AIChE WILEY InterScience DISCOVER SOMETHING GREAT

Quantification of Performance of Sensor Networks for Fault Diagnosis

Sridharakumar Narasimhan and Raghunathan Rengaswamy
Dept. of Chemical Engineering, Clarkson University, Potsdam, NY 13699

DOI 10.1002/aic.11105
Published online February 27, 2007 in Wiley InterScience (www.interscience.wiley.com).

Safety and optimality are crucial requirements in every industrial process. The success of any fault diagnosis technique depends critically on the sensors measuring the important process variables. Choosing an appropriate sensor network is a combinatorially difficult problem, especially when the number of potential measurements is large. There has been considerable amount of work that has been done on developing algorithms for sensor network design for fault diagnosis based on quantitative and qualitative models. Various objectives, such as cost, reliability and fault resolution have been used in the sensor network design. While these design algorithms can provide the best design locations for a given cost, the value of the sensor network for fault diagnosis or benefit accrued is usually not quantified in a manner that is transparent to the user. This is an important aspect that needs to be addressed if these algorithms have to be assimilated into industrial practice. An approach for characterizing the value of a sensor network from a fault diagnosis perspective is proposed. This notion of value can be used directly in sensor network design algorithms. The proposed concepts are explained through a simple example and numerical simulations of a CSTR. *AIChE J,* 53: 902–917, 2007
Keywords: fault diagnosis, sensor networks

Introduction

Fault detection and isolation (FDI) or fault diagnosis is the task of identifying the causal origins of a fault or malfunction in the process, given current and past data available through sensor measurements and *a priori* knowledge. The problem of identifying faults using a given set of sensor measurements has been extensively discussed in the literature. Much less attention, however, has been paid to the problem of locating sensors that provide maximum information for effective fault diagnosis. Some work has started to appear in the literature to solve this problem.[1–5] The success of a fault diagnosis strategy thus depends critically on the choice of sensors and sensor network design becomes an important task.

The hitherto proposed approaches for solving the sensor network design problem from fault diagnosis perspectives require the engineer to specify the level of resolvability of the faults that is required of the network. Here, it is implicitly assumed that safety concerns have been addressed and fault diagnosis strategies are primarily geared towards addressing operational issues. While partial resolvability is all that might be required for a given situation, it is difficult to justify such a choice. Hence, the default approach has been to look for maximum resolvability that is inherently possible for the system that is being studied. This is one shortcoming in the existing design approaches from the viewpoint of fault diagnosis.

Correspondence concerning this article should be addressed to R. Rengaswamy at raghu@clarkson.edu.

Researchers have worked in the area of sensor network design that are optimal with respect to various objectives. In addition to fault diagnosis, the sensed data can be used for (1) optimizing process operation for maximizing productivity, and (2) process control. Techniques for sensor selection based largely on state estimation and material accounting have been reviewed.[6] Reviews of the hardware selection literature for control systems have also been published.[7,8] However, there is only one sensor network that has to cater to all the process objectives, and, hence, sensor network design involves multiple objectives, designers and decision makers. Hence, development of an integrated sensor network design algorithm that addresses all these objectives simultaneously is an important activity. The current sensor network design algorithms for each individual objective are based on performance indices that are inherently incommensurate. One method of solving multiobjective optimization problems is the development of a utility function. Recently, some researchers have proposed to quantify the value of the sensor network from each of the perspectives, which can be used to develop a suitable utility function.[9] They have made progress in quantifying the value of sensor networks from the precision and accuracy perspectives. Simultaneously, value formulations need to be developed for the other perspectives also. There has been no work on value formulation from fault diagnosis viewpoint in the literature. This is a second gap in the area of sensor network design for fault diagnosis. While the final goal of this exercise in quantification is an integrated sensor network design, the focus of this contribution is sensor networks for fault diagnosis.

This article is geared towards addressing these problems. What is also lacking is a conceptually and theoretically sound framework for answering the questions posed above. This is the third major contribution in this work. Rational suggestions based on engineering judgement for the choice of various quantities that are required for the value calculation are provided. It is shown that these choices are theoretically and conceptually consistent with the proposed modelling approach. Hence, the proposed framework allows the necessary calculations to be performed with available data and methods. This is the fourth significant contribution.

Organization

Faults in a process can be classified as structural, parametric or sensor faults.[10] This article addresses the question of quantifying the value of a sensor network when parametric faults are assumed to occur. A further assumption is that only single faults can occur at a time. This assumption has a powerful intuitive appeal and will be justified with sound arguments. This only simplifies the succeeding analysis and does not limit the applicability of the proposed approach. The theoretical and conceptual tools that will be needed in this work are described in the Preliminaries section. The concept of using the expected operating profit for the value, and a valid approximation to determine the same is motivated subsequently. A two-step procedure for determining the value of a given sensor network from a fault diagnosis perspective is described. In the first step, the set of resolvable faults for the given sensor network is determined using the diagnostic strategy of choice. Subsequently, the value of the sensor network is determined by quantifying the value in terms of the expected instantaneous operating profit. These concepts are illustrated using a simple example, and more relevant and illuminating simulations of a CSTR.

Preliminaries

A concise statement of the problem that is being solved is the following: Given a process model and a definition of all the parametric faults that can occur, what is the value of any potential set of sensor measurements (the network)? The proposed solution strategy breaks down the question into two parts. The first part involves determining the diagnostic properties of the sensor network, and the second part involves the actual computation of the value of the sensor network. The tools and methods needed for the same will be developed in this section.

System specifications

In the following treatment, we assume a system $\mathbf{S}$ in which all processes occur in discrete time: $t = 1,2, \ldots, t \in \mathbb{N}$. The system is characterized by its state, $\mathbf{s}(t)$ (some of which may be measurable), measured inputs and outputs (which will be lumped together as a single variable $\mathbf{y}$), and other unmeasured disturbances $\mathbf{d}$. "Faults" are assumed to affect the system at these discrete times. The term fault shall be used in a generic manner to include traditionally defined faults, abnormal events, upsets, disturbances (which may be measured or unmeasured, and will be considered abnormal if they exceed the normal range), which are relevant from the point of diagnosis, that is, those abnormal events that need to be detected and isolated. Faults that can be parameterized by measurable quantities or those that can be suitably quantified or inferred will be considered as parametric faults. Examples of parametric faults include changes in flow rate, pressure, temperature, concentration, and so on. Faults such as heat exchanger fouling or catalyst deactivation can be quantified by change in overall heat-transfer coefficient and catalyst activity, respectively. Other kinds of faults, such as sensor faults or structural faults can also be considered. However, the analysis in this paper is restricted to parametric faults. A fault f_i is parameterized by the quantity x_i and is considered to occur, if $x_i \in X_{fi}$ and f_i is assumed not to occur if $x_i \in X_{ni}$, where X_{ni} and X_{fi} are expected normal and abnormal operating regions which can be suitably described. In addition, we assume that it is possible to specify an operating profit function $c(\mathbf{s}, \mathbf{y}, \mathbf{d}, \mathbf{x}, t)$ that is a time varying function of the states $\mathbf{s}$, inputs and outputs $\mathbf{y}$, fault quantities $\mathbf{x} = [x_1, \ldots, x_n]^t$, and the various costs involved. We neglect the effect of unmodelled disturbances in this study. We assume that the states $\mathbf{s}(t)$, the vector of fault parameters $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t$ and operating profit $c(t)$ (where the dependency of the instantaneous profit c with respect to the $\mathbf{s}$, $\mathbf{y}$, $\mathbf{x}$ has been dropped for brevity) can be modelled as stochastic processes.

Fault occurrence modes

It is common practice in the fault diagnosis literature to define the modes of occurrence of the faults *a priori*. In this analysis, it will be assumed that only one fault can occur at a time. This is not a restrictive assumption and is fairly common in literature. Intuitively, it can be argued that the probability of occurrence of simultaneous multiple faults is rather low. The succeeding analysis requires that the probability of two or more faults occurring simultaneously be strictly zero. This assumption, while aiding the analysis, is also motivated by the observation that in reality, multiple fault occurrences are rarer than single fault occurrences. For instance, consider three faults f_1, f_2, f_3 each occurring independently with probabilities p_1, p_2, p_3. Since the numbers p_i are probabilities of fault occurring, it is reasonable to assume that $p_i \ll 1$. The probability of a single fault occurring is $p_1(1 - p_2)(1 - p_3) + p_2(1 - p_1)(1 - p_3) + p_3(1 - p_1)(1 - p_2)$, while that of two faults occurring is $p_1p_2(1 - p_3) + p_2p_1(1 - p_3) + p_3p_1(1 - p_2)$. Since $p_i \ll 1$ and $1 - p_i \approx 1$, each term in the second expression is much smaller than the corresponding term in the first expression and so, the overall probability of two faults occurring is much lower than that of single faults occurring. This analysis can be extended to situations involving more than 3 faults with similar results.

Fault observability

The manner in which fault observability is defined is consistent with that of Raghuraj et al.,[1] but different from the traditional definitions of state observability. Intuitively, fault observability is related to the ability of the sensor network, and the diagnosis strategy to detect faults. A rigorous definition of fault observability is given in the following.

Definition 1. Fault observability: Given a set of faults $F = \{f_1, f_2, \ldots, f_n\}$ define a fault indicator vector $\mathbf{f}^i = [f_1^i, \ldots, f_n^i]$ where the jth component of $\mathbf{f}^i$ is 1 if the jth fault has occurred and conversely, it is 0 if it has not occurred. Given a sensor network, $S = \{S_1, S_2, \ldots, S_m\}$, the corresponding measurements $\mathbf{y} = [y_{S1}, y_{S2}, \ldots, y_{Sm}]^t$, and a fault diagnosis strategy, define a residual vector, $\mathbf{r} = [r_1, \ldots, r_n]'$ (the output of the diagnosis strategy). A fault f_j is observable if

$$f_j^i \neq 0 \Longrightarrow \mathbf{r} \neq \mathbf{0}$$

Fault resolvability

While the earlier property of fault observability is useful for fault detection, complete fault diagnosis requires being able to isolate the root cause of the fault. The property of resolvability as defined below follows that of Raghuraj et al.,[1] and will be useful in deciding the isolability properties of the sensor network. Intuitively, two observable faults f_i and f_j can be resolved if they produce "different" symptoms, the difference being appropriately defined. Conversely, the two faults are said to be unresolvable if they produce "identical" symptoms. It is sometimes more convenient to work with unresolvability rather than resolvability. Formally, in a single fault scenario, one can view the property of unresolvability as a relation or subset of $F \times F$, where $F = \{f_1, \ldots, f_n\}$.

Definition 2. Single fault unresolvability: Given two distinct observable faults, f_i and f_j, in a single fault scenario, that is, $f_i^i \wedge f_j^i = 0, f_i^i \vee f_j^i = 1$, f_i and f_j are unresolvable if, it can be determined that $r_i \vee r_j = 1$, but neither r_i nor r_j can be determined. When f_i and f_j are unresolvable, denote f_i UR f_j to mean $(f_i, f_j) \in UR$, where, $UR \subset F \times F$. In addition, define f_i UR f_i, $\forall i$. Conversely, f_i and f_j are resolvable if f_i and f_j are not unresolvable.

Obviously, determining whether f_i and f_j can be resolved from each other depends on the chosen sensor network and the diagnostic technique used. The fault modeling and diagnostic technique of choice is used to determine which faults can be resolved from each other. Formally, given a set of sensors $S = \{S_1, S_2, \ldots, S_m\}$, and a set of observable faults $F = \{f_1, f_2, \ldots, f_n\}$, it is necessary to determine $\mathcal{F}_{end} = \{F_1, F_2, \ldots, F_k\}$, where each F_i is in general a set of faults $\{f_{i_1}, \ldots, f_{i_l}\}$, and a subset of F. If any F_i is a singleton set, that is $F_i = \{f_{r_i}\}$, the implication is that the fault f_{r_i} is completely resolvable from all other faults. On the other hand, if F_i is not singleton, for example, $F_i = \{f_{i_1}, f_{i_2}, \ldots, f_{i_l}\}$, the implication is that, if any fault in F_i has occurred, based on the observed symptoms, it cannot be decided which of the faults $f_{i_1}, f_{i_2}, \ldots, f_{i_l}$ have occurred. For example, if $\mathcal{F}_{end} = \{\{f_1, f_3\}, \{f_2\}\}$, the implication is that f_2 is resolvable from f_1 and f_3. However, f_1 and f_3 are not resolvable from each other.

If for example, $\mathcal{F}_{end} = \{\{f_1, f_3, f_4\}, \{f_2\}\}$ the unresolvability relation is transitive, that is, f_1 and f_3 are not resolvable from each other, f_3 and f_4 are unresolvable from each other and so are f_1 and f_4. However, this is not always true. For example, it is very well possible that $\mathcal{F}_{end} = \{\{f_1, f_3\}, \{f_2, f_3\}\}$. It is clear that f_1 is not resolvable from f_3, f_2 is not resolvable from f_3, but f_2 is clearly resolvable from f_1. Formally, the following interesting property of the unresolvability relation can be proved.

Claim 1. *The unobservability relation UR as defined in Definition 2 is either a tolerance or equivalence relation.*

Proof. Refer to Appendix A for defintions of tolerance and equivalence relation and proof.

The reason for distinguishing between the two cases arises because a simpler algorithm, Algorithm URE (Appendix B) can be used to determine the set of resolvable faults when *UR* is an equivalence relation. When *UR* fails to be an equivalence relation, and in the light of the earlier claim is a tolerance relation, Algorithms URT and MAXIMAL (Appendix B) are used to determine the set of resolvable faults.

Expected operating profit as value

As mentioned earlier in the introduction, previously published strategies for optimal sensor placement have been guided towards determining the sensor network that is maximally informative from a fault diagnosis sense. The measure of optimality is defined in terms of a metric proposed by the respective author(s). The disadvantage of these strategies is that metrics proposed by different authors could be incommensurate and at times obscure to the plant designer and operator. In the context of fault diagnosis, it is not trivial to define a metric that enables a fair comparison of sensor networks. For example, consider an example where sensor networks consisting of some or all of sensors A, B, C, D are used to detect and isolate three faults f_1, f_2, f_3. Some of these sensor networks and their corresponding fault diagnostic properties are tabulated in Table 1. It is clear that it is not possible to compare the first four sensor networks $S_1, \ldots, S_4$, since the

Table 1. Sensors and Fault Resolvability

Label	Sensor Network	Fault res. Property
S_1	A	f_3 unobservable
S_2	C	$\{f_1, f_3\}$, $\{f_2\}$
S_3	B	$\{f_1\}$, $\{f_2, f_3\}$
S_4	A, D	$\{f_1, f_2\}$, $\{f_3\}$
S_5	C, D	$\{f_1\}$, $\{f_2\}$, $\{f_3\}$

fault observability and resolvability properties are incommensurate. It might be tempting to prefer S_5 consisting of C, D as all three faults are observable and resolvable from each other. However, at this stage, no further information about the process operation or economics have been taken into consideration, and it is difficult to justify such a choice. Hence, we eschew this paradigm and instead attempt to quantify the value of the sensor network in terms of the operating profit generated by the process. The reasons are summarized in the following:

(1) A monetized benefit like the operating profit is obviously transparent and easily accessible to all involved in the planning, design and operation of complex processes.

(2) Different sensor networks and the corresponding diagnostic strategies on which they were based are immediately comparable.

(3) Describing the value in terms of an economic metric immediately solves the problem of determining appropriate utility functions when solving real-world optimization problems like sensor design which involve multiple objectives and decision makers.

(4) In this analysis, the only factors affecting the process and contributing to a potential change in operating profit are the faults affecting the system and the corresponding diagnostic and corrective strategies employed. Thus, different sensor networks that vary in the diagnostic ability will limit the response of the plant operator to the affecting faults and therefore result in different economic returns. Hence, the benefit of fault diagnosis gets captured in this manner of characterization of the sensor network value.

Consider n parametric faults, $f_1, f_2, \ldots, f_n$, each of which is parameterized by variables $x_1, x_2, \ldots, x_n$. The vector of fault parameters, $\mathbf{x} = [x_1, \ldots, x_n]^t$ is a stochastic process and will be assumed to admit probability density function $p(x_1, x_2, \ldots, x_n, t)$. The functional forms of these distribution functions can be arbitrary and general. In addition to the probability distribution functions, the nominal operating data and relevant costs are specified. The plant can operate under a wide range of operating conditions. Correspondingly, the operating profit is a function of the fault variables x_i, other quantities including the states, inputs and outputs and the relevant costs.

Consider the nominal operating point and let us assume the profit function for operating the system, based on the available costs, at this nominal operating point to be $c^* = c^*(\mathbf{s}^*, \mathbf{y}^*, \mathbf{x}^*, t)$, where the superscript indicates the nominal values of the respective quantities. The instantaneous profit function $c(t)$ corresponding to any other point of operation is a function of the quantities at the respective operating points, $c = c(\mathbf{s}, \mathbf{y}, \mathbf{x}, \boldsymbol{t})$. It must be pointed out that $c(t)$ is the instantaneous operating profit before any diagnosis or correction strategy is employed. In the subsequent notation, the dependence of the profit on other quantities will be suppressed for brevity.

The aforementioned instantaneous operating profit is the "free running" instantaneous operating profit that is invariant of plant operation. The actual instantaneous operating profit is a function of the operating strategies, plant design envelopes, ability of the user/designer to ascribe appropriate costs, respective fault probabilities, control strategies and fault correction strategy. In real industrial situations, the loss of profit will depend on the operating procedures that are followed when complete or partial diagnosis is available. The operating profit after diagnosis and correction, $\overline{c(t)}$, is also a stochastic quantity that depends on the underlying profit function $c(t)$, and its distribution function $p(\mathbf{x}, t)$, and the diagnosis and correction strategies. In any realistic operation, over a reasonable period of time, it can be assumed that the net effect of faults is to decrease the profit over time. Formally, we assume that

$$E[\overline{c(t+1)}|\overline{c(t)}, \overline{c(t-1)}, \ldots, \overline{c(1)}] \leq \overline{c(t)} \tag{1}$$

where $E[.]$ is the expectation operator.

Remark 1. Physically, Eq. 1 implies that given the current (and past) operating profit, the future operating profit cannot increase (in an expected sense). This is in line with the expectation that with the passage of time, and the action of faults, performance is bound to degrade. Also, any substantial improvements in process operation, and the subsequent increase in operating profit are usually the result of major maintenance, overhauls or substantial changes in process operation, and so on. The time interval between such drastic changes is usually much larger as compared to the sampling time considered for modeling regular process operation. Also, it has been reported that annual losses in the US petrochemical industry amount to \$16 billion due to abnormal events or faults.[11] Thus, there is a strong intuitive appeal for this modeling approach.

Remark 2. It is possible that under the effect of certain abnormal events f_b, the operating profit is higher than the nominal profit, that is, $c(f_b) > c^*$. This is possible for instance, when the coolant water temperature decreases, or the reactant temperature in an exothermic reaction decreases resulting in a decrease in coolant requirement. If all other relevant quantities, for example, conversion, are unchanged, the overall profit can actually increase. Intuitively, one would conclude that it does not pay to diagnose such events as the overall profit would increase, while there would be a cost to diagnose such events (in terms of sensor costs). Likewise, even if diagnosed, it does not pay to correct these events and bring them to nominal operation, as the corrected profit (c^*) would only be lower than c. Hence, such events will be irrelevant from the perspective of sensor placement for fault diagnosis, and will not affect the final placement. However, the contribution of such events to the overall profit function will be considered to maintain consistency in deriving expectation values of the profit function. To allow for the fact that the profit function may not be a monotonic function of the deviation quantity, events that satisfy the following condition will be considered irrelevant from a fault diagnosis perspective and do not need to be corrected

$$\int_{\substack{x_b \in X_{fb} \\ x_j \in X_{nj}, j \neq b}} c(\mathbf{s}, \mathbf{y}, \mathbf{x}) p(\mathbf{x}, t) dx_1 \cdots dx_j \cdots dx_n$$

$$> c^* \int_{\substack{x_b \in X_{fb} \\ x_j \in X_{nj}, j \neq b}} p(\mathbf{x}, t) dx_1 \cdots dx_j \cdots dx_n. \quad (2)$$

Remark 3. It should be noted that at this stage, no further assumptions on the probability-distribution functions or detailed models have been made. In fact, in the absence of any such detailed models, it may not be possible to rigorously prove Eq. 1. Equation 1, is, thus, inspired by physical intuition and observation rather than the outcome of detailed mathematical models.

Taking further expectations in Eq. 1.

$$E[\overline{c(t+1)}] \leq E[\overline{c(t)}] \leq E[\overline{c(1)}] \quad (3)$$

In addition, in most physical processes, the process is actually stopped for routine maintenance after a time period T, where T is usually much larger than the sampling time used to model the process. The process is eventually restarted with either minor or major modifications subsequently and the analysis can be extended to include renewal of the process for $t \geq T$. But this is not attempted in this contribution, and, hence, we confine ourselves to the interval $1 \leq t \leq T$. Under these conditions, we define the sensor network value V, as the expected value of the cumulative operating profit as follows

$$V \equiv E[P_c(T)]/T \quad (4)$$

where $P_c(T)$ is the cumulative profit from $t = 1$ to $t = T$.

Remark 4. Expanding Eq. 4

$$V = \frac{1}{T}(E[\overline{c(1)}] + E[\overline{c(2)}] + \cdots + E[\overline{c(T)}]). \quad (5)$$

In principle, one can calculate $E[P_c(T)]$ given the initial state and evolution of the distribution functions or state transition probabilities. However, in practice, determining the sensor network value according to Eq. 4 would be very difficult as it would require the original probability-distribution functions for all times or stochastic-differential equations to be specified. Another method to evaluate $E[\overline{c(t+1)}]$ would be to evaluate the conditional expectation $E[\overline{c(t+1)}|\overline{c(t)}, \ldots, \overline{c(1)}]$, and then take expectations over $\overline{c(t)}, \ldots, \overline{c(1)}$. This requires that the transition probabilities be specified. Also, the fault diagnosis strategy has to be taken into consideration. Both of these the earlier approaches can be cumbersome. Another option is to calculate V using Monte-carlo methods to simulate process operation including the fault occurrence, diagnosis and correction. Monte-carlo simulations can, however, be computationally burdensome. Hence, further analysis would be greatly simplified if an analytical approximation to V can be determined instead.

Remark 5. On the other hand, if a data based approach is used to generate distribution functions, it would often be the case that joint operating data is not available, especially, in the case of a greenfield plant. However, it would usually be possible to determine $E[\overline{c(1)}]$, which can be viewed as the initial or starting condition. Thus, it would be greatly beneficial if the sensor network value can be characterized in terms of quantities involving $E[\overline{c(1)}]$ alone. This is the motivation behind the following result.

Theorem 2. *Let $P_c(T)$ be the cumulative profit for $1 \leq t \leq T$. Then,*

$$V \equiv E[P_c(T)]/T \leq E[\overline{c(1)}] \quad (6)$$

Proof. The expected value of the cumulative operating profit, $P_c(T)$ in a single operating cycle of length T is

$$E[P_c(T)] = \sum_{t=1}^{T} E[\overline{c(t)}] \quad (7)$$

From (3)

$$E[\overline{c(1)}] \geq E[\overline{c(t)}] \quad (8)$$

therefore

$$E[P_c(T)]/T \leq \frac{TE[\overline{c(1)}]}{T} = E[\overline{c(1)}] \equiv V_{ub} \quad (9)$$

Thus, an upper bound for the sensor network value V_{ub} is $E[\overline{c(1)}]$

Remark 6. As a result of the preceding discussion, $V_{ub} = E[\overline{c(1)}]$ is a valid upper bound for the sensor network value V, subject to the assumptions made earlier. Hence, V_{ub} can be used as an approximation for V, especially when it is not possible to compute V exactly.

Remark 7. One of the aims of determining the sensor network value is to be able to make clear and transparent decisions on the choice of the sensor network. More specifically, given sensor networks and the level of resolvability afforded by these networks (as in Table 1), one would like to rank these using a quantitative metric. The reasons for quantifying the network value in terms of the expected operating profit have been argued above. Hence, it is natural to question the errors produced in using V_{ub} as an approximation to V. Although it was proved that V_{ub} is indeed a valid upper bound for V, the quality of approximation was not discussed. Intuitively, one would expect the error ranges to be dependent on the actual process and the distribution functions. However, we speculate with some justification that the errors induced by the approximation would not invalidate any conclusions drawn about the networks themselves. For example, given two sensor networks $S1$ and $S2$, and network values V_1 and V_2 such that $V_1 > V_2$, one would prefer $S1$ to $S2$. However, if it is not possible to calculate V_1 and V_2 exactly, but only the respective upper bounds V_{ub1} and V_{ub2} are calculated, we believe that $V_{ub1} > V_{ub2}$. It will be shown in the numerical examples that conclusions drawn on the basis of the upper bounds V_{ub}, can also be backed by process knowledge and observation.

Determination of value of sensor network

In the previous section, an upper bound on the value was derived in a general manner that is applicable to all fault modeling and diagnosis strategies. The evaluation procedure depends on the framework for quantification and use of the information. In this section, the application of the above methods will be described in detail when a qualitative technique, signed directed graph (SDG) is used to model the cause-effect behavior of the faults. Since the objective is to quantify the value from a fault diagnosis perspective, the diagnosis and corrective strategy that is taken is the most important factor in deciding the network value. The actual determination of the value of the sensor network is a two-step process. The first step identifies the set of resolvable faults for a given sensor network and a particular diagnostic strategy. The value of the sensor network is calculated in the second step.

Signed Directed Graph

The signed directed graph (SDG) is a qualitative graph theoretic idea that can be used to model cause-effect behavior and fault propagation in a typical process. An SDG model assumes that all quantities, X_i, $i = 1, \ldots, k$ including measured and unmeasured, are in one of the following four states: Normal (0), High (+1), Low (−1) and Deviated, but direction of deviation unknown (±1).

The SDG consists of a set of vertices (nodes) and signed directed edges (branches). The nodes represent the process variables and includes measurable and unmeasurable quantities. The direction of the branches indicates the direction of causality, that is there is a branch from node i to node j if the former affects the latter. Furthermore, each branch has a sign (+) or (−) associated with it, which indicates whether the cause and effect variables tend to change in the same (+) or opposite directions (−). In addition, a branch with a ± sign indicates a causal effect, but with indeterminate direction of change which could be a result of ambiguity in determining the direction of deviation. This is shown in Figure 1.

Positive and negative deviations in the parameterized quantities will be considered as separate faults. However, some faults are unidirectional, for example, heat exchanger fouling, which results in a decrease in the overall heat-transfer coefficient or equivalently an increase in the fouling factor. Each fault is treated as a root node, that is, a node with no incident branches. At the first level, the effect of these faults on the other quantities is identified and a directed edge is drawn. The effect of these quantities on other process variables is then propagated. The SDG can be condensed into a bipartite matrix **A** between the set of faults, f_i, $i = 1, 2, \ldots, n$ on one side and process variables X_j, $j = 1,2, \ldots, k$ on the other side. The entries a_{ij} of the matrix **A** of size $n \times k$ can be interpreted as

$$\begin{aligned} a_{ij} &= +1 \text{ if fault } i \text{ affects variable } j \text{ in the positive direction} \\ &= -1 \text{ if fault } i \text{ affects variable } j \text{ in the negative direction} \\ &= \pm 1 \text{ if fault } i \text{ affects variable } j \text{ in an indeterminate direction} \\ &= 0 \text{ if fault } i \text{ does not affect variable } j \end{aligned} \quad (10)$$

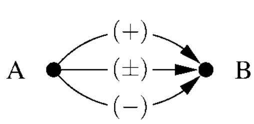

Figure 1. Signed directed graph: different possibilities.

Table 2. Cause-Effect Relationship

Faults	Effects
f_1	{A−, B−, C+}
f_2	{A−, B+, C−}
f_3	{B+, C+, D+}

As an illustration, consider the example in Table 2 consisting of three faults f_1, f_2, f_3 and four potential sensor locations A, B, C, D.

The bipartite matrix **A**, therefore, is

$$\mathbf{A} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \quad (11)$$

Techniques for determining the corresponding entries a_{ij} between a fault f_i, and quantities X_j, $j = 1, \ldots, k$ for a typical chemical process in the presence of multiple paths, multiplicative arcs, feedback loops, control loops, and so on, have been described by Bhushan and Rengaswamy.[2]

Fault Observability with SDG

In the example mentioned in Table 2, fault f_1 can be observed by considering the effect on A or B or C. Whenever f_1 occurs, it causes a negative deviation in A and B and a positive deviation in C. Given this the earlier modeling approach and, consequently, the generated bipartite matrix **A**, it is sufficient to choose a set of sensors, such that, in the equivalent bipartite representation, there is an arc from every fault node to at least one sensor node. In the example presented in Table 2, the choice of A, D is sufficient to ensure fault observability under a single fault assumption. However, this choice does not give the least number of sensors that are necessary. Determining the minimal set of sensors to ensure observability of all faults in a single fault scenario is a combinatorial problem equivalent to the well-known set cover problem.[1] Thus, for the above example, a minimal choice is C or B. Formally, given a set of faults $f_1, f_2, \ldots, f_n$ and the bipartite matrix A between the faults, and the affected quantities X_j, $j = 1, \ldots, k$ and the sensor network $S = \{S_1, \ldots, S_m\}$, define the set A_i as $A_i = \{X_j \mid a_{ij} \neq 0\}$. A_i can be interpreted as the set of quantities that are affected by f_i. The set of observable faults is $F_{obs} = \{f_i | A_i \cap S \neq \emptyset, i = 1, \ldots, n\}$. Thus, in the earlier example, given a network consisting of A only, the observable faults are f_1, f_2. f_3 is not observed since A is not affected by f_3.

Fault resolvability with SDG

The property of resolvability using SDG follows that of Raghuraj et al.[1] and will be useful in deciding the isolability properties of the sensor network. In the example described in Table 2, both f_1 and f_2 produce identical effects on A, that is, a negative deviation. B responds differently to the effects of f_1 and f_2 producing a negative and positive deviation respectively. Thus, a sensor that reacts identically to two faults cannot be used to distinguish them. Formalizing this idea requires the use of set-theoretic symmetric difference B_{ij} of two sets A_i and A_j defined as.

$$B_{ij} = B_{ji} \equiv A_i \cup A_j - A_i \cap A_j \tag{12}$$

When a directed graph (equivalent to an SDG without the signs) is used to model the cause-effect behavior, A_i and A_j would represent the quantities affected by the faults f_i and f_j respectively and the difference is the standard set-theoretic difference. If $S \cap A_i \neq \emptyset$, and $S \cap A_j \neq \emptyset$, the two faults can be observed as discussed in the previous section. If in addition $S \cap B_{ij} \neq \emptyset$, that is, at least one sensor is chosen from B_{ij}, the two faults f_i and f_j are resolvable in a single fault scenario.[2] In the current context of modeling and diagnosis using SDG, the notion of difference is the same as that considered by Bhushan and Rengaswamy.[2] If a signed quantity X is present either in A_i or A_j, but not in both, then the corresponding sensor is included in B_{ij}. For example, if S1+ is present in A_i and S1− is present in A_j (nodes S1+ and S1− would be considered as distinct though the sensed variable is the same), then S1 would be included in B_{ij}. If a sensor node with a positive or negative (+ or −) deviation is present in A_i, and the same sensor with an indeterminate effect (±) is present in A_j, then the corresponding sensor will not be included in B_{ij}. This is due to the fact that the particular sensor would not help distinguish between the faults based on an SDG analysis (because of qualitative ambiguity). Obviously, if a quantity (with the same sign) is present in both sets, it is not included in the symmetric difference.

Determining the set of sensors for single fault resolution is, thus, equivalent to solving an augmented set cover problem with the symmetric difference sets B_{ij} augmented to the original sets A_i.[1] In the example in Table 2, a minimal set that ensures fault observability and resolvability under a single fault scenario is *D, B*.

For the earlier example in Table 2, when the chosen sensor configuration is *C*, it can be verified that the previously described notion of unresolvability is an equivalence relation. The implementation of Algorithms URE and URT is described in Figures 2 and 3 when the chosen sensor network consists of *C* alone.

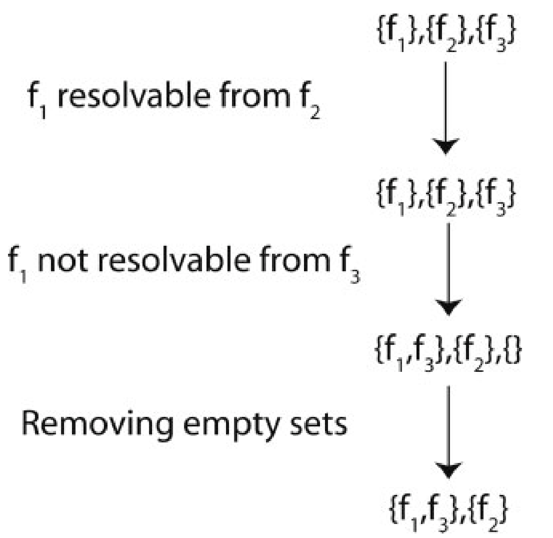

Figure 2. Algorithm URE to determine set of resolvable faults: illustrative example.

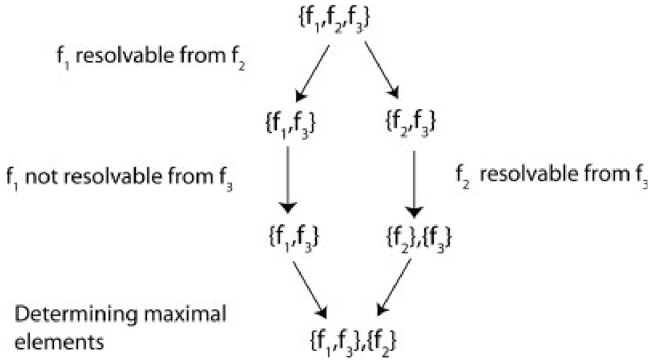

Figure 3. Algorithm URT to determine set of resolvable faults: illustrative example.

In Algorithm URE, at the start, we assume that all faults are resolvable from each other, that is, $\mathcal{F} = \{\{f_1\}, \{f_2\}, \{f_3\}\}$. f_1 is compared with f_2, and it can be seen that f_1 is resolvable from f_2 as f_1 and f_2 affect C differently (negative and positive deviations respectively). Next, f_1 is compared with f_3, and the conclusion is that f_1 is not resolvable from f_3 as C reacts identically to both f_1 and f_3 (positive deviation). Hence, f_1 and f_3 belong to the same equivalence class and $\mathcal{F} = \{\{f_1, f_2\}, \{f_3\}\}, \{\}\}$. Because of the properties of the equivalence relation, f_2 is also not resovlable from f_3, and, hence, $\mathcal{F}_{end} = \{\{f_1, f_2\}, \{f_3\}\}$. Thus, the partition of $\{f_1, f_2, f_3\}$ defined by the earlier equivalence relation given the sensor network consisting of only C is $\{f_1, f_3\}$, $\{f_2\}$, and in the language of Definition 3, f_i *UR* f_i, $i = 1, 2, 3$, f_1 *UR* f_3 and f_3 *UR* f_1. This is interpreted as f_1 not being resolvable from f_3, and, thus, $\{f_1, f_3\}$ forms an equivalence class. f_2, is, however, resolvable from both f_1 and f_3, and so, the sole other equivalence class is $\{f_2\}$.

In the earlier algorithm, the equivalence property of the unresolvability property was explicitly used. In the following implementation, we do not invoke this property. Rather, Alorithms URT and MAXIMAL (Figure 3) are used to determine $\mathcal{F}_{end}$ for the same example. Initially, it is assumed that all the three faults are not resolvable from each other, that is, $\mathcal{F}_{end} = \{\{f_1, f_2, f_3\}\}$. f_1 is resolvable from f_2 as explained previously, and $\mathcal{F}_{end}$ is updated to $\{\{f_1, f_3\}, \{f_2, f_3\}\}$. f_1 is not resolvable from f_3, and no change is made to $\mathcal{F}$. Subsequently, f_2 is compared with f_3, and so $\mathcal{F}$ is updated to $\{\{f_1, f_3\}, \{f_2\}, \{f_3\}\}$. Since $\{f_3\}$ is obviously a subset of $\{f_1, f_3\}$, $\{f_3\}$ is removed during implementation of Algorithm MAXIMAL.

However, because of the inherent ambiguities in the SDG representation, it may happen that for some cases and choices of the sensor networks, unresolvability is not an equivalence relation. Consider the following example of sensor A affected by faults f_1, f_2, f_3 as follows: f_1: $A+$, f_2: $A-$, f_3: $A\pm$. f_1 is not resolvable from f_3, f_3 is not resolvable from f_2. However, f_1 is resolvable from f_2. Hence, the unobservability relation is not transitive. Thus, the relation *UR* can be described as: f_1 *UR* f_3, f_2 *UR* f_3 apart from the trivial reflexive contributions. Here, the property of unresolvability is reflexive and symmetric (f_1 *UR* f_3 implies that f_3 *UR* f_1), but is clearly not transitive. Implementation of algorithms URT and MAXIMAL will result in the expected $\mathcal{F}_{end} = \{\{f_1, f_3\}, \{f_2, f_3\}\}$. Such a relation is symmetric and reflexive, but not transitive, and, hence, is a tolerance relation.

Correction strategy

The actual instantaneous operating profit is a function of the operating strategies, plant design envelopes, ability of the user/designer to ascribe appropriate costs, respective fault probabilities, control strategies and fault correction strategy. In real industrial situations, the loss of profit will depend on the operating procedures that are followed when complete or partial diagnosis is available. Having determined the set of resolvable faults using the earlier SDG based diagnosis ideas and the earlier algorithm, or equivalently any other fault modelling and diagnosis strategy, the next step is to assume a certain fault correction strategy.

We assume that if a resolvable fault occurs, then it will be immediately identified and corrected, and, thus, will result in no decrease in profit (since the process will continue to operate with properly functioning equipment). On the other hand, for a fault that cannot be resolved, the occurrence of such a fault will go uncorrected, and, thus, decrease profit (due to degraded performance of the process). Thus, in defining an overall profit function, we find that if the candidate sensor network is such that the all faults can be resolved, then no loss in profit will occur; the profit will be the nominal value. However, if for a sensor network, some faults cannot be resolved, then the profit function will fall from the nominal value. The set of relevant faults are identified based on Eq. 2. Given a sensor network $S = \{S_1, S_2, \ldots, S_m\}$, the earlier algorithms are implemented resulting in the set $\mathcal{F}_{end} = \{F_1, F_2, \ldots, F_k\}$. All completely resolvable faults are grouped together: $F_r = \{f_{ri}\}$, and all unresolvable faults are grouped together: $F_u = \{f_{ui}\}$. F_r is, thus, the union of all the singleton sets, and F_u is thus the union of all the unresolvable (nonsingleton) sets as determined by the algorithms. We assume that a resolvable fault f_{ri} can be detected, identified and corrected as and when it occurs, and the plant is brought back to the nominal mode of operation by a suitable choice of exogenous actions. The time and cost of diagnosis and correction will be neglected at this stage. Thus, the operating profit is restored to its nominal value of c^*. On the other hand, it is possible that a given sensor network S cannot resolve between some faults. Thus, when some fault $f_{ui} \in F_u$ occurs, one cannot categorically identify the actual fault and so in this case, it will be assumed that no corrective action can be taken. Thus, the system is uncorrected and the lowered operating profit is now a function of the unresolved fault and current, uncorrected operating conditions, that is $c(f_{ui}) < c^*$.

Explicit characterization for value

The value of the sensor network is calculated as the expected value of the operating profit. This is described in detail in this section. For the convenience of the reader, the assumptions made in the preceeding sections are listed below:

- Safety concerns have been addressed and fault diagnosis strategies and computation of the network value are primarily geared toward addressing operational issues.
- This analysis is concerned only with parametric faults.
- Only one fault can occur at a time and so the probability of two or more faults occuring simultaneously is zero.
- It is assumed that the operating profit (in an expected sense) satisfies Eq. 3, that is, the expected operating profit in the future cannot increase.
- If a resolvable fault occurs, then it is corrected immediately at no cost. If the fault is unresolvable, then no action is taken.

Given a sensor network $S = \{S_1, S_2, \ldots, S_m\}$, the earlier algorithms are implemented, resulting in the set $\mathcal{F}_{end} = \{F_1, F_2, \ldots, F_k\}$. The set of relevant faults are identified based on Eq. 2, and are grouped together as resolvable faults: $F_r = \{f_{ri}\}$ and unresolvable faults: $F_u = \{f_{ui}\}$. F_r, is, thus, the union of all the singleton sets, and F_u is, thus, the union of all the unresolvable sets as determined by the previously described algorithms. Rather than computing V exactly, it is more advantageous to compute V_{ub} for reasons discussed previously. Therefore

$$\begin{aligned} V &\equiv E[P_c]/T \\ &\leq E[\overline{c(1)}] \text{ from Theorem 2} \\ &= V_{ub} \end{aligned} \tag{13}$$

This procedure is explained in Figure 4.
The above expression can be evaluated by considering different cases

$$\overline{c(1)} = \begin{cases} c^* & \text{if no fault occurs} \\ c^* & \text{when any resolvable fault } f_{ri} \text{ occurs} \\ c(f_{ui}) & \text{when any unresolvable fault } f_{ui} \text{ occurs} \\ c(f_b) & \text{for any irrelevant event occuring} \end{cases} \tag{14}$$

Recalling that

$$V_{ub} = E[\overline{c(1)}] = \int p(x_1, \ldots, x_n, t = 1)\overline{c(1)}dx_1, \ldots, x_n \tag{15}$$

For brevity, the temporal argument of $p(x_1, \ldots, x_n, t)$ will be suppressed hereafter, and it will be assumed in the following analysis that $t = 1$. The evaluation of V_{ub} can be broken down into four parts or contributions. The first is the trivial no fault case, when the operating profit continues to be c^*.

$$V_1 = c^* \int\limits_{\substack{x_i \in X_{ni} \\ i=1,2,\ldots n}} p(x_1, \ldots, x_n)dx_1 \cdots dx_i \cdots dx_n \tag{16}$$

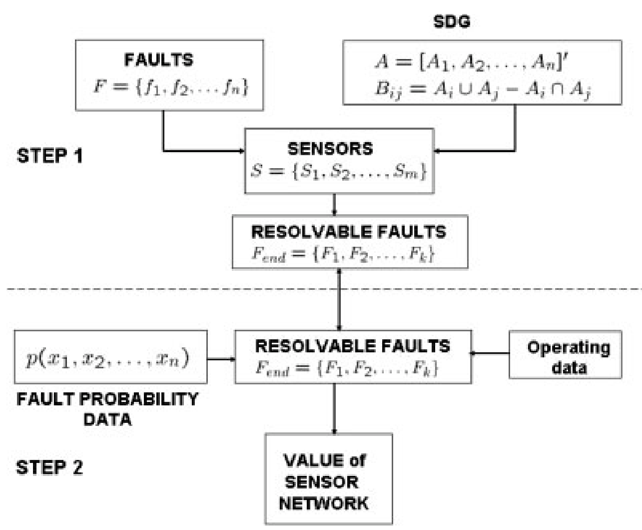

Figure 4. Two-step procedure to determine sensor network value.

The second contribution arises from the resolvable faults F_r, when the operating profit is restored to c^*.

$$V_r = c^* \sum_{i|f_i \in F_r} \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} p(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_j \cdots dx_n \quad (17)$$

The third contribution arises from the unresolvable faults F_u, when the operating profit is not restored to c^*.

$$V_u = \sum_{i|f_i \in F_u} \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} c(f_i) p(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_j \cdots dx_n \quad (18)$$

It is to be noted that these contributions are dependent on the respective costs involved. In addition, the set of resolvable and unresolvable faults, and, therefore, the earlier contributions depend on the sensor network and the fault diagnosis strategy employed. In addition to the above cases, a fourth case must be considered: that of events F_i, which satisfy Eq. 2, and are irrelevant from a fault diagnosis perspective. Thus, one would expect that only events that do not satisfy Eq. 2 are relevant for quantifying the value of the sensor network.

$$V_i = \sum_{i|f_i \in F_i} \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} c(f_i) p(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_j \cdots dx_n \quad (19)$$

Obviously, this contribution depends only on the respective costs considered, and independent of the chosen sensor network or fault diagnosis strategy, and one may choose to exclude this from the overall analysis. However, in order to be consistent with the definition of the value as the probabilistically calculated expected value of the operating profit, it is included in the value calculation. The sensor network value is the expected instantaneous operating profit for which an upper bound can be evaluated as the sum of the four contributions

$$V_{ub} = V_1 + V_r + V_u + V_i \quad (20)$$

Remark 8. In the earlier analysis, the cost and time of correction was neglected. If there is a significant cost of correction for a resolvable fault in the evaluation of in V_r described above, c^* is replaced by $c^* - c_{ri}$.

$$V_r = \sum_{i|f_i \in F_r} (c^* - c_{ri}) \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} p(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_j \cdots dx_n \quad (21)$$

Specification of pdf

The main reason in resorting to an approximation for the determination of the expected instantaneous operating profit, is the difficulty in determining and specifying the appropriate equations for evolution of the pdfs., and/or state transition probabilities. Hence, the determination of $p(\mathbf{x}, t)$ for $t = 1$ should be tractable and preferably possible by using a data based approach, if need be. Such a procedure is described in this section, and will be, subsequently, employed in the numerical example. As before, for brevity, the temporal argument of $p(x_1, \ldots, x_n, t)$ will be suppressed hereafter, and it will be assumed in the following analysis that $t = 1$.

In a single fault scenario, the joint probability distribution function should be consistently chosen so that the probability of two or more faults occurring simultaneously is zero (which implies that the following clause is true: $(\vee_i f_i^i) \wedge_k (\neg f_k^i \vee_{\substack{j \\ j \neq k}} f_j^i)$). Equivalently

$$p(x_1, \ldots, x_n) = 0 \text{ if } (\vee_i f_i^i) \wedge_k (\neg f_k^i \vee_{\substack{j \\ j \neq k}} f_j^i) \quad (22)$$

In addition, the joint pdf should be consistently chosen so that it sums to 1

$$\int_{x_i \in X_{ni} \cup X_{fi} i=1,\ldots,n} p(x_1, \ldots, x_n) dx_1 \cdots dx_i \cdots dx_n = 1 \quad (23)$$

There can be several different families of functions that satisfy these constraints. In this contribution, a consistent joint pdf shall be built starting from individual pdfs. using a renormalization approach. Given individual pdfs. $f_i(x_i)$, such that

$$\int_{x_i \in X_{ni} \cup X_{fi}} f_i(x_i) dx_i = 1, \qquad i = 1, 2, \ldots, n \quad (24)$$

The joint pdf. is now specified as

$$p(x_1, \ldots, x_n) = 0 \text{ if} (\vee_i f_i^i) \wedge_k (\neg f_k^i \vee_{\substack{j \\ j \neq k}} f_j^i)$$

$$= \frac{1}{N} \prod_{i=1}^{n} f_i(x_i) \text{ otherwise} \quad (25)$$

where N is a normalization factor, which can be evaluated by requiring that

$$\sum_{i=1}^{n} \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} p(x_1, \ldots, x_i, \ldots, x_n) dx_1 \cdots dx_i dx_n$$

$$+ \int_{\substack{x_i \in X_{ni} \\ i=1,\ldots,n}} p(x_1, x_2, \ldots, x_n) dx_1 \cdots dx_i \cdots dx_n = 1. \quad (26)$$

N is, thus, evaluated using the following expression

$$N = \sum_{i=1}^{n} \int_{\substack{x_i \in X_{fi} \\ x_j \in X_{nj}, j \neq i}} f_i(x_i) dx_i \prod_{j=1}^{n} f_j(x_j) dx_j + \int_{\substack{x_i \in X_{ni} \\ i=1\ldots,n}} \prod_{i=1}^{n} f_i(x_i) dx_i \quad (27)$$

Remark 9. The renormalization method can be quite readily applied to situations where the probability data is obtained from historical data. For instance, historical data of the quantities x_i can be available and can be converted to the corresponding probability distribution function $f_i(x_i)$ by several methods, the simplest being a simple histogram. However, the data for the quantities x_i could be obtained from several disparate sources over varying periods of time, for example, from different units in the same plant, different plants or different manufacturers. Hence, it would not be possible to obtain joint distribution functions. This method allows one to construct joint-distribution functions from individual historical data or individual distribution functions.

Remark 10. On certain occasions, the joint-distribution data for the various quantities may be available. The same can be used directly without the need for renormalization, if it is determined that the probability of multiple fault occurrence is zero. If the probability of multiple fault occurrence is small, the data may be renormalized to ensure theoretical consistency. If it happens that the probability of multiple fault occurrence is not negligible, the analysis must be suitably modified to explicitly handle multiple fault cases. However, because of the reasons mentioned earlier, this is very unlikely.

Numerical Example: Results and Discussions

The aforementioned described concepts to evaluate the sensor network value from a fault diagnosis perspective will be demonstrated using a simple numerical example.

The system consists of three faults f_1, f_2, f_3 and 4 potential sensor locations, each measuring A, B, C, D. Thus, the total number of nontrivial sensor configurations (not including redundancy) is $2^4 - 1 = 15$. Table 2 describes the cause-effect relationship between the faults and the sensors. The overall profit function is described by

$$c = \mathbf{x}'\mathbf{H}\mathbf{x} + c^* - C_S \tag{28}$$

where $\mathbf{x} = [x_1, x_2, x_3]'$ is a vector parameterizing the three faults f_1, f_2, f_3, respectively, in deviation form from nominal operation, $c^* = 20$ is the nominal operating profit at $\mathbf{x} = \mathbf{0}$ and C_S is the cost of the sensor network, and $\mathbf{H}$ is the following negative definite matrix

$$\mathbf{H} = \begin{bmatrix} -25.0 & 15.0 & -5.0 \\ 15.0 & -45.5 & -7.5 \\ -5.0 & -7.5 & -19.0 \end{bmatrix} \tag{29}$$

Thus, given a sensor network with cost C_S, the optimal profit is $c^* - C_S$ attained at $\mathbf{x} = \mathbf{0}$. The annualized cost of the sensors in two scenarios are tabulated in Table 3.

Starting with individual functions $f_i(x_i)$, the joint distribution probability function is determined by the procedure explained in the previous section. A truncated normal-distribution function is used to describe the individual $f_i(x_i)$ so as to explicitly handle the bounds on the variables x_i. The general form of a normal pdf with prior mean μ and variance σ^2 truncated in $R = [a_1, a_2]$ (where $[a_1, a_2]$ is the support of $f_i(x_i)$) is

$$f(x|R) = \frac{\phi(x, \mu, \sigma)}{\Phi(a_2, \mu, \sigma) - \Phi(a_1, \mu, \sigma)} \tag{30}$$

where $\phi(.)$ is the normal pdf, and $\Phi(.)$ is the cumulative normal distribution function for a random variable with mean μ and standard deviation σ. Thus, each of the $f_i(x_i)$ is parameterized by two quantities, the prior mean μ_i and variance σ_i^2. The prior mean μ_i is chosen to be the nominal value of the respective quantity x_i indicated in Column 6 of Table 4. The prior variance σ_i^2 is calculated so that

Table 3. Sensor Costs

	Sensor Cost			
Scenario	A	B	C	D
Case 1	5	5	5	5
Case 2	1	1	1	1

Table 4. Fault Quantification Parameters

Fault	a_1	a_2	LL	HL	Prior Mean
f_1	−1	1	−0.4	0.4	0
f_2	0	1	0	0.4	0
f_3	0	1	0	0.4	0

$$\int_{LL}^{HL} f_i(x_i)dx_i = 1 - p_i = 0.6, \qquad i = 1, 2, \ldots, n \tag{31}$$

where p_i is indicative of the probability of a single fault occurring. The normal operating range, the support of the distribution functions f_i, and the prior means μ_i used to describe the pdfs, are tabulated in Table 4.

Among the 15 possible sensor configurations, only 13 of them are considered so that all faults are observed. Table 5 contains the list of resolvable faults for each of the 13 permissible configurations determined by applying the algorithms described previously.

A naïve expectation would be that sensor networks with the best resolution properties, viz., S2, S4, S5, S6, S8–13 would be the most preferred sensor network from a fault diagnosis perspective. Indeed, such would be the conclusion that one would draw by applying the standard optimal sensor network designs in literature. However, such designs do not explicitly monetize the benefits of the diagnosis. An upper bound on the value of the sensor networks calculated as described in the previous sections is presented in the last two columns of Table 5 assuming sensor costs as given in Table 3.

It is observed that in Case 1, quantification of the value of the sensor network suggests that the preferred sensor network is S1, viz., that consisting of C alone. This, in spite of the fact that corresponding to this choice, all faults are not resolvable, viz., f_1 is not resolvable from f_3. This can be explained as follows: In order to attain complete fault resolution (as in S2, S4, S5, S6, S8–13), the benefits of fault detection, isolation and subsequent restoration of normalcy are offset by the additional hardware costs. However, the same cannot be argued for S3 or S7, both of which do not provide complete fault resolution. Indeed, the corresponding sets of resolvable faults are $\{f_1\}\{f_2, f_3\}$ and $\{f_1, f_2\}\{f_3\}$ respectively. However, the corresponding sensor network values are 9.05 and 3.05, respectively. This indicates that certain faults are more "valuable" than others, viz., f_2 is more valuable than f_1

Table 5. Sensor Networks Analysis

			Value-Upper Bound	
Tag	Sensor Network	Resolvable Faults	Case 1	Case 2
S1	C	$\{f_1, f_3\}$ $\{f_2\}$	10.94	14.94
S2	C D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	10	18.00
S3	B	$\{f_1\}$ $\{f_2, f_3\}$	9.05	13.05
S4	B D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	10	18.00
S5	B C	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	10	18.00
S6	B C D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	5	17.00
S7	A D	$\{f_1, f_2\}$ $\{f_3\}$	3.50	11.50
S8	A C	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	10	18.00
S9	A C D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	5	17.00
S10	A B	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	10	18.00
S11	A B D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	5	17.00
S12	A B C	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	5	17.00
S13	A B C D	$\{f_1\}$ $\{f_2\}$ $\{f_3\}$	0	16.00

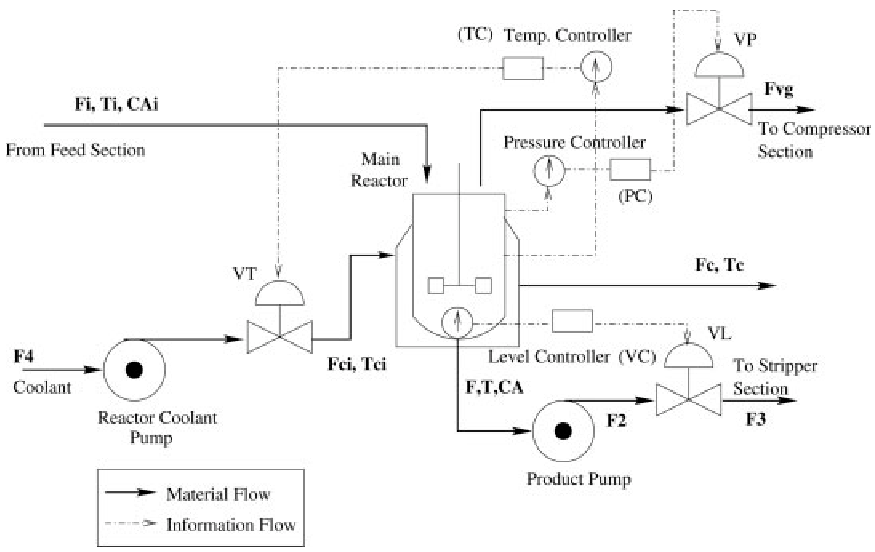

Figure 5. CSTR.

or f_3. Such a conclusion cannot be arrived at by analysis of resolvability properties of sensor networks alone.

That the value analysis is dependent on properties other than operating costs or resolvability properties is evident from the last column of Table 5 which corresponds to Case 2. In this particular scenario, all sensor costs are the same and sensor networks S2, S4, S5, S8 and S10 are the optimal choices. All these sensor networks have the best resolution properties, viz., all faults are resolvable. Even though $S9$, $S11$–$S13$ also provide complete fault resolution, they require more hardware expense than the others and hence are not preferred. In the changed scenario, it can be argued that it is truly beneficial to detect and isolate all the faults without exception.

CSTR Example: Results and Discussions

The earlier described concepts to evaluate the sensor network value from a fault diagnosis perspective, will be demonstrated through numerical simulations of a CSTR, which has been widely used as a test-bed for fault diagnosis. A highly exothermic reaction $A_l \rightarrow B_l + C_g$ is carried out in a jacketed CSTR (Figure 5). The nominal and steady state values of the quantities are presented in Table C1. Three PI controllers are used to control the reactor pressure, reactor liquid level and reactor temperature, respectively, by manipulating the outlet gas flow rate, outlet liquid flow and cooling water flow rate, respectively. The following disturbances: Cai^+, Cai^-, Ti^+, Ti^-, Fi^+, Fi^-, Tci^+, Tci^- (with correspondingly positive or negative deviations as indicated by the signs), which correspond to change in inlet concentration, inlet temperature and inlet flow rate of A, and inlet coolant temperature, respectively, and faults: Cd^-, U^-, which correspond to catalyst deactivation (as measured by catalyst activity) and heat exchanger fouling (reduction in overall heat-transfer coefficient) are modeled as faults in the system. The quantities x_i used to parameterize the faults f_i are physical quantities that can assume physically meaningful values only. The lower (a_1) and upper limits (a_2) for the individual quantities, and the normal operating range [LL, HL] are tabulated in Table C2.

The joint distribution probability function is determined by the procedure explained previously. As in the previous section, a truncated normal-distribution function is used to describe the individual $f_i(x_i)$, so as to explicitly handle the bounds on the variables x_i. The prior mean μ_i is chosen to be the nominal value of the respective quantity x_i indicated in Column 2 of Table C2. The prior variance σ_i^2 is calculated so that

$$\int_{LL}^{HL} f_i(x_i)dx_i = 1 - p_i = 0.6, \qquad i = 1, 2, \ldots, n \tag{32}$$

where p_i is indicative of the probability of a single fault occurring.

The model equations are presented in Appendix C. As described previously, only events that do not satisfy Eq. 2 are relevant from a fault diagnosis perspective. If a resolvable faults occurs, it can be detected and identified and corrected instantaneously. Conversely, when a fault that cannot be identified unambiguously (unresolvable faults as per current terminology) occurs, no corrective action is taken. The sets of quantities (A sets) affected by each event are identified based on the SDG model,[2] and presented in Table 6. Control-loop sensors measuring the pressure, reactant volume and temperature are assumed to be present by default. Hence, the contribution of these sensors can be excluded from the analysis. Sensors can be placed at some or all of the follow-

Table 6. A Sets based on SDG

Sr. No.	Tag	Event	Set A
1	F1	Cai^+	Ca^+, Fc^+, Tc^-, Fvg^+, Cai^+
2	F2	Cai^-	Ca^-, Fc^-, Tc^+, Fvg^-, Cai^-
3	F3	Ti^+	Fc^+, Tc^-, Ti^+
4	F4	Ti^-	Fc^-, Tc^+, Ti^-
5	F5	Fi^+	Ca^+, F^+, $Fc^\pm$, $Tc^\pm$, Fvg^+, Fi^+
6	F6	Fi^-	Ca^-, F^-, $Fc^\pm$, $Tc^\pm$, Fvg^-, Fi^-
7	F7	Tci^+	Fc^+, Tci^+
8	F8	Tci^-	Fc^-, Tci^-
9	F9	Cd^-	Ca^+, Fc^-, Tc^+, Fvg^-
10	F10	U^-	Fc^{-+}, Tc^-

Table 7. Cost Data

Case	Revenue R, \$/hr	Cost, \$/ft^3 Coolant	Vapor	Sensor Costs, \$/yr Temp	Conc	Press
1	$R = \begin{cases} 0.375P_B & \text{if } P_b \leq 10.61 \text{ mol/hr} \\ -0.2433P_B + 6.56 & \text{if } P_b > 10.61 \text{ mol/hr} \end{cases}$	0.015	0.00225	1000	1000	1000
2	$R = \begin{cases} 0.375P_B & \text{if } P_b \leq 10.61 \text{ mol/hr} \\ -0.2433P_B + 6.56 & \text{if } P_b > 10.61 \text{ mol/hr} \end{cases}$	0.015	0.00225	1000	2333	1000
3	$R = \begin{cases} 0.1P_B & \text{if } P_b \leq 10.61 \text{ mol/hr} \\ -0.147P_B + 2.63 & \text{if } P_b > 10.61 \text{ mol/hr} \end{cases}$	0.015	0.00225	100	100	100

ing 9 locations [Ca, F, Fc, Tc, Fvg, Cai, Fi, Tci, Ti], thus, resulting in 511 distinct sensor networks.

Cost data for various scenarios are presented in Table 7, where sensor costs are annualized, and P_B is production rate of product B in lb.mol/hr. Cases 1 and 2 refer to two scenarios which are identical except for sensor costs. Case 3 refers to a scenario when the product B is less valuable as compared to Cases 1 and 2, and corresponding sensor costs are also lower. Different scenarios are considered to illustrate that the preferred sensor network is strongly dependent on the operating scenario.

Computing the operating profits for Cases 1 and 2 reveals that events {F4, F8}, corresponding to a decrease in inlet temperature of feed and coolant, respectively, lead to an increase in the operating profit, and, hence, should not play a role in deciding placement of sensors from a fault diagnosis perspective. This agrees with the expectation that a decrease in either of these temperatures will result in a lowered coolant demand, and if it is possible to maintain conversion at desired levels, the overall profit should increase. The algorithms described previously are used to determine the set of resolvable and unresolvable faults which are tabulated in Table 8. Unobservable faults are treated in the same manner as unresolvable faults, with some abuse of notation and no further change in analysis. Sensor network S1 comprising of [Ca, F, Fc, Tc, Ti] (apart from other equally effective networks not shown here), gives the best resolution properties as far as fault diagnosis is concerned, viz., all faults can be resolved from each other. Sensor network S2 comprising of [Cai], that is, inlet concentration sensor can obviously observe and resolve only the deviations in inlet concentration. Likewise, sensor networks S3 [Fi: inlet flow rate of reactant] and S4 [F: outlet flow rate from reactor] can observe and resolve only deviations in inlet flow rates. A naïve expectation would be that since S1 gives best resolution properties, it would be the most preferred sensor network from a fault diagnosis perspective. Indeed, such would be the conclusion that one would draw by applying the standard optimal sensor network designs in literature. However, such designs do not explicitly monetise the benefits of the diagnosis. The sensor network values are calculated using the earlier information, and the procedure described and tabulated in Table 9 assuming 8,760 h of annual operation. Value is the profit calculated as revenue (R) less transport costs of coolant and vapor (Table 7). It turns out S2 [Cai] is the most preferred choice, even though its resolution properties (resolving only inlet concentration deviations) are inferior to that of S1. Sensor networks S3 and S4 with different diagnosis properties are also preferred to S1. The implication is that the benefits do not justify the extra expense incurred.

Sensor networks S2, S3 and S4 with slightly differing net values have different resolution properties. Sensor network S2 can detect and identify changes in inlet concentration while S3 and S4 can detect and identify changes in inlet flow rate. The previous figures were arrived at under the assumption that the annualized cost of all sensors is the same. However, if it is assumed that concentration sensors are more expensive at \$2,333/yr (a more realistic scenario) as against \$1,000/yr for other sensors (Case 2), it makes economic sense to choose sensor networks S7 or S8. This implies that in the changed scenario, being able to detect and identify inlet flow rate deviations is more attractive for process economics. The network values are tabulated in Table 10.

A common thread in the earlier analysis is that in the computation of the profit function, the contribution from sales of product B tends to dominate over coolant and vapor transportation costs. For instance, at the nominal production level, the contribution from sales is \$34,877/yr, while those due to coolant and vapour costs are \$7,442/yr and \$118/yr, respectively, bringing the net profit to \$27,317/yr. When the inlet concentration changes to 0.45 and 0.55 lb-mol/ft^3, the revenue due to sales changes to \$31,389/yr and \$32,614/yr, respectively. The corresponding figures for coolant costs are \$6,187/yr and \$8,757/yr, respectively. The vapor transportation costs change only marginally to \$106/yr and \$130/yr only. Similar, if not identical figures are observed in other events involving inlet flow rate changes. Thus, the conclusion is that faults or deviations that result in substantial change in

Table 8. Fault classification- Cases 1 and 2

Tag	Sensor Network	Sets of Resolvable Faults
S1	[Ca, F, Fc, Tc, Ti]	{F1}{F2}{F3}{F4}{F5}{F6} {F7}{F8}{F9}{F10}
S2	[Cai]	{F1}{F2}{F3,F4,F5,F6,F7,F8,F9,F10}
S3	[Fi]	{F1,F2,F3,F4,F7,F8,F9,F10}{F5}{F6}
S4	[F]	{F1,F2,F3,F4,F7,F8,F9,F10}{F5}{F6}

Table 9. Network Value- Case 1

Tag	Sensor Network	Value, \$/yr	Sensor Costs, \$/yr	Net Value = Value-Sensor Costs, \$/yr,
S1	[Ca, F, Fc, Tc, Ti]	27453	5000	22453
S2	[Cai]	26706	1000	25706
S3	[Fi]	26417	1000	25417
S4	[F]	26417	1000	25417

Table 10. Network Value- Case 2

Tag	Sensor Network	Value, $/yr	Sensor Costs, $/yr	Net Value = Value-Sensor Costs, $/yr,
S5	[Ca, F, Fc, Tc, Ti]	27453	6333	21120
S6	[Cai]	26706	2333	24373
S7	[Fi]	26417	1000	25417
S8	[F]	26417	1000	25417

Table 12. Network value- Case 3

Tag	Sensor Network	Value, $/yr	Sensor Costs, $/yr	Net Value = Value-Sensor Costs, $/yr.
S9	[Cai, Fi, Tci]	1867.4	300	1567.4
S10	[Fvg, Fi Tci]	1867.4	300	1567.4
S11	[F, Cai, Tci]	1867.4	300	1567.4
S12	[F, Fvg, Tci]	1867.4	300	1567.4

revenue (due to sales of product B) attract a huge premium ("valuable faults"), while other faults are relatively discounted. This would explain why, in the preferred sensor networks, inlet concentration and flow changes which affect production of B are detected and identified while other faults or deviations are not detected.

The earlier analysis was repeated at substantially lower revenues from sale of product, a scenario played out in Case 3. At nominal production levels, the revenue from product sales drops to $9,301/yr, while other costs obviously remain unchanged, thus, bringing down nominal profit to $1,740/yr. In such a situation, the contributions due to sales revenue and coolant costs are comparable, and this would make it economically attractive to be able to detect and identify other kinds of faults or deviations. Fault classification and sensor network values for Case 3 are tabulated in Tables 11 and 12. It is clear that substantially different sensor networks, with different resolution properties are preferred. In addition, events F2, F4, F8, F9 actually lead to an increase in operating profit and will play no role in deciding sensor placement.

The approximation discussed in the previous sections was used to determine an upper bound for the sensor network value. The hierarchy of sensor networks (in a decreasing order of preference) could be justified from an analysis of process economics also. Hence, it is unlikely that even with an accurate computation of the network value using Montecarlo simulations, the conclusions would be different. This justifies the use of the approximations.

Thus, it is clear that being able to attach an economic value to the sensor network from a fault diagnosis perspective clearly brings out the trade-offs between instrument costs and the utility value of diagnosis. It should be noted that the utility of the sensor network is strongly dependent on the operating conditions, as should be expected. This, however, is not possible in traditional sensor network designs which are impervious to economic factors (other than instrument cost) and operating or control strategies. This analysis clearly puts an economic value to diagnosing a fault, and, hence, some faults are more "valuable" (from a diagnosis perspective) than others. Hence, designs based on economic value would tend to pick networks that diagnose these "valuable" faults. Also, the operating, control, diagnosis and correction strategies will greatly affect the computation of the value, and, hence, sensor network design. For instance, in this work, it was assumed that one is able to correct a resolvable fault with negligible extra cost and within a negligibly short period of time, whereas in reality, this might note be possible. Alternatively, being able to detect and identify a fault might be useful in altering operating or control strategies. Hence, the analysis would then need to be modified. However, even in such a situation, the philosophy behind computing the sensor network value remains unchanged.

Table 11. Fault Classification- Case 3

Tag	Sensor Network	Resolvable Faults
S9	[Cai, Fi, Tci]	{F1}{F2}{F3, F4, F9, F10}{F5}{F6}{F7}{F8}
S10	[Fvg, Fi Tci]	{F1}{F2, F9}{F3, F4, F10}{F5}{F6}{F7}{F8}
S11	[F, Cai, Tci]	{F1}{F2}{F3, F4, F9, F10}{F5}{F6}{F7}{F8}
S12	[F, Fvg, Tci]	{F1}{F2, F9}{F3, F4, F10}{F5}{F6}{F7}{F8}

This work did not consider sensor faults or possible hardware redundancies. Gross errors in non control loop sensors can be included in this procedure. This methodology can also be extended to address simultaneous multiple faults.

Conclusions

A useful two-step procedure to quantify the value of a sensor network from a fault diagnosis perspective was described. In the first step, for the given sensor network, the set of faults that can be unambiguously resolved was determined. In the second step, various quantitative data was used to determine the value as a probabilistically calculated expected profit function subject to certain operation, control, diagnosis and correction strategies. Both ideas are quite general and can be applied to various other situations. The ideas were illustrated with a simple example and simulations on a CSTR. The trade-offs between instrument costs and utility of diagnosis and cost-benefit analysis were clearly evident when designing sensor networks for fault diagnosis in the numerical example. It must be noted that this procedure is very general and not tied to any specific fault modelling or diagnostic approach.

Structural faults, in general, lead to shutdown, and the value of a network that can diagnose structural faults can be calculated by the loss incurred during downtime. The value of detecting faults in control loop sensors can be calculated by quantifying unnecessary control effort that leads to loss in utility, and the loss incurred from the products being off-spec. Biases in noncontrol variables are related to loss incurred through loss of precision.[9] Gross errors in sensor faults can lead to loss of resolution property of the corresponding sensor network. The value of detecting this fault can be quantified through the loss that one will incur due to the loss in resolution property.

This earlier analysis used an upper bound for the value as an initial approximation. More sophisticated tools can be used to arrive at more accurate expressions. The analysis can be extended to include renewals after the end of one production cycle. Future work can include multiple fault occurrences, although rarer in practice.

It was pointed out in the introduction that design of an integrated sensor network can be formulated as a multiobjective optimization problem. Solution of the same involves optimizing a suitably defined utility function. The value of the sensor network as quantified by the method described in this article can be used in conjunction with the value functions from other perspectives, viz., material accounting, process control, and so on, to build the utility function.

Notation

$\boldsymbol{A}$	=	bipartite matrix between faults and process variables
$\boldsymbol{r}$	=	residual vector
$\mathcal{F}$	=	collection of subsets of F
μ	=	prior mean
$\overline{c(t)}$	=	instantaneous operating profit at time t after diagnosis and correction
$\Phi(\cdot)$	=	cumulative normal-distribution function
$\phi(\cdot)$	=	normal-density function
$\prec$	=	parital ordering
σ^2	=	prior variance
$\vee$	=	or operator
$\wedge$	=	and operator
A, B, C, D	=	sensors
A_i	=	set of process variables affected by fault f_i
$c(t)$	=	instantaneous operating profit at time t before diagnosis and correction
c^*	=	operating profit at nominal operation
$F \times F$	=	cartesian product of F
F	=	set of given faults, $F = \{f_1, \ldots, f_n\}$
f_i	=	ith fault
$f_i(x_i)$	=	individual pdf
F_r	=	union of all singleton sets in $\mathcal{F}_{\text{end}}$
F_u	=	union of all nonsingleton sets in $\mathcal{F}_{\text{end}}$
N	=	normalizing factor
$p(x_1, x_2, \ldots, x_n, t)$	=	probability-density function for $\mathbf{x} = [x_1, \ldots, x_n]^t$
$P_c(T)$	=	cumulative profit from $t = 1$ to $t = T$
$S = \{S_1, S_2, \ldots, S_m\}$	=	Sensor network consisting of sensors $S_1, \ldots, S_m$
T	=	time horizon
t	=	time index
V	=	sensor network value
V_1	=	contribution to V_{ub} from the no-fault case
V_i	=	contribution to V_i from irrelevant events
V_r	=	contribution to V_r from resolvable faults
V_{ub}	=	upper bound for sensor network value
V_u	=	contribution to V_u from unresolvable faults
x_i	=	quantity used to parametrize fault f_i
X_j	=	jth process variable
X_{fi}	=	abnormal operating region for fault f_i
X_{ni}	=	normal operating region for fault f_i
$\boldsymbol{d}$	=	unmeasured disturbances
$\boldsymbol{s}$	=	system states
$\boldsymbol{y}$	=	measured quantities

Literature Cited

1. Raghuraj R, Bhushan M, Rengaswamy R. Location of sensors in complex chemical plants based on fault diagnostic observability criteria. *AIChEJ*. 1999;45:310.
2. Bhushan M, Rengaswamy R. Design of sensor network based on the SDG of the process for efficient fault diagnosis. *Ind Eng Chem Res*. 2000;39:999.
3. Bhushan M, Rengaswamy R. Comprehensive design of a sensor network for chemical plants based on various diagnosability and reliability criteria: I. Framework. *Ind Eng Chem Res*. 2002;41:1826–1839.
4. Bhushan M, Rengaswamy R. Comprehensive design of a sensor network for chemical plants based on various diagnosability and reliability criteria: II. Applications. *Ind Eng Chem Res*. 2002;41:1840–1860.
5. Bhushan M, Narasimhan S, Rengaswamy R. Sensor Network Reallocation and Upgrade for Efficient Fault Diagnosis in *Fourth International Conference on Foundations of Computer-Aided Process Operations*, Coral Springs, Florida; 2003:443–446.
6. Bagajewicz M. *Process Plant Instrumentation. Design and Upgrade*. Technomic Publishing Company, Now CRC Press; 2000.
7. Kubrusly C. S, Malebranche H. Sensors and controllers location in distributed systems-a survey. *Automatica*. 1985;21:117–128.
8. Chmielevski DJ, Peng JK. Covariance based hardware selection—Part I: Globally optimal actuator selection. *IEEE Trans Contr Syst Technol*. 2006;14:355–61.
9. Bagajewicz M, Markowski M, Budek A. Economic value of precision in the monitoring of linear systems. *AIChE J*. 2005;51:1304–1309.
10. Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri S. Review of process fault detection diagnosis - Part I: Quantitative model-based methods. *Comp Chem Eng*. 2003;27:239–311.
11. Nimmo I. Adequately address abnormal operations. *Chem Eng Progress*. 1995:91(9):36–45.
12. Mylaraswamy D. DKit: *A Blackboard-based Distributed, Multi-Expert Environment for Abnormal Situation Management*. Purdue University; 1996. PhD thesis.

Appendix A. Mathematical Preliminaries

Definition 3. Equivalence Relation: A relation R on a set X is a subset of $X \times X$, that is a collection of ordered pairs of elements of X. Denoting $x\ R\ y$ to mean $(x, y) \in R$, we say x is related to y. A relation R is an equivalence relation if it satisfies the following properties:
Property 1. Reflexive: $a\ R\ a$, $\forall a \in X$.
Property 2. Symmetric: $a\ R\ b \Rightarrow b\ R\ a$, $\forall\ a, b \in X$.
Property 3. Transitive: $a\ R\ b$ and $b\ R\ c \Rightarrow a\ R\ c$, $\forall a, b, c \in X$.

An equivalence class is a subset of the form $X_a = \{x \in X, x\ R\ a, a \in X\}$. Any two equivalence classes are either equal or distinct. The collection of equivalence classes forms a partition of X.

Definition 4. Tolerance Relation: A tolerance relation R on a set X is a slightly weaker form of a relation as compared to the equivalence relation defined earlier. A relation R is a tolerance relation if it satisfies only Properties 1 and 2 listed earlier, that is, it is reflexive and symmetric, but not transitive.

Proof of Claim 1.

By definition, UR is reflexive. Also, it follows from the definition that $f_i\ UR\ f_j \Rightarrow f_j\ UR\ f_i$, and, hence, UR is symmetric. If UR is transitive, it satisfies all three properties listed in Definition 3 and is an equivalence relation. However, as it was demonstrated with examples, it cannot be guaranteed that UR is transitive. Hence, UR is not necessarily an equivalence relation. If UR is not transitive, then, it is a tolerance relation.

Appendix: B. Algorithms for Determining Set of Resolvable Faults

Algorithm 1. *ALGORITHM URE*

$S = \{S_1, S_2, \ldots, S_m\}$
$\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$, $F_i = \{f_i\}$, $i = 1, 2, \ldots, n$
for $i = 1$: n ***do***
  for $j = i + 1$: n ***do***
    if $(f_i\ UR\ f_j) \wedge (F_i \neq \{\}) \wedge (F_j \neq \{\})$ ***then***
      $F_i = F_i \cup F_j$
      $F_j = \{\}$
    end if
  end for
end for
$\mathcal{F}_{end} = \{F_i \mid F_i \neq \{\}\}$

In the earlier notation, $\mathcal{F}$ and $\mathcal{F}_{end}$ are families of sets. For example., $\mathcal{F}_{end} = \{\{f_1, f_5\}, \{f_3\}, \{f_4, f_2\}\}$. F_i can be thought of as those faults that are unresolvable from f_i. At the start of the algorithm. $F_i = \{f_i\}$, that is, we assume that there exists no fault that is unresolvable from f_i. The F_i are updated as the algorithm is executed. It is determined whether a particular pair of faults, f_i, f_j is resolvable, that is, whether f_i *UR* f_j. If the fault pair is unresolvable and the unresolvability relation is an equivalence relation, the two sets F_i and F_j are combined and either of the two is removed. In the earlier notation, F_i is the set of all faults that are unresolvable from f_i and F_j is the set of all faults that are unresolvable from f_j. Since *UR* is an equivalence relation, it is transitive, and, hence, all faults un- resolvable with f_i, are also unresolvable with f_j if f_i *UR* f_j. Hence, F_i and F_j may be combined to denote the set of all faults unresolvable from both f_i and f_j. Eventually, $\mathcal{F}_{end}$ is a partition of F, that is consistent with the unobservability relation.

Algorithm 2. *ALGORITHM URT*

$S = \{S_1, S_2, \ldots, S_m\}$
$\mathcal{F} = \{F\}, F = \{f_1, f_2, \ldots, f_n\}$
for $i = 1: n$ ***do***
  for $j = i + 1: n$ ***do***
    $\mathcal{F} = \{F_1, \ldots, F_k\}, k = card(\mathcal{F})$
    $\mathcal{F}_{temp} = \{\}$
    if $\neg\ (f_i\ UR\ f_j)$ ***then***
      for $l = 1: k$ ***do***
        if $\{f_i, f_j\} \subset F_l$ ***then***
          $\mathcal{F}_{temp} = \mathcal{F}_{temp} \cup \{F_l - f_i\} \cup \{F_l - f_j\}$
          $F_l = \{\}$
        end if
      end for
    end if
    $\mathcal{F} = \mathcal{F} \cup \mathcal{F}_{temp}$
  end for
end for

Unlike in the previous algorithm, Algorithm URT is a bottom-up approach, that is, at the start, we assume that all faults are unresolvable from each other. Since *UR* is not an equivalence relation, but only a tolerance relation, the property of transitivity does not hold. Hence, every pair of faults f_i and f_j has to be compared and if they are determined to be resolvable from each other, the sets are accordingly updated to ensure that if $\{f_i, f_j\} \subset F_l$, F_l is split into two sets, one that contains f_i but not f_j, and the other that contains f_j, but not f_i. Termination of the above algorithm yields $\mathcal{F}_{end} = \{F_1, F_2, \ldots, F_k\}$, whence, it may occur that $F_i \subset F_j$ for some i and j, $j \neq i$. Hence, Algorithm Maximal is carried out to collapse sets which are subsets into their parent sets. Equivalently, only the maximal elements of $\mathcal{F}_{end}$ are of interest, where the idea of maximal element follows from the partial ordering induced by subset inclusion, that is, $F_i \prec F_j$ if $F_i \subset F_j$. For example, if $\mathcal{F}_{end} = \{\{f_1, f_5\}, \{f_3\}, \{f_4, f_2, f_6\}, \{f_2, f_6\}\}$, the maximal elements are $\{f_4, f_2, f_6\}$, $\{f_3\}$ and $\{f_1, f_5\}$.

Algorithms URT and MAXIMAL can also be used when *UR* is an equivalence relation.

Table C1. Nominal Values and Operating Data for CSTR

Notation	Variable	Nominal Value
V	Volume of liquid in reactor	48 ft^3
C_A	Reactant concentration in reactor	0.2345 lb.mol A/ft^3
T	Reactor temperature	600° R
n	Number of moles in gas phase of reactor	28.3656 lb.mol C
V_g	Volume of gas phase (constant)	16 ft^3
F_i	Inlet feed flow rate	40 ft^3/h
C_{Ai}	Inlet reactant concentration	0.50 lb.mol A/ft^3
T_c	Jacket temperature	590.51° R
F_c	Coolant flow rate	56.626 ft^3/h
T_i	Inlet feed temp.	530° R
V_j	Volume of jacket	3.85 ft^3
k_o	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$^{2\circ}$ R
A	Heat transfer area	150 ft^2
T_{ci}	Inlet coolant temp.	530° R
ΔH	Heat of reaction	−30,000 Btu/lb.mol
C_p	Heat capacity (process side)	0.75 Btu/lbm° R
C_{pj}	Heat capacity (coolant side)	1.0 Btu/lbm° R
ρ	Density of process mixture	50 lbm/ft^3
ρ_j	Density of coolant	62.3 lbm/ft^3
K_v, T_v	PI Controller parameters for volume	1
K_t, T_t	PI Controller parameters for temperature	4.3
K_p, T_p	PI Controller parameters for pressure	0.5

Algorithm 3. *ALGORITHM MAXIMAL*

$\mathcal{F}_{end} = \{F_1, F_2, \ldots, F_k\}$
$V_i = 1, i = 1,2, \ldots, k$
for $i = 1: n$ ***do***
  for $j = i + 1: k$ ***do***
    if $F_j = F_i$ ***then***
      $V_j = 0$
    else if $F_j \subset F_i$ ***then***
      $V_j = 0$
    else if $F_i \subset F_j$ ***then***
      $V_i = 0$
    end if
  end for
end for
$\mathcal{F}_{end} = \{F_i \mid V_i = 1\}$

Appendix C. CSTR Model Equations

The CSTR system is shown in Figure 5. The process involves a highly exothermic catalytic reaction $A_{(l)} \rightarrow B_{(l)} + C_{(g)}$. The temperature in the reactor is controlled by manipulating the flow rate of the coolant flowing through the jacket. The level in the reactor is controlled by manipulating the outlet flow rate from the reactor. The pressure in the reactor is controlled by changing the vent gas flow rate. PI controllers ($K_i + \frac{T_i}{s}$) are used to control the temperature ($i = t$), volume ($i = v$), and pressure ($i = p$) of the reactor. Both the reactor and the jacket are modeled with perfectly mixed-tank dynamics. The reactor holdup V, at any time is given by

$$\frac{dV}{dt} = F_i - F \tag{C1}$$

Table C2. Fault Variable Limits

Fault	Nominal Value	Units	Physical Limits		Normal Range X_n		Abnormal Region X_f	
			Lower	Upper	Lower Limit (LL)	Higher Limit (HL)	Lower Limit	Upper Limit
Cai^+	0.5	lb.mol/ft^3	0.5	0.65	0.5	0.55	0.55	0.65
Cai^-	0.5	lb.mol/ft^3	0	0.5	0.45	0.5	0	0.45
Ti^+	530	° R	530	560	530	535	535	560
Ti^-	530	° R	500	530	525	530	500	525
Fi^+	40	ft^3/hr	40	100	40	45	45	100
Fi^-	40	ft^3/hr	0	40	35	40	0	35
Tci^+	530	°R	530	545	530	535	535	545
Tci^-	530	°R	500	530	525	530	500	525
Cd^-	1		0.35	1	0.95	1	0.35	0.95
U^-	1		0.45	1	0.95	1	0.45	0.95

where F_i and F are the inlet and outlet liquid flow rates, respectively. The reactant concentration C_A is given by

$$\frac{dC_A}{dt} = \frac{F_i}{V}(C_{Ai} - C_A) - r_A \tag{C2}$$

where C_{Ai} is the inlet reactant concentration, and r_A is the specific reaction rate which is first-order in C_A, and depends on the catalyst activity c_d:

$$r_A = c_d C_A k_o e^{-E/RT} \tag{C3}$$

Assuming constant heat capacity C_p, density ρ, and heat of reaction ΔH, an overall heat balance on the reactor gives the reactor temperature T as

$$\frac{dT}{dt} = \frac{F_i}{V}(T_i - T) + \frac{r_A(-\Delta H)}{\rho C_p} - \frac{UA(T - T_c)}{V\rho C_p} \tag{C4}$$

where T_i is the inlet temperature of the reactant, V is the liquid hold-up, and U and A are the overall heat-transfer coefficient and heat exchanger area, respectively. Given constant coolant heat capacity C_{pj} and density ρ_{pj}, the coolant temperature T_c, is given by a similar heat balance on the jacket:

$$\frac{dT_c}{dt} = \frac{F_c}{V_j}(T_{ci} - T_c) + \frac{UA(T - T_c)}{V_j \rho_j C_{pj}} \tag{C5}$$

where V_j is the jacket volume. The pressure in the reactor P depends on the number of moles of vapor n. This in turn depends on the rate of reaction, and vent (molar) flow rate F_{vg}. The vapour space V_g, is assumed to be constant and the vapor is assumed to behave ideally,

$$\frac{dn}{dt} = r_A V - F_{vg} \tag{C6}$$

$$PV_g = nRT \tag{C7}$$

Assuming no accumulation in the pumps, valves and jacket, the following relations are obtained

$$F_3 = F_2 = F \tag{C8}$$

$$F_4 = F_c = F_{ci} \tag{C9}$$

Relevant nominal values and operating data for the CSTR case study that are used in determining network value presented in Table C1.[12]

Manuscript received Feb. 6, 2006, revision received Aug. 9, 2006, and final revision received Dec. 12, 2006.