IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 13 June 2022, accepted 10 July 2022, date of publication 14 July 2022, date of current version 20 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3190882

RESEARCH ARTICLE

Construction of Double-Layer Knowledge Coordination Network for Product Quality Problem Solving

YUBIN WANG, ZHAOGUANG XU, AND YANZHONG DANG
Institute of System Engineering, Dalian University of Technology, Dalian 116024, China

Corresponding author: Zhaoguang Xu (zhgxu@dlut.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 72001034 and Grant 71871041.

ABSTRACT Product quality problems are becoming increasingly complex with the increasing richness of product types and functions, making it challenging to solve quality problems with only personal knowledge effectively. Thus, we propose an approach to extract solution knowledge from quality problem-solving data to improve the efficiency of quality problem-solving. Specifically, by analyzing the characteristics of the data, a method for calculating the membership degree of quality knowledge nodes is first designed. The corresponding quality knowledge triads are constructed based on the relationship between the data structures. Subsequently, we build a double-layer knowledge coordination network (DL-KCN) composed of a problem layer and a solution layer. Finally, using the quality problem-solving data generated during the stamping production process of the body-in-white of an automobile manufacturing enterprise, a DL-KCN with practical application significance is constructed. The DL-KCN is beneficial for problem-solvers to address product quality problems.

INDEX TERMS Quality management, quality problem solving, double-layer knowledge coordination network, knowledge management, automobile manufacturing.

I. INTRODUCTION

With the improvement of people's living standards, consumers have put forward higher requirements for higher quality and various products, which makes the quality of products more and more concerned, and the quality problems become more complex, changeable, and challenging [18]. Therefore, to conform to the change of the times and promote manufacturing development, enterprises must pay more attention to product quality and constantly improve the efficiency of quality problem-solving.

Quality problem-solving (QPS) is to improve the existing quality status to achieve the expected quality status through related quality tools and methods [19]. Solutions to product quality problems are generally divided into containment measures and long-term solutions [14]. A containment measure is used to prevent new problems or cause more severe consequences while temporarily containing quality problems. A long-term solution involves carefully analyzing the cause behind the problem to solve the problem permanently. The key to containment measures is formulating temporary solutions to address the problems quickly. However, long-term solutions rely on an in-depth analysis of the root cause. In practice, providing interim containment measures and long-term solutions requires a wealth of experience and knowledge.

With the development of the economy and the change in the concept of consumption, the richness of product types and functions is increasing daily. Product quality problems are becoming increasingly more complex and challenging, presenting higher requirements for the richness of required knowledge to solve quality problems. However, the experience and knowledge of enterprise employees are greatly restricted by education and experience. With the informatization transformation of manufacturing enterprises, front-line employees usually use the quality management

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan Bu.

information system to record the QPS process and form QPS data. This type of data contains the rich experience of problem-solvers, which integrates group experience. If this part of experience knowledge can be effectively managed and used, it will play an essential role in solving new quality problems, a knowledge-intensive activity. The QPS process will go beyond personal experience limitations to help problem-solvers improve QPS efficiency and reduce the impact of organizational knowledge loss due to staff turnover.

With the advent of the era of quality 4.0, more and more scholars have realized the importance of data and knowledge in quality problem-solving [5]. For example, Xu *et al.* [21] introduced a data-driven method to automatically construct the component-failure mode matrix to increase the efficiency of creating failure mode and effects analysis. Taking advantage of the availability of big operations data, Ma, Li, and Thorstenson *et al.* [12] designed an extensive data-driven root cause analysis system utilizing Machine Learning techniques to improve the performance of root cause analysis.

However, few scholars have fully considered the characteristics of QPS data and thoroughly mined the knowledge in QPS data to help improve the efficiency of solving quality problems. Therefore, this paper intends to study from this perspective.

The front-line employees record QPS data according to the actual situation of the problem-solving process. Therefore, this kind of data has the following characteristics:

a) Objectivity and subjectivity coexistence: QPS data is an objective record, so the data have a certain degree of objectivity. However, problem data are manually recorded, simultaneously making the data subjective; different people describe the same problem differently.

b) Structured and unstructured data are intertwined: QPS data are usually recorded following the QPS process, so the data have a structured feature, mainly including structured fields such as "problem description", "cause analysis", and "solution." However, each field is filled by the quality problem-solver in a text description, so the data field also presents unstructured characteristics.

c) The relationship between data structures mainly includes the causal relationship between the two fields of "problem description" and "cause analysis" and the "problem-solution" relationship between the two fields of "problem description" and "solution."

d) The source and purpose of QPS data result in a high content of knowledge in the data.

In summary, QPS data are semi-standardized, semi-structured, subjective, and objective complex data containing a lot of experience knowledge, but it is challenging to use for problem-solvers and managers.

A Knowledge Graph is a data structure based on a graph model, effectively organizing and reusing scattered knowledge [13]. Hence, we use the knowledge triad to express knowledge in QPS data. The knowledge contained in each field can be expressed in the form of nodes, namely Quality Knowledge Nodes (QKNs), including "problem knowledge nodes," "cause knowledge nodes," and "solution knowledge nodes." At the same time, the relationship between fields can be expressed as edges, such as the causal relationship and the problem-solution relationship. Two nodes and one edge connecting them form a Quality Knowledge Triad (QKT).

QKTs consist of two parts: the first part is a causal knowledge triad based on causal relationship, and the second part is a problem-solution knowledge triad based on the problem-solution relationship. The two parts are independent but closely related to each other, and they are connected to form the Double-layer Knowledge Coordination Network (DL-KCN).

The coordination of DL-KCN is mainly reflected in three aspects. Firstly, the formation of DL-KCN results from comprehensive coordination of the experience knowledge of many subjects, and each subject has a contribution to the network. Secondly, DL-KCN can be used as an essential platform for communication and consensus when multiple relevant subjects cooperate to solve problems. Thirdly, the components of DL-KCN are independent but inseparable; mutual coordination and mutual action in the efficient solution of quality problems will not only help the quality problem-solver to uncover the root cause of quality problems and formulate long-term measures. Still, they will also help the quality problem-solver to quickly formulate the corresponding containment measures based on the status of the problem.

Therefore, we intend to discover knowledge from QPS data with the abovementioned characteristics. Firstly, we identify, find, and uncover QKNs in QPS data and then build QKTs based on the causal relationship and the problem-solution relationship between the data structures. On this basis, the DL-KCN for QPS is constructed.

DL-KCN is the externalization and expression of the knowledge in QPS data, which can provide extensive and comprehensive support for the quality problem-solvers to solve problems, transcend personal experience limitations, and improve the efficiency of solving quality problems effectively. In addition, this paper also enriches the relevant theories and methods of quality management, knowledge management, text analysis and knowledge mining.

II. LITERATURE REVIEW

The below scholars' research laid the foundation for this paper, providing research ideas and technical support.

In the 1870s, Taylor advocated establishing a "full-time inspection" link and, thus, published the book The Principles of Scientific Management" [17]. In the subsequent development process, the "full-time inspection" link gradually evolved into the quality control and quality assurance department [3]. Subsequently, quality management science developed rapidly and went through three stages: quality inspection, statistical quality control, and total quality management [11].

With the information transformation of manufacturing enterprises, the amount of data in quality management has increased exponentially. Data, information, and knowledge have increasingly become essential resources for quality management. Scholars have begun to focus on better-applying knowledge to optimize quality management processes and improve product quality. The intellectualization of quality management tools is also a research stream. For example, Yuen *et al.* [23] combined quality function deployment with fuzzy cognitive networks to improve criterion evaluation and analysis in product development. By mining unstructured and short-quality problem texts, Jiang *et al.* [8] introduced the fuzzy analytic hierarchy process and chaotic fuzzy regression when using quality function deployment for new product planning; Srikanth *et al.* [15] proposed an operational knowledge discovery and analysis framework of an Intelligent Quality Management expert system. Tang *et al.* [16] combined Taguchi methods in quality management with artificial intelligence techniques to create a hybrid method to improve overall system performance. Aslan *et al.* [1] used exponentially weighted moving averages to build a set of intelligent quality control charts. When the out-of-control signal in the intelligent control chart is triggered, the machine can be controlled and repaired in time. Some studies combine fuzzy evidence reasoning with belief rules to identify potential failure modes and their effects [9], [10]. It also introduces heuristic optimization methods and fuzzy methods in the design of experiments to carry out the best experimental design and determine the optimization criteria [2], [4].

In recent years, scholars have begun to study how to collect and use the knowledge in the discovery and solution process of quality problems, providing essential research ideas for this paper. For example, Xu *et al.* [20] proposed an intelligent QPS system applied to the automotive industry. Huang and Dang [7] aimed at manufacturing enterprises' demand for intelligent decision support in production practice starting from decision support for multi-problem processing and, based on the problem knowledge representation model in the domain, transforming decision support into graph search problems. Gai *et al.* [6] started from the connotation and essential characteristics of tacit knowledge, based on the practice theory of "the unity of knowledge and action," took the problems encountered in actual work and the process of solving as the starting point, and proposed a problem-driven method for acquiring tacit knowledge. Xu *et al.* [21] proposed a bipartite graph clustering method and used this method to mine practical knowledge applied to temporary containment problems.

While the above scholars' research laid the foundation for this paper, QPS data are a type of semi-standardized, semi-structured, subjective and objective complex data. How to identify and mine the relevant knowledge in the data according to the characteristics of this type of data and to construct DL-KCN for QPS still need further study.

III. CONCEPTUAL MODEL OF DL-KCN FOR QPS

The conceptual model of the quality knowledge network is illustrated in Figure 1:

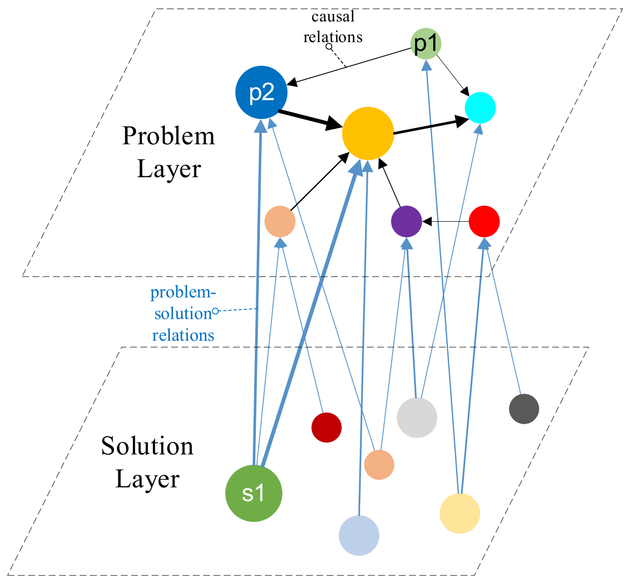

FIGURE 1. The conceptual model of DL-KCN for QPS. The figure shows the basic structure, expression form and the relationship between nodes of DL-KCN.

DL-KCN is a three-dimensional network composed of a problem layer and a solution layer. The problem layer is a network formed by quality problem nodes and their directed edges. The solution layer is a collection of discrete solutions. The relationship between quality problems is called a causal relation in the problem layer. The relationship between quality problem nodes p1 and p2 is a causal relation. Since node p1 points to node p2, p1 represents the cause, and p2 represents the effect. Each causal relation and both problem nodes form a causal knowledge triad. The problem layer and the solution layer are related to each other based on the problem-solution relationship. For example, s1 in the solution layer points to p2 in the problem layer, indicating that s1 is the solution for p2. Each problem-solution relationship and the problem and solution nodes form a problem-solution knowledge triad.

In the DL-KCN, node size represents the frequency of the node, while the edge thickness represents the frequency of the edge. The larger the node, the higher the frequency of the problem. The thicker the edge in the problem layer, the higher the frequency of the corresponding causal relationship. The wider the edge between the two layers, the higher the number of times the solution connected by the edge is used to solve the corresponding problem.

The DL-KCN can effectively help improve the efficiency of QPS. We can first identify the problem node in the problem layer when formulating containment measures according to the actual problems encountered. Second, we can directly discern suitable containment measures in the solution layer, highlighting the problem node. When formulating long-term solutions, problem-solvers can first uncover the root cause of the problem through the causal connection in the problem layer and then find appropriate long-term solutions by mapping between the solution layer and the problem layer.

In summary, nodes represent quality problems and solutions, and edges represent the relationship between nodes in DL-KCN. Simultaneously, due to different overlapping accumulation frequencies, nodes and edges have corresponding parameters. Therefore, DL-KCN can be expressed by equation (1).

$$G = (N, E, P) \tag{1}$$

In equation (1), G represents DL-KCN; N represents the collection of nodes; E is the collection of edges; P represents the set of parameters in DL-KCN.

A. NODES IN DL-KCN

Nodes in DL-KCN represent an objective phenomenon, which can be divided into two categories: the nodes in the problem layer, which represent quality problems, denoted as $n_i^p, n_i^p \in N, i = 1, 2, \cdots, m_p$; and the node in the solution layer, which represents solutions to the quality problem and is denoted as $n_j^s, n_j^s \in N, i = 1, 2, \cdots, m_s$.

B. EDGES IN DL-KCN

Edges in DL-KCN can represent both causality and problem-solution relationships. Directed edges between nodes represent both.

The directed edges in the problem layer indicate causality, denoted as $e_{ij}^p = \langle n_i^p, n_j^p \rangle, e_{ij}^p \in E$. In the equation, e_{ij}^p represents the directed edge from n_i^p to n_j^p; n_i^p represents a cause for n_j^p; and n_j^p represents a problem caused by n_i^p.

The directed edge from the solution level to the problem level indicates the problem relationship, denoted as $e_{ij}^{sp} = \langle n_i^s, n_j^p \rangle, e_{ij}^{sp} \in E$. In the equation, e_{ij}^{sp} represents the directed edge from n_i^s to n_j^p; n_i^s represents a solution for n_j^p; and n_j^p p represents a quality problem that n_i^s can solve.

C. PARAMETERS IN DL-KCN

Both nodes and edges have numerical attributes, and these values can be obtained through statistics of QKTs. These parameters are an expression of DL-KCN from a quantitative perspective. The meaning of each parameter is as follows:

Node Frequency represents the number of times a node appears in the DL-KCN, represented by G. For node n_i^p, its frequency is denoted as $f(n_i^p)$. The larger the value of $f(n_i^p)$, the more frequent the problem, which is represented by the node n_i^p. Therefore, more attention should be placed on the quality n_i^p. Correspondingly, the greater the solution node's frequency, the more frequent the solution appears and, therefore, more attention should be placed on it.

Edge Frequency represents the number of times an edge is repeated in the DL-KCN, represented by G. For edge e_{ij}^p, the frequency is denoted as $f(e_{ij}^p)$. The larger value of $f(e_{ij}^p)$, the higher the frequency of the relationship represented by the edge e_{ij}^p. Therefore, more attention should be placed on the edge.

IV. EXTRACTION OF QKNS

A. MEMBERSHIP FUNCTION OF QKNS

QPS data is recorded by employees of the enterprise, which makes the candidate words of QKNs mostly vague, unclear and inaccurate, such as: "Engine Hood" may be described as "ENG hood", "EH", which is difficult to extract directly from the text description.

Fuzzy Mathematics is a mathematical method to study Fuzzy phenomena or concepts, which can effectively analyze and learn uncertain things with the help of membership functions [24]. Therefore, this paper draws on the theoretical method of fuzzy mathematics to design the membership function to help extract QKNs.

The source and purpose of the QPS data determine that the data is knowledge-intensive data, with a high rate of knowledge contained in the data.

Therefore, there are many descriptive words related to quality problem-solving knowledge in this kind of data, such as descriptive words of problems and causes. So, a word or phrase frequently occurring in a single text is more likely to be a QKN (Characteristic 1). Not only that, in the production process of enterprises, quality problems are repeated. So, the more frequently a word or phrase that frequently occurs in multiple texts is more likely to be a QKN (Characteristic 2).

Combining the above characteristics, the membership function of QKNs is proposed as follows:

$$TF_w = \frac{\sum n_{w,d}}{\sum_k n_{k,d}} \tag{2}$$

$$PDF_w = arcsin \frac{\sum D_w}{\sum\limits_m D_m} \tag{3}$$

$$MEM_w = f(MEM_w) * TF_w * PDF_w \tag{4}$$

In equation (2), TF_w is the word frequency of the word or phrase w in the single text; n_w is the number of times of the word or phrase w appears in the document, denoted by d; and $n_{k,d}$ is the number of the word or phrase in the document d.

In equation (3), PDF_{word} is the calculation method of the frequency of the word or phrase w in the multiple texts designed according to Characteristic 2 of QPS data; D_w is the document containing the word or phrase w; and D_m is the number of documents.

In equation (4), MEM_w is the membership degree of QKNs of the word or phrase w; $f(MEM_{word})$ is the reconciliation formula designed according to the membership degree of QKNs, and its purpose is to prevent the membership of QKNs from becoming too large or excessively small.

B. EXTRACTION METHOD OF QKNS

QPS data is recorded by all front-line employees who found or solved the quality problem. As a result, different employees describe the same problem in different ways. Most of

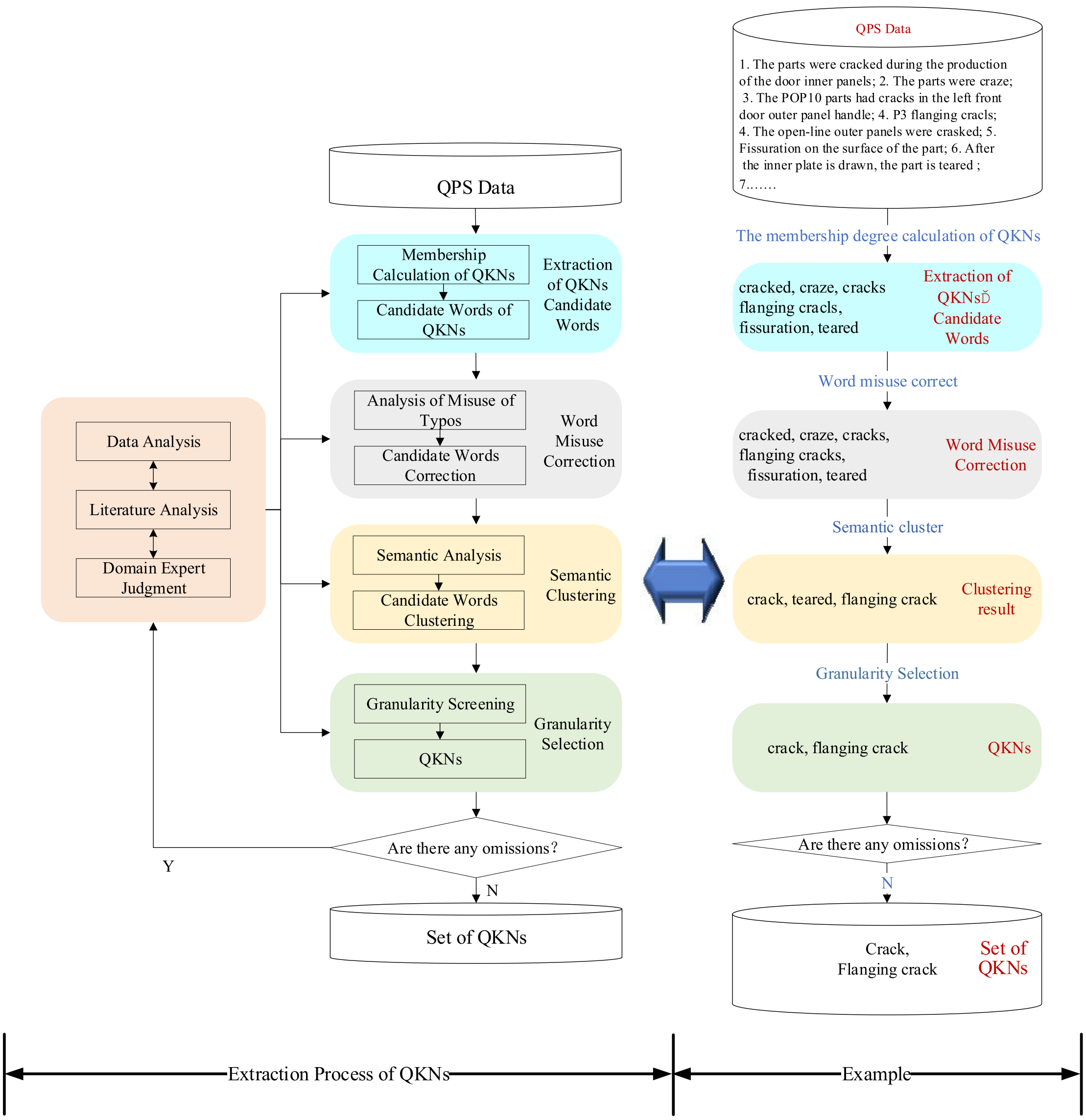

FIGURE 2. **The extraction process of QKNs. The left part of the figure is the extraction process of QKNs, and the right part of the figure is an example of the extraction process.**

these descriptions tend to be colloquial, with typos and misuses. The QPS data of various industries also has specific domain characteristics. All of these factors make it challenging to extract QKNs from the QPS data at once through the existing text analysis methods. Therefore, this paper uses a multi-stage method to extract QKNs from the QPS data layer by layer. The specific process is shown in Figure 2:

The specific extraction steps are as follows:

Step 1 (Extract Candidate Words of QKNs): The aim of this step is mainly to find the initial vocabulary describing the quality problem, the cause, and its solution from the QPS data through data preprocessing and the calculation of the membership degree of QKNs, and use it as a candidate word for the node. For example: find the word "cracked," which describes the quality problem from the problem description "The product cracked during the production of the door inner panel." The specific method is:

The text description of each field in QPS data is composed of vocabulary; suppose each field in QPS data S can be expressed as a collection of vocabulary w_i, that is $S = \{w_1 w_2 \cdots w_i \cdots w_n$. After the field S undergoes data preprocessing, word segmentation, and calculation of the membership degree of QKNs of the vocabulary w_i, the vocabulary is sorted according to the degree of membership, and the original field S can be expressed as equation (5).

$$S' = \{w_{i_1}|mem_1, w_{i_2}|mem_2, \cdots w_{i_n}|mem_n \tag{5}$$

In equation (5), $w_{i_n}|mem_n$ indicates the membership degree of the QKNs of a vocabulary in a field.

According to the data structure and characteristics of the specific QPS data, some vocabularies before sorting can be selected as the candidate words of QKNs.

Step 2 (Word Misuse Correction): Many problem-solvers form the QPS data. As a result, there are many typos and misuses in the data, such as "necking down" versus "nacking donw," and "cracks" versus "carcks". Typos significantly affect the word meaning similarity calculation of candidate words of QKNs. Therefore, it is necessary to use typo correction tools, such as Enchant, Check, and PyCorrector, to correct typos and misused words in the candidate words of QKNs.

Step 3 (Semantic Clustering of Node Candidate Words): The initial description vocabulary of QKNs that are extracted from the QPS data varies, such as "cracks," which may be described as "cracking," "part of the crack," "parts of the crack," and "cracked." It is, therefore, necessary to use text analysis methods, such as similarity calculation, synonym analysis, and word meaning analysis, to analyze and cluster different descriptions and tenses of the same problem. Ensuring that the descriptions are unified, and semantic clustering is realized. The specific method is:

Candidate words of QKNs are compared in pairs, the semantic similarity is calculated, and the semantic similarity matrix SM is constructed, as shown in equation (6).

$$SM = \begin{bmatrix} 1 & m_{12} & \cdots & m_{1n} \\ - & 1 & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ - & - & \cdots & 1 \end{bmatrix} \tag{6}$$

In equation (6), m_{ij} is the similarity of the vocabulary i and the vocabulary j.

According to the specific characteristics of QPS data, set the similarity threshold ϑ, and classify the words with a similarity greater than the threshold ϑ into one category to complete the semantic clustering of node candidate words.

Step 4 (Granularity Selection of QKNs): There are categorical relationships among candidate words of QKNs, such as "cracking," "tearing" and "flanging cracks," which are all crack problems. Specifically, the crack problems expressed by each word are different. It is then necessary to perform inductive analysis on the candidate words of QKNs according to the granularity of the constructed DL-KCN, that is, to select the granularity to complete the extraction of QKNs.

Step 5 (Verification of QKNs): Validate the extracted QKNs by checking whether there are any omissions or missing descriptions of problems, causes, and solutions. If there are no omissions, then prove that the extracted QKNs can be accurate and complete expressing all the quality problems, causes, and solutions in the current QPS data. In other words, through the extracted QKNs, DL-KCN can be constructed to describe the quality problems, causes, and solutions accurately. Otherwise, it can be supplemented and perfected by data analysis, literature analysis, domain expert judgment, and other methods until there are no omissions.

V. CONSTRUCTION OF DL-KCN FOR QPS

A. CONSTRUCTION OF QKTS

To construct QKTs, first of all, the text descriptions of "problem description field," "cause analysis field," and "solution field" need to be replaced with the extracted QKNs; then using the relationship between the fields as the basis, the QKTs are constructed.

QKNs are extracted layer by layer from QPS data. Therefore, QKNs and the candidate words in the corresponding QPS data can be correlated, and an associated dictionary set can be constructed, as shown in equations (7) and (8).

$$D^W = \left\{D_1^W, D_2^W, \cdots D_i^W, \cdots D_n^W\right\} \tag{7}$$

$$D_i^W = W_i : \left[w_1^i, w_2^i, \cdots, w_m^i\right] \tag{8}$$

In equations (7) and (8), D^W is the associated dictionary set; D_i^W is the related dictionary of the QKN and its initial description vocabulary; W_i is one of the QKNs; and $\left[w_1^i, w_2^i, \cdots, w_m^i\right]$ is the description vocabulary in QPS data corresponding to the QKN W_i.

Supposing the QPS data $\mathbb{S} = \left\{S_1, S_2 \cdots, S_j, \cdots S_u\right\}$, S_j is one of the quality problems records; the record mainly contains the problem description field, the cause analysis field and the solution field. Thus, this record can be expressed as $S_j = \left\{S_j^{Pb}, S_j^{Cau}, S_j^{Sol}\right\}$. $\left\{S_j^{Pb}, S_j^{Cau}, S_j^{Sol}\right\}$ are the corresponding problem description field, cause analysis field, and solution field in the problem record S_j. The construction algorithm of QKTs is shown in Table 1.

B. CONSTRUCTION OF DL-KCN FOR QPS

The construction of DL-KCN for QPS is mainly divided into two steps: firstly, the causal knowledge triads are analyzed and processed, connected to construct a causal knowledge network, and then the problem knowledge triads and the causal knowledge network are mapped and combined to construct the DL-KCN.

Step 1 (Construction of Causal Knowledge Network): Suppose that G^μ, G^η are two causal knowledge triads, as shown in equations (9) and (10).

$$G^\mu = (N^\mu, E^\mu, P^\mu) \tag{9}$$

$$G^\eta = (N^\eta, E^\eta, P^\eta) \tag{10}$$

In equation (9), N^μ represents the set of nodes in G^μ; E^μ represents the set of edges in G^μ; and P^μ represents the set of parameters in G^μ. In equation (10), N^η represents the set of nodes in G^η; E^η represents the set of edges in G^η; and P^η represents the set of parameters in G^η.

Suppose n_i^μ and n_t^η are the nodes in G^μ and G^η, respectively. Wherein $n_i^\mu \in N^\mu, i = 1, 2; n_t^\eta \in N^\eta, t = 1, 2$.

Suppose e_{ij}^μ and e_{tk}^η are the edges in G^μ and G^η, respectively. Wherein $e_{ij}^\mu \in E^\mu, e_{ij}^\mu = \langle n_i^\mu, n_j^\mu \rangle$, e_{ij}^μ means directed edges from n_i^μ to n_j^μ in G^μ; $e_{tk}^\eta \in E^\eta, e_{tk}^\eta = \langle n_t^\eta, n_k^\eta \rangle$, e_{tk}^η means directed edges from n_t^η to n_k^η in G^η.

The nodes in G^μ, G^η are not merged, so the node frequency and edge frequency are the initial values 1, that is $f^\mu\left(n_i^\mu\right) = 1, f^\eta\left(n_t^\eta\right) = 1; f^\mu\left(e_{ij}^\mu\right) = 1, f^\eta\left(e_{tk}^\eta\right) = 1$. Wherein $f^\mu\left(n_i^\mu\right) \in P^\mu, f^\mu\left(e_{ij}^\mu\right) \in P^\mu; f^\eta\left(n_t^\eta\right) \in P^\eta, f^\eta\left(e_{tk}^\eta\right) \in P^\eta$.

TABLE 1. Algorithms for constructing QKTs.

Input: QPS data $\mathbb{S}$, Related Dictionary Set D^W
Output: Causal knowledge triad RDF^{PC}, problem-solution knowledge triad RDF^{PS}

for S_j in $\mathbb{S}$	#Read a record of QPS data $\mathbb{S}$
for D_i^W in D^W	# Read a record of the related dictionary set D^W
for w_q^i in D_i^W.values	#Read the initial description vocabulary of the QKN W_i
if w_q^i in S_j^{Pb} **then**	#Determine whether the initial description vocabulary of W_i is in the problem description field of S_j
Build $W^{Pb} = \{D_i^W.\text{keys} : S_j^{Pb}\}$	#Establish the associated dictionary of QKNs W_i and the problem description of S_j
end if	
if w_q^i in S_j^{Cau} **then**	#Determine whether the initial description vocabulary of W_i is in the reason analysis field of S_j
Build $W^{Cau} = \{D_i^W.\text{keys} : S_j^{Cau}\}$	#Establish the associated dictionary of QKNs W_i and the cause analysis of S_j
end if	
if w_q^i in S_j^{Sol} **then**	#Determine whether the initial description vocabulary of W_i is in the solution field of S_j
Build $W^{Pb} = \{D_i^W.\text{keys} : S_j^{Pb}\}$	#Establish the associated dictionary of QKNs W_i and the solution of S_j
end if	
end for	
end for	#Traverse QKNs and find all the QKNs related to each field of S_j
for W_δ^{Pb} in W^{Pb}.keys	#Read one of QKNs W_δ^{Pb} which is related to the problem description of S_j
for W_ξ^{Cau} in W^{Cau}.keys	#Read one of QKNs W_ξ^{Cau} which is related to the cause analysis of S_j
Build $RDF^{PC} \Leftarrow (W_\xi^{Cau}, W_\delta^{Pb})$	#Establish causal knowledge triples
end for	
for W_ζ^{Sol} in W^{Sol}.keys	#Read one of QKNs W_ζ^{Sol} which is related to the solution of S_j
Build $RDF^{PS} \Leftarrow (W_\zeta^{Sol}, W_\delta^{Pb})$	#Establish problem-solution knowledge triples
end for	
end for	
end for	

If node n_i^μ is the same as the node n_t^η, then these two nodes can be merged to form a new node $n_{it}^{\mu\eta}$. The frequency $f^{\mu\eta}\left(n_{it}^{\mu\eta}\right)$ of the new node changes, as shown in equation (11).

$$\begin{cases} n_{it}^{\mu\eta} = n_i^\mu \oplus n_t^\eta \\ f^{\mu\eta}\left(n_{it}^{\mu\eta}\right) = f^\mu\left(n_i^\mu\right) + f^\eta\left(n_t^\eta\right) \end{cases} \tag{11}$$

In equation (11), The symbol $\oplus$ indicates the merger of nodes.

If edge e_{ij}^μ and edge e_{tk}^η are the same, that is, the start node n_i^μ of edge e_{ij}^μ is the same as the start point n_t^η of edge e_{tk}^η and the end node n_j^μ of edge e_{ij}^μ is the same as the end node n_k^η of side e_{tk}^η, then edge e_{ij}^μ and edge e_{tk}^η can be merged to form an edge $e_{ij}^{\mu\eta}$ and the frequency $f^{\mu\eta}\left(e_{ij}^{\mu\eta}\right)$ of the new edge also changes, as shown in equation (12).

$$\begin{cases} e_{ij}^{\mu\eta} = e_{ij}^\mu \oplus e_{tk}^\eta \\ f^{\mu\eta}\left(e_{ij}^{\mu\eta}\right) = f^\mu\left(e_{ij}^\mu\right) + f^\eta\left(e_{tk}^\eta\right) \end{cases} \tag{12}$$

The above method is used to merge all the same nodes and edges in the causal knowledge triads and count their node frequency and edge frequency to construct the causal knowledge network $G^{ces} = (N^{ces}, E^{ces}, P^{ces})$.

Step 2 (Construction of DL-KCN): Suppose that G^τ is a problem-solution knowledge triad, and G^{ces} is the causal knowledge network constructed, shown in equations (13) and (14).

$$G^\tau = (N^\tau, E^\tau, P^\tau) \tag{13}$$

$$G^{ces} = (N^{ces}, E^{ces}, P^{ces}) \tag{14}$$

In equation (12), N^τ represents the set of nodes in G^τ; E^τ represents the set of edges in G^τ; and P^τ represents the set of parameters in G^τ. In equation (13), N^{ces} represents the set of nodes in G^{ces}; E^{ces} represents the set of edges in G^{ces}; and P^{ces} represents the set of parameters in G^{ces}.

Suppose that n_s^τ, n_p^τ are the solution node and the problem node in G^τ respectively, and e_{sp}^τ is s directed edge in G^τ.

Suppose that n_c^{ces} is the node in the causal knowledge network G^{ces}, $n_c^{ces} \in N^{ces}$, $c = 1, 2, \cdots, m$.

Nodes in G^τ are not merged, and the node frequency and edge frequency are both the initial value 1, that is $f^\tau\left(n_s^\tau\right) = 1$, $f^\tau\left(e_{sp}^\tau\right) = 1$.

If the problem node n_p^τ in the problem-solution knowledge triad is the same as the node n_c^{ces} in the causal knowledge network, then the problem node n_p^τ is integrated into the causal knowledge network node n_c^{ces}, as shown in equation (14).

$$n_c^{ces} = n_c^{ces} \odot n_p^\tau \tag{15}$$

In equation (15), the symbol $\odot$ indicates the integration of nodes; two nodes are integrated into one, and the frequency does not change.

The above method merges other problem-solution knowledge triads with the causal knowledge network. If another solution node n_s^τ is the same as another solution node n^*, then these two nodes are merged to form a new node $nn_s^{\tau *}$. The frequency $f\left(n_s^{\tau *}\right)$ of the new node changes, as shown in equation (16).

$$\begin{cases} n_s^{\tau *} = n_s^\tau \oplus n^* \\ f\left(n_s^{\tau *}\right) = f^\tau\left(n_s^\tau\right) + f\left(n^*\right) \end{cases} \tag{16}$$

As with the same solution nodes processing method, if the edge e^* and edge e_s^τ are the same edge; that is, the starting node and pointing node of the edge e^* is the same as the starting node and pointing node of the edge e_{sp}^τ; then edge e^* and edge e_s^τ can be merged to form side $e_s^{\tau *}$, and the frequency $f\left(e_s^{\tau *}\right)$ of the new edge also changes, as shown in equation (17).

$$\begin{cases} e_s^{\tau *} = e_s^\tau \oplus e^* \\ f\left(e_s^{\tau *}\right) = f^\tau\left(e_s^\tau\right) + f\left(e^*\right) \end{cases} \tag{17}$$

All the problem-solution knowledge triads can be transversed, and their node frequency and edge frequency can be counted. Then DL-KCN for QPS can be constructed.

VI. CASE STUDY

A. CASE BACKGROUND

This paper uses the quality problems that occur during the stamping production process of the body-in-white of an automobile manufacturing enterprise as the case and constructs DL-KCN for solving actual product quality problems. The validity, rationality, and scientificity of the method proposed in this paper are verified.

In the process of stamping and manufacturing the body-in-white, quality problems such as "cracks," "wrinkles," and "napping" often occur. To solve product quality problems and improve production efficiency and product quality, front-line employees in the stamping workshop usually record the quality problem resolution process and form "Record of Product Quality Problems in the Stamping Workshop," a type of QPS data.

The table contains many fields, such as problem description, cause analysis, and solutions, and contains much experience knowledge. However, due to the differences in each employee's writing styles, the record tables are expressed differently, which has caused great difficulties for the construction of DL-KCN. Therefore, this paper selects the 2927 QPS data recorded in the "Record of Product Quality Problems in Stamping Workshop" and uses the theoretical methods proposed in this paper to uncover QKNs and, based on the structural relationship between the field in data, constructs DL-KCN of practical application significance; and preliminary verification of the effectiveness of DL-KCN with the remaining 216 pieces of data.

B. EXTRACTION OF QKNS

According to the QKNs extraction method described in this paper, firstly, QPS data should be preprocessed, such as removing stop words and word segmentation. According to the calculation method of membership degree of QKNs, each knowledge membership degree of the vocabulary is calculated. The calculation results of some vocabulary items are shown in Table 2.

The complexity of the problem records in the "Record of Product Quality Problems in Stamping Workshop" is different, and the vocabulary in each field varies greatly, most of which range from 5 to 100. The QKNs contained in texts of the various fields will also be different.

Therefore, this paper selects the value of λ stepwise according to the difference in the vocabulary content in each field. The specific method is shown in equation (18).

$$\lambda = \begin{cases} 1 & n_w \in (0, 5] \\ 2 & n_w \in (5, 20] \\ 4 & n_w \in (20, 50] \\ 5 & n_w \in (50, 90] \\ 6 & n_w \in (90, +\infty] \end{cases} \tag{18}$$

In equation (18), n_w indicates the vocabulary contained in the field.

By using the above method, 19,219 vocabularies were selected as candidate vocabulary for KQNs, and word frequency statistics were performed on them. Some results are shown in Table 3.

With the help of text analysis tools, the misuse and correction of QKNs candidate vocabulary, such as "carck" should be corrected to "crack," and "nack" should be corrected to "neck." Then, we compared the semantic similarity of the candidate words of QKNs and constructed the semantic similarity matrix, part of which is shown in Table 4.

According to the characteristics of QPS data, the similarity threshold $\vartheta = 0.50$ is selected for semantic clustering. Then, through the granular selection and combining common vocabulary and expert opinions, QKNs is constructed and

TABLE 2. Schematic table for calculating the membership degree of QKNs.

Original Data	Data preprocessing	Membership Degree of QKNs
There is a pit near the skylight when the skylight is produced.	'pit', 'skylight', 'skylight', 'produced'	('produced', 0.0135), ('skylight', 0.0095), ('pit', 0.1204)
Scrap in the middle of the door frame cannot be cut off.	'scrap', 'door frame', 'cannot be cut off'	('scrap', 0.1735), ('door frame', 0.0031), ('cannot be cut off', 0.2464),
⋮	⋮	⋮
The bottom of the tank opening is cracked.	'bottom', 'tank opening', 'cracked'	('tank opening', 0.0031), ('bottom', 0.0027), ('cracked', 0.5027)

TABLE 3. Candidate vocabulary list of QKNs.

Candidate vocabulary of QKNs	Frequency
Cracks	638
Wrinkles	481
Necking	376
Napped	342
Folds	295
⋮	⋮
Teared	290

TABLE 4. Semantic similarity matrix.

—	Cracks	Part craze	Flanging cracks	Crarizonae	⋯	Stifled-scrap	Scrap does not slide	Scrap bunged up
Cracks	1	0.919	0.491	0.888	⋯	0.049	0.079	0.099
Part craze	-	1	0.411	0.519	⋯	0.089	0.057	0.045
Flanging cracks	-	-	1	0.417	⋯	0.042	0.079	0.051
Crarizonae	-	-	-	1	⋯	0.096	0.075	0.034
⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮
Stifled-scrap	-	-	-	-	⋯	1	0.427	0.917
Scrap does not slide	-	-	-	-	⋯	-	1	0.49
Scrap bunged up	-	-	-	-	⋯	-	-	1

TABLE 5. Schematic table for extracting QKNs.

QKNs	Granularity selection	Semantic clustering	Node description vocabulary revision	Node initial description vocabulary
cracks flanging cracks	cracks flanging cracks	crack, tear, flanging crack	cracks, crack, part crack, tears, teared, tear, craze, crarizonae, flanging cracks, flanging gap, flanging craze	cracks, carck, part crcak, tear, tears, taered, crase, craze, craza, crarizonae, carrizonae, flanging cracks, flanging gap, flanging craze, flanging carcks
wrinkles	wrinkles	wrinkles	wrinkle, wrinkles, part wrinkles, crinkle, crease, crumpled, crumple	wrinkle, wrinkles, wirnkle, part wrinkles, crinkle, cirnkle, crease, craese, crumpled, crumple
⋮	⋮	⋮	⋮	⋮
necking	necking	necking, reduced section	necking, necking down, collaring, reduced section	necking, nacking, necking down, collaring, colaring, reduced section

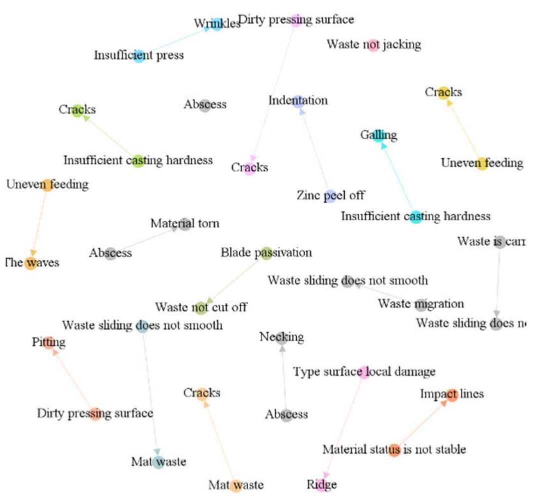

FIGURE 3. Some examples of causal knowledge triads. The cause node points to the problem node in the causal knowledge triads.

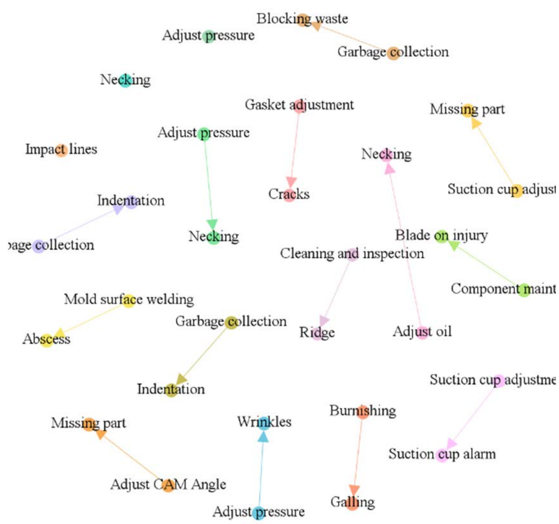

FIGURE 4. Some examples of problem-solution knowledge triads. The solution node points to the problem node in the problem-solution knowledge triads.

verified. Finally, 543 QKNs were extracted. The extraction of some QKNs and their results are shown in Table 5.

C. CONSTRUCTION OF QKTS

According to the method of constructing QKTs described in this article, causal knowledge triads can be constructed based on the causal relationship between structures, such as "impact lines cause wrinkles." Many causal knowledge triads can be aggregated together to build the causal knowledge network. Since the extracted causal knowledge triads are too large, with a total of 3408 triads, it is not easy to display them, so only part of them are selected for display, as shown in Figure 3.

Based on the problem-solution relationship between the structures, the problem-solution knowledge triads can be constructed, such as polishing as a solution to galling. A large number of problem-solution knowledge triads aggregated with causal knowledge networks can complete DL-KCN. Since the extracted problem-solution knowledge triads are too large, with a total of 2789 triads, it is not easy to display

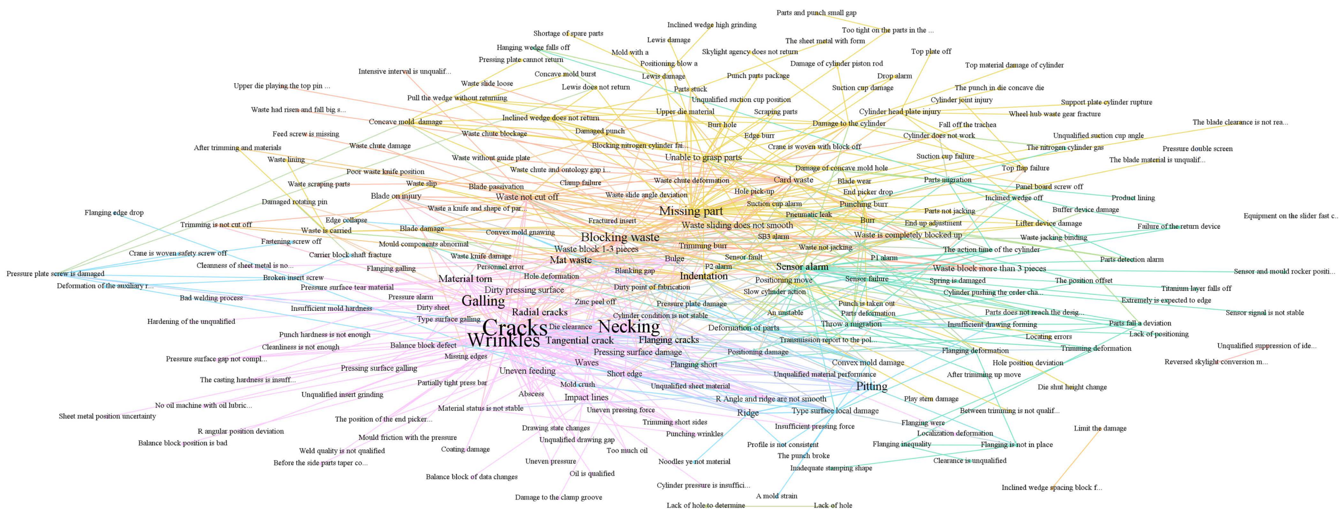

FIGURE 5. The causal knowledge network. The larger the vertex is, the more times the problem/cause occurs.

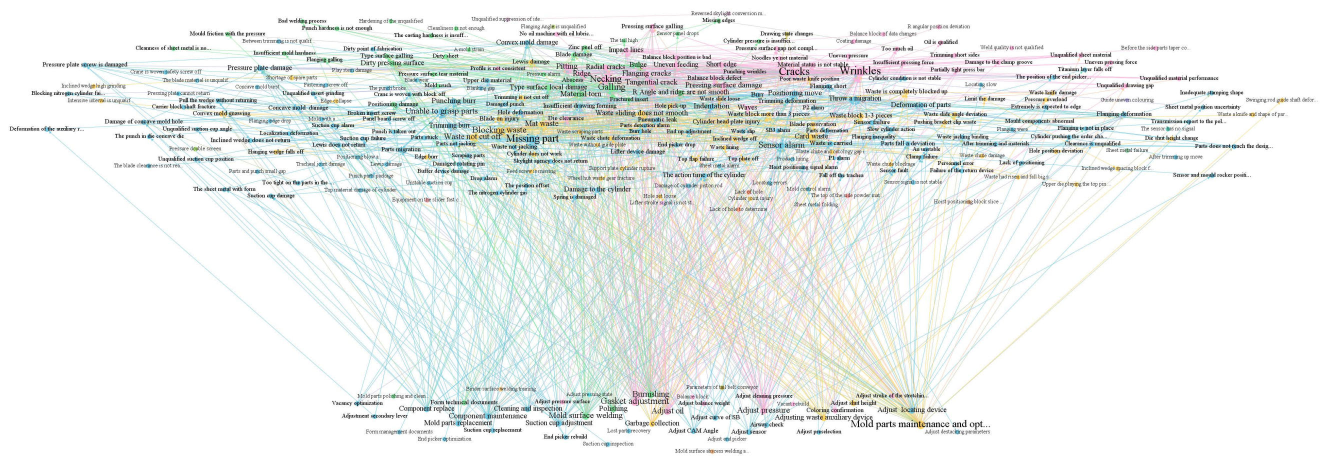

FIGURE 6. DL-KCN for QPS. The bottom half of the graph is the set of solutions and the top half is the causal network.

them, so only part of them are selected for display, as shown in Figure 4.

D. CONSTRUCTION OF DL-KCN

According to the construction of DL-KCN described in the paper, the same nodes in the causal knowledge triads should be first merged, and the node frequency and edge frequency are counted to obtain the causal knowledge network, as shown in Figure 5.

Then, the problem-solution knowledge triad should be mapped into the causal knowledge network, and the corresponding node frequency and edge frequency are counted to obtain the DL-KCN for QPS, as shown in Figure 6.

In the problem layer of DL-KCN, the frequency of nodes such as cracks, wrinkles, and necking are relatively high, indicating that these problems frequently appear in the stamping production process of automobiles body-in-white, and more attention should be placed on it. Furthermore, in the solution layer of DL-KCN, the frequency of solution nodes such as gasket adjustment, pressure adjustment, and polishing is relatively high, suggesting that these solutions are often used and can be focused on.

The DL-KCN can also effectively support the problem-solvers to solve product quality problems. We take the problem ''wrinkles'' as an example to illustrate how the DL-KCN works. As shown in Figure 7, a total of 15 methods can be found through the DL-KCN to solve the problem of ''wrinkles''. It is also apparent that ''polishing'' and ''gasket adjustment'' are commonly used solutions to tackle this issue, which can be given priority when containment measures are required.

Through analyzing the causal relationship in the problem layer of Figure 7, it is not difficult to find that ''Uneven feeding'' and ''Dirty press surface'' are the leading causes of ''wrinkles''. If it is determined that the cause of the problem is ''Uneven feeding'', ''Adjust oil'' and ''Adjust pressure'' can be prioritized as solutions. If ''Dirty press surface'' is the cause, ''Burnishing'' and ''Mold surface welding'' should be given priority. It also can be analyzed through the causal relationship that the ''wrinkles'' problems

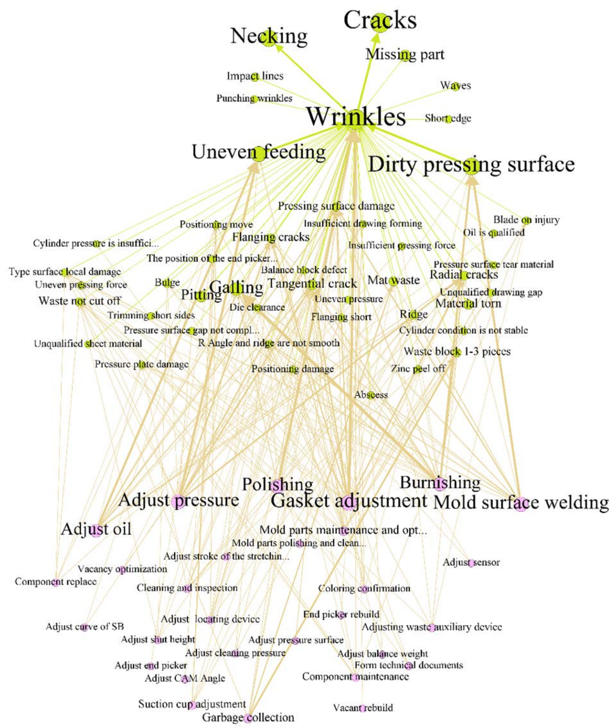

FIGURE 7. Take the problem "wrinkles" as an example to show DL-KCN.

may further lead to "cracks" and "necking," which can be prevented.

There are a few discrete knowledge triads in DL-KCN, for example, "Reversed Skylight Conversion Mechanism" due to "Unqualified Suppression of Identification." Some problems have no corresponding solutions, such as "Pressing Plate Cannot Return" and "Waste Scraping Parts." This situation may be caused by the occasional occurrence of such problems and rarely encountered. Moreover, insufficient data may also be a cause. As production continues, the amount of data gradually increases, and the quality knowledge network will be further improved.

Finally, the experts and front-line employees of the example enterprise evaluated DL-KCN through the actual situation of production, and they all agreed that DL-KCN could improve the problem-solving efficiency by at least 15%, and the effectiveness of DL-KCN would gradually be enhanced with the increase of QPS data volume caused by production going on.

However, the specific change in quality problem-solving efficiency still needs to be determined by the change of production data after DL-KCN is popularized and used by the example enterprise.

VII. CONCLUSION & CONTRIBUTION

QPS data is text recorded by problem-solvers in QPS, and it contains a large amount of experience knowledge. However, it is difficult to use directly. Based on the characteristics of QPS data, this paper proposes a method for mining theoretical and practical knowledge from QPS data. In this method, we first calculate the membership degree of QKNs. On this basis, we use a multi-stage method to extract QKNs from QPS data layer by layer.

Moreover, based on the causal relationship and the problem-solution relationship between the data structures, the QKTs are constructed. Then we merge the QKTs to form the DL-KCN for QPS. Finally, we use QPS data generated from the stamping process of the body-in-white of an automobile manufacturing enterprise to construct DL-KCN with practical application significance to verify the method's validity, rationality, and scientificity.

The contribution of this paper is as follows:

a) DL-KCN was proposed and constructed using the knowledge in QPS data, providing new thinking and methods for quality problem-solving. DL-KCN contains the problem layer and the solution layer. The problem layer is a causal network composed of causal knowledge triads, through which employees can deeply analyze quality problems. After finding the cause of the problem, the corresponding problem solution can be quickly found in the solution layer through the problem-solution relationship between the problem layer and the solution layer. DL-KCN provides extensive and comprehensive support for employees to solve the problem and contribute to improving the efficiency of product quality problem-solving.

b) According to the characteristics of QPS data, using the theoretical method of fuzzy mathematics for reference, the new membership function is designed. On this basis, this paper presents a multi-stage way to extract QKNs from QPS data layer by layer and depending on the structural characteristics of QPS data, QKTs are constructed. This method is not only suitable for QPS data but also has a particular reference value and significance for other text data with similar characteristics. This method also enriches the relevant theories and methods of text analysis and knowledge mining.

However, we still need to employ the scientific and reasonable effectiveness verification method to determine whether the DL-KCN can help enterprises solve product quality problems in the actual application according to the substantial production data feedback from enterprises. Meanwhile, the causal relationship in DL-KCN is complex and changeable. Future studies may identify the root cause of product quality problems and provide appropriate reasoning based on DL-KCN.

ACKNOWLEDGMENT

The authors would like to thank the editors and the anonymous reviewers whose insightful comments have helped to improve the quality of this paper considerably.

REFERENCES

[1] B. Aslan, Y. Ekinci, and A. Toy, *Special Control Charts Using Intelligent Techniques: EWMA Control Charts* (Intelligent Decision Making in Quality Management). Cham, Switzerland: Springer, 2016, pp. 101–125.

[2] S. Basu, P. K. Dan, and A. Thakur, "Experimental design in soap manufacturing for optimization of fuzzified process capability index," *J. Manuf. Syst.*, vol. 33, no. 3, pp. 323–334, Jul. 2014.

[3] J. Beckford, *Quality: A Critical Introduction*, 4th ed. London, U.K.: Routledge, 2016.
[4] S. Breban, C. Saudemont, S. Vieillard, and B. Robyns, ''Experimental design and genetic algorithm optimization of a fuzzy-logic supervisor for embedded electrical power systems,'' *Math. Comput. Simul.* vol. 91, no. 5, pp. 91–107, 2013.
[5] C. A. Escobar, M. E. McGovern, and R. Morales-Menendez, ''Quality 4.0: A review of big data challenges in manufacturing,'' *J. Intell. Manuf.*, vol. 32, no. 8, pp. 2319–2334, 2021.
[6] Y. Gai, Y. Dang, and D. Tian, ''Research on a problem-driven method of acquiring tacit knowledge,'' *Chin. J. Manage.* vol. 11, no. 10, pp. 1528–1535, 2014.
[7] J. Huang and Y. Dang, ''Research on decision support method for multi-problem processing,'' *Manuf. Automat.* vol. 37, no. 10, pp. 93–96, 2015.
[8] H. Jiang, C. K. Kwong, and X. G. Luo, *Intelligent Quality Function Deployment* (Intelligent Decision Making in Quality Management: Theory and Applications), C. Kahraman and S. Yanik, Eds. Cham, Switzerland: Springer, 2016, pp. 327–362.
[9] H. Liu, L. Liu, and Q. Lin, ''Fuzzy failure mode and effects analysis using fuzzy evidential reasoning and belief rule-based methodology,'' *IEEE Trans. Rel.*, vol. 62, no. 1, pp. 23–36, Mar. 2013.
[10] H.-C. Liu, L. Liu, N. Liu, and L.-X. Mao, ''Risk evaluation in failure mode and effects analysis with extended VIKOR method under fuzzy environment,'' *Exp. Syst. Appl.*, vol. 39, no. 17, pp. 12926–12934, 2012.
[11] S. Luthra, D. Garg, A. Agarwal, and S. Mangla, *Total Quality Management (TQM): Principles, Methods, and Applications*. Boca Raton, FL, USA: CRC Press, 2020.
[12] Q. Ma, H. Li, and A. Thorstenson, ''A big data-driven root cause analysis system: Application of machine learning in quality problem solving,'' *Comput. Ind. Eng.*, vol. 160, Oct. 2021, Art. no. 107580.
[13] H. Paulheim, ''Knowledge graph refinement: A survey of approaches and evaluation methods,'' *Semantic Web*, vol. 8, no. 3, pp. 489–508, 2017.
[14] R. Simmons, ''Introduction to 8D problem solving: Including practical applications and examples,'' *Qual. Manage. J.*, vol. 25, no. 4, p. 201, 2008.
[15] K. Srikanth, A. Harish, C. Heymaraju, N. Kumar, and J. Anitha, ''Intelligent quality management expert system using PA-AKD in large databases,'' *Int. J. Eng. Sci. Technol.*, vol. 2, no. 4, pp. 632–636, 2010.
[16] K. Tang, K. Tan, and T. Lee, *Taguchi Method Using Intelligent Techniques* (Intelligent Decision Making in Quality Management). Cham, Switzerland: Springer, 2016, pp. 389–419.
[17] F. W. Taylor, *The Principles of Scientific Management*. Manhattan, NY, USA: Harper & Brothers, 1919.
[18] Y. Wang, Y. Dang, and Z. Xu, ''A causal network construction method based on Chinese quality problem solving data,'' *Chin. J. Manage. Sci.*, pp. 1–14, 2022, doi: 10.16381/j.cnki.issn1003-207x.2021.2414.
[19] Z. Xu and Y. Dang, ''Automated digital cause-and-effect diagrams to assist causal analysis in problem-solving: A data-driven approach,'' *Int. J. Prod. Res.*, vol. 58, no. 17, pp. 5359–5379, 2020.
[20] Z. Xu, Y. Dang, and P. Munro, ''Knowledge-driven intelligent quality problem-solving system in the automotive industry,'' *Adv. Eng. Inform.*, vol. 38, pp. 441–457, Oct. 2018.
[21] Z. Xu, Y. Dang, P. Munro, and Y. Wang, ''A data-driven approach for constructing the component-failure mode matrix for FMEA,'' *J. Intell. Manuf.*, vol. 31 no. 1, pp. 249–265, 2020.
[22] Z. Xu, Y. Dang, Z. Zhang, and J. Chen, ''Typical short-term remedy knowledge mining for product quality problem-solving based on bipartite graph clustering,'' *Comput. Ind.*, vol. 122, Nov. 2020, Art. no. 103277, doi: 10.1016/j.compind.2020.103277.
[23] K. Yuen, ''A hybrid fuzzy quality function deployment framework using cognitive network process and aggregative grading clustering: An application to cloud software product development,'' *Neurocomputing*, vol. 142, pp. 95–106, Oct. 2014.
[24] A. Syropoulos and T. Grammenos, *A Modern Introduction to Fuzzy Mathematics*. Hoboken, NJ, USA: Wiley, 2020.

YUBIN WANG received the B.S. and M.S. degrees from the School of Management Engineering, Zhengzhou University, China, in 2011 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Dalian University of Technology. His research interests include information technology, data mining, and knowledge management.

ZHAOGUANG XU received the Ph.D. degree from the Dalian University of Technology, in 2019. From July 2019 to July 2021, he was a Postdoctoral Researcher at the Dalian University of Technology. He is currently an Associate Professor with the School of Economics and Management, Dalian University of Technology. His research interests include data mining and knowledge management with applications to quality management in the manufacturing industry.

YANZHONG DANG is currently a Professor with the School of Economics and Management, Dalian University of Technology. His research interests include knowledge management, information system and information management, system engineering, and computational experiment method. He has published extensively in internationally recognized journals and received over five million RMB in research funding.