electronics

MDPI

Article

Intelligent Diagnosis Method of Data Center Precision Air Conditioning Fault Based on Knowledge Graph

Jinsong Wu [1,2,*], Xiangming Xu [1], Xiao Liao [2], Zhuohui Li [1,2], Shaofeng Zhang [2] and Yong Huang [2]

[1] School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China
[2] China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd., Guangzhou 510663, China
* Correspondence: wujinsong@gedi.com.cn

Abstract: This study first digitizes, rules and structures complex unstructured data such as massive historical operation and maintenance data and fault judgment experience of operation and maintenance engineering based on semi-automatic entity extraction method; annotate the association or indirect relationship between 63,724 types of faults among triads by means of decision trees. Bayesian algorithm is used to further explore the relationship between triples, the realizes knowledge fusion, knowledge reasoning and knowledge update, and completes knowledge graph construction; combines with fault intelligent diagnosis method, realizes fault prediction, fast discovery, locates fault, type and business impact reasoning, and provides solutions to assist decision making.

Keywords: knowledge graph (KG); precision air conditioning (PAC); data center (DC); intelligent diagnosis; fault prediction

check for updates

Citation: Wu, J.; Xu, X.; Liao, X.; Li, Z.; Zhang, S.; Huang, Y. Intelligent Diagnosis Method of Data Center Precision Air Conditioning Fault Based on Knowledge Graph. *Electronics* **2023**, *12*, 498. https://doi.org/10.3390/electronics12030498

Academic Editors: Yusen He, Huajin Li, Tinghui Ouyang, Xun Shen and Zhenhao Tang

Received: 21 December 2022
Revised: 11 January 2023
Accepted: 14 January 2023
Published: 18 January 2023

Copyright: © 2023 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

Data center is the core place for storing a large amount of information and business processing, and its importance is self-evident, so it attaches great importance to the operation and maintenance management of data center. In order to improve the stability, reliability and energy saving of data center services, cooling systems consisting of precision air conditioners are usually equipped to strictly control the temperature and humidity in the data center, but in the long-term operation process, precision air conditioners can also exist with disease operation, high temperature operation, load operation and other situations, and there are safety risks. Incidents in critical data centers can lead to digital disasters, such as the failure of the Google Cloud data center in London and Oracle's data center in July 2022. The accidents were "simultaneous failure of the air conditioning cooling system" and the machine was unable to maintain a safe operating temperature.

However, in the real environment, usually 3–6 operation and maintenance engineers manage dozens of server rooms and hundreds of precision air conditioners, how to conduct real-time status inspection of equipment, safety confirmation, reduce energy consumption, improve energy utilization, rapid diagnosis and location of faults, has become an objective demand.

Knowledge Graphs are widely used in Wikipedia, network diagnosis, disease diagnosis and auto repair diagnosis. Due to the large number of precision air conditioning equipment in data centers and the diversity and hidden nature of the causes of failure, it takes a lot of time and material resources to make accurate diagnosis, similar to the difficulties of network diagnosis, disease diagnosis and auto repair diagnosis, so it is inferred that knowledge graph technology can also be applied to the fault diagnosis of precision air conditioning. In this paper, through the method of manual experience summary, historical operation and maintenance data structuring and analysis, we will build a knowledge map and fault diagnosis process, invoke real-time measurement point data of precision air

conditioners, realize intelligent fault diagnosis, reason out the causes of fault formation and the scope of impact, and display them in 3D to achieve rapid elimination of safety hazards.

In this paper, the knowledge graph model is used as the basis for the diagnosis of faults combined with Bayesian networks to improve the model performance. Experiments show that the knowledge graph model combined with Bayesian theory proposed in this paper can detect faults. Meanwhile, compared with the comparison algorithms, the accuracy of fault judgment of the algorithm in this paper is higher, which indicates that the algorithm model in this paper has better performance.

2. Methods

2.1. Research Methods

To implement a knowledge graph-based intelligent diagnosis method for precision air conditioning faults in data centers, this study is divided into three steps for validation.

(a) Overall planning. Determining the performance characteristics of precision air conditioners and the definition of the knowledge graph, designing the building process and the fault intelligent diagnosis process.

(b) Build graph. Based on 12 million precision air conditioner historical monitoring data, combined with conventional fault diagnosis methods and causes, transforming entities into attributes, labeling the relationships between entities and attributes, completing the five processes of knowledge representation, knowledge extraction, knowledge fusion, knowledge inference and knowledge update, constructing knowledge graph, and substituting into the fault intelligent diagnosis process.

Precision air conditioning historical monitoring data include: cascade function, dehumidification lock release, dehumidification running time, current air supply humidity, floor overflow alarm, low pressure sensor lock alarm, low pressure alarm, low pressure lock alarm, low pressure pressure sensor fault, electric heating fault alarm, electric heating output, electric heating running time, power loss alarm, power over voltage alarm, power frequency, power under voltage alarm, power out of phase alarm, etc.

(c) Method validation. Substitute into the actual operation data set, realize inference of fault cause, location and solution suggestion based on fault phenomena through knowledge graph.

2.2. Research Subjects

Precision air conditioners (constant temperature and humidity units) that is a special air conditioner that precisely controls the temperature and humidity of the data center environment. Precision air conditioners have microprocessor control systems to achieve integrated control of room temperature, air cleanliness and air humidity. In order to ensure the stability and reliability of data center server operation, a dedicated precision air conditioning system is required. Precision air conditioning and ordinary air conditioning performance and operating parameters are compared in Table 1 [1,2].

2.3. Research Steps

2.3.1. Knowledge Graph Definition

Knowledge graphs are part of the underlying artificial intelligence technology and depict relationships between entities through semantic networks. Each knowledge graph development tool comes with semantics, logical meanings and rules to describe the relationships between things in the form of triples, i.e., "entity x relationship x entity" or "entity x attribute x attribute value" sets. Knowledge graphs can achieve functions that cannot be achieved by ordinary relational databases, such as structuring, aiding inference, prediction, and categorization of knowledge information in a non-linear world [3–6]. Please see (Figure 1) for the definition of knowledge graph.

Table 1. Precision Air Conditioning (PAC) VS General Air Conditioner (GAC).

Precision Air Conditioning (PAC) vs. General Air Conditioner (GAC)			
Description		**PAC**	**GAC**
Operating Temperature	−40~45 °C	−40~45 °C	15~35 °C
Heat Density	300~500 W/m^2	300~800 W/m^2	100~150 W/m^2
Apparent Heat Ratio	0.80~0.90	0.9~1.0	0.6~0.7
Temperature Control Accuracy	≤±1 °C	±1 °C	±3~5 °C
Ventilation Capacity	≥30 TIMOS/HOUR	30~60 T/H	10~15 T/H
Air Delivery Volume	Air delivery volume/Cooling capacity >1:4	1:3.5	1:5
Air Supply Temperature	13~15 °C	13~15 °C	6~8 °C
Air Filtration Capacity	ASHRAE52–76 Standards: 0.5 μm/L < 18,000	0.5 μm/L < 16,000	0.5 μm/L < 20,000
Outlet Air Temperature	13~15 °C	13~15°C	6~8 °C
Humidity Control Capability	Humidity Control: 45%~50%	Humidity Regulation ±5%	NONE
Power Function	Power Failure Recovery Function	Support	NONE
Running Time	>8000 h	8760 h	1000~2500 h
Service Life	10~15 years	10 years	2~3 years

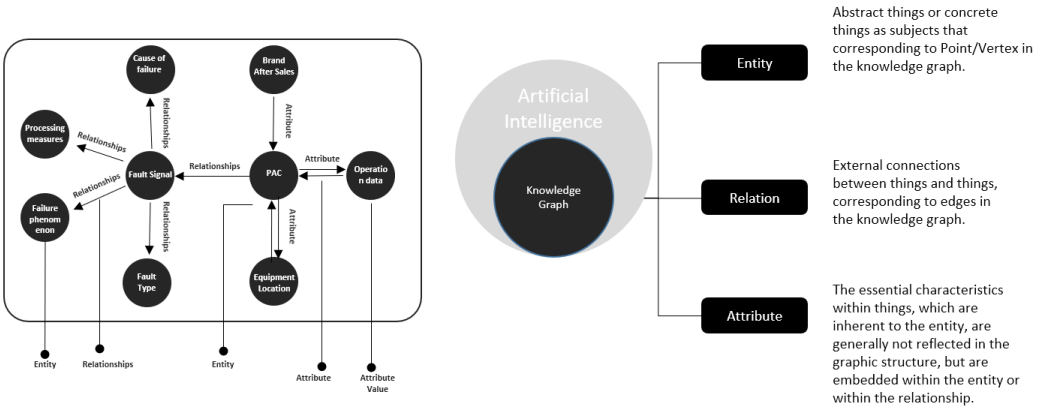


Figure 1. Knowledge Graph Definition at a Glance.

2.3.2. Knowledge Graphs Building

Obtain 12 million structured data of the operation and maintenance status of precision air conditioners in data centers for nearly one year, and perform entity extraction, relationship extraction, and attribute extraction. The 179 common fault diagnosis triads of precision air conditioners are marked out, and the knowledge map (knowledge base) created by fusing knowledge of 63,724 concurrent or indirectly caused correlations among the triads. Establishing data anomaly threshold criteria for precision air conditioners, using CNN (Convolutional Neural Network) to identify data anomalies in abnormal precision air conditioners matched to the response fault phenomenon. Provide inference model for fault diagnosis process, and display intelligent fault diagnosis report containing fault location, parts, real-time status, business impact and solution in the form of BIM digital twin model to achieve intelligent diagnosis (Figure 2). Complete the process from Knowledge Representation, Knowledge Extraction, Knowledge Integration, Knowledge Reasoning, Knowledge Update and Knowledge Application [7,8].

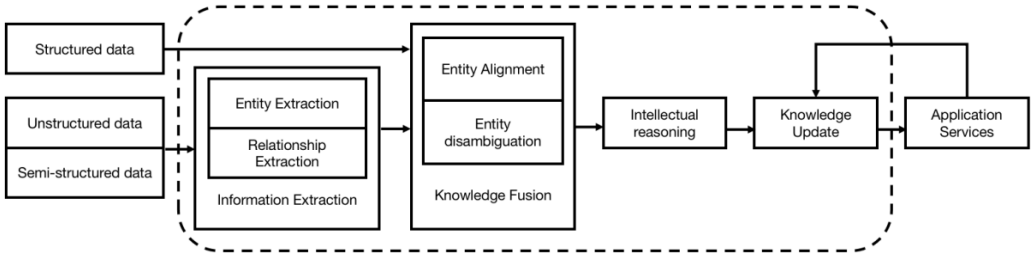


Figure 2. Knowledge Graph Architecture.

Knowledge Representation

Based on nearly one year 12 million data center precision air conditioning operation and maintenance data and operation and maintenance diagnosis manual new entity list, relationship list, attribute list, to achieve unstructured data structured. Knowledge graphs can be stored using graph databases, and common graph databases include Neo4j [9] and FlockDB [10].

Knowledge Extraction

Due to the complexity of precision air conditioning systems, the influence between each entity is not a single factor influence, the factors are usually multifaceted. Therefore, a composite representation of RDF and OWL combined with multi-level knowledge Graph is used, there are two kinds of triples in the knowledge base, "entity x relationship x entity" and "entity x attribute x attribute value", and 179 triples of direct relationships or attributes are constructed by extracting "entity", "relationship" and "attribute" about fault diagnosis from the entity list, relationship list and attribute list. Therefore, please see (Figure 3) for the multi-level knowledge graph structure diagram.

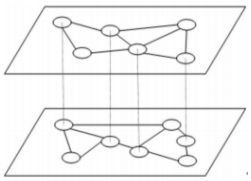


Figure 3. Multi-Level Knowledge Graph.

Knowledge Integration

Through extensive manual work with KGCLOUD knowledge graph annotation tool from the 63,724 concurrent or indirect relationships among 179 groups of triads, the ambiguous relationship items are mapped to the triads they actually refer to, so as to achieve the alignment and fusion of multiple associated triads, solve the problem that it is difficult to relate "fault hidden danger" and "cause of fault", achieve knowledge fusion, and build a knowledge base. To solve the problem of difficulty in correlating "fault hazards" and "fault causes", the realize knowledge integration [11,12].

Knowledge Reasoning

By using Bayesian theory, it consists of inference based on precision air conditioning ontology, inference based on energy consumption and cooling capacity cycle rules, inference based on component monitoring model, similarity factor graph calculation, link prediction, and inconsistency detection. The Bayesian model calculates the posterior probability of each inferred failure cause and finally obtains the failure cause with the highest probability. Bayesian probabilistic inference is shown as follows.

Let the discrete set $U = [x_1, x_2, \ldots, x_n, C]$, and the element C of the discrete set is the multiple possible causes of the fault occurrence, which belongs to the fault type variable. And these causes may be caused by more than one factor, so multiple samples can be formed into a set and expressed as:

$$Y_i = [x_1, x_2, \ldots, x_n]$$

The variable value x_n is the fault attribute, and using the knowledge of probability, it is known that the probability of the fault occurring is the following [13,14]:

$$P(c_j|x_1, x_2, \ldots, x_n) = \frac{P(x_1, x_2, \ldots, x_n|c_j) \cdot P(c_j)}{P(x_1, x_2, \ldots, x_n)}$$

The probability P on the left side of the above equation is the conditional probability, and the right side is the detailed calculated value of Bayesian probability, where $P(x_1, x_2, \ldots, x_n|c_j)$ is the probability of occurrence of fault c_j when the sample value is Y_i, $P(c_j)$ is the full probability of occurrence of fault c_j when it occurs, and $P(x_1, x_2, \ldots, x_n)$ is the joint probability of attribute values. It can be derived:

$$P = \alpha \cdot P(c_j) \cdot P(x_1, x_2, \ldots, x_n|c_j)$$

From the above equation, it can be seen that the calculation is performed using using the Bayesian theory to derive the probability of fault occurrence i.e., pairs. The calculation is then performed using the plain Bayesian theory with the following topology (Figure 4).

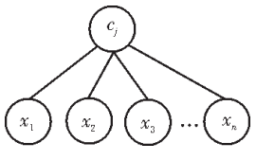


Figure 4. Plain Bayesian theory.

From the topology of the plain Bayesian network, it can be seen that the relationships between the fault attributes x_n are independent, and the probability value is obtained by the ratio of the number of faults occurring. Firstly, $P(c_j)$ is calculated assuming that there are N sample data, the number of occurrence of faults is, and the probability value is the following:

$$P(c_j) = \frac{N_{C_J}}{N}$$

$$P(Y_j = x_i|c_j) = \frac{N_{C_J}^{(x_i)}}{N_{C_J}}$$

If $N_{C_J}^{(x_i)}$ is 0 at this point, the probability value is

$$P(x_i|c_j) = \frac{1/N}{N_{C_J} + N_{X_j}/N}$$

The probability of the failure value of sample Y_i can be calculated using the above method, and the maximum value of the probability value corresponds to the cause of the failure presumed by the system.

In the case of massive monitoring data calculation, the paper uses Bayesian theory.

Bayesian neural networks are constructed, which usually consist of input layer, hidden layer, and output layer [15–18]. The Bayesian network is used as a training and testing model for the samples, and the constructed Bayesian network model is shown in (Figure 5):

The Bayesian model proposes a constraint function, which in turn allows the solution of the model to converge, and the chosen constraint function is:

$$F = \beta E_D + \alpha E_\omega$$

Based on the above constraint function, the values of the scale coefficients α and β can be set according to the Bayesian network self-regulation theory as:

$$\alpha = \frac{y}{2E_\omega}$$

$$\beta = \frac{n-y}{2E_D}$$

In Bayesian models, the constraint function is commonly used to converge the model's solution to the optimal solution. Therefore, the parameters α and β in the constraint function are used to adjust the convergence of the model.

In this constraint function, α and β respectively represent the weights of E_ω and E_D in the model. E_ω refers to the sum of the weights of the nodes in the Bayesian network, and E_D refers to the entropy of the samples in the dataset. According to the Bayesian network self-adjustment theory, the values of α and β can be set to $y/(2E_\omega)$ and $(n\text{-}y)/(2E_D)$, respectively. This means that the model will pay more attention to the weights of the nodes and the entropy of the samples during the solution process, leading to better convergence.

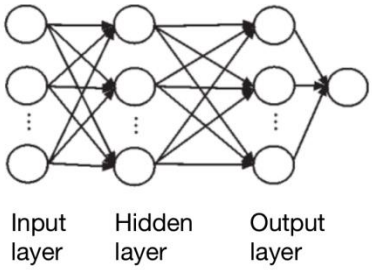


Figure 5. Bayesian Network Model.

Knowledge Update

Based on the feedback from the application results of knowledge graph, new "entities", "attributes" and "relationships" are created, and created the preliminary "triplet", the actual fault diagnosis results are judged, and if they are true, the knowledge base is returned and new knowledge supplement is executed to realize automatic updates driven by data business applications.

Knowledge Application

The knowledge graph model is used as the basis for fault diagnosis combined with Bayesian theory to improve the model performance. The final established algorithm model consists of three modules, which are multi-level knowledge graph module, state detection module and fault diagnosis module, and the algorithm model is shown in the following (Figure 6). Firstly, the online data of each equipment is obtained, and the state is judged by the historical data when the weapon and equipment are working normally; then the multi-level knowledge graph is constructed and the deep correlation path between each factor is calculated, and whether the system is working normally is judged by setting the judgment coefficient; finally, the symptoms of the fault are found based on the knowledge graph, and the cause of the fault is analyzed by the plain Bayesian theory. The cause of the fault is analyzed by the plain Bayesian theory, and the cause of the fault is obtained.

Knowledge Graph Application Algorithm Process: The specific working process of the knowledge mapping module is to establish the information extraction module first, and the data acquisition can be completed using the crawler model. After acquiring the data, knowledge fusion is performed, and the fusion coefficients need to be determined for knowledge fusion, and the fusion coefficients W determined in this paper can be expressed as:

$$W = Qxc_fx\left(\frac{E_{sim} + R_{esim}}{2}\right)$$

In the formula, Q is the confidence level of the real-time monitoring data source, and c_f is the confidence level of the adjacent entities of the atlas. E_{sim} and R_{esim} are the text

similarity and relational similarity, respectively, and the use of fused relational coefficients converts uncertain information into definite and format-correct information.

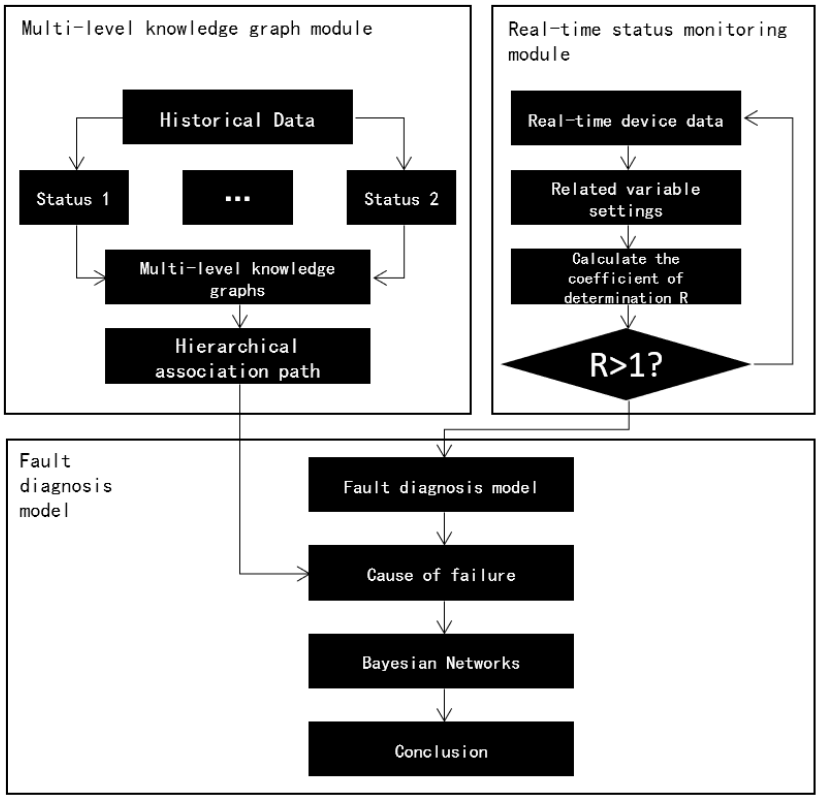


Figure 6. Knowledge Graph Application Algorithm Process.

When there is a fault signal, the correlation coefficient of the knowledge graph will be changed significantly, so the detection threshold of the changed value needs to be set, i.e., [19–23], the determination factor R needs to be added to judge the current stable state of the system, and R can be defined as:

$$R = \frac{\sum_{k=1}^{n} T_k + \sum_{k=1}^{n} T_k + t}{n} + W = \frac{|n-t| + |m-t| + t}{n} + W = \frac{n+m-t}{n} + W$$

When the value of R is 1, the equipment is in normal working condition; when the value of R is not 1, the fault diagnosis module is activated and fault detection is carried out.

When a fault is detected by the system state detection module, the variable values of the fault are initially determined. Then a multi-level knowledge mapping model is built to find all possible causes of the fault and the probability value of the fault occurrence is calculated using the plain Bayesian theory. The final fault probability detection formula is.

$$P(r_1|E) = \frac{P(r_i, E_1, \ldots, E_n)}{P(E_1, \ldots, E_n)} = \frac{P(r_i \cap E_1 \cap \ldots \cap E_n)}{P(E_1 \cap \ldots \cap E_n)}$$

In this equation, $P(r_1|E)$ denotes the probability of r_1 given that E is known. E denotes the set of all events. r_i denotes the probability of the ith fault. $E_1, \ldots, E_n$ denote the probability of the occurrence of n events, respectively.

So, this equation is to calculate the probability of occurrence of the ith fault r_i, given that all events E are known to occur.

It means that what we need to calculate is the probability of the occurrence of fault r_i given that all events E are known to occur. This probability can be calculated by the Bayesian formula, which is

$$\frac{P(r_i \cap E_1 \cap \ldots \cap E_n)}{P(E_1 \cap \ldots \cap E_n)}$$

3. Results

3.1. Experimental Tests and Results Analysis

The validity of the model is first verified, and the information related to the equipment shows that when there is a sudden change in the power consumption of the equipment, that is, it may be caused by a short circuit. Therefore process variables and fault variable values C4 are sequentially corresponded to each other. Inputting process variables to the Bayesian model, the final training results of the model are shown in the following Table 2.

Table 2. Validity validation results.

Cause of Failure	Model Output Probability (%)
C4	89.1
C2	32.1
C1	5.3
C3	3.6

From the above table, it can be seen that when the system detects the output of four causes of failure, the probability of C4 (circuit short-circuit) is 89.1%. Therefore, the system finally judges the fault state as circuit short circuit, which is in line with the actual situation, so the model in this paper can effectively judge the faults caused by equipment state change. The above is the fault caused by a single state change, while the fault in real equipment is usually caused by multiple states. The more complete the data is, the more accurate the detection result will be. Next, experiments on model accuracy are conducted, and the algorithms selected for comparison in the paper are BP neural network, judgment tree algorithm, and support vector machine algorithm. These algorithms are common fault detection algorithms, and the results of the comparison algorithms are shown in the following Table 3.

Table 3. Comparison of algorithm detection results.

Model Algorithms	Accuracy (%)
BP Neural Network Algorithm	87.1
Judgment tree algorithm	88.5
Support vector machine algorithm	90.2
Algorithms in the paper	92.7

As can be seen from the results in the above table, the accuracy measured by the algorithm in the paper is the highest among the compared algorithms. Compared with other algorithms, the accuracy rate of the algorithm in the paper is improved by 2.5%, 4.2% and 5.6%. It proves that the algorithm of the paper has a high accuracy rate of fault determination, which indicates that the algorithm model constructed in the paper can determine the fault of precision air conditioning system.

3.2. Precision Air Conditioning System Troubleshooting Example

3.2.1. Knowledge Graph of Precision Air Conditioning Fault Information

Through digitization, regularization and structuring of complex unstructured data such as massive historical operation and maintenance data and fault judgment experience of operation and maintenance engineering, 179 common fault diagnosis triads are marked out (Figure 7), and knowledge fusion of 63,724 concurrent or indirect associations among

the triads is carried out to complete the construction of knowledge graph for precision air conditioning fault diagnosis (Figure 8), realize the jump-level retrieval of various fault information of precision air conditioning, and quickly obtain the knowledge reasoning results of fault phenomena and solutions [24–27].

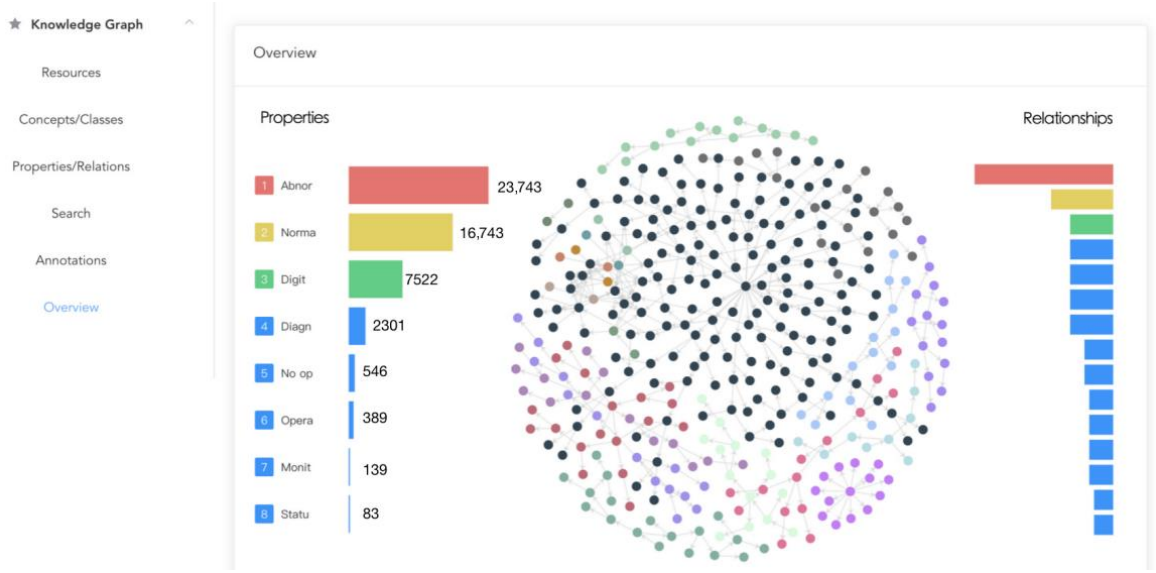


Figure 7. Knowledge Graph Overview.

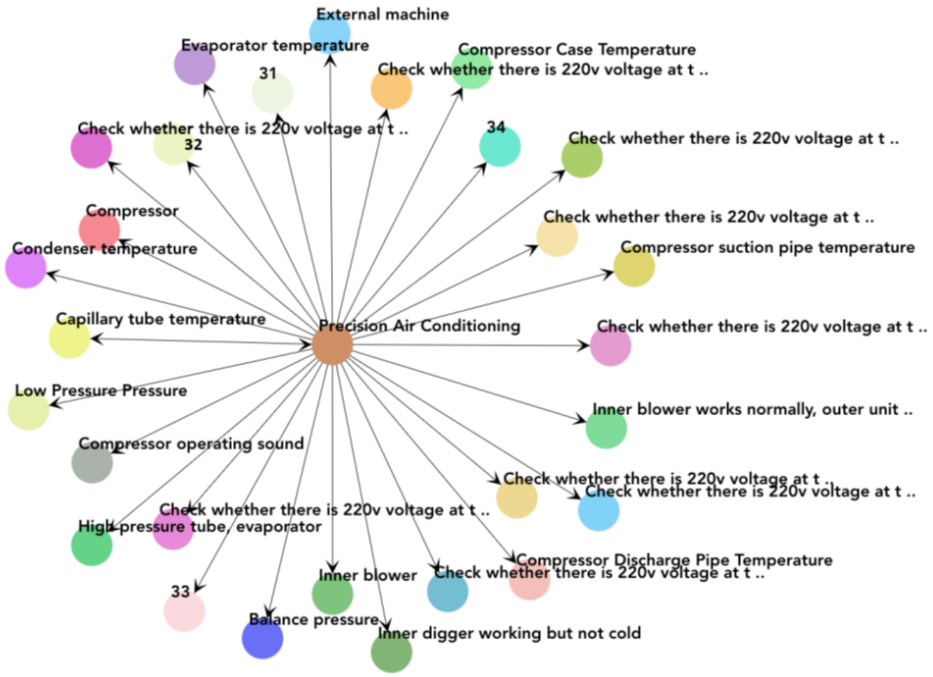


Figure 8. The Search of Precision Air Conditioning from Knowledge Graph.

Knowledge Gragh Description: Box Represents Class, Circle Represents lnstance, Arrow Represents Relation.

3.2.2. Application of Intelligent Diagnosis of Precision Air Conditioning Fault Based on Knowledge Graph

Obtain 12 million historical data from the edge data center of China Guangdong Park for data analysis and data mining, develop a threshold model for early warning data, correlate

it with the fault result entity of the triad in the knowledge graph library, and import it into the model of the 1:1 digital twin of BIM technology to realize fault intelligent diagnosis based on knowledge graph for timely warning of precision air conditioning faults, 3D rapid positioning, and intelligent diagnosis, etc., function.

The following steps are followed to import intelligent diagnostic methods into a Building Information Modeling (BIM) system.

(1) Determine the specific intelligent diagnostic method to be used. There are many different methods, choose one that fits the needs.
(2) Determine the data and input parameters required for the intelligent diagnostic method. This may include information about the building itself, as well as any relevant operational data.
(3) Determine how the required data will be accessed and entered into the digital twin. This involves setting up interfaces or connectors to collect data from different sources, or it may involve manually inputting data into the system.
(4) Implement intelligent diagnostics within the digital twin. This involves writing code or using tools to integrate the method into the system.
(5) Testing the implementation (Figure 9) to ensure that it functions correctly and produces accurate results Table 4.



Figure 9. Precision Air Conditioning Troubleshooting Application Interface for Data Center at the Edge of Guangdong Campus in China.

Table 4. Monitoring model based on 12 million O&M data analysis and mining of precision air conditioning failure thresholds.

	Normal Value	**Corresponding Value**						
	Check Items	**Refrigerant Lack of Fluoride**	**All Refrigerant Leaks**	**Poor Heat Dissipation of External Unit**	**Poor Ventilation of The Internal Unit**	**Too Much Refrigerant**	**Air in The System**	**Clogged Filter**
Compressor working sound	60–100 db	Less than 60 db	Less than 60 db	Above 90 dB	Less than 60 decibels	Above 90 dB	Above 90 dB	Less than 60 db
Compressor suction pipe temperature	Cold, about 13 °C, the return air tube to the reservoir tube frost	Over 25 °C, little or no frost condensation	More than 40 °C	More than 25 °C, little or no frost condensation	Less than 7 °C	Below 7 °C	More than 25 °C, little or no frost condensation	Over 25 °C, little or no frost condensation
Compressor discharge pipe temperature	Ambient temperature plus 55 °C, not exceeding 95 °C	Over 95 °C	Over 40 degrees Celsius	Above 95 °C	Below ambient temperature plus 55 °C	More than 95 °C	Above 95 °C	Over 95 °C
Compressor Case Temperature	88–92 °C	Over 90 °C	More than 90 degrees	Over 90 °C	Less than 90 °C	Below 90 °C	Over 90 °C	More than 90 °C
Low Pressure Pressure	4.5–5.5 kg	Below normal pressure	0–2 kg	More than 5.5 kg	Less than 4.5 kg	More than 5.5 kg	Unstable, beating	Less than 4.5 kg
Balance pressure	Saturation pressure value at ambient temperature	Severely below saturation pressure	/	/	/	/	/	/
Evaporator temperature	Cold, frost, ambient temperature minus 15 °C	Local frosting or icing phenomenon	Over 30 °C	Above ambient temperature minus 15 °C	Less than 5 °C	Below 5 °C	Above ambient temperature minus 15 °C	Above ambient temperature minus 15 °C
Condenser temperature	Hot, ambient temperature plus 15 °C (45~55 °C)	Hot, warm	Over 30 °C	Above ambient temperature plus 15 °C	Below ambient temperature plus 15 °C	Above ambient temperature plus 15 °C	Above ambient temperature plus 15 °C	Above ambient temperature plus 15 °C
Capillary tube temperature	Room temperature	Cold, even condensation and ice	Over 30 °C	More than 45 °C	/	/	/	Less than 7 °C

3.3. Comparison of Existing Technology Applications

In the traditional manual diagnosis method, the operation and maintenance engineer will determine the diagnosis path related to the fault situation based on experience, and use the elimination method to gradually narrow down the diagnosis and finally locate the problem. When the problem cannot be located, the diagnostic process is bottlenecked and the fault can only be solved by expanding the diagnostic scope. In a data center troubleshooting process with many devices, the time taken for locating and diagnosing is much higher than the solution time.

Through the practical application of the fault intelligent diagnosis method based on knowledge graph in this study, the knowledge graph constructed by various data sources and rules such as expert knowledge, case accumulation, and historical data, combined with measurement point data, can reduce most of the manual inspection work, and we can monitor medium and large data center assets and equipment, hundreds of precision air conditioners, with the assistance of digital technology, by only a few people, and only 3–5 people are required to maintain them. Among them, the intelligent diagnosis uses correlation analysis, which can converge quickly on the diagnosis path and superimpose the machine's arithmetic power, which can find and locate the problem more quickly and accurately relative to the manual diagnosis, The new diagnostic experience will also be fed back to the system to form a new triad, which will be integrated with the original knowledge to improve the knowledge graph and provide more accurate fault intelligence diagnosis capability for the application.

4. Discussion

4.1. Inference of Results

If a traditional relational database is used to build a knowledge base without using knowledge graph technology, the following problems will arise in the implementation process.

(1) How to bridge spatial data, operational data and real-time data, establish unified relationships with entities, and complete triad construction.
(2) When multiple anomalies trigger alarm thresholds, how to perform complex correlation, fault inference, and fast location.
(3) How to apply the knowledge graph to the intelligent diagnosis of precision air conditioning faults and realize the autonomous update of the knowledge graph.

Under the conventional technology, it is impossible to achieve intelligent diagnosis of precision air conditioning faults. If we use neural network analysis, machine learning, big data analysis and other technologies, we will pay higher R&D cost and longer R&D cycle, and the hardware requirements for operation and R&D cost will also become higher. Therefore, the choice of knowledge graph technology in artificial intelligence to implement the data center precision air conditioning fault intelligent diagnosis method is the most in line with the practical needs, which can realize multi-source heterogeneous data fusion, knowledge base construction and complementation, unification and digital relationship construction of data with entity attributes and meanings, and provide the function of simple and complex association reasoning.

4.2. Reason

Knowledge graph technology is the key technology for the realization of intelligent diagnosis of precision air conditioning faults, because knowledge graph has the following characteristics in the intelligent diagnosis of faults:

4.2.1. Efficient Knowledge Retrieval Capabilities

The relational database query is done through tables to optimize the query for the huge amount of data already stored; while the lookup mode of the knowledge graph is to find the required content from a triad. For multi-hop search, knowledge graph can quickly

locate and obtain fault diagnosis results through knowledge fusion annotation among triads, which greatly improves query efficiency.

4.2.2. Entity Extraction Capability for Semi-Structured Data

Relational databases require data to be strictly structured, and usually build the structure of data with multi-table relationships. In contrast, Knowledge Graph uses a triadic approach to flexibly construct a knowledge base, and in the construction process, it directly extracts entity x relationship x entity data from documents or historical data, and fuses one independent triad for knowledge fusion and knowledge inference, thus realizing support for semi-structured data retrieval. In the application of fault intelligent diagnosis, the retrieval of unstructured data can be realized by simple processing of natural language through CNN technology [28].

4.2.3. Intuitive Intellectual Reasoning and Analysis Skills

Knowledge graphs provide the most effective and intuitive representation of the relationships between entities. Connecting a large number of different types of entities together to obtain a relational network provides the system with the ability to analyze problems from a "relational" perspective.

4.3. Comparison with the Results of Other Researchers

Relative to diagnosis in other industries (such as intelligent diagnosis of diseases based on knowledge graph, automotive fault diagnosis [29], network fault diagnosis [30], etc.), knowledge graph technology, when applied to the intelligent diagnosis of faults in data center precision air conditioners, can also give good play to the advantages of knowledge graph technology application and the ability of knowledge graph to construct and reason about complex relationships of entities. However, there are also technical implementation difficulties, such as the need to spend a lot of manual labor in the process of triad construction and fusion in order to build a more perfect knowledge graph.

4.4. Meaning of the Results

Through the study of intelligent diagnosis of precision air conditioning faults in data centers, it was found that the seemingly simple fault diagnosis of precision air conditioning has up to more than 3724 fault relationships for each module and component, which requires a lot of manual effort to transform manual experience into knowledge graph. In the actual environment, due to the stability of precision air conditioning special equipment, sometimes there is no apparent failure for several years, and it is difficult for operations and maintenance staff to find potential hidden problems, especially in large data center scenarios where hundreds of precision air conditioners are running at the same time, making it almost impossible to achieve a daily inspection for each unit. The application of data center precision air conditioner fault intelligent diagnosis method based on knowledge graph not only provides intelligent real-time status monitoring and alarm for precision air conditioners, but also provides fault intelligent diagnosis to achieve fast fault discovery, location and solution, and provides auxiliary decision for operation and maintenance engineers.

4.5. Limitations of the Study in This Paper

From the data center precision air conditioning data samples observed that the more frequent failure for the cabinet host heat is too much, not timely increase the cooling capacity of precision air conditioning led to the failure of the cabinet server; or idle cabinets when more, the continuous cooling of precision air conditioning, resulting in a waste of energy consumption situation. If the limitations of this study are broken, the triad of precision air conditioner status and other precision equipment status data is constructed, and the triad of precision air conditioner fault identification is constructed, a more complete knowledge map is formed, which can further enrich the content and scope of fault intelligent diagnosis report, and will better assist operation and maintenance to provide more accurate fault diagnosis, energy consumption diagnosis and operation diagnosis services.

4.6. Directions for Future Research

The data center precision air conditioning fault intelligent diagnosis method based on knowledge graph is introduced into the data center operation and maintenance system for practical application to further optimize the knowledge graph and gradually establish the relationship and attributes between new entities and original entities, such as UPS, diesel and generators, precision cabinets and other equipment, so as to provide more and more comprehensive fault intelligent diagnosis services for data center operation and maintenance.

5. Conclusions

5.1. Overview of the Research on This Topic

5.1.1. Knowledge Graph of Precision Air Conditioning Troubleshooting

Based on the operation structure of precision air conditioners, it took 6 months of labeling time to label 179 common fault diagnosis triads, fuse the knowledge of 63,724 concurrent or indirect associations between the triads, build knowledge, combine the warning data thresholds calculated from 12 million historical data, build associations with the corresponding historical solutions, and complete the construction of the knowledge graph for precision air conditioner fault diagnosis. Based on the unified data standard and the marked triad relationship, it realizes universal search of fault knowledge and provides support for any data search interface such as fuzzy search, correlation search and real-time monitoring for fault intelligent diagnosis applications.

5.1.2. Application Flow of Fault Intelligent Diagnosis Based on Knowledge Graph

Through the research, it is realized that when the real-time status data triggers the threshold value and an alarm is issued, the fault intelligent diagnosis process is triggered, the data of the data center precision air conditioner is packaged as input, and the intelligent diagnosis report is formed through knowledge graph retrieval, and after pushing to the operation and maintenance staff, the staff's choice of solution measures and the record of actual operation are tracked and recorded, which is fed back to the professional knowledge base of the knowledge graph, and the fault risk reassessment is executed to guarantee that the fault phenomenon is correctly solved, escorting the data center, improving the efficiency and accuracy of fault diagnosis, and reducing the accidents caused by the failure of the cooling system.

5.1.3. Intelligent Diagnosis Based on Fault Prevention Diagnosis Model

With the help of knowledge graph technology, a precision air conditioning fault prevention and diagnosis model is constructed. It realizes fault intelligent diagnosis through real-time monitoring of measurement point data of each collection module, which not only assists operation and maintenance engineers to diagnose the faults that have occurred, but also completes the diagnosis and early warning of faults to be discovered, reducing the possibility of digital disasters caused by refrigeration system faults.

5.2. Specify the Results of the Study

Proving that knowledge graph technology applied to industries such as medical, auto repair and networking can be well applied to data center operations and maintenance work. This study takes data center precision air conditioning operation and maintenance as the demand scenario, establishes fault diagnosis triad, and completes the construction of professional fault diagnosis knowledge map by correlating and labeling its 63,724 combinations through the cooperation of manual and tools. Deploy fault diagnosis process to realize knowledge update, so as the number of fault diagnosis cases increases, gradually improve the knowledge map and increase the accuracy of intelligent diagnosis. Linkage data center precision air conditioning data collector to achieve real-time monitoring of precision status, fault alarms and alarms, generation of work orders, rapid positioning and diagnosis reports and other fault intelligent diagnosis functions. The method can theoretically also be applied to data center UPS, precision cabinets, diesel and generators and other equipment to assist the operation and maintenance work to protect the data center.

Precision air conditioning systems generate great uncertainty due to their complex structures. The traditional fault diagnosis method confirms whether a fault occurs by analyzing the output information of a mathematical model, but this cannot meet the needs of precision air conditioning system maintenance. This paper uses the knowledge mapping model as the basis and combines Bayesian networks for fault diagnosis to improve the model performance. Experiments show that the knowledge graph model combined with Bayesian theory proposed in the paper can detect faults. Also compared with the comparison algorithm, the algorithm fault judgment accuracy of the paper is higher, which indicates that the algorithm model in the paper has better performance.

Author Contributions: Conceptualization, J.W. and X.X.; methodology, J.W. and X.X.; validation, X.L., Z.L. and S.Z.; investigation, S.Z. and Y.H.; writing–original draft preparation, X.L. and Z.L.; writing–review and editing, S.Z. and Y.H.; project administration, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work receives no external funding.

Data Availability Statement: The data used to support the findings of this study are included within the article.

Conflicts of Interest: The authors declare that they have no competing financial or non-financial interest.

References

1. Luo, Y.-J. A theoretical study on the fault intelligent diagnosis method of precision air conditioning system in broadcasting integrated media data center. *Ind. Control Comput.* **2020**, *33*, 3.
2. Zhao, Y.-L.; Yu, Q.; Deng, B.-O. Equipment fault diagnosis technology based on knowledge graph. *Electron. Des. Eng.* **2022**, *30*, 5.
3. Cheol, Y.-G. Knowledge map (K-map) modeling for improving construction project performance with the integration of key construction project resources. *Constr. Eng. Manag.* **2017**, *18*, 151–162.
4. Li, M.; Lu, X.-Z.; Chen, L.-S.; Wang, J. Knowledge map construction for question and answer archives. *Expert Syst. Appl.* **2020**, *14*, 39–50. [CrossRef]
5. Huang, H.-Q.; Yu, J.; Liao, X. A review of knowledge graph research. *Comput. Syst. Appl.* **2019**, *28*, 5–12.
6. Li, Z.-W.; Ding, Y.; Hua, Z.-Y. A knowledge graph complementation model combining triadic importance. *Comput. Sci.* **2020**, *47*, 231–236.
7. Liu, X.-Z.; Song, C.-Y.; Sun, L. Research on the means of fault intelligent diagnosis based on knowledge graph. *Shandong Commun. Technol.* **2019**, *39*, 3.
8. Shu, N.; Ge, Z.-J.; Luo, J.-W. A fault diagnosis system based on knowledge mapping. *Electron. Prod. Reliab. Environ. Test.* **2021**, *39*, 3.
9. Shenshouer. neo4j [EB/OL]. 2016. Available online: http://neo4j.com/ (accessed on 5 November 2022).
10. FlockDB Official. flockDB [EB/OL]. 2016. Available online: http://webscripts.softpedia.com/script/Database-Tools/FlockDB-66248.html (accessed on 8 November 2022).
11. Yang, B.; Cai, D.-F.; Yang, H. Advances in open information extraction research. *J. Chin. Inform.* **2014**, *28*, 10–19.
12. Bao, J.-P.; Liu, X.-D.; Shen, J.-Y. XML-based knowledge fusion and knowledge base organization. *Comput. Eng.* **2003**, *60*, 142.
13. Tan, W.-K.; Yang, Q.-F.; Chen, W. A Bayesian network-based fault location method for communication networks. *Comput. Appl.* **2018**, *38*, 217–220.
14. Wang, L.; Song, B.; Zhang, Y.-W. Fault path tracing algorithm for Bayesian networks considering parent nodes. *Comput. Sci. Explor.* **2018**, *12*, 1796–1805.
15. Zhai, S.; Tian, S.; Chen, Q.-Q. Research and application of Bayesian network-based reliability analysis method. *Comput. Meas. Control* **2020**, *28*, 262–266.
16. Fang, B.-W.; Huang, C.-K.; Li, Y. A Bayesian network-based quantitative analysis method for dynamic fault trees of complex systems. *J. Electr.* **2016**, *44*, 1234–1239.
17. Yu, J.-S.; Shen, L.; Tang, D.-Y. Performance evaluation of fault diagnosis system based on Bayesian network. *J. Beijing Univ. Aeronaut. Astronaut.* **2016**, *42*, 35–40.
18. Wang, C.-Y.; Xu, J.-J.; Yan, Z.-J. A dynamic Bayesian network fault diagnosis method based on timeliness analysis. *J. Dalian Univ. Technol.* **2019**, *59*, 201–210.
19. Cieslak, J.; Efimov, D.; Zolghadri, A.; Gheorghe, A.; Goupi, P.; Dayre, R. A method for actuator lock-in-place failure detection in aircraft control surface servo-loops. *IFAC Proc. Vol.* **2014**, *47*, 89–96. [CrossRef]
20. Peng, K.-X.; Ma, L.; Zhang, K. A review of fault detection and diagnosis techniques related to the quality of complex industrial processes. *J. Autom.* **2017**, *43*, 349–365.
21. Wu, K.; Wang, X.-Y.; Sun, J. A deep learning-based fault detection method. *Comput. Meas. Control* **2017**, *25*, 43–47.
22. Liu, X. Research on knowledge graph construction technology for fault analysis. *Beijing Beijing Univ. Posts Telecommun.* **2019**, *12*, 58–78.

23. Wen, F.; Cao, X.; Huang, H.-X. Research on recommendation algorithm based on knowledge graph. *J. Shenyang Univ. Technol.* **2021**, *40*, 5.
24. Deng, X.-B.; Fei, W.-L. Cable fault diagnosis analysis based on knowledge mapping. *Mod. Inf. Technol.* **2020**, *4*, 148–151.
25. Xiao, F.-D.; Wu, Y.-Z.; Shen, X.-H. Intelligent diagnosis of substation equipment faults based on deep learning and knowledge graph. *Power Constr.* **2022**, *43*, 9.
26. Tian, T.; Du, Y.; Yuan, Z.-X. Literature measurement and analysis software application–mechanical fault diagnosis as an example. *Softw. Guide* **2019**, *18*, 6.
27. Bai, Y.-S.; Xu, H.-S.; Wei, M. Research on generic learning path generation based on knowledge graph. *J. Mianyang Norm. Coll.* **2022**, *41*, 8.
28. Feng, C.-W. An important resource for natural language processing: "knowledge graph". *J. Foreign Lang.* **2021**, *5*, 99–125.
29. Xie, R.-Y. A review of new energy vehicle research in social science based on knowledge graph. *Glob. Sci. Technol. Econ. Outlook* **2020**, *35*, 11.
30. Gu, F.-J.; He, C.-M.; Pan, Q.-Y. A knowledge graph-based fault analysis method for 5G networks. *Radio Commun. Technol.* **2022**, *15*, 32–50.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.