Diagnosability evaluation of systems using bipartite graph and matrix approach

M.F. WANI AND O.P. GANDHI

Industrial Tribology, Machine Dynamics and Maintenance Engineering Centre (ITMMEC), Indian Institute of Technology, Delhi, New Delhi, India

(RECEIVED October 5, 1999; ACCEPTED January 7, 2000)

Abstract

This paper presents a methodology for assessment of diagnosability of mechanical and hydraulic systems. The method is developed on the basis of relationships between system performance parameters and physical objects, that is, components of the system. These relationships are identified by system functional domain and are modeled in terms of a bipartite graph, called Diagnosability Bipartite Graph (DBG). A matrix called Diagnosability Matrix (DM) represents the DBG. Various diagnosability parameters of the system are derived from the DBG and the DM and these are useful in evaluation and comparison of design variants of the system. These parameters: are maximum number of set conflicts (MNS), maximum number of components in a set conflict (MNCS), diagnosability effort and cost (DEC), and average merit of diagnosability (AMD). The design having the lowest value of MNCS, AMD, and DEC; and highest value of MNS has the highest diagnosability. On the basis of these, a best design alternative is selected from diagnosability point of view. Moreover, components, which have poor diagnosability, are also identified. Maximum number of set conflicts (MNS) also guides in system fault diagnosis. The proposed procedure aids in the design and development of maintainable systems from diagnosability consideration. The method can also be used for evaluating and comparing the diagnosability of the systems. This method is illustrated with the help of two examples.

Keywords: Diagnosability; Bipartite Graph; Performance Parameter; System Design; Fault Diagnosis; Artificial Intelligence

1. INTRODUCTION

Fault identification and localization of a system constitute a major portion of downtime due to poor its diagnosability. System diagnosability is characteristic of the system design and is facilitated by its features, such as malfunction annunciation and fault isolation. In addition, it is also dependent upon accessibility, modularization, support equipment, skill of maintenance person, and documentation. All these factors contribute to enhance system diagnosability at its operational stage. In recent years, expert systems, neural network and condition monitoring in conjunction with artificial intelligence (AI) have been applied to improve diagnosability for various systems. But, identification and

Reprint requests to: M.F. Wani, Industrial Tribology, Machine Dynamics and Maintenance Engineering Centre (ITMMEC), Indian Institute of Technology, Delhi, New Delhi 110016, India. E-mail: wani@itmmec.iitd.ernet.in

localization of faulty line replaceable unit is still a major cause of concern. Two main reasons for this are: increasing complexity of the systems, and the nonincorporation of diagnosability at design stage. In addition, diagnosability procedures are not well documented and are not user friendly to the maintenance person. It is possible to enhance diagnosability of a mechanical system, if the designer takes into consideration diagnosability at its design stage considering all its features. Various methodologies have been applied for evaluation of diagnosability of the systems. Some of these are discussed under the literature review.

AI model-based diagnosis has been suggested for electronic devices (De Kleer, 1979; Davis et al. 1982; Davis, 1984; Forbus, 1984; Genesereth, 1984; De Kleer & Williams, 1987; Reiter, 1987; Tzafestas, 1989; Larsson, 1994, 1996). In general, computational complexity of the model-based diagnosis grows rapidly with the complexity of the system model. Moreover, model-based diagnosis is suited for diagnosis of electronic circuits and similar domains and not for mechanical systems. Failure mode and effects analy-

sis (FMEA) has been applied for reliability and safety analysis of systems at design stage (Bellinger, 1966). FMEA is useful to identify critical components and assemblies considering their failure effects. Although it provides directions to incorporate malfunctions annunciation for the critical components and assemblies of the system, yet it cannot be used for the diagnosability evaluation of the system. Moreover, this is applied when hardware details of the system are known.

Diagnosability evaluation for the system has been initiated based on design checklists and scoring criteria under maintainability consideration (DOD, 1984). In this method, diagnosability is evaluated in terms of score value 0, 2, or 4 based on capability of fault or malfunction localization of the system including its Built-In-Test Equipment (BITE) feature. This scoring method is subjective, and diagnosability of the system cannot be ascertained. Mathematical models have been used to define false-alarm probability, alarm defect, and fault localization for mechanical and electronic systems (IEC, 1994). These models are based on probability approach and are applicable only to BITE. Three principles of testability characteristics of equipment have been proposed to evaluate the diagnosability (Kowalski, 1988). These are fault-detection capability, fault-isolation capability, and fault-alarm rate for measuring the failures detected by the system, ambiguity associated with the fault-isolation activities, and rate of deceleration of detection of failure, respectively. These measures are more suitable for detecting and isolating faults to some reasonable subset of the system, so that repairs can be accomplished in a reasonable time. However, this methodology is not suitable for measuring the diagnosability of components and systems. A robust model has been proposed for diagnosability evaluation, using hierarchical model scheme and failure probability of components and assemblies (Nakakuki et al., 1992). The authors claim its applicability to electronic and mechanical systems. However, its application to an electronic system has only been demonstrated. Moreover, as this depends upon the probability of failure of component and systems, and can, therefore, be applied to the existing designs only. Diagnosis based on explicit means-end models has been developed and applied to process plants (Larsson, 1994, 1996). Diagnosability-evaluation methodologies have been developed for mechanical systems for new and existing designs (Clark & Paasch, 1996; Murphy & Paasch 1997; Paasch & Ruff, 1997). These methodologies use the concept of relationship between functional hierarchy and physical hierarchy of the system to determine the various parameters for diagnosability evaluation. Although, this approach addresses the basic requirement of evaluation of diagnosability in mechanical systems, the proper modeling of the diagnosability has not been suggested. Moreover, it is not convenient to use this methodology when the number of relations between functions and components increases.

It is apparent, that the diagnosability evaluation of mechanical systems is still far from complete. The main requirement is to provide a metric to designer, which can be used at design stage for the evaluation of diagnosability, in particular for a new design, in a convenient and efficient way. As such, there is no appropriate procedure available to the designer. Diagnosability can be modeled if the relationships between system performance parameters and components/assemblies of the system are represented in an appropriate and convenient way. This can be accomplished by using graph-theoretic concepts.

In this paper, diagnosability modeling and diagnosability evaluation of mechanical systems is suggested. The suggested approach is not only valid at its conceptual and hardware design stages of the system, but also at the operating stage of the system. This is achieved by identifying relationships between functions and physical objects. These relationships determine the connections between performance parameters and physical objects. This relationship is modeled in terms of Diagnosability Bipartite Graph (DBG) and also in terms of Diagnosability Matrix (DM). Various diagnostic parameters are obtained from these models, which aid the designer in evaluating and comparing the diagnosability of various design variants of the system.

2. DIAGNOSABILITY MODELING

A design process is initiated by transferring customer requirements into design parameters or functional requirements (Bracewell & Sharpe, 1996; Iwasaki & Chandrasekaran, 1992; Qian & Gero, 1996). The main function is decomposed into various subfunctions depending upon, its complexity, and then transformed to design objects. The design object is physical object, that is, component or assembly, also called physical domain (Suh, 1990). The physical domain represents components or assemblies of the system, which fulfil the functional objectives to achieve the desired output. As such, the desired output of the system depends upon the behavior of every individual component. A performance parameter is related to functioning and behavior of the physical object, that is, it is in fact a feature or specified measurable attribute of a physical object, assembly or system, such as temperature, pressure, vibration, noise, wear etc.

The relations among functional requirements and physical objects are established first at the system conceptual-design stage. Therefore, the relationship between performance parameters (PPs) attributed to functions and physical objects (POs) is developed. If the functional independence between the physical objects is maintained, that is, no functional sharing exists, diagnosability is forthright. If the independence of the function is not sustained, a physical object, which takes part in the highest number of functions is likely to be a candidate in locating the fault. However, this needs further testing for its confirmation. This reduces the diagnosability of this component, in particular, and the system in general. Therefore, the relationship between PPs and POs need to be modeled to develop a procedure for diagnosability evalua-

tion of components and assemblies of the system. This, however, needs appropriate and efficient representation of PPs and POs and their relationship with one another. Graph-theoretical models provide a suitable framework for representing these relationships (Deo, 1974).

It has already been discussed that main function of the system is decomposed into various subfunctions for understanding and development of its design. The functional requirements are transformed into POs to build an artifact. These functions are measured in terms of PPs. Assuming the number of PPs is equal to the number of POs, which means each performance parameter is able to monitor the functioning of each component independently to understand the behavior of the system.

Let us say the number of PPs is a set N, that is,

$$N = \{n_1 | n_i \in \mathbb{N}, i = 1, 2, \dots, y\}$$
 (1)

and POs is a set V, that is,

$$V = \{v_j | v_j \in V, j = 1, 2, \dots, z\},\tag{2}$$

where n_i and v_j represent the PP and PO of a system under consideration, and y and z are its total number of PPs and POs, respectively. There exists a relationship between the PPs and POs depending upon the functional sharing of POs. This functional sharing relationship (R_{NV}) is represented by a set:

$$R_{NV} \subset N \times V \tag{3}$$

$$R_{NV} = \{ (n_i, v_i) | n_i \in N, v_i \in V \}.$$
 (4)

The expression $\langle n_i, v_j \rangle \in R_{NV}$ means relationship exists between n_i and v_j and this relationship is represented in terms of a diagnosability bipartite graph, G_{DB} , that is,

$$G_{DB} = \langle M, D, R_{MD} \rangle \tag{5}$$

$$R_{MD} \subset M \times D$$
 (6)

$$M = \{m_p : p = 1, 2, \dots\}$$
 (7)

$$D = \{d_q : q = 1, 2, \dots\},\tag{8}$$

where M is the set of nodes representing PPs and D is the set of nodes representing the POs. The expression $\langle m_p, d_q \rangle \in R_{MD}$ means that node m_p representing performance parameter is related to node d_q representing physical object of the DBG.

A DBG for a simple case is shown in Figure 1, in this case p=q. In this DBG, the relationship between PPs and POs is subset of nodes of PPs, $m_p \in M$ adjacent to set of nodes POs $d_a \in D$ and is represented as:

$$R_{MD} = \{ (m_1, d_1), (m_2, d_2), (m_3, d_3), (m_4, d_4), (m_5, d_5).$$
 (9)

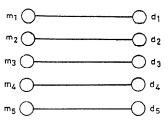


Fig. 1. Diagnosability bipartite graph—simple case.

In this case, the diagnosability is forthright, as the performance parameters are able to isolate the faulty component directly, that is, without further testing. This is substantiated by the adjacency in the DBG, that is, every node in M is matched against only one node in D. Therefore, adjacency of the DBG indicates the diagnosability of the system design.

Let 'r' represents subset of nodes of PPs in M and 's' the number of adjacent nodes of subset of M in D. The value of (r-s) indicates the adjacency for any value of r between the nodes M representing the PPs, to nodes in D representing the POs. The value of (r-s) is adapted to compare the diagnosability of the component and the system. For better diagnosability, r=s, that is, the value of (r-s) is equal to 0. A lower value of (r-s) (i.e., <0) indicates a higher relationship between the nodes of M and D. It means functional independence is not sustained, and the PPs cannot isolate the faulty component without further testing. For subset of r=1 in M, that is, $\{m_1\}$, $\{m_2\}$, $\{m_3\}$, $\{m_4\}$, and $\{m_5\}$ has only one node adjacent to them in D, that is, $\{d_1\}$, $\{d_2\}$, $\{d_3\}$, $\{d_4\}$, and $\{d_5\}$, respectively, as shown in Figure 1.

For r = 1, for the above graph, the values of s and (r - s) are obtained from Eq. (9) and these are as:

Subset of M	Value of <i>r</i>			Value of $r - s$	
$\{m_1\}$	1	$\{d_1\}$	1	0	
$\{m_2\}$	1	$\{d_2\}$	1	0	
$\{m_3\}$	1	$\{d_3\}$	1	0	
$\{m_4\}$	1	$\{d_4\}$	1	0	
$\{m_5\}$	1	$\{d_5\}$	1	0	

In this case, the minimum value of (r - s) = 0, which indicates higher diagnosability, because the PPs can isolate the faulty component directly.

It is also inferred that if r = 1 and r - s is less than 0.

$$r - s < 0, r = 1. (10)$$

This means for r = 1 subset of M, that is, m_p has more than one adjacent node in D, that is, d_q 's and it indicates poor diagnosability of the system.

The relationship between PPs and POs is not always as indicated in Figure 1. In general, number of PPs is less than

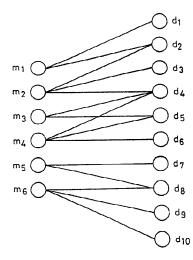


Fig. 2. A typical diagnosability bipartite graph.

POs, and, as such, functional independence cannot be maintained, that is, functional sharing exists. In such cases, the PPs cannot explicitly isolate a faulty component, without further testing a set of suspected system components. This is shown in Figure 2 for a more general case. For r=1, the subset of M adjacent to vertices of D indicates the relationship between PPs and POs. The set of R_{MD} is equal to, that is,

$$\begin{split} R_{MD} &= \big[\{m_1,(d_1,d_2)\},\{m_2,(d_2,d_3,d_4)\},\{m_3,(d_4,d_5)\},\\ &\quad \{m_4,(d_4,d_5,d_6)\},\{m_5,(d_7,d_8)\},\{m_6,(d_8,d_9,d_{10})\}\big]. \ (11) \end{split}$$

This indicates functional interdependence exists. There can be more than one suspected component (POs) attributed to the faulty PPs. This may lead to poor diagnosability. From Eq. (11) the value of s and (r - s) are:

Subset of M	Value of <i>r</i>	Adjacent Nodes in <i>D</i> for Subset of <i>M</i>	Value of s	Value of $r - s$
$\overline{\{m_1\}}$	1	$\{d_1, d_2.\}$	2	-1
$\{m_2\}$	1	$\{d_2, d_3, d_4\}$	3	-2
$\{m_3\}$	1	$\{d_4, d_5\}$	2	-1
$\{m_4\}$	1	$\{d_4, d_5, d_6.\}$	3	-2
$\{m_5\}$	1	$\{d_7, d_8.\}$	2	-1
$\{m_6\}$	1	$\{d_8, d_9, d_{10}.\}$	3	-2

The minimum value of (r - s) is -2, which is <0 in this case. This means PPs cannot isolate the faulty component without the further testing, and the diagnosability is evidently poor compared to simple case (Fig. 1). This case, is inferred as:

$$\{m_1\} \cap \{m_2\} = \{d_2\} \Rightarrow m_1 \wedge m_2 \underline{\mathbb{R}} d_2$$

$$\{m_2\} \cap \{m_3\} \cap \{m_4\} = \{d_4\} \Rightarrow m_2 \wedge m_3 \wedge m_4 \underline{\mathbb{R}} d_2$$

$$\{m_3\} \cap \{m_4\} = \{d_5\} \Rightarrow m_3 \wedge m_4 \underline{\mathbb{R}} d_5$$

$$\{m_5\} \cap \{m_6\} = \{d_8\} \Rightarrow m_5 \wedge m_6 \underline{\mathbb{R}} d_8$$

$$(12)$$

and

$$m_1 \underline{R} d_1, \text{ for } d_1 \notin \bigcap_{i=1}^6 m_i$$

$$m_2 \underline{R} d_3, \text{ for } d_3 \notin \bigcap_{i=1}^6 m_i$$

$$m_4 \underline{R} d_6, \text{ for } d_6 \notin \bigcap_{i=1}^6 m_i$$

$$m_5 \underline{R} d_7, \text{ for } d_7 \notin \bigcap_{i=1}^6 m_i$$

$$m_6 \underline{R} d_9, d_{10}, \text{ for } d_9, d_{10} \notin \bigcap_{i=1}^6 m_i.$$

$$(13)$$

This helps to find out a set of suspected components for a performance parameter and is called a set conflict. This is represented as:

$$\langle \dots, \dots, \dots \rangle$$
. (14)

The set conflict also helps to determine the total number of set conflicts. One can identify number of all possible set conflicts, and is called maximum number of set conflicts (MNS) for a system. This is the first diagnosability parameter to evaluate system-design alternatives. The value of this parameter is obtained from Eqs. (12) and (13). The maximum number of set conflicts for the above two cases, that is, Figures 1 and 2 are obtained and are as below:

$$\langle d_1 \rangle$$
 $\langle d_2 \rangle$
 $\langle d_3 \rangle$
 $\langle d_4 \rangle$
 $\langle d_5 \rangle$ (15)

and

$$\langle d_1 \rangle$$

$$\langle d_2 \rangle$$

$$\langle d_3 \rangle$$

$$\langle d_4 \rangle$$

$$\langle d_5 \rangle$$

$$\langle d_6 \rangle$$

$$\langle d_7 \rangle$$

$$\langle d_8 \rangle$$

$$\langle d_9, d_{10} \rangle. \tag{16}$$

The number of set conflicts in first case (Figure 1) is equal to five. These form the MNS. Similarly, the MNS for the second case, that is, Figure 2 are obtained using Eq. (16) and are equal to nine.

The maximum number of elements in any set conflict indicates the maximum number of suspected components for performance parameter PPs. If the number of candidates in the set conflict is more than one, it means two more parameters are required to isolate the faulty component, this means poor diagnosability. Therefore, the candidates in a set conflict also determine the diagnosability of the system. This is defined as maximum number of candidates in set conflict (MNCS). The value of MNCS gives the second parameter for evaluation and comparison of diagnosability. In the two cases discussed the MNCS is obtained from Eqs. (15) and (16) for DBG Figures 1 and 2, respectively. The value of MNCS for first case is obtained from Eq. (15) and is equal to one. Similarly, the value of MNCS equal to two for second case indicated by set conflict $\langle d_9, d_{10} \rangle$, obtained from Eq. (16).

The diagnosability is also indicated by the effort required to measure the parameters in terms of time and test equipment required. This is defined as diagnosability effort and cost (DEC). The value of DEC gives the third parameter to evaluate the diagnosability of the system. If the data for the cost and effort for the system parameters are available, it is preferred in DEC determination. Assume that the cost of measuring is the same for all the parameters for the above two cases. The value of DEC for the two cases is equal to the number of parameters to be measured 5 and 10.

The fourth parameter of diagnosability is obtained by considering the affect of diagnostic cost and maximum number of set conflicts. This is defined as average merit of diagnosability (AMD), and is equal to the ratio of diagnostic effort and cost to the maximum number of set conflicts in a system, that is,

$$AMD = \frac{DEC}{MNS}.$$
 (17)

This gives an idea of ambiguity the diagnostician may face in diagnosing the system. The lower the value of AMD, close to one, indicates better diagnosability. This means the maximum number of set conflicts is equal to diagnostic cost and effort. However, when the number of set conflicts decreases, the value of AMD increases and diagnostician needs more prior knowledge about the system. The value of AMD for Figures 1 and 2 is obtained from Eq. (17), and is equal to 1 and 1.1, respectively. System with low values of MNCS and AMD require less parameter measurement to isolate a fault for worst and average cost. The system with low values of DEC requires minimum effort in terms of time and test to isolate the faulty component.

To rate the design variants of a system, the first priority is given to the design with minimum value of MNCS, as its value represents need for further testing to identify the faulty component. The second priority is given to designs with a low value of AMD, as its value gives average cost required to isolate the faulty component. The third with low value of DEC, as its value gives the diagnostic cost of a system, and the fourth with highest value of MNS. The higher value of

MNS means more number of set conflicts. The DBGs represented by Figures1 and 2 are not the design variant of the same system, and these cannot be compared.

The discussion so far clearly shows that DBG helps in defining the various parameters quantitatively, evaluating the diagnosability of the system at conceptual design stage. However, when the number of nodes of M and D increases for complex systems due to the increase in the number of functional requirements and POs. The relationship between the nodes of the DBG also changes and culminates into more diagnosability problems. Therefore, it is appropriate to represent DBG by a matrix for dealing it conveniently using computer. This is discussed in the next section.

3. MATRIX REPRESENTATION OF DBG

The relationship between the set of vertices of PPs and POs in DBG is represented as:

$$R_{MD} = \{0 \to \langle m_p, d_q \rangle \notin R_{MD}\} \tag{18}$$

$$R_{MD} = \{1 \to \langle m_p, d_q \rangle \in R_{MD}\},\tag{19}$$

that is, if m_p is connected to d_q , the value of matrix entity is $r_{pq} = 1$, or otherwise the value is $r_{pq} = 0$. The matrix thus developed for a DBG is called DM and is represented by Expression (20).

$$\mathbf{D_{m}} = \begin{bmatrix} d_{1} & d_{2} & d_{3} & d_{4} & d_{5} \\ m_{1} & r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ m_{2} & r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \\ m_{5} & r_{51} & r_{52} & r_{53} & r_{54} & r_{55} \end{bmatrix}.$$
(20)

Diagnosability matrix in Expressions (21) and (22) as therefore represents the matrices for DBGs Figures 1 and 2:

$$\mathbf{D_{m1}} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 & d_5 \\ m_1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ m_5 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$(21)$$

(22)

Expressions (21) and (22) are convenient and quick for measuring the various diagnosability parameters discussed and obtained from the DBG in Section 2.

The set conflicts are determined by Entity 1 in the matrix. For identification of set conflict, check if there exists Entity 1 in one or more than one row of a column of the matrix. It is considered a set conflict. In addition, if there exists Entity 1 in one or more than one consecutive columns of the same row. It is also considered a set conflict. In Expression (21), set conflicts are equal to 5. In Expression (22) d_1 related to m_1 , d_2 related to m_1 and m_2 , d_3 related to m_2 , d_4 related to m_2 , m_3 and m_4 , d_5 related to m_3 and m_4 , d_6 related to m_4 , d_7 related to m_5 , and d_8 related to m_5 and d_8 related to d_8 relat

An algorithm has been developed on the basis of Eqs. (12) and (13) to obtain the set conflicts of a system design in C++ and using bubble-sort algorithm for sorting the matrix. This algorithm is more useful for sorting the matrix of higher orders when manually sorting becomes inconvenient. The MNCS is determined by the maximum number of elements in the consecutive columns of the matrix in the same row, that is, d_9 and d_{10} in the matrix Expression (22) and forms the set conflict related to m_6 . Hence, the value of MNS = 9 and MNCS = 2. The values are the same as obtained from the DBG in Section 2. The value of DEC is again equal to the 5 and 10, that is, equal to total number of parameters. The value of AMD is obtained from the Eq. (17) by the ratio of DEC to MNS.

3.1. Diagnosability evaluation

The procedure suggested above is used to evaluate diagnosability of the system at the design stage considering the PPs and POs and on the basis of DBG and DM. The procedure is illustrated by considering various hypothetical cases, Cases 1 to 5 representing various design alternatives of a system.

3.1.1. Case 1.

In this case, the DBG model shown in Figure 3 represents relationship between ten PPs and 15 POs. The Diagnosability Matrix is shown in Expression (23). The various diagnosability parameters for the case I are obtained from diagnosability matrix $\mathbf{D_{m3}}$, that is, Expression (23)

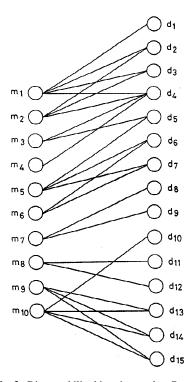


Fig. 3. Diagnosability bipartite graph—Case 1.

The various parameters are obtained on the basis of entity 1 in the matrix as explained in the previous section. These are as:

MNS = 9, MNCS = 3, DEC = 15, and AMD = 1.66.

The value of MNCS indicates that three parameters are to be further measured for the isolation of the faulty component. The value of AMD is more than one and this indicates poor diagnosability.

3.1.2. Case 2.

In this case, the number of relationship is increased between PPs, and POs to find the effect of functional sharing

of the components on diagnosability. This is shown in Figure 4 and the Diagnosability Matrix is represented by Expression (24)

The values of various parameters of diagnosability are:

MNS = 9,

MNCS = 4,

DEC = 15, and

AMD = 1.66.

The value of MNCS is increased by 4, although the value of AMD remains the same (1.66). This shows poor diagnosability, as compared to the Case 1, as four performance parameters are measured to isolate the faulty component. This is attributed to the increase in the functional sharing.

3.1.3. Case 3.

In this case, functional sharing is decreased as is indicated by lower number of edges between PPs and POs in Figure 5. The Diagnosability Matrix $(\mathbf{D_{m5}})$ of Figure 5 is represented by Expression (25)

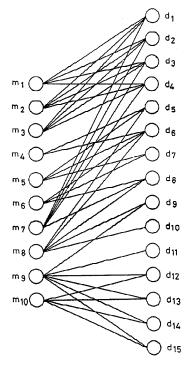


Fig. 4. Diagnosability bipartite graph—Case 2.

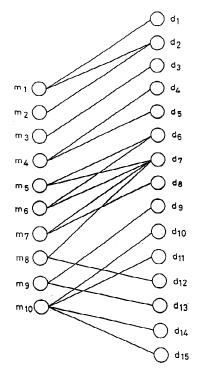


Fig. 5. Diagnosability bipartite graph—Case 3.

200 M.F. Wani and O.P. Gandhi

The various parameters of diagnosability obtained are:

MNS = 12, MNCS = 2, DEC = 15, and AMD = 1.25.

The value of MNCS and AMD is low, compared to the Cases 1 and 2. Thus, diagnosability is better. This shows diagnosability is improved by sustaining the functional independence.

3.1.4. Case 4.

In this case, the number of components is increased to 20 and the number of PPs is maintained, that is, equal to 10 to ascertain the influence of the increase in the number of components. This is represented in DBG in Figure 6. The Diagnosability Matrix is represented in Expression (26)

The values of various diagnosability parameters are:

MNS = 14, MNCS = 3, DEC = 20, and AMD = 1.42.

There is marginal increase in the value of MNCS and AMD as compared with Case 3, however, it should be noted that the number of components is greater. This means if the functional independence is sustained the diagnosability is better.

3.1.5. Case 5.

In this case the number of components is same due to functional packaging number of parameters that are to be measured after the performance parameter go out of design stage are low. The relationship between the PPs and POs is shown in Figure 7. The Diagnosability Matrix is given in Expression (27)

$$\mathbf{D_{m7}} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & d_7 & d_8 \\ m_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ m_5 & m_6 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ m_7 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$
 (27)

The value of various parameters obtained is as:

MNS = 8, MNCS = 1,

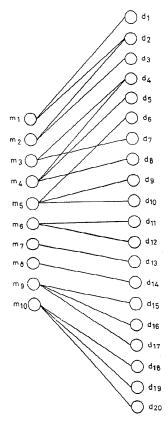


Fig. 6. Diagnosability bipartite graph—Case 4.

DEC = 8, and AMD = 1.

The values of these parameters indicate better diagnosability as the value of MNCS and AMD is equal to 1.

Design variants can now be compared on the basis of values of MNS, MNCS, DEC, and AMD individually to select the best design alternative from diagnosability point of view.

As discussed, in Section 2, the first priority is given to MNCS, that is, the variant, which possesses the minimum number of candidates in a set conflict, as having higher diagnosability as compared to others. The second priority is given to AMD, that is, the variant having the minimum value of AMD is preferred over others. The third priority goes to variant having minimum DEC, and the design having the highest value of MNS is given least priority. The value of diagnosability parameters MNS, MNCS, DEC, and AMD are shown in Table 1. It is inferred from the values of MNS, MNCS, DEC, and AMD that the variant 5 has the highest diagnosability and is given first priority, as the values of MNCS and AMD and DEC are low as compared to the other alternatives. The variant 3 is given second priority, as the values of MNCS and AMD are low as compared to remaining alternatives. The variants 4 and 1 have the same values of first priority parameter, that is, MNCS, however; the value of second priority parameter, that is, AMD is low for variant 4. Hence, the variant 4 is preferred over the variant 1.

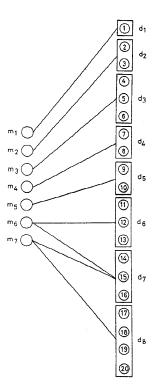


Fig. 7. Diagnosability bipartite graph—Case 5.

The last priority is given to variant 2 having highest value of MNCS.

The component with poor diagnosability can also be identified with the help of DBG and DM. If the connectivity between a node represents a physical object, that is, d_q in D and PPs is maximum, then it indicates poor diagnosability. This shows that the component will create complication during the diagnosability, as it is always considered a suspected component for PPs. This is also indicated by the maximum number of one entry in the column of the D_m . For example, in Figure 3 the node d_4 is connected to m_1 , m_2 , m_3 , and m_4 , and has maximum connectivity, which is indicated by the fourth column in the matrix Expression (23). This means component d_4 will always create complications during the fault isolation, as compared with the other components of the system.

Table 1. Comparison of design variants on the basis of diagnosability

System Diagnosability Parameters			
MNS	MNCS	DEC	AMD
9	3	15	1.66
9	4	15	1.66
12	2	15	1.25
14	3	20	1.42
8	1.0	8	1.0
	MNS 9 9 12 14	MNS MNCS 9 3 9 4 12 2 14 3	MNS MNCS DEC 9 3 15 9 4 15 12 2 15 14 3 20

4. STEPS FOR DIAGNOSABILITY EVALUATION

Our methodology is used for diagnosability evaluation and comparison of system design alternatives. The procedure is given as:

- 1. Consider all the system design concepts or alternatives.
- Identify for the first alternative, set of relationships between performance parameters (PPs) and components/assemblies, that is, POs, on the basis of relationship between functional requirements and components/assemblies.
- 3. Develop the DBG as per the procedure given in Section 2.
- 4. Develop the DM on the basis of DBG obtained in Step 3.
- 5. Obtain the values of the diagnosability parameters MNS, MNCS, DEC, and AMD (Refer to Section 3.1).
- 6. Repeat steps 2 to 5 for all other alternatives.
- 7. Compare the design alternatives on the basis of the values of parameters and priority to identify and select best design from diagnosability point of view. (Refer to Section 3.1 for details.)

5. EXAMPLES

The proposed methodology can be used for both new and existing designs for evaluation and improvement of diagnosability. Two examples are considered in this Section. Example 1 is a hydraulic system with two design alternatives, and is meant for illustrating the procedure. Example 2 has been selected from the literature (Paasch & Ruff, 1997) to validate the methodology on existing system.

5.1. Example-hydraulic system

A hydraulic system to lift and support heavy loads in a plant is considered. This system has to perform three main subfunctions:

- 1. to lift the load to desired level with a controlled lift velocity (L_n) ;
- 2. to hold the load at desired level and maintain that level with out sagging (L_s) ; and
- 3. return of load lifting arm to original position (R_o) .

These three subfunctions are directly observed by the maintenance person to understand the behavior of the system. These observations, therefore, constitute the PPs. Assume the PPs are observed to be either good (the functionality is present) or bad (not satisfied).

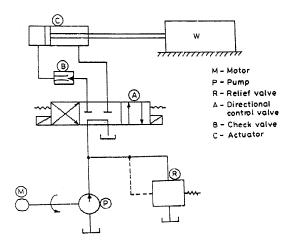


Fig. 8. Hydraulic system with flow control meter in check valve (FCMI).

Two design concepts of hydraulic system are shown in Figures 8 and 9, and form the two design alternatives. In Figure 8, the first alternative shows a flow control valve in meter in circuit to control the flow to the actuator, and is referred as FCMI. In the second alternative (Fig. 9), a pilot-operated check valve is incorporated, and is referred as POCV. For each system alternative, there are three performance parameters, L_v , L_s , and R_o . These are related to the various POs based on functional relationships. For the FCMI, the PPs are related to POs as follows:

- 1. If the motor (M) were to fail, all performance parameters $(L_v, L_s, \text{ and } R_o)$ would be affected.
- 2. If the pump (P) were to fail, all the performance parameters $(L_v, L_s, \text{ and } R_o)$ would be affected.
- 3. If the relief valve (R) were to fail, all the performance parameters L_v , L_s , and R_o would be affected.
- 4. If the actuator (C) were to fail, all the performance parameters L_v , L_s , and R_o would be affected.

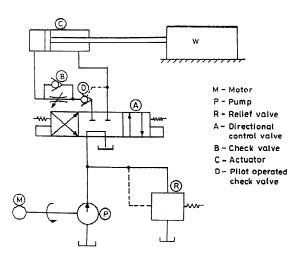


Fig. 9. Hydraulic system with pilot-operated check valve (POCV).

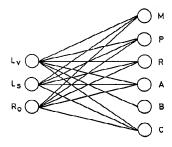


Fig. 10. Diagnosability bipartite graph (FCMI).

- 5. If the check valve (B) were to fail, only L_s and R_o would be affected.
- 6. If the control valve (A) were to fail, all the performance parameters L_v , L_s , and R_o would be affected.

Similarly, for the POCV, the PPs are related to POs as follows:

- 1. If the motor (M) were to fail, all performance parameters $(L_n, L_s, \text{ and } R_o)$ would be affected.
- 2. If the pump (P) were to fail, all the performance parameters $(L_v, L_s, \text{ and } R_o)$ would be affected.
- 3. If the relief valve (R) were to fail, all the performance parameters L_v , L_s , and R_o would be affected.
- 4. If the actuator (C) were to fail, all the performance parameters L_v , L_s , and R_o would be affected.
- 5. If the check valve (B) were to fail, only L_v would be affected.
- 6. If the POCV (D) were to fail, only L_s would be affected.
- 7. If the control valve (A) were to fail, only R_o would be affected.

These two system alternatives are represented in terms of DBGs in Figures 10 and 11.

These DBGs are represented by two DMs by expressions (28) and (29), and are as:

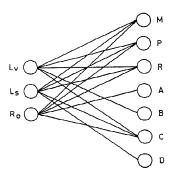


Fig. 11. Diagnosability bipartite graph (POCV).

Table 2. Comparison of design variants on the basis of diagnosability

	System Diagnosability Parameters			
Nomenclature	MNS	MNCS	DEC	AMD
FCMI	3	5	6	2
POCV	4	4	7	1.75
	FCMI	Nomenclature MNS FCMI 3	Nomenclature MNS MNCS FCMI 3 5	Nomenclature MNS MNCS DEC FCMI 3 5 6

$$\mathbf{D_{m9}} = \begin{bmatrix} M & P & R & A & B & C & D \\ L_v & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ R_6 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$
(29)

The various parameters of diagnosability are obtained from the DBGs and are shown in Table 2. For the two system alternatives, the Design Variant 2, that is, POCV has the lowest value of first priority MNCS, the design possesses the minimum number of candidates in the set conflict, and will not create complication during the isolation of fault. The value of second-priority parameter AMD is also lower than the first alternative. Hence, it possesses maximum diagnosability, although the value third-priority parameter DEC is slightly higher. The value of the lowest priority parameter, that is, MNS is higher for the second alternative. Therefore, the first alternative will create more confusion during the fault isolation. Hence, we conclude that the second alternative possesses better diagnosability.

In this example, diagnosability evaluation and comparison of two alternatives of hydraulic system design was carried out at conceptual design stage.

5.2. Example—Bleed air system

A bleed air control System shown in Figure 12 supplies compressed air supply to various compartments in an aircraft during flight operations. The bleed air control system performs three main functions, supply compressed air for air conditioning, engine starting system, and pneumatic components. Pneumatic control valves regulate air temperature and pressure to ensure that pressure is not lost through bleed air control system. The high-pressure shut-off valve (SOV) is used to control bleed airflow from the turbine into the ducts. A pressure check valve (CK) prevents airflow into the compressor. The pressure relief valve (PRV), which is before the precooler (PC), acts if the system becomes over pressurized. The fan air-modulating valve (MV) controls the rate of cooling air through PC. The pressure reducing and shutoff valve (PRSV) limits the air pressure supplied

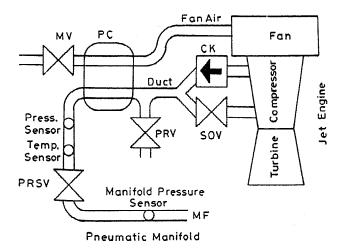


Fig. 12. Bleed air control system of an aircraft.

to pneumatic manifold (MF). The PRSV also provides overtemperature protection for the MF by reducing flow if the bleed air temperature is too high, and provides a checking function to prevent manifold pressure loss through the bleed air control system. The system has three sensors: one for temperature and two for pressure monitoring. There are also switches that indicate when the PRV is closed, and when SOV is open.

The performance parameters of the bleed air control system are temperature (T_1) and the pressure (P_2) measured at the MF and the pressure (P_1) measured before PRSV. In addition, PRV open, PRSV closed, and bleed high stage (BHS) directly indicate status of the system components.

The relations between PPs and POs are established and are shown in a DBG (Fig. 13), and its Diagnosability Matrix is written as:

$$\mathbf{D_{m10}} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 5 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 6 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$
(30)

It is evident that the PC and PRSV have maximum relationship with performance parameters, and as indicated by one entity in the third and fifth columns in Expression (30). These are likely to reduce diagnosability. It is possible to minimize this problem by shifting one or more of the functions performed by the PRSV to other POs already performing these functions, or to the physical object having the least or minimum functional relationship. One function, which the PRSV shares, maximum with other POs is control temperature. This can be shifted from PRSV to SOV. The high-pressure SOV can be used to protect the system

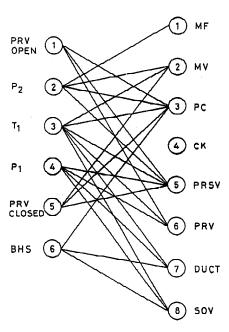


Fig. 13. Diagnosability bipartite graph of bleed air control system.

from overheating by restricting flow from high-pressure port. The second modification can be done by shifting the pressure-control function from PRSV to PRV. The third modification involves shifting of the pressure-control function from PRSV. For this modification, the pressure is regulated directly at high and low values by the CK. The DBG of the modified system is shown in Figure 14, and its Diagnostic Matrix is written as:

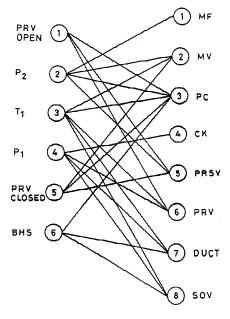


Fig. 14. Diagnosability bipartite graph of modified bleed air control system.

Table 3. Bleed air control system-diagnosability comparison

	System Diagnosability Parameters				
Design Variant	MNS	MNCS	DEC	AMD	
Existing design	6	1	8	1.3	
Modified design	7	1	8	1.1	

$$\mathbf{D_{m11}} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 5 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$
(31)

As are shown in Table 3, the value of MNCS and DEC are the same (one). However, the value of AMD is low and MNS is high for the modified case. Hence, the modified design has higher diagnosability.

This example shows that the methodology clearly identifies the component(s) having poor diagnosability in the system, which can be improved further through modifications in the system without adding any further function to the system. The results obtained also match published results. Therefore, the example demonstrates the proposed methodology. Moreover, this exercise also shows that the methodology can effectively be applied to an operating systems as well. The methodology is, however, limited to binary relations between PPs and POs.

6. SYSTEM FAULT DIAGNOSIS

A fault-localization methodology based on information captured from the DBG or the DM is suggested here. Set conflicts, (MNS) are obtained from Eq. (15). MNS contains maximum number of set conflicts. Each set conflict contains candidate(s), for example, d_1 is the candidate of first set conflict. Each set conflict is related to performance parameter(s), for example, d_1 is related to performance parameter m_1 . The set conflict and its relation with performance parameter(s) represent a rule, which can help in conflict resolution for fault diagnosis. The performance parameter m_1, m_2, \ldots, m_5 form antecedents (conditions), and the candidates d_1, d_2, \ldots, d_5 form consequents (actions) of the rules for the fault diagnosis. The rules of the fault diagnosis are derived from Eq. (15) and these are as:

IF: m_1 is observed abnormal

THEN: d_1 is faulty,

IF: m_2 is observed abnormal

THEN: d_2 is faulty,

IF: m_3 is observed abnormal

THEN: d_3 is faulty,

IF: m_4 is observed abnormal

THEN: d_4 is faulty and

IF: m_5 is observed abnormal

THEN: d_5 is faulty.

In this case, the diagnosability is forthright and each set conflict contains only one candidate. However, if the set conflict contains more than one component, then rule has two or more consequents. In the second hypothetical example, the set conflict (Section 2) $\langle d_9, d_{10} \rangle$ has more than two components and the rule for the fault diagnosis is derived as:

IF: m_6 is abnormal THEN: d_9 is faulty

ELSE: d_{10} is faulty.

A set of rules has been derived for the example of hydraulic system, shown in Figure 9. In this case, the MNS is equal to four.

Rule1 is obtained from first set conflict, that is, $\langle M, P, R, C \rangle$ and its relation with the performance parameters L_v, L_s , and R_o , where L_v, L_s , and R_o indicate lift-forward velocity, load-sagging, and return velocity of load, respectively.

Rule 1:

IF:

(lift-forward-velocity abnormal)

(load sagging)

(lift-return -velocity abnormal)

THEN: (motor faulty)

ELSE: (pump faulty)

ELSE: (relief-valve faulty)

ELSE: (actuator faulty).

Similarly, two to four are also derived from set conflicts and their relations with performance parameters.

Rule 2:

IF:

(lift-forward-velocity abnormal)

THEN: (check-valve faulty).

Rule 3:

IF:

(load sagging)

THEN: (pilot-operator-check- valve faulty).

Rule 4:

IF:

(load-return-velocity abnormal)

THEN: (control-valve faulty).

This shows that the set conflicts assist in developing the rules for system fault diagnosis, and also for fault resolution. This can help designer to develop rules for expert system for fault diagnosis at the conceptual design stage. This

can be improved further by incorporating the knowledge or experience gained during the service period of the system. This system fault-diagnosis methodology can help to combat the unpredicted failures at initial stage of operation, when the diagnosis experience is insufficient.

7. SUMMARY

Diagnosability of mechanical systems is modeled in terms of DBG considering relationship between performance parameters and physical objects. The DBG is represented by a Diagnosability Matrix. Quantitative values of proposed parameters MNS, MNCS, DEC, and AMD are obtained from DBG and DM. A design variant is compared on the basis of these parameters. This helps the designer select the best alternative of the design from diagnosability point of view for better maintainability during service period. Moreover, this will also ease of identification and localization of the faulty line replaceable unit in complex system design. The individual components or subsystems are also identified which are having poor diagnosability. The maximum number of set conflicts (MNS) also help a designer in deriving rules for system fault diagnosis.

Main utilization of suggested diagnostic evaluation approach is its application at conceptual-design stage.

REFERENCES

- Bellinger, D.Q. (1966). Failure analysis techniques and application. In *Handbook of Reliability Engineering and Management*, (Ireson, W.G. and Coombs, C.F., Eds.), pp. 13.1–13.41. McGraw Hill, New York.
- Bracewell, R.H., & Sharpe, J.E.E. (1996). Functional description used in computer support for qualitative scheme generation—"Scheme builder." Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 10, 333–345.
- Clark, G.E., & Paasch, P.K. (1996). Diagnostic evaluation modeling and diagnostic evaluation of mechanical system. *Journal of Mechanical De*sign, Transactions of the ASME 118(3), 425–431.
- Davis, R. (1984). Diagnostic reasoning based on structure and behavior. Artificial Intelligence 24, 347–410.
- Davis, R., Shrobe, H., Hamsher, W., Wieckert, K., Shirly, M., & Polit, S. (1982). Diagnosis based on structure and function. *Proc. National Conf.*, *Arti. Intell.*, 137–142.
- De Kleer, J. (1979). The origin and resolution of ambiguities in casual arguments. *Proc. Sixth Int. Joint Conf. on Artificial Intell.*, 197–203.
- De Kleer, J., & Williams, B.C. (1987). Diagnosing multiple faults. *Artificial Intelligence* 32, 87–130.
- Deo, N. (1974). Graph Theory. Prentice Hall, Englewood Cliffs, NJ.
- DOD. (1984). Military Standardization Handbook. MIL-HDBK-472, Maintainability Prediction. U.S. Department of Defense, Washington, DC.

- Forbus, K.D. (1984). A qualitative process theory. *Artificial Intelligence* 24, 85–168.
- Genesereth, M.R. (1984). The use of design descriptions in automated diagnosis. Artificial Intelligence 24, 411–436.
- IEC. (1994). Standard 706-5, Guide on maintainability of equipment parts: Section 4, Diagnostic testing, 3 ruede Varembe Geneva, Suisse.
- Iwasaki, Y., & Chandrasekaran, B. (1992). Design verification through function-and behavior-oriented representation: Bridging the gap between function and behavior. In *Artificial Intelligence in Design* '92, (Gero, J.S. Ed.), pp. 597–616. Kluwer, Dordrecht.
- Kowalski, R. (1988). Maintainability and reliability. In *Handbook of Reliability and Management* (Ireson, W.G., Clydre, F.C., Jr. and Moss, R.Y., Eds.), McGraw-Hill, New York.
- Larsson, J.E. (1994). Diagnostic reasoning strategies for means end models. Automatica 830(5), 29–93.
- Larsson, J.E. (1996). Diagnosis based on explicit means-end models. Artificial Intelligence 8, 29–93.
- Murphy, M.D., & Paasch, P.K. (1997). Reliability centered prediction techniques for diagnostic modeling and improvement. Research in Engineering Design 9(1), 35–45.
- Nakakuki, Y., Koseki, Y., & Tanaka, M. (1992). Adaptive model-based diagnostic mechanism using a hierarchical model scheme. *Proc. 10th National Conf. on Artificial Intelligence*, 554–559.
- Paasch, P.K., & Ruff, D.N. (1997). Evaluation of failure diagnosis in conceptual design of mechanical systems. *Journal of Mechanical Design*, *Transactions of the ASME*, Vol. 119(1), 57–67.
- Qian, L., & Gero, J.S. (1996). Function behavior structure and their role in analogy-based design. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 10(4), 289–312.
- Reiter, R. (1987). A theory of diagnosis from the first principle. *Artificial Intelligence*, 32, 57–95.
- Suh, N.M. (1990). The Principals of Design. Oxford University Press, New York.
- Tzafestas, S.G. (1989). System fault diagnosis using the knowledge-based methodology. In *Fault Diagnosis in Dynamic Systems*, (Palton, R., Frank, P. and Clark, R. Eds.), pp. 509–594. Prentice Hall, London.

M.F. Wani is now with R.E.C. Srinagar working as an Assistant Professor in the Mechanical Engineering Department. Previously, he was at the Industrial Tribology, Machine Dynamics, and Maintainance Engineering Center (ITMMEC), Indian Institute of Technology, New Delhi, India. His areas of interest are AI design, maintainability, and life cycle cost. His teaching and research experience spans over 10 years.

O.P. Gandhi is an Associate Professor at the ITMMEC, Indian Institute of Technology, New Delhi, India. He has teaching and research experience of over 25 years. His areas of interest inlude AI maintainance, reliability, and risk analysis.