Optimal safety system design using fault tree analysis

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This article was published in the journal, Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering [© IMechE] and is also available at: http://archive.pepublishing.com/home/main.mp

Metadata Record: https://dspace.lboro.ac.uk/2134/3822

Publisher: © IMechE / Professional Engineering Publishing

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

**Creative Commons License**

** Attribution-NonCommercial-NoDerivs 2.5**

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

** Attribution.** You must attribute the work in the manner specified by the author or licensor.

**Noncommercial.** You may not use this work for commercial purposes.

**No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the **Legal Code** (the full license).

Disclaimer: [icon]

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
Optimal safety system design using fault tree analysis

J D Andrews, BSc, PhD, CEng, MIMechE, MSaRS, FRSS, AFIMA
Department of Mathematical Sciences, Loughborough University of Technology, Leicestershire

This paper describes a design optimization scheme for systems that require a high likelihood of functioning on demand. The final design specification is achieved by solving a sequence of optimization problems. Each of these problems is defined by assuming some form of the objective function and specifying a sub-region of the design space over which this function will be representative of the system unavailability. An example of a high-pressure protection system has been used to demonstrate the practicality of the technique. Design parameters for this system include redundancy levels, the number of elements required for a voting system to function, component selection options and maintenance inspection intervals. Both implicit and explicit constraint forms have been used in the example. The implicit constraints require a full system analysis to determine whether the current design is feasible or not. All system assessments have been carried out using fault tree analysis.

1 INTRODUCTION

For potentially hazardous industrial systems or processes it is now common practice to carry out a quantitative assessment of the safety features. By predicting the probability that the safety systems will fail to perform the intended task of either preventing or reducing the consequences of hazardous situations the acceptability of the design can be judged. This type of assessment is best carried out at the design stage when the system is more easily modified. Once the system is installed and operational, design modifications can be prohibitively expensive. Over the past two decades several methods have been developed for predicting the likelihood that systems will fail. Such methods include fault tree analysis (1), simulation (2), Markov analysis (3) and cause–consequence analysis (4). Of these methods fault tree analysis is the most frequently used in evaluating the performance of safety protection systems. As with the majority of engineering disciplines, the approach to the design of a safety system follows the traditional design stages of preliminary design, analysis, appraisal and redesign. The initially proposed design specification is analysed to predict its likelihood of failure to perform according to the design intention. Its predicted performance is then compared to that which is considered acceptable. If the system performance is not acceptable then deficiencies in the design are removed by redesign, and the analysis and appraisal stages repeated. When it is considered that the system performance is adequate then the design process stops and the design is adopted.

The criterion used to determine the adequacy of the system is usually a comparison with some pre-set target value for its probability to function on demand. With all the choices of selection of component type, levels of redundancy, levels of diversity, system structure and maintenance options there is little chance that the final design achieved by this approach would utilize the available resources to the full. For a safety system whose consequences of failure could be catastrophic, it is important to look for the best performance possible within the resources available, not just adequate performance. The target figure should be used as a minimum acceptance level and there should be an aim to produce the optimal performance attainable within the constraints imposed on resources.

The analysis methods are now well developed and attention can be focused on incorporating these methods into a computerized, automatic design cycle which is terminated when optimal system performance is achieved. It is with such a design process that this paper is concerned. An algorithm is presented that automates the system design. Its practicality is demonstrated by application to a high-pressure protection system.

2 SYSTEM DESIGN OPTIONS

Safety systems are designed to operate when certain conditions occur and act to prevent their development into a hazardous situation. As such there are certain features common to all safety protection systems. All safety systems have sensing devices which monitor for the occurrence of the triggering event. These sensors usually measure some process variable and transmit its current level to a controlling device. The controlling device determines whether the current situation is acceptable by comparing the input signal to a set point. When the sensed variable violates the set point the protective action is initiated. The protective action may either prevent a hazardous situation occurring or reduce its consequences.

The design engineer has a number of choices to make regarding the structure and operation of a safety system. These design options are described below.

2.1 Redundancy and diversity levels

The safety system must be designed to have a high likelihood of working on demand. Thus single component failures should not be able to prevent the system from functioning. One means of achieving this is by incorporating redundancy or diversity into the system structure. Redundancy duplicates elements within a system while diversity involves the addition of a totally different means of achieving the same function. Both redundancy and diversity can be used at component level or sub-system level.

Increased levels of duplication in the form of fully redundant and fully diverse elements will also increase
the number of spurious system trips. To counteract this, partial redundancy is commonly utilized. This is in the form of k-out-of-n voting structures where, for example, the trip function requires k sensors of the n sensors fitted to indicate the trip condition.

It is a design decision where to incorporate redundancy or diversity and whether fully or partially redundant systems are appropriate.

2.2 Component selection

Each component selected for the design will be chosen from a group of possible alternatives. For every valve, relay, pressure sensor, etc., for which a selection is to be made there will be several choices, each with associated characteristics such as failure rate in each failure mode, cost and time taken for their scheduled maintenance. The design engineer has to decide how to trade-off these characteristics to give the most effective opinion for the overall system performance.

2.3 Maintenance interval

Generally the time interval between preventive maintenance activities is assigned on an ad hoc basis. This is now changing with the more frequent application of reliability centred maintenance. Even so, the allocation of available maintenance effort is only considered after the system design has been finalized. Since component selection determines the time taken to maintain the system there are significant gains to be made by considering the maintenance frequency at the design stage.

3 LIMITATIONS ON THE DESIGN CHOICES

As outlined above, there are many options open to the design engineer. To produce the most effective system these parameters need to be selected to give optimal system performance. The choice of design is not, however, unrestricted. Some of the possible design variations will not be feasible. Practical considerations of limits placed on resources will prevent a completely free choice of system design.

4 HIGH-PRESSURE PROTECTION SYSTEM

The scheme developed in this paper is demonstrated by showing its application to the design of a high-pressure protection system. The basic features of the high-pressure protection system are shown in Fig. 1. Its function is to prevent a high-pressure surge passing through the system. In this way protection is provided for processing equipment whose pressure rating would be exceeded. The high pressure originates from a production well of a not normally manned offshore plat-

form and the pieces of equipment to be protected are vessels located downstream on the processing platform.

The first level of protection is to be the ESD (emergency shutdown) sub-system. Pressure in the pipeline is monitored using pressure transmitters (PTs). When the pipeline pressure exceeds the permitted value then the ESD system acts to close the wing and master valves on the well together with any ESD valves that have been fitted.

To provide an additional level of protection a second level of redundancy can be incorporated by the inclusion of a HIPS (high-integrity protection system). This works in a similar manner to the ESD system but is completely independent in operation. Again pressure transmitters are used to determine when the pipeline pressure exceeds the allowed limit. The pipeline is closed down using the HIPS valves.

4.1 Design options

Even with a relatively simple system such as this there are a vast number of options for the designer to consider. In this example it is required to determine values for the design variables that represent the following:

\[
\begin{align*}
\text{Designer options} & \quad \text{Design variable} \\
\text{How many ESD valves are required (0, 1, 2)?} & \quad E \text{ integer} \\
\text{How many HIPS valves are required (0, 1, 2)?} & \quad H \text{ integer} \\
\text{How many pressure transmitters for each sub-system (0, 1, 2, 3, 4)?} & \quad N_1, N_2 \text{ integers} \\
\text{How many transmitters are required to trip?} & \quad K_1, K_2 \text{ integers} \\
\text{Which of two possible ESD/HIPS valves to select?} & \quad V_1, V_2 \text{ Boolean} \\
\text{Which of two possible PT to select?} & \quad P_1, P_2 \text{ Boolean} \\
\text{Maintenance test interval for each sub-system (1 week–2 years)?} & \quad \theta_1, \theta_2 \text{ (in practice integer)}
\end{align*}
\]

In all there are 42,831,360 combinations of the twelve design variables. It is not possible to evaluate each potential design! Nor is it possible for the design engineer to understand the effect changes in each design parameter will have on the system performance.

4.2 Constraints

Limitations have been placed on the design such that:

1. The total system cost must be less than 1000 units. Hardware costs for each item in the system are given in Table 1.
2. The average time each year that the system resides in the down state due to preventive maintenance must be less than 130 hours. Times taken to service each

![Fig. 1 High-integrity protection system (HIPS)](image-url)
component at each maintenance test are shown in Table 1.
3. The number of times that a spurious system shutdown occurs would be unacceptable if it was more than once per year.

5 DESIGN OPTIMIZATION PROBLEM

The objective of the design optimization is to minimize system unavailability (the probability of system failure on demand) by manipulating the design variables such that limitations placed on them by constraints are not violated. Commonly with mathematical optimization problems there will be an explicit function (objective function) which defines how the characteristic to be minimized is related to the variables. An explicit objective function cannot be formulated and the system performance is assessed using a fault tree. Some form of the objective function must therefore be assumed and the region defined over which this approximate function can be considered accurate.

The nature of the design variables also adds difficulty to the problem, as do the constraints. Design variables that represent the levels of duplication for fully or partially redundant systems, selections of component types or the number of weeks between maintenance activity are all integer. A numerical scheme is required that produces integer values for these variables since it will not be appropriate to utilize a method where real numbers are rounded to the nearest whole number.

Constraints fall into two categories: those that can be determined from an explicit function of the design variables, and are therefore easily evaluated, and those that cannot be expressed as a function and can only be evaluated by a full analysis.

6 OPTIMIZATION METHOD

The method that has been used to solve the design optimization problem is an iterative scheme which produces a sequence of system designs gradually improving the safety system performance. When the design can no longer be improved due to restrictions placed by the constraints the optimization procedure terminates.

6.1 Initial design and analysis

An initial design is produced by the design engineer; for this system it is specified by the first row in Table 2. As can be seen from the table, pressure transmitter type 1 and isolation valve type 1 have been selected. A single ESD valve has been added to the wing and master valves on sub-system 1, the ESD system. Sub-system 2 has been included in the design and features one HIPS valve. Both the ESD and HIPS systems are fitted with two pressure transmitters and a trip is initiated when either transmitter indicates an unacceptably high pressure. Maintenance intervals have been specified as annually and six monthly for sub-systems 1 and 2 respectively.

To assess the performance of the preliminary design a fault tree was constructed and analysed for two system failure modes: (a) failure to work on demand and (b) spurious system trip. The basic event data for component failures are given in Table 1.

Analysis of the fault trees for the initial design produced a demand failure probability of $2.023 \times 10^{-3}$ and a spurious trip frequency of 0.8143 times per year. The initial system design would cost 902 units and if maintained at the intervals shown in the table would average downtime for maintenance activity of 109 hours per year.

6.2 Formulation of the optimization problem

To progress from the initial design to a new design that improves the system performance requires a means of

<table>
<thead>
<tr>
<th>Component</th>
<th>Failure rate</th>
<th>Mean repair time</th>
<th>Failure rate</th>
<th>Mean repair time</th>
<th>Cost</th>
<th>Test time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wing valve</td>
<td>$1.14 \times 10^{-5}$</td>
<td>36.0</td>
<td>$1 \times 10^{-6}$</td>
<td>36.0</td>
<td>100</td>
<td>12</td>
</tr>
<tr>
<td>Master valve</td>
<td>$1.14 \times 10^{-5}$</td>
<td>36.0</td>
<td>$1 \times 10^{-6}$</td>
<td>36.0</td>
<td>100</td>
<td>12</td>
</tr>
<tr>
<td>HIPS valve 1</td>
<td>$5.44 \times 10^{-6}$</td>
<td>36.0</td>
<td>$5 \times 10^{-7}$</td>
<td>36.0</td>
<td>250</td>
<td>15</td>
</tr>
<tr>
<td>HIPS valve 2</td>
<td>$1 \times 10^{-3}$</td>
<td>36.0</td>
<td>$1 \times 10^{-4}$</td>
<td>36.0</td>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>ESDV valve 1</td>
<td>$5.44 \times 10^{-6}$</td>
<td>36.0</td>
<td>$5 \times 10^{-7}$</td>
<td>36.0</td>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>ESDV valve 2</td>
<td>$1 \times 10^{-5}$</td>
<td>36.0</td>
<td>$1 \times 10^{-6}$</td>
<td>36.0</td>
<td>200</td>
<td>10</td>
</tr>
<tr>
<td>Solenoid valve</td>
<td>$5 \times 10^{-6}$</td>
<td>36.0</td>
<td>$5 \times 10^{-7}$</td>
<td>36.0</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Relay contacts</td>
<td>$0.23 \times 10^{-4}$</td>
<td>36.0</td>
<td>$2 \times 10^{-4}$</td>
<td>36.0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Pressure transmitter 1</td>
<td>$1.5 \times 10^{-6}$</td>
<td>36.0</td>
<td>$1.5 \times 10^{-4}$</td>
<td>36.0</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>Pressure transmitter 2</td>
<td>$7 \times 10^{-6}$</td>
<td>36.0</td>
<td>$7 \times 10^{-5}$</td>
<td>36.0</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Computer logic</td>
<td>$1 \times 10^{-5}$</td>
<td>36.0</td>
<td>$1 \times 10^{-5}$</td>
<td>36.0</td>
<td>20</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Pressure transmitter type</th>
<th>Isolation valve type</th>
<th>Number of ESDs</th>
<th>Number of HIPSs</th>
<th>Trip configuration ESD</th>
<th>Trip configuration HIPS</th>
<th>Test interval $\beta_1$</th>
<th>Test interval $\beta_2$</th>
<th>Unavailability</th>
<th>Trip rate</th>
<th>Cost</th>
<th>Maintenance down time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>52</td>
<td>26</td>
<td>2.023 $\times 10^{-3}$</td>
<td>0.8143</td>
<td>902</td>
<td>109.00</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>36</td>
<td>31</td>
<td>1.04 $\times 10^{-3}$</td>
<td>0.683</td>
<td>882</td>
<td>129.935</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>34</td>
<td>26</td>
<td>7.347 $\times 10^{-4}$</td>
<td>0.9807</td>
<td>802</td>
<td>129.369</td>
</tr>
</tbody>
</table>

© IMechE 1994
expressing the system performance criterion as a function of the design variables:

\[ Q_{\text{SYS}} = f(E, H, N_1, K_1, N_2, K_2, V_1, V_2, P_1, P_2, \theta_1, \theta_2) \]

that is

\[ Q_{\text{SYS}} = f(x) \]

where

\[ x = (E, H, N_1, K_1, N_2, K_2, V_1, V_2, P_1, P_2, \theta_1, \theta_2) \]

There is no function that can be written down for \( Q_{\text{SYS}} \) for all design options. However, if a Taylor series is expanded about the current design \( x \) the following is obtained:

\[ f(x + \Delta x) = f(x) + g^T \Delta x + \frac{1}{2} \Delta x^T H \Delta x + \cdots \quad (1) \]

where

\[ \Delta x = \text{the change in the design vector} \]

\[ g = \text{the gradient vector} \]

and

\[ H = \text{the Hessian matrix} \]

Truncating equation (1) after the linear term in \( \Delta x \) means that \( f(x + \Delta x) \) can be evaluated providing that the gradient vector can be obtained, that is \( \partial f/\partial x \), for each design parameter. Since integer variables are being dealt with \( \partial f/\partial x \), cannot be strictly formulated, but if consideration is taken of the fact that a smooth curve has been used to link all discrete points to give the marginal distribution of \( f \) as a function of \( x_i \), then \( \partial f/\partial x \), can be obtained for the smooth curve. The partial derivatives can be used to determine what values of \( f \) can be improved by updating each \( x_i \) by \( \Delta x_i \). A fault tree can be developed to obtain \( f(x_i + \Delta x_i) \) for each \( x_i \) providing \( x_i + \Delta x_i \) is integer; finite differences can then be used to estimate \( \partial f/\partial x \). This would require a large number of fault trees to be produced and analysed which would usually mean that this option would not be pursued from a practical viewpoint.

Since truncating the series at a finite number of terms provides only an approximation to \( f(x + \Delta x) \) the solution space over which this approximation is regarded as being acceptably accurate also needs to be defined. This is accomplished by setting up a restricted solution space in the neighbourhood of the current design point.

This procedure results in an iterative scheme where the optimal solution is approached by solving a sequence of optimization problems. Each problem in the sequence produces a new improved solution and the next problem to solve is defined by moving the feasible solution space to the neighbourhood surrounding this new solution and re-evaluating the objective function coefficients.

The optimization procedure requires fault trees to be evaluated at each iteration for the current design point. In addition, within each iteration fault trees are analysed to obtain numerical estimates of the partial derivatives of system performance with respect to each design variable. This information is required to produce the objective function coefficients. Fault tree construction is a time consuming task. A requirement to draw fault trees to represent causes of the two system failure models for several potential system designs would make the optimization method impractical. Manual development of a new tree for each assessment could not be contemplated. One possible way to resolve this difficulty would be to utilize a computerized automatic fault tree synthesis program (5–8). Unfortunately, these are not at present developed to the stage where they could confidently be used to accomplish this task.

Therefore, if the optimization scheme for the pressure protection system was to be successful an alternative approach was required to construct a fault tree for any specified system design. This was achieved using house events. House events can be included in the fault tree structure and either occur with certainty (event set to TRUE) or do not occur with certainty (event set to FALSE). Their inclusion in a fault tree model has the effect of turning on or off branches in the tree. A single fault tree was constructed, which by defining the status of house events could represent the causes of system failure on demand for any of the potential designs. A section of this fault tree which develops causes of a dormant failure of the HIPS sub-system is illustrated in Fig. 2. House events represent the following conditions:

- H1: HIPS valve 1 fitted
- NH1: HIPS valve 1 not fitted
- H2: HIPS valve 2 fitted
- NH2: HIPS valve 2 not fitted
- V1: Valve type 1 selected
- V2: Valve type 2 selected

As can be seen from the fault tree, if no HIPS option is included in the system design house events NH1 and NH2 will both be set. Once these events are set (that is TRUE) the output event from the OR gates into which they feed will also be true. At the next level up in the tree structure both inputs to the AND gate will have occurred and therefore the HIPS system will not provide protection.

Where HIPS valves are fitted the appropriate house events NH1 or NH2 will be set to FALSE and it will require component failure events to render the HIPS sub-systems inactive. Component selections were treated in the same way. Consider the bottom left-hand branch in Fig. 2 which represents 'HIPS valve 1 fails stuck'. This will depend on which type of valve has been selected in the design. If type 1 has been fitted V1 is set to TRUE (that is set to 1). If type 2 is fitted then V2 is set to TRUE. This provides the correct causes of the event being developed whichever valve is fitted. One of either V1 or V2 must be set.

By using house events in this manner all design options can be represented in a single fault tree. A second fault tree was constructed using the same technique to represent causes of spurious system failure for each potential design.

A feasible design space for a two-variable problem is illustrated in Fig. 3. Limitations on the design variables define the upper and lower bounds of the design space. Constraint functions will also remove other design alternatives as shown. Within the design space the feasible design points can be represented by the intersection points on the lattice since variables must take only integer values.

The objective function has been derived by truncating the Taylor series at the linear term. This truncation will
mean that the objective function is only a valid representation of \( Q_{\text{sys}} \) within the neighbourhood of the current design. Additional constraints are therefore included to place bounds on values taken by the updated design variable using this approximation; that is

\[ x_j^l - \Delta x_{i0} \leq x_j^{+1} \leq x_j^l + \Delta x_{i0} \] (2)

where \( x_j^l \) and \( x_j^{+1} \) represent design variables at the \( j \)th and \((j + 1)\)th iterations and \( \Delta x_{i0} \) are the lower and upper limits by which \( x_j \) is allowed to change. The objective function is then evaluated at each feasible joint in the restricted design space and the optimal point is selected.

The constraint forms for the cost and maintenance downtime are:

\[
\text{COST} = \text{COST(SUBSYS1)} + \text{COST(SUBSYS2)} \leq 1000 \quad (3)
\]

\[
\text{COST(SUBSYS1)} = E(V_1 C_{V1} + V_2 C_{V2} + C_0) + N_1(P_1 C_{P1} + P_2 C_{P2}) + 261 \quad (4)
\]

\[
\text{COST(SUBSYS2)} = H(V_1 C_{V1} + V_2 C_{V2} + C_0) + N_2(P_1 C_{P1} + P_2 C_{P2}) + 21 \quad (5)
\]
where
\[ C_{V1}, C_{V2} = \text{costs of the two valve types} \]
\[ C_{P1}, C_{P2} = \text{costs of the two pressure transmitter types} \]
\[ C_S = \text{cost of the solenoid valves} \]
and 261 [equation (4)] is the cost of the wing valve, master valve, their solenoid valves, the computer and control relay for sub-system 1. The constant in equation (5), 21, is the cost of the computer and control relay for sub-system 2.

**Average Maintenance Downtime (MDT)**

\[
\text{MDT} = \text{MDT(SUBSYS1) + MDT(SUBSYS2)} \leq 130
\]  

(6)

\[
\text{MDT(SUBSYS1)} = \frac{52}{\theta_1} [E(V_1 M_{V1} + V_2 M_{V2} + M_S) + N_1(p_1 M_{P1} + p_2 M_{P2}) + 47]
\]  

(7)

\[
\text{MDT(SUBSYS2)} = \frac{52}{\theta_2} [H(V_1 M_{V1} + V_2 M_{V2} + M_S) + N_2(p_1 M_{P1} + p_2 M_{P2}) + 13]
\]  

(8)

where
\[ M_{V1}, M_{V2} = \text{test times of the two valve types} \]
\[ M_{P1}, M_{P2} = \text{test times of the two pressure transmitter types} \]
\[ M_S = \text{test time of the solenoid valve} \]
and the constant 47 in equation (7) is the test time for the wing and master valves, their solenoids, the computer and control relay for sub-system 1. The test time for the computer and control relay for sub-system 2 [the constant in equation (8)] is 13.

These constraints are explicit functions of the design variables and are therefore easily evaluated. The third and final constraint limits the spurious trip rate of the system; this is an implicit constraint requiring a full evaluation of a fault tree to determine whether it is less than the acceptable value of once per year.

**7 Optimization Algorithm**

The full sequential optimization algorithm is therefore:

1. Construct a fault tree which by the use of house events is capable of representing the causes of dormant failure for each possible system design. This fault tree will be utilized to determine the system unavailability for any design vector \( x \).
2. Construct a fault tree which by use of house events is capable of representing the causes of spurious trip for each possible system design. This fault tree will be utilized to determine the frequency of spurious system failure for any design vector \( x \).
3. Select some feasible initial design vector \( x^0 \). Set the relevant house events in the system unavailability fault tree. Incorporate the selected maintenance inspection intervals in the component failure models and analyse the fault tree to determine \( Q_{SYS}^0 \); the unavailability of the initial design. Set the relevant house events in the spurious trip fault tree and analyse to get \( \omega^0 \) (check that this value satisfies the trip rate constraint).
4. Choose the form of the objective function that can be used to represent \( Q_{SYS} \) in the neighbourhood of the current design vector. For a linear objective function
evaluate the partial derivatives to specify the terms in
the function using the central difference formula if
possible:
\[
\frac{\partial Q_{\text{SYS}}}{\partial x_i} = \frac{Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i + dx_i, x_{i+1}, \ldots, x_n) - Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i - dx_i, x_{i+1}, \ldots, x_n)}{2dx_i}
\]

the system unavailability terms being evaluated by
perturbing the variable \(x_i\), setting the correct house
events and re-analysing the fault tree.

If constraints are violated by the new values for \(x_i\),
that is \(x_i + dx_i\) or \(x_i - dx_i\) then use either the back-
ward difference scheme:
\[
\frac{\partial Q_{\text{SYS}}}{\partial x_i} = \frac{Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)}{dx_i}
\]
or the forward difference scheme:
\[
\frac{\partial Q_{\text{SYS}}}{\partial x_i} = \frac{Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i + dx_i, x_{i+1}, \ldots, x_n) - Q_{\text{SYS}}(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)}{dx_i}
\]
to resolve this.

When each design derivative has been evaluated
the objective function is given by
\[
Q_{\text{SYS}}(x + \Delta x) = Q_{\text{SYS}} + \sum_{i=1}^{n} \left( \frac{\partial Q_{\text{SYS}}}{\partial x_i} \right) dx_i
\]

5. Incorporate the constraints limiting the cost and
maintenance downtime for the design together with
the constraints restricting the current feasible design
space (move limit constraints) to formulate the opti-
mization problem.

6. For implicit constraints such as the system spurious
trip rate \(\omega_i\), formulate a function to represent \(\omega_i\) in
terms of the design variables over the current feasible
region. This involves assuming some form of the
function and expanding in a Taylor series as for the
objective function. Derivatives required to specify
coefficients in this function are again evaluated by
analysing the spurious trip fault tree with the rele-
vant house events set and utilizing the finite differ-
ence formulae.

7. Minimize the system unavailability \(Q_{\text{SYS}}(x^j)\) over
the current design space such that the imposed con-
straints are satisfied. The spurious trip rate will be
evaluated using its approximation function. This will
produce a new updated design vector \(x^{j+1}\).

8. If \(x^j = x^{j+1}\) STOP, since the optimization scheme
cannot improve on the current design. The optimal
design achieved is specified by \(x^j\).

9. Otherwise re-set the relevant house events and com-
ponent inspection intervals in the system fault trees
and reanalyse to give \(Q_{\text{SYS}}^{j+1}\). Check for feas-
ibility \(\omega_{\text{min}} - \omega^{j+1} \geq 0\). Check for conver-
gence \(Q_{\text{SYS}}^{j+1} - Q_{\text{SYS}}^j \leq 0\). If either condition is violated
the approximation functions did not represent \(Q_{\text{SYS}}\) and
\(\omega\) with sufficient accuracy over the feasible region.
Retune move limits constraints. If move limits cannot be
reduced further then the limit of resolution is
achieved and the optimization procedure stops with
\(x^{j+1}\) optimal. Else return to step 4 with the new
design vector \(x^{j+1}\).

8 MINIMIZATION OF THE OBJECTIVE FUNCTION

Choosing an algorithm to minimize the objective func-
tion proved to be a difficult task owing to the presence
of the integer variables and the implicit constraint. An
approach whereby a standard method for real variable
optimization is used and the result rounded to the
nearest whole number could not be used due to the
Boolean variables which represented the component
selections. Rounding in this case could well produce an
infeasible design.

The method used was a simple grid-sampling method.
The lattice of possible design points was constructed for
the current design space and the values of the objective
function for each point on the lattice were then evalu-
ated and directly compared. Each constraint was also
evaluated to ensure feasibility (17 constraints in all).
With modern computers and the restriction placed on
the design space for each iteration, this proved to be
quite a fast method which guaranteed that the global
minimum was obtained.

9 OPTIMIZATION RESULTS (LINEAR OBJECTIVE
FUNCTION)

Carrying out the sequential design optimization pro-
cedure detailed above results in the design specification
given in iteration 1 of Table 2. Design changes have
been made from the initial design which remove the
ESD valve and include an additional HIPS valve. A trip
on the HIPS now takes place when a high pressure is
indicated by a single transmitter. In addition, main-
tenance activity of the ESD and HIPS sub-systems will
take place every 36 and 32 weeks respectively. As a
result of these modifications the system failure proba-
bility has been reduced by almost 50 per cent to
1.04 \times 10^{-5}. Other benefits obtained are the reduction
in both cost and spurious trip rate compared with the
preliminary design. Inspection intervals have been
adjusted so that downtime due to maintenance is just
within the maximum limit imposed.

The optimization process is repeated using the
updated design as the current design point. Again an
improvement was obtained. System unavailability was
now reduced to 7.347 \times 10^{-4} at the same time as
reducing cost (iteration 2, Table 2). This proved to be
the optimal design.

10 OTHER FORMULATIONS OF THE OBJECTIVE
FUNCTION

To investigate the potential to improve the efficiency of
the technique by taking extra terms in the Taylor
The first variation investigated was to incorporate the pure quadratic contributions from an additional term in the Taylor expansion. The objective function was then of the form:

$$Q_{SYS}(x + \Delta x) = Q_{SYS} + \sum_{i=1}^{n} a_i \frac{\partial Q_{SYS}}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{i=1}^{n} a_{i+i} (\Delta x_i)^2$$

where

$$a_i = \left( \frac{\partial Q_{SYS}}{\partial x_i} \right)_i, \quad i = 1, \ldots, n$$

and

$$a_{i+i} = \left( \frac{\partial^2 Q_{SYS}}{\partial x_i \partial x_i} \right)_i, \quad i = 1, \ldots, n$$

The first derivatives are calculated numerically as for the linear scheme. To determine the estimates for the second-order derivatives successive design points $j$ and $j + 1$ are used. Then

$$a_{i+i} = \frac{(Q_{SYS}(\Delta x_i)^j - Q_{SYS}(\Delta x_i)^{j+1})}{x_i^j - x_i^{j+1}}$$

This means that the first iteration step in the sequential procedure is linear. Once the first iteration has been accomplished information is available to evaluate the second-order terms which enables the use of the quadratic objective function. Since a more accurate approximation was being made to the objective function the region over which it was assumed to be accurate was extended. The move limit constraints were relaxed by a factor of 2. The solution to the problem achieved employing this objective function is shown in row 2 of Table 3. It can be seen that the optimal design varies significantly from that achieved with a linear function and its resulting system unavailability is relatively poor in comparison. This design was produced after only two iterations, at which point the optimization routine failed to produce an improved design vector.

The second variation in the objective function was to then include the cross-product terms in the Taylor expansion, that is terms in $\frac{\partial^2 Q_{SYS}}{\partial x_i \partial x_j}$, again relaxing the bounds. This resulted in the design vector represented in row 3 of Table 3 with system unavailability of $9.894 \times 10^{-4}$. This was again a poorer result than had been achieved with the simpler linear objective function. Its only difference to the linear function design was the omission of a HIPS valve. Although this slightly degraded the system performance it had trade-offs in an improved spurious trip rate and cost. Five iterations were required to achieve optimality.

The final variation to the objective function was to retain all elements of the quadratic terms of the Taylor series but to restrict the region over which it was regarded as being sufficiently accurate to that used with the linear approximation. This form of the objective function produced the optimal design vector in three iterations. Row 4 of Table 3 shows the resulting design vector. An unavailability of $7.447 \times 10^{-4}$ results. This performance level is very similar to that with the linear objective function. The only difference between the designs is that a trip for the ESD sub-system is initiated by two of the three pressure transmitters, indicating a high pressure. Although this costs more than the best design produced (row 1) the difference between the two system unavailabilities is negligible and the second design may be preferred due to its lower spurious trip rate.

11 DISCUSSION

From Table 3 it can be seen that the most effective optimization procedure was based on the simple linear formulation of the objective function representing the system unavailability. Bounds governing the feasible design space for each iteration over which the optimization is to be carried out are then placed fairly tight on the current design. For example, variables representing the redundancy levels are only permitted to increase or decrease by one level from the current design.

This proved to be the best way to model the unavailability since it was found that incorporation of the quadratic terms added little accuracy to the objective function when the domain was extended. The reason for this is the way that the system unavailability is affected by changes in redundancy levels. If a particular component or sub-system is the most likely cause of system failure it will make a contribution to the lower order minimal cut sets. Adding an extra level of redundancy increases the order of these minimal cut sets by one and usually means that other components or sub-systems now provide the more significant contributions to system failure. Increasing the redundancy levels beyond this does not significantly reduce the system unavailability further. The rate of improvement achieved therefore does not match that predicted by the objective function, even when the quadratic terms are included.

The linear objective function is also the easiest to determine since fewer fault tree evaluations are required for its formulation.

Table 3  Optimal solutions

<table>
<thead>
<tr>
<th>Objective function</th>
<th>Pressure transmitter type</th>
<th>Isolation valve type</th>
<th>Number of ESDs</th>
<th>Number of HIPS</th>
<th>Trip configuration ESD</th>
<th>Trip configuration HIPS</th>
<th>Test interval $\delta_1$</th>
<th>Test interval $\delta_2$</th>
<th>Unavailability</th>
<th>Trip rate</th>
<th>Cost</th>
<th>Maintenance delay time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1/2</td>
<td>1/2</td>
<td>34</td>
<td>26</td>
<td>7.347 $\times 10^{-4}$</td>
<td>0.9807</td>
<td>802</td>
<td>129.369</td>
</tr>
<tr>
<td>Quadratic (pure)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1/2</td>
<td>1/2</td>
<td>33</td>
<td>19</td>
<td>1.02 $\times 10^{-4}$</td>
<td>0.806</td>
<td>632</td>
<td>129.455</td>
</tr>
<tr>
<td>Quadratic (cross)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1/2</td>
<td>1/2</td>
<td>31</td>
<td>16</td>
<td>9.894 $\times 10^{-4}$</td>
<td>0.8887</td>
<td>582</td>
<td>129.9</td>
</tr>
<tr>
<td>Quadratic (bounded)</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1/2</td>
<td>1/2</td>
<td>32</td>
<td>28</td>
<td>7.447 $\times 10^{-4}$</td>
<td>0.7185</td>
<td>822</td>
<td>129.737</td>
</tr>
</tbody>
</table>

12 CONCLUSIONS

A method has been presented that enables safety systems to be designed making the best use of the resources available. The design optimization technique is an iterative procedure that utilizes the fault tree analysis technique to assess the performance of any given design. To demonstrate the practicality of the approach it has been applied to a high-pressure protection system. As a result, the system unavailability was reduced from $2.023 \times 10^{-3}$ to $7.347 \times 10^{-4}$ in only two iterations. This improvement was accomplished within the stated constraints and as an additional benefit the cost of the system hardware was reduced.

The method presented here has the following features:

1. The design procedure can be automated, providing less of a reliance on engineering judgement.
2. The resulting design makes the best use of available resources rather than achieving just an 'adequate' design.
3. While the method has been applied here to a safety system the procedure is generic and could equally be applied to optimize the performance of any engineering system whose failure causes can be represented by a fault tree.
4. Constraint forms do not inhibit the application of the sequential method. Implicit constraints or explicit constraints of the equality or inequality form can be incorporated.
5. Maintenance activity can be incorporated into the optimization scheme if required to provide maximum benefit from the procedure or the design can be considered in two stages, as is traditional, with the system structure being formulated first and the maintenance activity specified afterwards.

REFERENCES
