

Article

An Industrial Fault Diagnostic System Based on a Cubic Dynamic Uncertain Causality Graph

Xusong Bu ^{1,†} , Hao Nie ^{2,†}, Zhan Zhang ² and Qin Zhang ^{1,2,*}

¹ Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China; buxusong@mail.tsinghua.edu.cn

² Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing 100084, China; nh17@mails.tsinghua.edu.cn (H.N.); zhangzhan190@tsinghua.edu.cn (Z.Z.)

* Correspondence: qin Zhang@mail.tsinghua.edu.cn

† These authors contributed equally to this work.

Abstract: This study presents an industrial fault diagnosis system based on the cubic dynamic uncertain causality graph (cubic DUCG) used to model and diagnose industrial systems without sufficient data for model training. The system is developed based on cloud native technology. It contains two main parts, the diagnostic knowledge base and the inference method. The knowledge base was built by domain experts modularly based on professional knowledge. It represented the causality between events in the target industrial system in a visual and graphical form. During the inference, the cubic DUCG algorithm could dynamically generate the cubic causal graph according to the real-time data and perform the logic and probability calculations based on the generated cubic DUCG models, visually displaying the dynamic causal evolution of faults. To verify the system's feasibility, we rebuild a fault-diagnosis model of the secondary circuit system of No. 1 at the Ningde nuclear power plant based on the new system. Twenty-four fault cases were used to test the diagnostic accuracy of the system, and all faults were correctly diagnosed. The results showed that it was feasible to use the cubic DUCG platform for fault diagnosis.

Keywords: industrial fault diagnosis; cubic DUCG; causal inference; expert knowledge



Citation: Bu, X.; Nie, H.; Zhang, Z.; Zhang, Q. An Industrial Fault Diagnostic System Based on a Cubic Dynamic Uncertain Causality Graph. *Sensors* **2022**, *22*, 4118. <https://doi.org/10.3390/s22114118>

Academic Editors: Erika Ottaviano, Jose Machado, Katarzyna Antosz, Dariusz Mazurkiewicz, Yi Ren, Pierluigi Rea, Rochdi El Abdi, Marina Ranga, Vijaya Kumar Manupati and Emilia Villani

Received: 26 March 2022

Accepted: 23 May 2022

Published: 28 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the increasing complexity of industrial systems, the systems' safety has attracted extensive attention. Suppose some minor faults in the system cannot be detected and eliminated in time. In that case, it may cause the failure and paralysis of the entire system, and even lead to substantial disastrous consequences [1–4]. Improving the safety and reliability of the system and preventing and eliminating the occurrence and development of faults that affect the regular operation of the system has become a crucial problem to be solved. Fault monitoring and diagnostic technology is an effective method to improve the safety and reliability of complex systems [5–7]. Early fault-diagnosis expert systems were mostly rule-based and case-based [8–10]. The systems' diagnostic rules and cases were constructed by experts based on experience. The advantage is that they are less dependent on data and are interpretable. When the industrial system is relatively simple, it is practical to use those methods. When the industrial system is complex, those methods are prone to knowledge conflict, repetition, and circulation [11]. In addition, the increase in knowledge reduces the reasoning efficiency of the system [12,13]. With the development of machine learning, some machine-learning algorithms are applied in fault-diagnosis systems, including SVM [14–16], ANN [17,18], and DNN [19–21]. Those methods use machine-learning theory to adaptively learn the diagnostic knowledge from the collected data, rather than using the experience and knowledge of engineers. The fault diagnostic system based on machine learning has high diagnostic efficiency. When the training data are sufficient, the diagnostic accuracy of the model is high. However, the models constructed

based on these methods are black-box models, which lack interpretability of the diagnostic results [22].

In some customized complex systems, such as nuclear power plants or spacecraft [23], these systems are usually unique. They have high complexity, safety, and reliability, and almost no failures. So, the data used to train the diagnostic model are insufficient [24]. The fault diagnosis of such systems requires high reliability and the entire fault diagnostic process to be interpretable, helping the operator understand the fault transmission process to take measures to eliminate the fault quickly. Obviously, in the fault diagnosis of such systems, the fault diagnostic method based on machine learning is not feasible. However, the construction of such systems condenses a large amount of expert knowledge, experience, and statistical data. Therefore, using domain expert knowledge, experience, and statistical data to construct a diagnostic model intuitively and realize an efficient and reliable diagnosis is a feasible method to solve such a diagnostic system [25].

The cubic DUCG is a probabilistic graphical model [26] developed from the dynamic uncertain causality graph theory [27]. It can intuitively describe the uncertain causality of events using probabilities and different graphical symbols, and has a solid ability to express the propagation of causal uncertainty, as well as the advantages of visualization, interpretability, and high computational efficiency. It has a complete mathematical foundation and theoretical system, and provides a concise expression and reasoning method of uncertain knowledge in the form of causal graphs. It can realize the causal inference problems of discrete, continuous, and fuzzy variables. At the same time, it can handle causal loops, allowing directed loops in the model [28–30]. These features enable the cubic DUCG to implement accurate inference of multiconnected causality efficiently. The cubic DUCG's reasoning knowledge base can be constructed by domain experts based on their knowledge, experience, and statistical data [31,32]. Therefore, it is suitable for scenarios in which causal inference is required but there are insufficient data to train the diagnostic model.

In this study, we built a time-series-based industrial fault diagnosis platform based on cubic DUCG theory. The following sections of this paper are arranged as follows: Section 2 introduces the cubic DUCG theory and describes the reasoning algorithm used in this diagnosis system; Section 3 presents the design and implementation of the diagnosis system; Section 4 provides the system verification; and Section 5 summarizes the study and suggests directions for future research.

2. The Cubic DUCG Modeling and Fault-Diagnosis Method

2.1. Cubic DUCG Modeling

The cubic DUCG theory was developed from DUCG theory. It used the DUCG model as the inference knowledge base. The DUCG model represents the causal relationship between events, and can be built graphically and visually by domain experts using their knowledge, experience, and statistical data. Figure 1 is an example of the DUCG; its parameters are shown in Appendix A. As shown in Figure 1, in the DUCG model, different variables are represented by different shapes with numbers as indexes. Various variables give the DUCG a complete causal expression ability and make the causal expression more accurate. When constructing a DUCG model, domain experts only need to determine the events to be expressed and the causal relationship between events. Then, the appropriate variables and relationships to represent them are selected according to the construction rules of the DUCG model. The variables and their physical meanings are shown in Table 1.

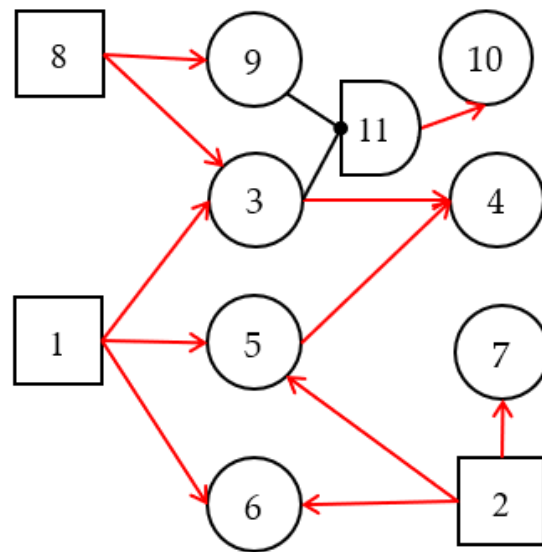


Figure 1. An example of the DUCG model. It describes the fault-propagation process through graphical symbols.

Table 1. The physical meanings of variables in the DUCG.

Type	Shape	Description
B_i		The B -type variable is the root-cause variable used to represent the root cause/fault that causes other variables to occur.
X_i		The X -type variable is the consequence or process variable used to represent the result caused by the root-cause variable, and can also be used as the cause of other variables.
G_i		The G -type variable is the logic-gate variable. It is used to describe the logical relation combination of parent variables.
D_i		The D -type variable is the unknown cause or default-cause variable. When the cause of a variable's occurrence is unknown, then the D -type variable is used to represent the root cause that causes it to occur.
F_{ij}		The F -type variable is the weighted functional-event variable. It is used to represent and quantify the causalities between parent variables and child variables.

The causal mechanism between variables of the DUCG is shown in Figure 2. The child event X_{nk} may be caused by one or more of its parent events V_{ij} ($V \in \{B, X, D, G\}$ -type variable). The parent events cause the child event through the weighted functional events $F_{nk;ij}$. $F_{nk;ij} = (r_{n;i}/r_n) A_{nk;ij}$, where $A_{nk;ij}$ is the virtual random functional event representing the causal mechanism that V_{ij} independently causes X_{nk} . $a_{nk;ij} = Pr \{A_{nk;ij}\}$ is the probability that V_{ij} causes X_{nk} to occur independently. The virtual random functional events between the

parent V_i and the child X_n are represented by the matrix $A_{n;i} = \begin{pmatrix} A_{n0;i0} & \cdots & A_{n0;iJ} \\ \vdots & A_{nk;ij} & \vdots \\ A_{nK;i0} & \cdots & A_{nK;iJ} \end{pmatrix}$.

The matrix allows incomplete expression; if there is no causality between V_{ij} and X_{nk} , then $A_{nk;ij}$ does not exist, and $A_{nk;ij}$ is replaced with the symbol “-” in the matrix. When constructing causalities, we only need to give the parameters of the states that we care about to reduce the difficulty of knowledge-base construction. $r_{n;i}/r_n$ ($r_n = \sum_i r_{n;i}$) is the weight parameter used to normalize the effect of parent variables on child variables. The relationship between parents is the logical weighted exclusive OR of the DUCG weighted set theory [33]. The weighted functional event and the logical weighted exclusive OR enable the DUCG to freely modify (add, delete, or update) the influence of parent variables on child variables.

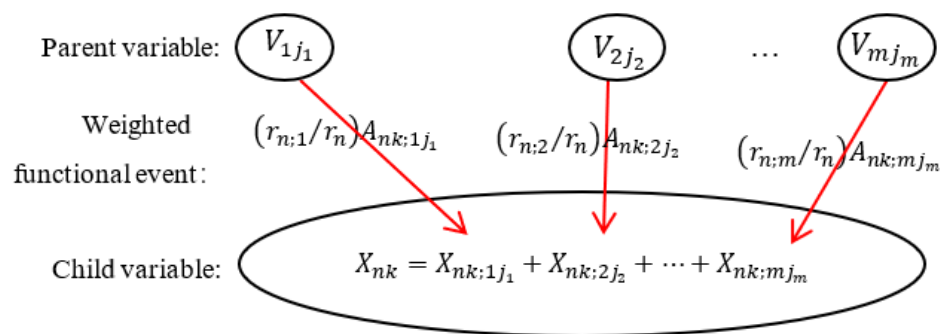


Figure 2. The causal mechanism between variables of the DUCG, where X_{nk} is the child variable and V_{ij} is the parent variable.

In order to infer to which root causes the child event is related, the child event performs logic expression expansion along the opposite direction of the causal chain. After expression expansion, the child event is expressed by its parent events, and the expansion process can be executed recursively until the parent events are the B -type variables. The B -type variables are the root causes of other variables and the objects of inference calculation. The expansion expression is shown in Equation (1) [27]:

$$X_{nk} = \sum_i \sum_{j_i} X_{nk:ij_i} = \sum_i \sum_{j_i} F_{nk:ij_i} V_{ij_i} = \sum_i \sum_{j_i} (r_{n:i}/r_n) A_{nk:ij_i} V_{ij_i} \tag{1}$$

For simplicity, Equation (1) can be briefly written as Equation (2):

$$X_{nk} = \sum_i \sum_j F_{nk:ij} V_{ij} = \sum_i \sum_j (r_{n:i}/r_n) A_{nk:ij} V_{ij} \tag{2}$$

Logic expression expansion is an essential step of cubic DUCG reasoning. It can recursively expand the observation evidence E ($E = E'E''$, $E' = \{X_{ij}, j \neq 0\}$ is the collection of abnormal evidence, and $E'' = \{X_{i0}\}$ is the collection of normal evidence) to the root faults. The result of the logic-expanded expression is used to calculate the conditional probability of each hypothesis under the current evidence.

2.2. The Inference Method of the Cubic DUCG

The reasoning process of the cubic DUCG reflects the temporal correlations among events, and is capable of representing the sequential causality interactions in fault-spreading processes. The reasoning step of the cubic DUCG is described as follows: (1) DUCG decomposition. The original DUCG is decomposed into several sub-DUCGs; each sub-DUCG contains one root event B_i , remarked as $DUCG(B_i)$; (2) Obtain $Slice_DG(B_i, t_m)$. The $Slice_DG(B_i, t_m)$ is the intraslice causality graph at time t_m . According to the evidence $E(t_m)$ at t_m , the $DUCG(B_i)$ s are decomposed into several $Slice_DG(B_i, t_m)$ s based on the cubic DUCG simplification rules [26,34]. Each $Slice_DG(B_i, t_m)$ contains one root fault, and it describes the causality between the root fault B_i and the evidence $E(t_m)$ at t_m ; (3) $Cubic_DG(B_i, t_m)$ generation. The $Cubic_DG(t_m)$ is the generated cubic causality graph at t_m . It is generated by merging the $Cubic_DG(B_i, t_{m-1})$ generated at t_{m-1} and $Slice_DG(B_i, t_m)$ obtained at t_m . $Cubic_DG(t_m)$ describes the propagation process of the evidence related to root fault B_i from time t_1 to time t_m . In particular, when $t_m=t_1$, $Cubic_DG(B_i, t_1) = Slice_DG(B_i, t_1)$; (4) Select a valid $Cubic_DG(B_i, t_m)$ s. The reasoning of the cubic DUCG is based on the unary assumption that there can only be one root fault at the same time. Therefore, if the $Cubic_DG(B_i, t_m)$ cannot explain all the evidence $E(t_m)$ at t_m , it is regarded as an invalid $Cubic_DG(B_i, t_m)$ and discarded. Only the $Cubic_DG(B_i, t_m)$ that can explain all abnormal evidence $E(t_m)$ is regarded as a valid $Cubic_DG(B_i, t_m)$; (5) Obtain the hypothesis space $S_H(t_m)$. The root fault B_i in a valid $Cubic_DG(B_i, t_m)$ is regarded as the hypothesis H_{kj} , and forms the hypothesis space, $S_H(t_m) = \{H_{kj}\}$; (6) The logical expansion of expression. Logical expansion

of $E(t_m)$ and $H_{kj}E(t_m)$ are based on each respective $Cubic_DG(B_i, t_m)$, until the expression contains only the form of the sum of variable products of $\{B-, D-, A-, r-\}$ -type variables. Those expressions are used for the probability calculation; (7) Probability calculation. Calculate the conditional probability $\Pr\{H_{kj}(t_m)\}$ of each H_{kj} in $S_H(t_m)$ to evaluate which root fault is more likely to occur. The probability is calculated using Equation (3):

$$\Pr\{H_{kj}(t_m)\} = h_{kj}^s(t_m) = \zeta_i(B_i, t_m) \frac{\Pr\{H_{kj}E(t_m)\}}{\Pr\{E(t_m)\}} \quad (3)$$

$$\zeta_i(B_i, t_m) = \frac{\Pr\{E(t_m)\}}{\sum_i \Pr\{E(t_m)\}} = \frac{\zeta(B_i, t_m)}{\sum_i \zeta(B_i, t_m)} \quad (4)$$

In Equation (3), $E(t_m)$ is the abnormal evidence at time t_m , and $E(t_m) = \prod_i X_{i,j_i}$. $\Pr\{E(t_m)\} = \zeta(B_i, t_m)$ is the joint probability of $E(t_m)$ on $Cubic_DG(B_i, t_m)$. $\Pr\{H_{kj}E(t_m)\}$ is the joint probability of $H_{kj}E(t_m)$ on $Cubic_DG(B_i, t_m)$. $\zeta_i(B_i, t_m)$ is the weight factor of $E(t_m)$ on different $Cubic_DG(B_i, t_m)$ s calculated using Equation (4). If there is only one $Cubic_DG(B_i, t_m)$, then $\zeta_i(B_i, t_m) = 1$. The sorted results are the reasoning result of the cubic DUCG at t_m . Those $Cubic_DG(B_i, t_m)$ s are used as the graphical explanations for inference results; (8) Repeat the procedures until the root fault is confirmed or no more abnormal evidence is received. The pseudo algorithm description of the calculation process is shown in Algorithm 1.

Algorithm 1: The Inference of the Cubic DUCG. (Note: The pseudo algorithm for inference process of the cubic DUCG.)

Input: the original DUCG, the evidence $E(t_m) = \prod_i X_{i,j_i}$ at time t_m .

Steps:

1. Decompose the original DUCG to get the list of $DUCG(B_i)$;
2. **for** $DUCG(B_i)$ **in** list of $DUCG(B_i)$
3. $Slice_DG(B_i, t_m) \leftarrow GenerateValidSlice_DG(DUCG(B_i), E(t_m))$;
4. $Cubic_DG(B_i, t_m) \leftarrow GenerateCubic_DG(Slice_DG(B_i, t_m), Cubic_DG(B_i, t_{m-1}))$;
5. $S_H(t_m) \leftarrow GetHypothesis(Cubic_DG(B_i, t_m))$;
6. **end**
7. **for** $Cubic_DG(B_i, t_m)$ **in** list of $Cubic_DG(B_i, t_m)$
8. Logical expand $E(t_m)$ and $H_{kj}E(t_m)$ based on each $Cubic_DG(B_i, t_m)$;
9. **end**
10. **for** $H_{kj}(t_m)$ **in** $S_H(t_m)$
11. $\zeta_i(B_i, t_m) = \frac{\Pr\{E(t_m)\}}{\sum_i \Pr\{E(t_m)\}} = \frac{\zeta(B_i, t_m)}{\sum_i \zeta(B_i, t_m)}$
12. $\Pr\{H_{kj}(t_m)\} = h_{kj}^s(t_m) = \zeta_i(B_i, t_m) \frac{\Pr\{H_{kj}E(t_m)\}}{\Pr\{E(t_m)\}}$
13. **end**
14. Sort the list of $h_{kj}^s(t_m)$ in descending order.

Output: The sorted list of $h_{kj}^s(t_m)$ and $Cubic_DG(B_i, t_m)$ at t_m .

An example is used to exemplify the dynamic inference process of the cubic DUCG in a continuous time series. The original DUCG knowledge base is shown in Figure 1. It contains three root faults: B_1 , B_2 , and B_8 . Other variables include some intermediate processes or results caused by root faults. The inference process contains three moments: t_1 , t_2 , and t_3 .

At time t_1 , suppose the received evidence is $E(t_1) = X_{3,0}X_{5,1}X_{6,0}$.

Step 1.1. DUCG decomposition. By decomposing the original DUCG shown in Figure 1, we obtain three sub-DUCGs: $DUCG(B_1)$, $DUCG(B_2)$, and $DUCG(B_8)$. Each $DUCG(B_i)$ describes the relationships between the root fault B_i and its related variables; the results are shown in Figure 3. Usually, the green circle stands for the normal evidence,

the circle with other colors stands for the abnormal evidence, and the circle without colors stands for the state of the variable is known.

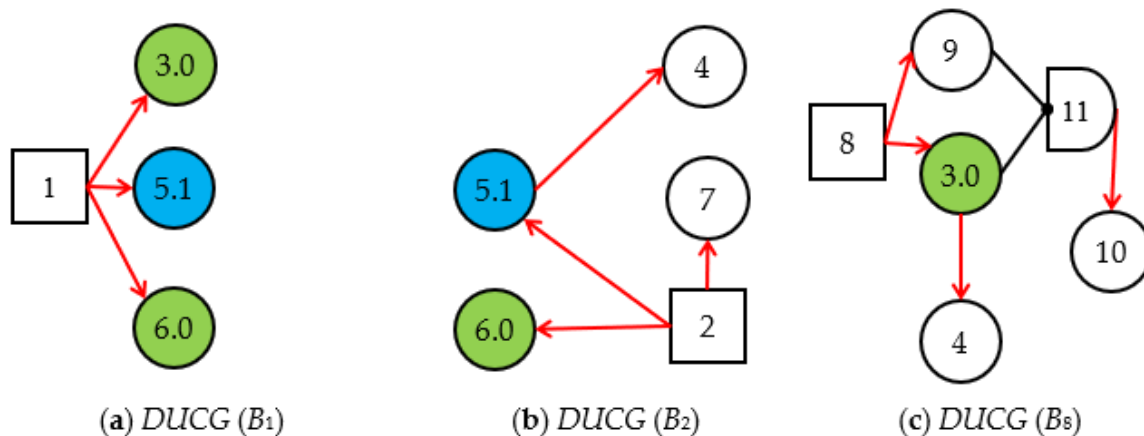


Figure 3. The original DUCG decomposed into three $DUCG(B_i)$ s.

Step 1.2. Obtain the valid $Slice_DG(B_i, t_1)$ s and generate $Cubic_DG(B_i, t_1)$ s at t_1 . By simplifying $DUCG(B_i)$ s according to the simplification rules of the cubic DUCG, we obtain the three $Slice_DG(B_i, t_1)$ s shown in Figure 4; The $Slice_DG(B_i, t_1)$ shows the relationship between the root fault B_i and the current evidence $E(t_1)$ at t_1 . The $Slice_DG(B_1, t_1)$ and the $Slice_DG(B_2, t_1)$ can explain the abnormal evidence, so they are the valid $Slice_DG(B_i, t_1)$ s. In $Slice_DG(B_8, t_1)$, the abnormal evidence $X_{5,1}$ cannot be explained by the root fault B_8 , so it is the invalid $Slice_DG(B_i, t_1)$ and is deleted.

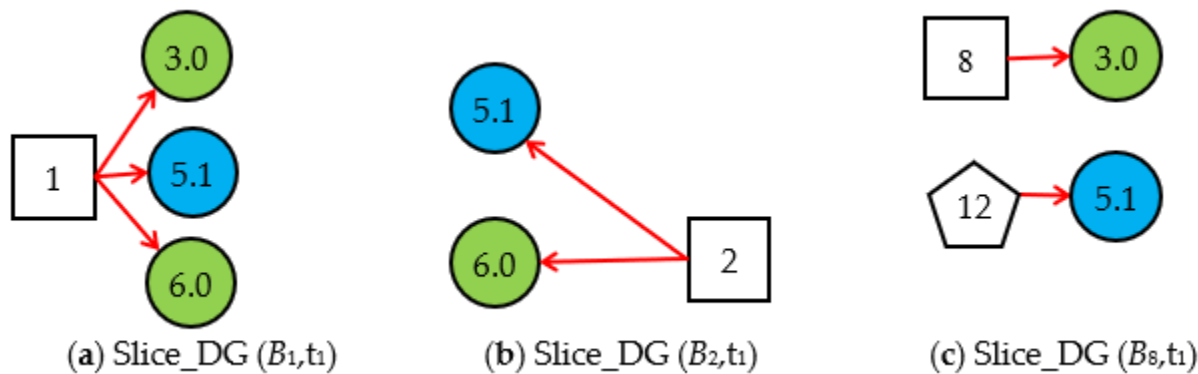


Figure 4. Three $Slice_DG(B_i, t_1)$ s based on the evidence $E(t_1) = X_{3,0}X_{5,1}X_{6,0}$.

Step 1.3. $Cubic_DB(B_i, t_1)$ generation. At time t_1 , the valid $Slice_DG(B_i, t_1)$ s are used as $Cubic_DG(B_i, t_1)$ s, as shown in Figure 5. From the two $Cubic_DG(B_i, t_1)$ s, we obtain the hypothesis spaces $S_H(t_1) = \{H_{1,1}, H_{2,1}, H_{2,2}\} = \{B_{1,1}, B_{2,1}, B_{2,2}\}$.

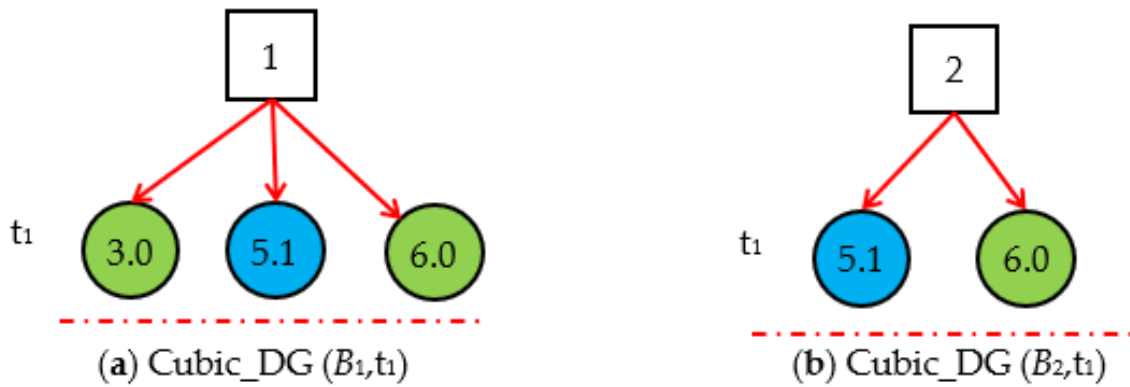


Figure 5. The Cubic_DG (B_i, t_1)s generated by the valid Slice_DG (B_i, t_1)s at t_1 .

Step 1.4. Logical expression expansion and probability calculation. The hypothesis $H_{1,1}$ is included in $Cubic_DG(B_1, t_1)$, and the hypotheses $H_{2,1}$ and $H_{2,2}$ are included in $Cubic_DG(B_2, t_1)$. In order to calculate the $\Pr\{H_{1,1}(t_1)\}$, we should use expression expansion of $E(t_1)$ and $H_{1,1}E(t_1)$ based on $Cubic_DG(B_1, t_1)$, and then calculate $\Pr\{E(t_1)\}$, $\Pr\{H_{1,1}E(t_1)\}$. Equations (5) and (6) are the logic expansion of $E(t_1)$ and $H_{1,1}E(t_1)$ according to Equation (2). Equations (7) and (8) are the numerical calculation process of the expanded expressions; the parameters used in calculation can be seen in Appendix A.

$$E(t_1) = X_{3,0}X_{5,1}X_{6,0} = F_{3,0;1,1}B_{1,1}F_{5,1;1,1}B_{1,1}F_{6,0;1,1}B_{1,1} = F_{3,0;1,1}F_{5,1;1,1}F_{6,0;1,1}B_{1,1} \quad (5)$$

$$H_{1,1}E(t_1) = B_{1,1}X_{3,0}X_{5,1}X_{6,0} = F_{3,0;1,1}F_{5,1;1,1}F_{6,0;1,1}B_{1,1} \quad (6)$$

$$\begin{aligned} \zeta(B_1, t_1) &= \Pr\{E(t_1)\} = \Pr\{X_{3,0}X_{5,1}X_{6,0}\} = \Pr\{F_{3,0;1,1}F_{5,1;1,1}F_{6,0;1,1}B_{1,1}\} \\ &= f_{3,0;1,1}f_{5,1;1,1}f_{6,0;1,1}b_{1,1} \\ &= (r_1/r_3)a_{3,0;1,1}(r_1/r_5)a_{5,1;1,1}(r_1/r_6)a_{6,0;1,1}b_{1,1} \\ &= (1/1) \times (1 - 0.5) \times (1/1) \times 0.1 \times (1/1) \times (1 - 0.4) \times 0.2 \\ &= 0.006 \end{aligned} \quad (7)$$

$$\Pr\{H_{1,1}E(t_1)\} = \Pr\{B_{1,1}X_{3,0}X_{5,1}X_{6,0}\} = \Pr\{F_{3,0;1,1}F_{5,1;1,1}F_{6,0;1,1}B_{1,1}\} = 0.006 \quad (8)$$

Similarly, for $H_{2,1}$ and $H_{2,2}$ in $Cubic_DG(B_2, t_1)$, the results of expression expansion for $E(t_1)$, $H_{2,1}E(t_1)$, and $H_{2,2}E(t_1)$, as well as the results of $\Pr\{E(t_1)\}$, $\Pr\{H_{2,1}E(t_1)\}$, $\Pr\{H_{2,2}E(t_1)\}$, are shown in Equations (9)–(14).

$$E(t_1) = X_{5,1}X_{6,0} = F_{5,1;2,1}F_{6,0;2,1}B_{2,1} + F_{5,1;2,2}F_{6,0;2,2}B_{2,2} \quad (9)$$

$$H_{2,1}E(t_1) = B_{2,1}X_{5,1}X_{6,0} = F_{5,1;2,1}F_{6,0;2,1}B_{2,1} \quad (10)$$

$$H_{2,2}E(t_1) = B_{2,2}X_{5,1}X_{6,0} = F_{5,1;2,2}F_{6,0;2,2}B_{2,2} \quad (11)$$

$$\zeta(B_2, t_1) = \Pr\{E(t_1)\} = \Pr\{F_{5,1;2,1}F_{6,0;2,1}B_{2,1} + F_{5,1;2,2}F_{6,0;2,2}B_{2,2}\} = 0.04 \quad (12)$$

$$\Pr\{H_{2,1}E(t_1)\} = \Pr\{F_{5,1;2,1}F_{6,0;2,1}B_{2,1}\} = 0.025 \quad (13)$$

$$\Pr\{H_{2,2}E(t_1)\} = \Pr\{F_{5,1;2,2}F_{6,0;2,2}B_{2,2}\} = 0.015 \quad (14)$$

According to Equation (3), we obtain the posterior probability of $H_{1,1}$, $H_{2,1}$, and $H_{2,2}$ shown in Equations (15)–(17):

$$\Pr\{H_{1,1}(t_1)\} = h_{1,1}^s(t_1) = \zeta_1(B_1, t_1) \frac{\Pr\{H_{1,1}E(t_1)\}}{\Pr\{E(t_1)\}} = \frac{\zeta(B_1, t_1)}{\zeta(B_1, t_1) + \zeta(B_2, t_1)} \times \frac{\Pr\{H_{1,1}E(t_1)\}}{\Pr\{E(t_1)\}} = 0.13 \quad (15)$$

$$\Pr\{H_{2,1}(t_1)\} = h_{2,1}^s(t_1) = \zeta_1(B_2, t_1) \frac{\Pr\{H_{2,1}E(t_1)\}}{\Pr\{E(t_1)\}} = 0.543 \quad (16)$$

$$\Pr\{H_{2,2}(t_1)\} = h_{2,2}^s(t_1) = \zeta_1(B_2, t_1) \frac{\Pr\{H_{2,2}E(t_1)\}}{\Pr\{E(t_1)\}} = 0.326 \tag{17}$$

In the sorted results for $h_{2,1}^s(t_1) > h_{2,2}^s(t_1) > h_{1,1}^s(t_1)$, we can see that the hypotheses $H_{1,1}$, $H_{2,1}$, and $H_{2,2}$ may cause the occurrence of the abnormal evidence $X_{5,1}$ at time t_1 , and $H_{2,1}$ is the most likely. The root fault B_8 is excluded because it could not explain abnormal evidence.

At time t_2 , suppose the received abnormal evidence is $X_{6,1}$, and combined with the evidence at t_1 , the total evidence at t_2 is $E(t_2) = X_{3,0} X_{5,1} X_{6,1}$. The reasoning calculation process at t_2 is as follows.

Step 2.1. By simplifying $DUCG(B_i)$ s based on the evidence $E(t_2)$, we obtain the $Slice_DG(B_1, t_2)$ and $Slice_DG(B_2, t_2)$ shown in Figure 6.



Figure 6. $Slice_DG(B_i, t_2)$ s at t_2 .

Step 2.2. Generate $Cubic_DG(B_i, t_2)$ s. The $Cubic_DG(B_i, t_2)$ s at t_2 are generated by synthesizing the $Cubic_DG(B_i, t_1)$ s at t_1 and the $Slice_DG(B_i, t_2)$ s at t_2 ; we then obtain the $Cubic_DG(B_i, t_2)$ s shown in Figure 7. From $Cubic_DG(B_i, t_2)$ s, we can see that X_3 and X_5 did not change, but the state of X_6 changed from normal to abnormal.

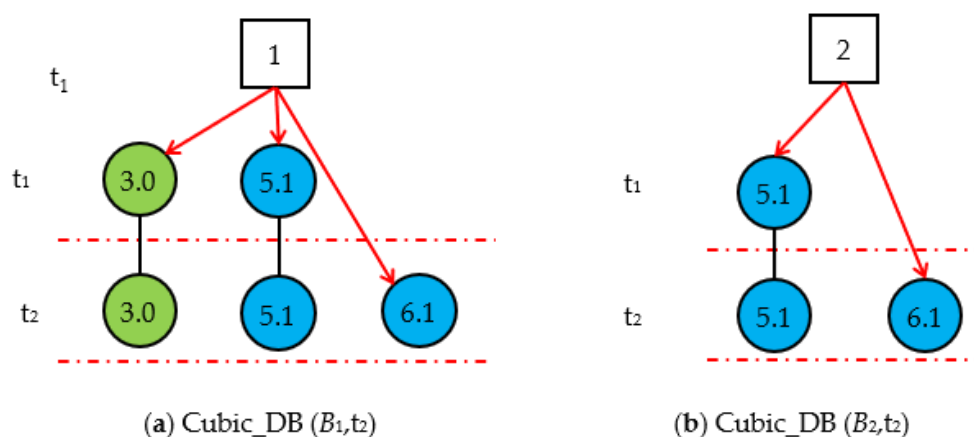


Figure 7. $Cubic_DG(B_i, t_2)$ s at t_2 .

Step 2.3. At time t_2 , due to the emergence of the abnormal evidence $X_{6,1}$, $Cubic_DB(B_i, t_2)$ s are regenerated, so the posterior probability of each hypothesis needs to be recalculated under the evidence $E(t_2)$ based on $Cubic_DB(B_i, t_2)$ s. For $H_{1,1}$ in $Cubic_DB(B_1, t_2)$:

$$E(t_2) = X_{3,0} X_{5,1} X_{6,1} = F_{3,0;1,1} F_{5,1;1,1} F_{6,1;1,1} B_{1,1} \tag{18}$$

$$\zeta(B_1, t_2) = \Pr\{E(t_2)\} = \Pr\{F_{3,0;1,1} F_{5,1;1,1} F_{6,1;1,1} B_{1,1}\} = 0.004 \tag{19}$$

$$H_{1,1}E(t_2) = B_{1,1} X_{3,0} X_{5,1} X_{6,1} = F_{3,0;1,1} F_{5,1;1,1} F_{6,1;1,1} B_{1,1} \tag{20}$$

$$\Pr\{H_{1,1}E(t_2)\} = \Pr\{B_{1,1}X_{3,0}X_{5,1}X_{6,1}\} = 0.004 \tag{21}$$

For $H_{2,1}$ and $H_{2,2}$ in *Cubic_DB* (B_2, t_2):

$$E(t_2) = X_{5,1}X_{6,1} = F_{5,1;2,1}F_{6,1;2,1}B_{2,1} + F_{5,1;2,2}F_{6,1;2,2}B_{2,2} \tag{22}$$

$$\zeta(B_2, t_2) = \Pr\{E(t_2)\} = \Pr\{F_{5,1;2,1}F_{6,1;2,1}B_{2,1} + F_{5,1;2,2}F_{6,1;2,2}B_{2,2}\} = 0.16 \tag{23}$$

$$H_{2,1}E(t_2) = B_{2,1}X_{5,1}X_{6,1} = F_{5,1;2,1}F_{6,1;2,1}B_{2,1} \tag{24}$$

$$\Pr\{H_{2,1}E(t_2)\} = 0.025 \tag{25}$$

$$\Pr\{H_{2,2}E(t_2)\} = 0.135 \tag{26}$$

Calculate the posterior probability of each hypothesis according to Equation (3), and we get:

$$\Pr\{H_{1,1}(t_2)\} = h_{1,1}^s(t_2) = \zeta_1(B_1, t_2) \frac{\Pr\{H_{1,1}E(t_2)\}}{\Pr\{E(t_2)\}} = 0.024 \tag{27}$$

$$\Pr\{H_{2,1}(t_2)\} = h_{2,1}^s(t_2) = \zeta_1(B_2, t_2) \frac{\Pr\{H_{2,1}E(t_2)\}}{\Pr\{E(t_2)\}} = 0.152 \tag{28}$$

$$\Pr\{H_{2,2}(t_2)\} = h_{2,2}^s(t_2) = \zeta_2(B_2, t_2) \frac{\Pr\{H_{2,2}E(t_2)\}}{\Pr\{E(t_2)\}} = 0.823 \tag{29}$$

The sorted result is $h_{2,2}^s(t_2) > h_{2,1}^s(t_2) > h_{1,1}^s(t_2)$. With the emergence of new evidence at time t_2 , the ranking probability of each hypothesis changes. From the ranking results, it can be seen that $H_{2,2}$ is most likely to cause the abnormal evidence, and the probability of $H_{1,1}$ and $H_{2,1}$ decreases.

At time t_3 , suppose the received new abnormal evidence is $E = X_{4,1} X_{7,1}$. Then, we get $E(t_3) = X_{3,0}X_{5,1}X_{6,1}X_{4,1}X_{7,1}$.

Step 3.1. Simplify the DUCG (B_i)s under $E(t_3)$, then we obtain the two *Slice_DG* (B_i, t_3)s shown in Figure 8. In *Slice_DG* (B_1, t_3), the abnormal variable $X_{7,1}$ is isolated evidence. $H_{1,1}$ cannot explain the occurrence of $X_{7,1}$, so *Slice_DG* (B_1, t_3) is regarded as an invalidated *Slice_DG* (B_i, t_3) and deleted. Only *Slice_DG* (B_2, t_3) is valid.

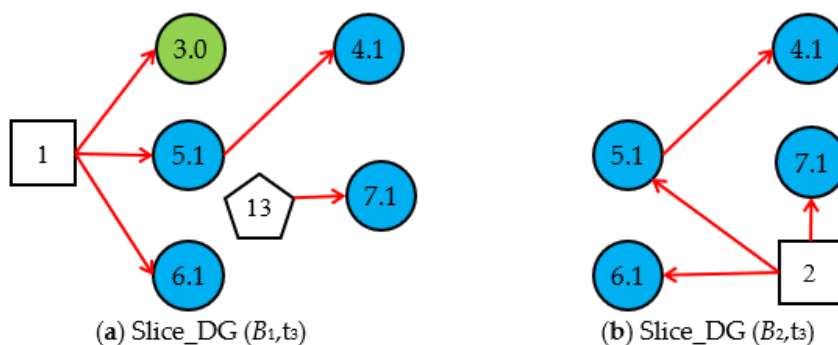


Figure 8. *Slice_DG* (B_i, t_3) at t_3 .

Step 3.2. According to *Cubic_DG* (B_2, t_2) at t_2 and *Slice_DG* (B_2, t_3), generate the *Cubic_DG* (B_2, t_3) at t_3 , as shown in Figure 9.

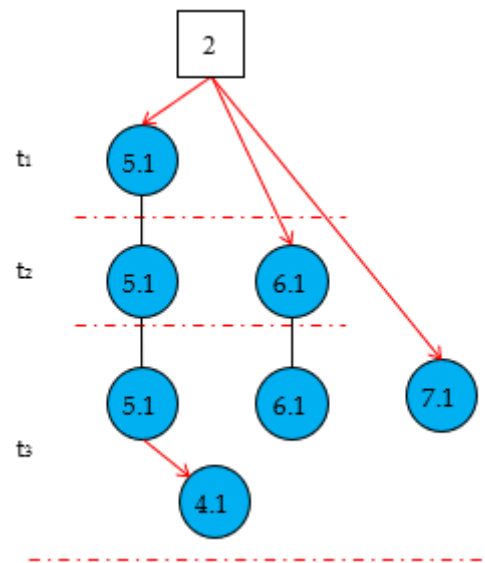


Figure 9. Cubic_DB (B_{2,t_3}) at t_3 .

Step 3.3. Logical expand $E(t_3)$, $H_{2,1} E(t_3)$, and $H_{2,2} E(t_3)$ based on $Cubic_DB(B_{2,t_3})$ and calculate the joint probability of $\Pr\{E(t_3)\}$, $\Pr\{H_{2,1} E(t_3)\}$, $\Pr\{H_{2,2} E(t_3)\}$ shown in Equations (30)–(35):

$$E(t_3) = F_{4,1;5,1}F_{5,1;2,1}F_{6,1;2,1}F_{7,1;2,1}B_{2,1} + F_{4,1;5,1}F_{5,1;2,2}F_{6,1;2,2}F_{7,1;2,2}B_{2,2} \quad (30)$$

$$\Pr\{E(t_3)\} = \Pr\{F_{4,1;5,1}F_{5,1;2,1}F_{6,1;2,1}F_{7,1;2,1}B_{2,1} + F_{4,1;5,1}F_{5,1;2,2}F_{6,1;2,2}F_{7,1;2,2}B_{2,2}\} = 0.08085 \quad (31)$$

$$H_{2,1}E(t_3) = X_{4,1}X_{5,1}X_{6,1}X_{7,1}B_{1,1} = F_{4,1;5,1}F_{5,1;2,1}F_{6,1;2,1}F_{7,1;2,1}B_{2,1} \quad (32)$$

$$\Pr\{H_{2,1}E(t_3)\} = \Pr\{F_{4,1;5,1}F_{5,1;2,1}F_{6,1;2,1}F_{7,1;2,1}B_{2,1}\} = 0.00525 \quad (33)$$

$$H_{2,2}E(t_3) = X_{4,1}X_{5,1}X_{6,1}X_{7,1}B_{2,2} = F_{4,1;5,1}F_{5,1;2,2}F_{6,1;2,2}F_{7,1;2,2}B_{2,2} \quad (34)$$

$$\Pr\{H_{2,2}E(t_3)\} = \Pr\{F_{4,1;5,1}F_{5,1;2,2}F_{6,1;2,2}F_{7,1;2,2}B_{2,2}\} = 0.0756 \quad (35)$$

Because there is only one $Slice_DG$ at t_3 , $\xi_2(B_2, t_3) = 1$, and according to Equation (3), we get the inference results of $H_{2,1}$ and $H_{2,2}$ shown in Equations (36) and (37):

$$\Pr\{H_{2,1}(t_3)\} = \frac{\Pr\{H_{2,1}E(t_3)\}}{\Pr\{E(t_3)\}} = 0.065 \quad (36)$$

$$\Pr\{H_{2,2}(t_3)\} = \frac{\Pr\{H_{2,2}E(t_3)\}}{\Pr\{E(t_3)\}} = 0.935 \quad (37)$$

The sorted result is $h_{2,2}^s(t_1) > h_{2,1}^s(t_1)$. Since $H_{1,1}$ cannot explain the occurrence of $X_{7,1}$ at time t_3 , it is considered an invalid hypothesis and excluded. Both $H_{2,2}$ and $H_{2,1}$ can explain all abnormal evidence. The probability of $H_{2,2}$ causing the abnormal evidence at t_3 is 0.935, while the probability of $H_{2,1}$ causing abnormal evidence is 0.065. According to the results, we can preliminarily judge that the abnormal evidence may be caused by $H_{2,2}$.

This was the reasoning process of the cubic DUCG, and the inference was based on a time series. The algorithm reconstructed the current cubic DUCG based on the evidence received at the current moment and the cubic DUCG at the last moment, showing the causal propagation process based on the time series. The DUCG simplification could simplify the complex original DUCG into a set of simple $Slice_DG(B_i, t_m)$ according to the evidence $E(t_m)$ at t_m . $Slice_DG(B_i, t_m)$ described the relationship between the evidence $E(t_m)$ and hypothesis B_i . The DUCG simplification could remove the impossible causalities and irrelevant variables on the condition of the evidence $E(t_m)$. Meanwhile, the computation scale was reduced exponentially without losing accuracy. The $Cubic_DG(B_i, t_m)$ reflected

how evidence changed over time slices, and was used as the graphical explanation of the hypothesis B_i and to enhance the interpretability of the inference results. The logical expansion of expression then expressed the logical relationship between the evidence and hypothesis, which is the premise of probability calculation. The reasoning calculation calculated the conditional probability of each hypothesis under the current evidence according to the causal effect between variables.

3. System Design

According to the reasoning mode of cubic DUCG and the characteristics of industrial diagnostic systems, the cubic-DUCG-based industrial fault diagnostic system was divided into four parts: the communication module, real-time monitoring and diagnosis module, inference engine, and knowledge-editing tool. The summary of each functional module is shown in Figure 10.

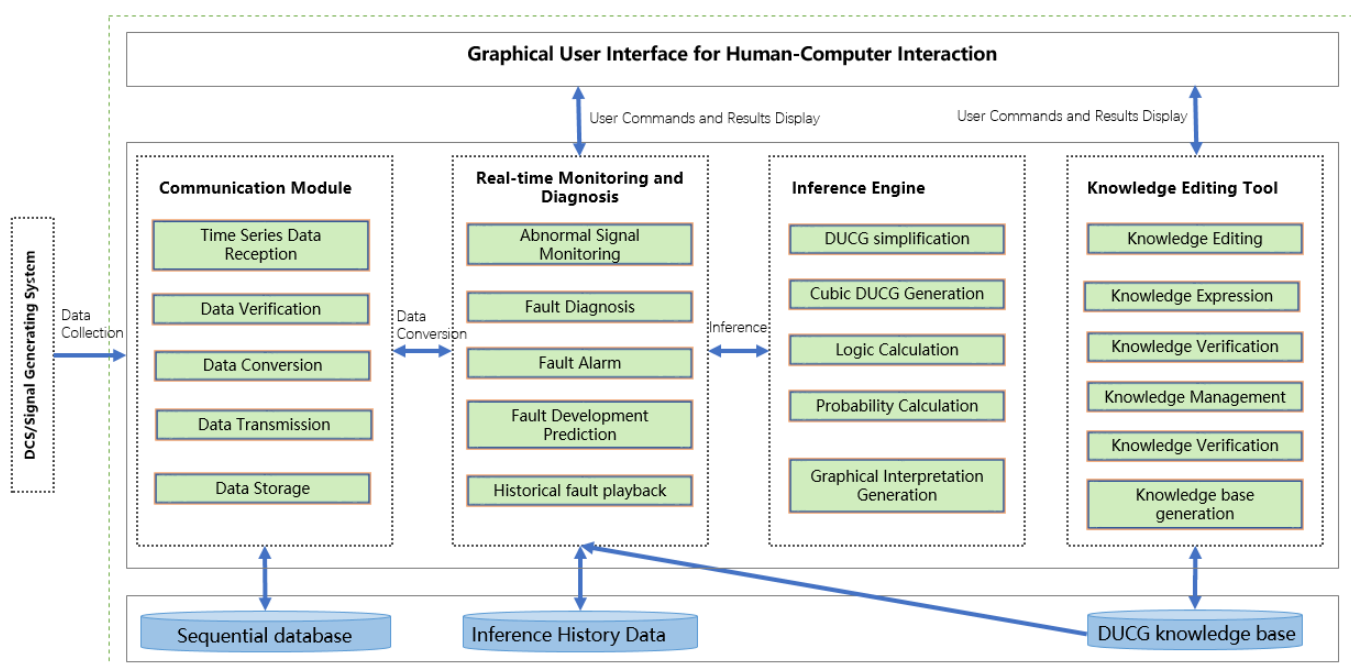


Figure 10. System block diagram of the cubic-DUCG-based industrial fault diagnostic system.

The knowledge-editing tool was used by domain experts to design the DUCG knowledge base. The DUCG knowledge base can be built in a modular way, and a whole DUCG can be divided into several sub-DUCGs. Generally, each sub-DUCG contains one fault, and represents causal relations between the fault and its related monitoring signals. This modular method of knowledge-base construction could reduce the difficulty in constructing a large complex knowledge base. Figure 11 shows a complete DUCG knowledge base. It was used for the fault diagnosis of the secondary circuit of No. 1 at the Ningde nuclear power plant. This DUCG knowledge base is reconstructed from the knowledge base in paper [35]. It contains 24 B -type variables that represent 24 different root faults in the secondary circuit of the nuclear power water reactor; the detail of the faults are shown in Appendix B. 141 X -type variables were used to describe the intermediate process or results arising from a root fault; a total of 1192 F -type variables (the direct red line) were used to describe the causal relations among variables.

The communication module was used for signal processing. The communication module received the monitoring data from the industrial system. It transformed the data to conform to the data format requirements of DUCG according to the mapping relationship between the measure points and variables, then transmitted the data to the real-time monitoring and diagnosis module.

The inference engine was the core module of the fault diagnostic system. It could generate cubic DUCG and engage in continuous causal reasoning based on the abnormal evidence. Its diagnostic results were presented in probabilistic form, and the generated cubic DUCG was used to explain the results. The inference engine was an independent service. Its data resulted from the real-time monitoring module, and its inference results were sent back to the real-time monitoring module for user decision making.

The real-time monitoring and diagnosis module was the control and human–computer interaction center. In this module, users could choose the DUCG knowledge base based on monitoring requirements, and then the instruction was sent to the communication module to receive and process signals associated with the current DUCG knowledge base. The monitoring module displayed and monitored signals in real-time.

The four functional modules of the system completed the functions of building a knowledge base, receiving and processing data, diagnosing faults, and displaying results. Through this system, users could translate knowledge and experience into diagnostic models. Furthermore, the model was used for real-time fault diagnosis. The system was a web application, the web client of the system was implemented with jquery+html, and the server–client was implemented using Java; the framework adopted by the system was the spring boot. The system is developed based on cloud native technology. Compared with the DUCG system based on traditional web technology [35]. The system has good scalability and can dynamically increase computing power according to task requirements.

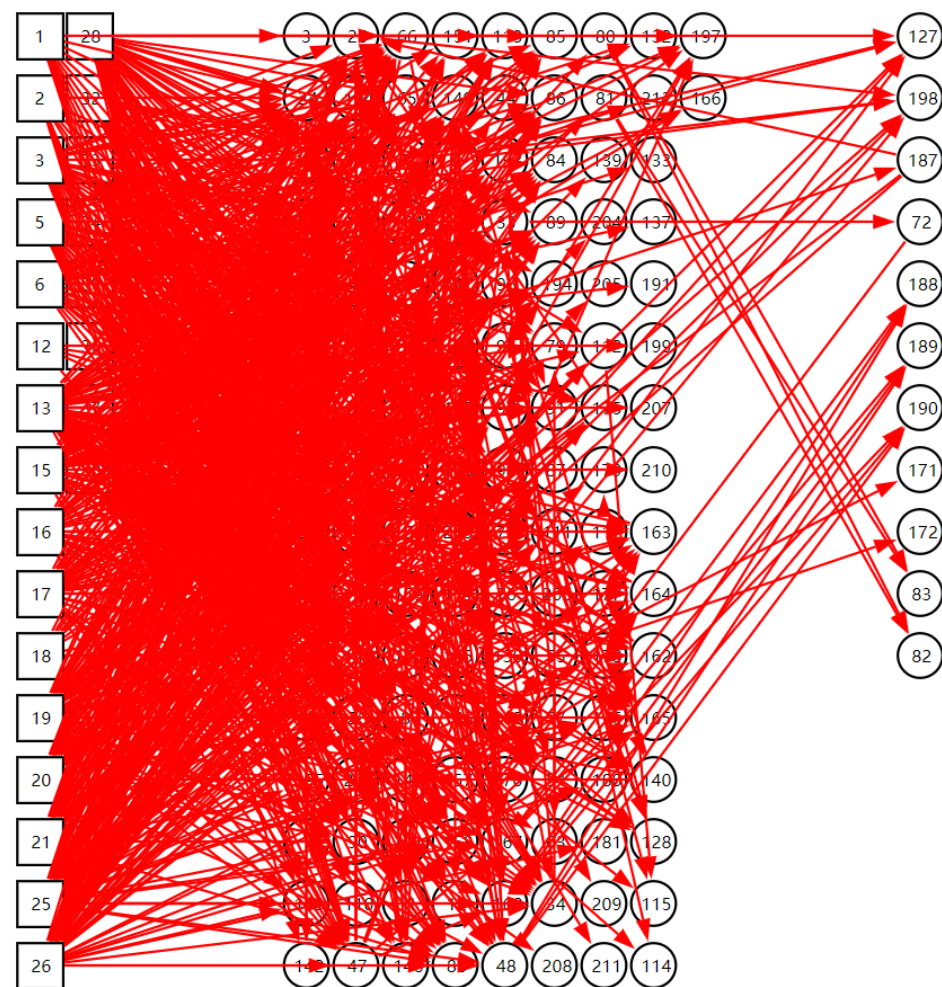


Figure 11. The DUCG knowledge base for the secondary circuit of No. 1 at the Ningde nuclear power plant.

4. Experiment

In order to validate the feasibility and diagnostic accuracy of the system, an experiment was done based on the secondary circuit of No. 1 at the Ningde nuclear power plant. The DUCG knowledge base is shown in Figure 11. The system was deployed on the computer cluster. The inference engine was deployed on one machine of the cluster with an AMD Ryzen 7 5700G CPU at 4.45 GHz, an 8-core processor, and 128 GB of RAM. The test data were collected from the simulator of the secondary circuit of No. 1 at the Ningde nuclear power plant. A total of 24 fault cases were used to test the system. Each fault case contained several time slices, and each time slice contained 141 signal data. The data types of the signals included switching value, continuous data, and discrete data. These signal data corresponded to 141 variables in the model one by one. During the system test, the simulator sent a group of signal data to the system every other second, which was recorded as a time slice. An example of condensate extraction pump failure was used to demonstrate the diagnostic process of the system.

The fault “condensate extraction pump fault (CEX001PO)” was inserted at the 13th second after the simulator operated stably, and the opening of the pump CEX001PO gradually decreased; then, when the communication module received the real-time data at the 14th second, one of the variables was in its abnormal state (the intake pressure of ABP401RE was low (ABP004MP), $X_{71,2}$). Because this was the first time the system received the abnormal evidence, this time was marked as t_1 and the system started the inference. The inference results and graphical explanation at t_1 are shown in Figures 12 and 13.

ID	Fault	State	Probability
36	Condenser vacuum pump (CVI101PO)	failure	61.07%
1	Condensate extraction pump status (CEX001PO)	closed	35.87%
2	Condensate extraction pump status (CEX002PO)	closed	3.05%
26	Water supply valve of electric main water supply system (APA113VL)	fail closed	<0.001%
3	Feedwater flow control valve status (ARE031VL)	all-open	<0.001%
5	Feedwater flow control valve status (ARE032VL)	all-open	<0.001%
28	Condensate valve (CEX108VL)	fail closed	<0.001%
6	Feedwater flow control valve status (ARE033VL)	all-open	<0.001%
32	Turbine bypass valve (GCT115VV)	fail opened	<0.001%
33	Status of A-way steam pipeline	rupture	<0.001%
35	Steam turbine regulating valve (GRE001VV)	fail opened	<0.001%
25	Pump (MFPA) failure (APA102PO)	fault	<0.001%
38	SG feedwater status	lose	<0.001%
15	Low pressure heater pipe (ABP401RE)	rupture	<0.001%
16	Feedwater heater bypass valve status (ABP011VL)	Unexpected open	<0.001%
19	Feedwater heater bypass valve (AHP009VL)	unexpected open	<0.001%

Figure 12. The inference results at t_1 according to the abnormal evidence $E = X_{71,2}$.

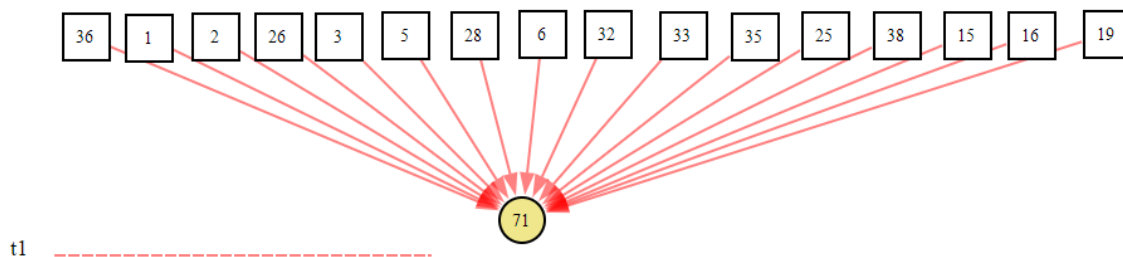


Figure 13. The graphical interpretation of the inference hypothesis and abnormal evidence at t_1 .

Figure 13 shows that because multiple fault sources could cause $X_{71,2}$ to occur, the fault source cannot be diagnosed at t_1 , but the scope of the fault and the probability of each fault could be preliminarily inferred. Sixteen faults could cause $X_{71,2}$ to occur, and the top three faults in the result list were more likely to cause $X_{71,2}$ to occur.

At the 15th second, the system did not receive the new abnormal evidence, so it did not perform the reasoning calculation. At the 16th second, the communication module received the new abnormal signal (condensate extraction pump (CEX003PO) failure, $X_{195,1}$). This time was marked as t_2 . The inference results and graphical explanation are shown in Figures 14 and 15. Comparing the inference results in Figures 12 and 14, we can see that hypothesis space was further reduced. The hypotheses in the first inference that could not explain the new evidence were excluded. Only the hypotheses that could explain all abnormal evidence were valid. According to the results, we could infer that the abnormal signals were possibly caused by the condensate extraction pump status (CEX001PO) ($B_{1,1}$) or the condensate extraction pump status (CEX002PO) ($B_{2,1}$), and $B_{1,1}$ was more likely.

Real Time Fault

ID	Fault	State	Probability
1	Condensate extraction pump status (CEX001PO)	closed	54.02%
2	Condensate extraction pump status (CEX002PO)	closed	45.98%
25	Pump (MFPA) failure (APA102PO)	fault	<0.001%

Figure 14. The inference results at t_2 according to the abnormal evidence $E = X_{71,2} X_{195,1}$.

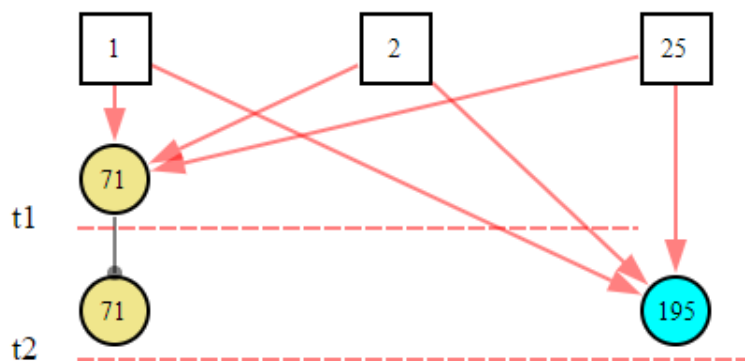


Figure 15. The graphical interpretation of the inference hypothesis and abnormal evidence at t_2 .

At the 17th second, the communication system received more abnormal signals. This time was marked as time t_3 . The third inference results and graphical interpretation are shown in Figures 16 and 17. Because only $B_{1,1}$ could explain all of the known abnormal evidence, according to the current evidence, $B_{1,1}$ was diagnosed.

Real Time Fault

ID	Fault	State	Probability
1	Condensate extraction pump status (CEX001PO)	closed	99.99%

Figure 16. The third inference results when more abnormal evidence was received.

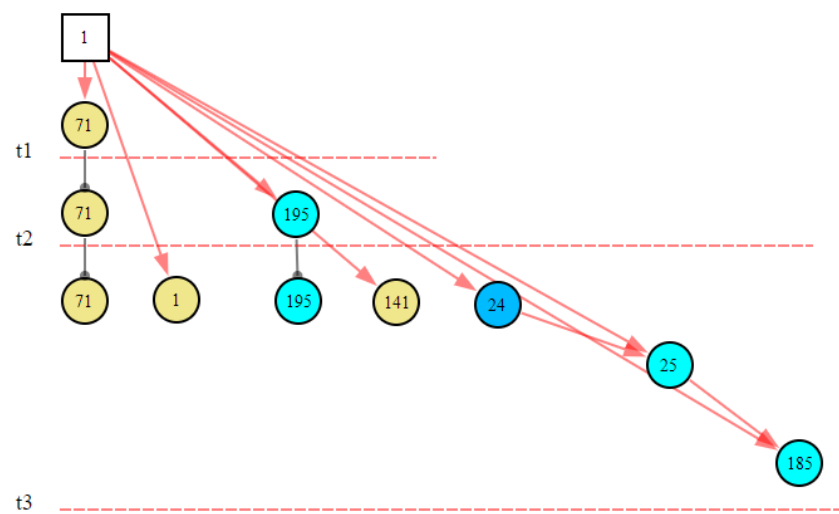


Figure 17. The graphical interpretation of the inference hypothesis and abnormal evidence at t_3 .

This example showed the inference process of the cubic DUCG. It performed inference calculations based on the time series. This fault diagnosis included three moments, and the moment that first showed the abnormal evidence was marked as t_1 . At t_1 , there was less abnormal evidence, so the specific diagnosis result could not be determined, but the range of possible failures could be roughly determined. At t_2 , new abnormal evidence was added to the diagnosis, which further narrowed the scope of the fault diagnosis. At t_3 , new evidence increased, and the fault was uniquely determined. In the next few moments, new evidence continued to be received. However, since there was only one diagnostic fault left and the fault could explain all abnormal evidence, the diagnosis result did not change, and the diagnosis was completed. At each moment, the system dynamically generated a new cubic DUCG for the diagnosis combined with the cubic DUCG obtained from the diagnosis at the last moment and the new evidence, giving the reasoning diagnosis results. The graphical interpretations could demonstrate the development and evolution process of the fault with time. It was convenient for the operator to understand the development of the fault for troubleshooting.

Table 2 shows the test results of the 24 fault cases. The “Fault” is the fault code in the DUCG knowledge base. “Rank First” indicates the first moment when the fault was ranked first in the diagnostic results. “Confirmed Diagnosis” indicates the moment when the fault was confirmed. “Time Consumption” indicates the total reasoning time for diagnosing the fault. “Average Time” represents the average time of each diagnosis. We can see that the 24 fault cases were all correctly diagnosed from the results. This proved that the fault-diagnosis model of the secondary circuit system constructed in this study was accurate. It also proved that using the cubic DUCG to construct a complex diagnostic

model was feasible. Among the 24 fault cases, 17 of them could be diagnosed at time t_1 . The remaining seven faults needed multiple time slices to be diagnosed, but they could all be sorted to the first place in the list of diagnostic faults within four diagnostic time slices, which proved that the system could quickly locate the root faults in a short time. In terms of inference calculation time, although the inference calculation time increased with an increase in the diagnosis time slice, the average time (from the start of the diagnosis to the fault being confirmed) for each diagnosis was within 400 ms. Compared with the DUCG-based diagnostic system proposed by Zhao [35], the diagnostic efficiency of the system in this study is more efficient. This showed that the system could complete a real-time and efficient reasoning diagnosis and meet the real-time task requirements of industrial diagnosis.

Table 2. The diagnostic results of the 24 faults of the secondary circuit of No. 1 at Ningde nuclear power plant.

Fault	Rank First	Confirmed Diagnosis	Time Consumption	Average Time
$B_{1,1}$	t_2	t_3	145 ms	48 ms
$B_{2,1}$	t_1	t_1	19 ms	19 ms
$B_{3,2}$	t_1	t_1	18 ms	18 ms
$B_{5,2}$	t_1	t_1	19 ms	19 ms
$B_{6,2}$	t_1	t_1	18 ms	18 ms
$B_{13,1}$	t_1	t_1	313 ms	313 ms
$B_{15,1}$	t_1	t_1	26 ms	26 ms
$B_{16,1}$	t_3	t_8	437 ms	55 ms
$B_{17,1}$	t_1	t_1	21 ms	21 ms
$B_{18,1}$	t_1	t_1	20 ms	20 ms
$B_{19,1}$	t_3	t_8	433 ms	55 ms
$B_{20,1}$	t_1	t_1	274 ms	274 ms
$B_{21,1}$	t_1	t_1	276 ms	276 ms
$B_{25,1}$	t_1	t_1	81 ms	81 ms
$B_{26,1}$	t_2	t_2	188 ms	94 ms
$B_{28,1}$	t_1	t_1	23 ms	23 ms
$B_{32,1}$	t_1	t_1	142 ms	142 ms
$B_{33,1}$	t_1	t_1	230 ms	230 ms
$B_{34,1}$	t_4	t_4	496 ms	124 ms
$B_{35,1}$	t_1	t_{12}	1874 ms	157 ms
$B_{35,2}$	t_1	t_1	31 ms	31 ms
$B_{36,1}$	t_1	t_1	18 ms	18 ms
$B_{37,1}$	t_1	t_1	21 ms	21 ms
$B_{38,1}$	t_2	t_2	398 ms	199 ms

5. Conclusions

With the rapid development of science and technology in industrial systems, such as space systems and nuclear power plants, those systems have shown the characteristics of increasingly complex structures, increasing scales, and increasingly rich functions, as well as gradual improvements in the levels of integration and automation. Once an accident occurs in a complex industrial system, it may cause casualties, damage to equipment and facilities, damage to the ecological environment, and loss of economic property. These systems have high security, and the fault data used for model training is insufficient. In addition, these systems are usually customized according to different task requirements, so the fault data between different systems lacks universality. Therefore, it is not feasible to use big data to train models for fault diagnosis. However, experts accumulate a great deal of experience and knowledge when designing and constructing such systems. At the same time, the fault data of other similar systems can also provide a reference for the fault analysis of the system. Therefore, some diagnostic methods based on expert knowledge are more suitable for the fault diagnosis of these systems. Traditional fault-diagnosis methods, such as the rule-based method, fault tree, and event tree, have the disadvantages of a low

reasoning efficiency, an inability to deal with the logical cycle or uncertain causalities, and difficulty in managing the growing knowledge.

This study proposed a fault-diagnosis method for a unique industrial system based on the cubic DUCG. The model was built based on expert knowledge, experience, and statistical data, and it described the causal mechanism of faults and abnormal signals. The system could carry out a continuous fault diagnosis according to the time sequence, display the results in the form of probability, and graphically represent the propagation of faults over time. The modular construction method of the cubic DUCG knowledge base reduced the modeling difficulty for large and complex knowledge bases and facilitated knowledge management and maintenance. The cubic DUCG could express logical cycles and uncertain causal relationships to express expert knowledge accurately. The reasoning process of the cubic DUCG included model simplification, logical calculation, and probability calculation. This inference method reduced the computational complexity of inference without losing the accuracy of the results. It solved the problem of the high computational complexity of large and complex knowledge base reasoning. These characteristics of the cubic DUCG made it more suitable for industrial system modeling and fault diagnosis based on expert knowledge.

In order to verify the feasibility and effectiveness of the system, we cooperated with nuclear experts to build the fault diagnostic model of the secondary circuit system of No. 1 at the Ningde nuclear power plant. The model's 24 root faults represented the operation status of 24 leading components of the secondary circuit system, such as the steam turbines and electric generators. A total of 141 variables represented the abnormal signals that root faults may have caused. The data types of variables included switching value, discrete type, and continuous type. Experts determined the values of the causal strength between variables based on experience or statistical data. The variables in the model could be mapped one-to-one with the detection points in the secondary circuit system. They could reflect the operation state of the secondary circuit system. The secondary system simulator generated 24 groups of fault cases to test the model. Each case tested 1 root fault, and 24 faults were correctly diagnosed. The test results showed that using the cubic DUCG for fault diagnosis in unique industrial systems was feasible. At the same time, it should be pointed out that the diagnostic accuracy of the diagnosis system based on expert knowledge depended on the accuracy of the model. Therefore, multiexpert joint modeling and third-party auditing of the model is one of the methods to ensure the model's accuracy. Since the current verification data were only provided by the nuclear power plant simulator, the verification model was also limited to 24 faults in the secondary circuit system of the nuclear power plant. Therefore, the verification of the system with only one application scenario was not complete and systematic. In a following work, we will continue to extend the model to test it using more data. In addition, we will verify the performance of the system in more application scenarios.

Author Contributions: Conceptualization, X.B., H.N., Z.Z. and Q.Z.; methodology, X.B., H.N. and Q.Z.; software, X.B., H.N. and Z.Z.; validation, X.B. and H.N.; formal analysis, Z.Z. and Q.Z.; investigation, X.B., H.N. and Z.Z.; resources, Z.Z. and Q.Z.; writing—original draft preparation, X.B.; writing—review and editing, X.B. and Q.Z.; visualization, X.B. and H.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank all authors of previous papers for approving the use of their published research results in this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. The Parameters of the Example DUCG Knowledge Base

$$\begin{aligned}
 b_1 &= (- \ 0.2)^T, \ b_2 = (- \ 0.1 \ 0.3)^T, \ b_8 = (- \ 0.3)^T, \ a_{3;1} = \begin{pmatrix} - & - \\ - & 0.5 \end{pmatrix}, \\
 a_{4;4} &= \begin{pmatrix} - & - \\ - & 0.2 \end{pmatrix}, \ a_{5;1} = \begin{pmatrix} - & - \\ - & 0.1 \end{pmatrix}, \ a_{4;5} = \begin{pmatrix} - & - \\ - & 0.7 \end{pmatrix}, \ a_{6;1} = \begin{pmatrix} - & - \\ - & 0.4 \end{pmatrix}, \\
 a_{6;2} &= \begin{pmatrix} - & - & - \\ - & 0.5 & 0.9 \end{pmatrix}, \ a_{5;2} = \begin{pmatrix} - & - & - \\ - & 0.5 & 0.5 \end{pmatrix}, \ a_{7;2} = \begin{pmatrix} - & - & - \\ - & 0.3 & 0.8 \end{pmatrix}, \\
 a_{3;8} &= \begin{pmatrix} - & - \\ - & 0.1 \end{pmatrix}, \ a_{9;8} = \begin{pmatrix} - & - \\ - & 0.9 \end{pmatrix}, \ LGS_{11} = \begin{Bmatrix} 0 & \text{Remnant} \\ 1 & X_{3,1}X_{9,1} \end{Bmatrix}, \\
 a_{10;11} &= \begin{pmatrix} - & - \\ - & 0.9 \end{pmatrix}, \ r_i = 1.
 \end{aligned}
 \tag{A1}$$

Appendix B. Description of 24 Faults in the Knowledge Base, Including Fault Name, State Description, and Its Number in DUCG

Table A1. The root faults in the cubic DUCG of the secondary circuit of No. 1 at Ningde nuclear power plant.

ID	Fault Description	State	State Description
B ₁	Condensate extraction pump status (CEX001PO)	0	Working order
		1	Closed
B ₂	Condensate extraction pump status (CEX002PO)	0	Working order
		1	Closed
B ₃	Feedwater flow control valve status (ARE031VL)	0	Normal
		1	All closed
		2	All opened
B ₅	Feedwater flow control valve status (ARE032VL)	0	Normal
		1	All closed
		2	All opened
B ₆	Feedwater flow control valve status (ARE033VL)	0	Normal
		1	All closed
		2	All opened
B ₁₂	Water supply pipeline of route A	0	Normal
		1	Leaked
B ₁₃	Turbine operation status	0	Working order
		1	Tripping state
B ₁₅	Low pressure heater pipe (ABP401RE)	0	Normal
		1	Rupture
B ₁₆	Feedwater heater bypass valve status (ABP011VL)	0	Normal
		1	Accidental opening
B ₁₇	Main steam header status	0	Normal
		1	Leaked
B ₁₈	Steam generator heat transfer tube	0	Normal
		1	Leaked
B ₁₉	Feedwater heater bypass valve (AHP009VL)	0	Normal
		1	Accidental opening
B ₂₀	Water supply pipeline of route B	0	Normal
		1	Leaked
B ₂₁	Water supply pipeline of route C	0	Normal
		1	Leaked
B ₂₅	Pump (MFPA) failure (APA102PO)	0	Normal
		1	Fault condition
B ₂₆	Water supply valve of electric main water supply system (APA113VL)	0	Normal
		1	Accidental shutdown

Table A1. Cont.

ID	Fault Description	State	State Description
B ₂₈	Water supply valve of electric main water supply system (APA113VL)	0	Normal
		1	Accidental shutdown
B ₃₂	Turbine bypass valve (GCT115VV)	0	Normal
		1	Accidental opening
B ₃₃	Status of A-way steam pipeline	0	Normal
		1	Rupture
B ₃₄	Turbine bypass valve (GCT131VV)	0	Normal
		1	Accidental opening
B ₃₅	Steam turbine regulating valve (GRE001VV)	0	Normal
		1	Accidental all closed
B ₃₆	Condenser vacuum pump (CVI101PO)	2	Accidental all opened
		0	Normal
B ₃₇	Status of three main steam isolation valves (VVP001/002/003VV)	1	Accidental failure
		0	Normal
B ₃₈	SG feedwater status	1	All closed
		0	Normal
		1	Loss
		2	Moderate

References

- Fernandes, M.; Corchado, J.M.; Marreiros, G. Machine learning techniques applied to mechanical fault diagnosis and fault prognosis in the context of real industrial manufacturing use-cases: A systematic literature review. *Appl. Intell.* **2022**, 1–35. [[CrossRef](#)] [[PubMed](#)]
- Liu, G.; Shen, W.; Gao, L.; Kusiak, A. Knowledge transfer in fault diagnosis of rotary machines. *IET Collab. Intell. Manuf.* **2022**, 4, 17–34. [[CrossRef](#)]
- Hongm, W.; Tian-You, C.; Jin-Liang, D.; Brown, M. Data Driven Fault Diagnosis and Fault Tolerant Control: Some Advances and Possible New Directions. *Acta Autom. Sin.* **2009**, 35, 739–747. [[CrossRef](#)]
- Nasser, A.R.; Azar, A.T.; Humaidi, A.J.; Al-Mhdawi, A.K.; Ibraheem, I.K. Intelligent Fault Detection and Identification Approach for Analog Electronic Circuits Based on Fuzzy Logic Classifier. *Electronics* **2021**, 10, 2888. [[CrossRef](#)]
- He, Y.L.; Ma, Y.; Xu, Y.; Zhu, Q.X. Fault Diagnosis Using Novel Class-Specific Distributed Monitoring Weighted Nave Bayes: Applications to Process Industry. *Ind. Eng. Chem. Res.* **2020**, 59, 9593–9603. [[CrossRef](#)]
- Estima, J.O.; Cardoso, A. A New Approach for Real-Time Multiple Open-Circuit Fault Diagnosis in Voltage Source Inverters. *IEEE Trans. Ind. Appl.* **2011**, 47, 2487–2494. [[CrossRef](#)]
- Wang, J.; Wang, F.; Chen, S.; Wang, J.; Hu, L.; Yin, Y.; Wu, Y. Fault-tree-based instantaneous risk computing core in nuclear power plant risk monitor. *Ann. Nucl. Energy* **2016**, 95, 35–41. [[CrossRef](#)]
- Rojas-Guzman, C.; Kramer, M.A. Comparison of belief networks and rule-based expert systems for fault diagnosis of chemical processes. *Eng. Appl. Artif. Intell.* **1993**, 6, 191–202. [[CrossRef](#)]
- Grant, P.W.; Harris, P.M.; Moseley, L.G. Fault diagnosis for industrial printers using case-based reasoning. *Eng. Appl. Artif. Intell.* **1996**, 9, 163–173. [[CrossRef](#)]
- Mustapha, F.; Sapuan, S.M.; Ismail, N.; Mokhtar, A.S. A computer-based intelligent system for fault diagnosis of an aircraft engine. *Eng. Comput.* **2004**, 21, 78–90. [[CrossRef](#)]
- Nan, C.; Khan, F.; Iqbal, M.T. Real-time fault diagnosis using knowledge-based expert system. *Process Saf. Environ. Prot.* **2008**, 86, 55–71. [[CrossRef](#)]
- Bernard, J.A. Use of a rule-based system for process control. *IEEE Control Syst. Mag.* **1987**, 8, 3–13. [[CrossRef](#)]
- Geus, A.J.D.; Cohen, W. A Rule-Based System for Optimizing Combinational Logic. *IEEE Des. Test Comput.* **1985**, 2, 22–32. [[CrossRef](#)]
- You, D.; Gao, X.; Katayama, S. WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM. *IEEE Trans. Ind. Electron.* **2015**, 62, 628–636. [[CrossRef](#)]
- Chiang, L.H.; Kotanchek, M.E.; Kordon, A.K. Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Comput. Chem. Eng.* **2004**, 28, 1389–1401. [[CrossRef](#)]
- He, S.; Liu, X.; Wang, Y.; Xu, S.; Lu, J.; Yang, C.; Zhou, S.; Sun, Y.; Gui, W.; Qin, W. An effective fault diagnosis approach based on optimal weighted least squares support vector machine. *Can. J. Chem. Eng.* **2017**, 95, 2357–2366. [[CrossRef](#)]
- Malik, H.; Ashish, Y.; Kumar, S.A. EMD and ANN based intelligent model for bearing fault diagnosis. *J. Intell. Fuzzy Syst.* **2018**, 35, 5391–5402. [[CrossRef](#)]
- Wang, Z.; Liu, Y.; Griffin, P.J. A combined ANN and expert system tool for transformer fault diagnosis. *IEEE Trans. Power Deliv.* **2000**, 13, 1224–1229. [[CrossRef](#)]

19. Han, T.; Liu, C.; Yang, W.; Jiang, D. Deep Transfer Network with Joint Distribution Adaptation: A New Intelligent Fault Diagnosis Framework for Industry Application. *ISA Trans.* **2018**, *97*, 269–281. [[CrossRef](#)]
20. Khorram, A.; Khalooei, M.; Rezghi, M. End-to-end CNN+LSTM deep learning approach for bearing fault diagnosis. *Appl. Intell.* **2021**, *51*, 736–751. [[CrossRef](#)]
21. Talebi, N.; Sadrnia, M.A.; Darabi, A. Robust Fault Detection of Wind Energy Conversion Systems Based on Dynamic Neural Networks. *Comput. Intell. Neurosci.* **2014**, *2014*, 580972. [[CrossRef](#)] [[PubMed](#)]
22. Zhao, X.; Jia, M.; Bin, J.; Wang, T.; Liu, Z. Multiple-Order Graphical Deep Extreme Learning Machine for Unsupervised Fault Diagnosis of Rolling Bearing. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 3506012. [[CrossRef](#)]
23. Vaddi, P.K.; Pietrykowski, M.C.; Kar, D.; Diao, X.; Zhao, Y.; Mabry, T.; Ray, I.; Smidts, C. Dynamic bayesian networks based abnormal event classifier for nuclear power plants in case of cyber security threats. *Prog. Nucl. Energy* **2020**, *128*, 103479. [[CrossRef](#)]
24. Feng, L.; Zhao, C. Fault Description Based Attribute Transfer for Zero-Sample Industrial Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2020**, *17*, 1852–1862. [[CrossRef](#)]
25. Martin, T.P.; Glasgow, J.I.; Féret, M.; Kelley, T. *A Knowledge-Based System for Fault Diagnosis in Real-Time Engineering Applications*; Springer: Vienna, Austria, 1991. [[CrossRef](#)]
26. Dong, C.; Zhang, Q. The Cubic Dynamic Uncertain Causality Graph: A Methodology for Temporal Process Modeling and Diagnostic Logic Inference. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 4239–4253. [[CrossRef](#)] [[PubMed](#)]
27. Zhang, Q. Dynamic Uncertain Causality Graph for Knowledge Representation and Reasoning: Discrete DAG Cases. *J. Comput. Sci. Technol.* **2012**, *27*, 3–25. [[CrossRef](#)]
28. Zhang, Q. Dynamic Uncertain Causality Graph for Knowledge Representation and Probabilistic Reasoning: Directed Cyclic Graph and Joint Probability Distribution. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 1503–1517. [[CrossRef](#)]
29. Zhang, Q. Dynamic Uncertain Causality Graph for Knowledge Representation and Reasoning: Continuous Variable, Uncertain Evidence, and Failure Forecast. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *45*, 990–1003. [[CrossRef](#)]
30. Qin, Z.; Zhan, Z. Dynamic Uncertain Causality Graph Applied to Dynamic Fault Diagnoses and Predictions with Negative Feedbacks. *IEEE Trans. Reliab.* **2016**, *65*, 1030–1044. [[CrossRef](#)]
31. Zhang, Q.; Dong, C.; Cui, Y.; Yang, Z. Dynamic Uncertain Causality Graph for Knowledge Representation and Probabilistic Reasoning: Statistics Base, Matrix, and Application. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 645. [[CrossRef](#)]
32. Zhang, Q.; Yao, Q. Dynamic Uncertain Causality Graph for Knowledge Representation and Reasoning: Utilization of Statistical Data and Domain Knowledge in Complex Cases. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 1637–1651. [[CrossRef](#)] [[PubMed](#)]
33. Dong, C.L.; Zhang, Q. Research on weighted logical inference for uncertain fault diagnosis. *Acta Autom. Sin.* **2014**, *40*, 2766–2781. [[CrossRef](#)]
34. Dong, C.; Zhou, Z.; Qin, Z. Cubic dynamic uncertain causality graph: A new methodology for modeling and reasoning about complex faults with negative feedbacks. *IEEE Trans. Reliab.* **2018**, *67*, 920–932. [[CrossRef](#)]
35. Zhao, Y. Research on DUCG Theory and Application for Fault Diagnosis and Procedure Improvement of Nuclear Power Plant. Ph.D. Thesis, Tsinghua University, Beijing, China, 2017.