) Bagajewicz, M. Design and Retrofit of Sensor Networks in Process Plants. *AIChE J.* **1997**, *43*, 2300.

(9) Bagajewicz, M.; Sánchez, M. Duality of sensor network design models for parameter estimation. *AIChE J.* **1999**, *45*, 661.

(10) Bagajewicz, M.; Sánchez, M. Sensor Network Design and Upgrade for Plant Parameter Estimation. *Comput. Chem. Eng.* **1999**, *S23*, 585.

(11) Sen, S.; Narasimhan, S.; Deb, K. Sensor Network Design of Linear Processes using Genetic Algorithms. *Comput. Chem. Eng.* **1998**, *22*, 385.

(12) Nemhauser, G. L.; Wolsey, L. A. *Integer and Combinatorial Optimization*; John Wiley & Sons: New York, 1988.

(13) Romagnoli, J.; Sánchez, M. *Data Processing and Reconciliation for Chemical Process Operations*; Academic Press: San Diego, 1999.

(14) Papadimitriou, C.; Steiglitz, K. *Combinatorial Optimization—Algorithms and Complexity*; Prentice Hall: Upper Saddle River, NJ, 1982.

(15) Maquin, D.; Luong, M.; Ragot, J. Observability Analysis and Sensor Placement. Presented at SAFE PROCESS '94, IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Process, Espoo, Finland, June 13−15, 1994.

(16) Bansal, P.; Ali, Y.; Narasimhan S. Sensor Network Design for Linear Processes. Presented at IFAC Workshop on Integration of Process Design and Control, Baltimore, MD, June 27−28, 1994.

(17) Rao, S. *Engineering Optimization*; John Wiley & Sons: New York, 1996.

(18) Michalewicz, Z. *Genetic Algorithms + Data Structures = Evolution Programs*; Springer-Verlag: New York, 1992.

(19) Srinivas, N.; Deb, K. Multiobjetive Optimization Using Nondominated Sorting In Genetic Algorithms. *J. Evol. Comput.* **1994**, *2* (3), 221..

(20) Deb, K. Goldbeg, D. An Investigation of Niches and Species Formation in Genectic Function Optimization. *Ann. Math. Artif. Intell.* **1989**, *10*, 385.

(21) Serth, R.; Heenan, W. Gross Error Detection and Data Reconciliation in Steam-Metering Systems. *AIChE J.* **1986**, *32*, 733.

Received for review November 2, 2000
Revised manuscript received August 8, 2001
Accepted August 23, 2001

IE000941K