



Construction Method of Equipment Defect Knowledge Graph in IoT

Huafei Yang¹, Wenqing Yang¹, Nan Zhang¹, Shanming Wei^{2,*} and Yingnan Shang¹

¹NARI Group Corporation (State Grid Electric Power Research Institute), Nanjing, 211106, China

²School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China

*Corresponding Author: Shanming Wei. Email: weishanming@njust.edu.cn

Received: 06 October 2022; Accepted: 13 December 2022; Published: 11 September 2023

Abstract: Equipment defect detection is essential to the security and stability of power grid networking operations. Besides the status of the power grid itself, environmental information is also necessary for equipment defect detection. At the same time, different types of intelligent sensors can monitor environmental information, such as temperature, humidity, dust, etc. Therefore, we apply the Internet of Things (IoT) technology to monitor the related environment and pervasive interconnections to diverse physical objects. However, the data related to device defects in the existing Internet of Things are complex and lack uniform association hence building a knowledge graph is proposed to solve the problems. Intelligent equipment defect domain ontology is the semantic basis for constructing a defect knowledge graph, which can be used to organize, share, and analyze equipment defect-related knowledge. At present, there are a lot of relevant data in the field of intelligent equipment defects. These equipment defect data often focus on a single aspect of the defect field. It is difficult to integrate the database with various types of equipment defect information. This paper combines the characteristics of existing data sources to build a general intelligent equipment defect domain ontology. Based on ontology, this paper proposed the BERT-BiLSTM-Att-CRF model to recognize the entities. This method solves the problem of diverse entity names and insufficient feature information extraction in the field of equipment defect field. The final experiment proves that this model is superior to other models in precision, recall, and F1 value. This research can break the barrier of multi-source heterogeneous knowledge, build an efficient storage engine for multimodal data, and empower the safety of Industrial applications, data, and platforms in multi-clouds for Internet of Things.

Keywords: Knowledge graph; IoT; equipment defect; named entity recognition

1 Introduction

In recent years, with the implementation of the Internet of Things, traditional manufacturing equipment has been put on the cloud one after another. The information on various links between production and manufacturing has been put on the industrial Internet platform for users to query



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

and use [1]. However, the exponentially increasing number of devices deployed in IoT to satisfy various demands from users raise the big challenges of data fusion, processing, and protection [2]. To ensure the stable operation of production and manufacturing, more attention should be given to the inquiry and traceability of equipment defect data. With the help of these IoT sensors [3], we can get the environmental data and running status of the electric equipment in real-time from temperature, humidity, solar radiation, air pressure, and CO² to real-time videos and RFID reports [4]. How to effectively integrate sensory data and other types of equipment defect information is the key to the traceability of equipment defects [5,6]. The application of equipment defect domain ontology in the field of knowledge graph construction is the basis of building an equipment defect domain database [7,8], which can effectively sort out equipment defect information [9]. It provides a solution for the traceability of industrial Internet security problems. Ontology design and knowledge extraction are the important parts of construction of knowledge graph [10].

Ontology refers to the formal, explicit, and shared conceptual explanation of real world [11]. An ontology is a formal representation of a set of concepts and their relationships in a particular domain [12,13], which is suitable for semantic information representation and inference. Ontology defines constraints that control the simultaneous training of classifiers [14]. In other words, it employs bootstrapping approach along with a set of constraints that tend to couple different relations. Meanwhile, a well-constructed ontology can facilitate machine-processable definitions and help develop knowledge-based information search and management systems more effectively and efficiently [15]. The construction of a domain knowledge graph mainly includes two parts: domain ontology construction and domain knowledge construction [16]. Domain ontology construction plays an important role in the whole process of knowledge graph construction. It is the skeleton of the final knowledge graph. Based on the current domain ontology construction methods, combined with the characteristics of existing knowledge sources [17]. This paper constructs an ontology construction process suitable for the field of equipment defects. We extract the important concept sets in the equipment defect field as the core concept sets of the equipment defect field ontology to be constructed, draw lessons from the concept relationship in the ontology, combine the structural characteristics of the existing knowledge, determine the relationship between the concepts in the core concept set, and constructs an ontology suitable for the equipment defect field [18,19]. Different from the traditional construction method, our construction method aims at the characteristics of defect data, such as strong complexity, high heterogeneity, and more unstructured data, and closely combines ontology evaluation with the construction process to improve the efficiency and accuracy of ontology construction.

Named Entity Recognition (NER) technology can be used to identify specific entity information in text, such as names, places, organizations, etc. It is widely used in knowledge extraction. Named entity recognition tasks are typically formalized as sequential labeling tasks, which predict entity boundaries and entity types jointly by predicting labels for words. Named entity recognition enables different entities to be identified from many semi-structured and unstructured data at a small cost. In this paper, the main task is to locate different types of device defects. The first problem in this NER task is that traditional recognition technology relies on artificial feature extraction and professional knowledge, which wastes a lot of human resources. Secondly, in the process of model training for Chinese named entity recognition based on a neural network model, there is a problem of too single vector representation in the process of word vector representation, which cannot handle the ambiguity of words very well. This paper proposed the BERT-BiLSTM-CRF method to solve these problems. The experiment proved that this method is effective.

2 Related Work

Recently, the open knowledge graph construction method has become increasingly mature, and many mature large-scale knowledge bases have been built, such as DBpedia [20], freebase [21], etc., but the domain-oriented knowledge graph construction method is still in the research stage. For the domain knowledge graph, we need to build domain ontology and extract knowledge based on the constructed domain ontology.

Considering the complexity of power systems, knowledge graphs are needed to integrate different factors related to equipment defect detection. Kamsu-Foguem et al. [22] proposed a graph structure for knowledge management for industry maintenance supporting knowledge sharing and integration. Le-Phuoc et al. [23] proposed a knowledge graph-based IoT management, and a concept called the Graph of Thing. Ahmed et al. [24] presented a graph-based structure of a given surveillance scene and learned of relevant features including origin, destination, path, speed, size, etc., to support abnormal behavior detection. However, in electric system equipment detection fields, the knowledge graph is not widely used. Wang et al. [25] introduced a text-mining method to transform unstructured defect descriptions into structured and deep-analyzing models of power equipment defects [26]. Gaha et al. [27] used the combination of common information model (CIM) ontology to solve the exchange of information between heterogeneous applications remains a complex task [28].

Domain ontology describes the characteristics and laws of a specific domain formally [29]. At present, the methods of Constructing Ontology can be divided into manual construction, semi-automatic construction, and automatic construction according to the degree of automation.

The following describes several mature methods of manually building ontologies. Seven steps [30–32] is developed by the team at Stanford University. This method is mainly used to build domain ontology. The Methodology method [33] is similar to the software engineering development method and proposed by the Artificial Intelligence Laboratory of the Technical University of Madrid. The construction of ontology is divided into two parts: the ontology life cycle and the ontology development process. The ICAM DEFinition method (IDEF-5) method [34,35] is often used for the construction of enterprise ontology. This method mainly includes five steps: organizing and scoping, data collection, data analysis, initial ontology development, and ontology refinement and validation.

There are mainly two kinds of semi-automatic ontology construction, one is based on thesaurus [36], and the other is based on top-level ontology. Thesaurus can also be called glossary, which mainly includes the set of terms and the relationship between terms. Thesaurus can reflect the collection of related concepts in a specific field to a certain extent. In addition, the thesaurus also contains rich domain concepts and semantic relations, which is a relatively complete summary of the term set in a specific domain. In the specific structure of knowledge expression, thesaurus and ontology are closely related [37]. Therefore, we can construct ontology according to a domain-related thesaurus. The method of constructing domain ontology based on top-level ontology [38] is to integrate thesaurus and top-level ontology based on reference to seven step method and skeleton method. The key to this method is to use the basic idea of ontology engineering to standardize the selected words through the thesaurus method, then find the appropriate top-level ontology, and finally integrate the domain ontology into the appropriate top-level ontology.

NER (Named entity recognition) means labeling the location and type of the related entity from a natural language text. NER generally uses sequence labeling to jointly identify entity boundaries and determine entity types. NER is widely used in such tasks as knowledge graph construction, information extraction, information retrieval, machine translation, public opinion monitoring, and social systems construction. It is an important application of natural language processing. There

are three main methods for named entity recognition tasks, which are based on pattern matching, traditional machine learning, and deep learning.

The named entity recognition method based on neural network is usually treated as a sequential labeling task, and the text is an entity recognized by establishing a sequential labeling model. For the sequence labeling model, Collobert et al. [39] used CNN for feature extraction and propose a sentence-level logarithmic likelihood function, which achieves good recognition results by fusing other features. RNN solves the problem of variable length input and how to obtain long-term dependency before and after sequence. A variety of variants derived from RNN can well capture and preserve the context information of a sequence when processing time series data. Alzaidy et al. [40] proposed the BiLSTM-CRF model, which also incorporates other linguistic features to improve model performance. Jason et al. [41] proposed to use convolution neural network to learn character-level features automatically, which alleviates the dependence of the model on Feature Engineering to a certain extent, and combines dictionary features constructed by two open external resources, resulting in leading recognition results in English named entity recognition corpus.

With further research in the field of natural language processing, Word2Vec word vector training tools can no longer meet the needs of scholars. Peters et al. [42] proposed the ELMo model. The basic structure of ELMo is a two-layer Bi-LSTM, which uses the two-way stitching of forward and reverse encoders to extract feature information so that context-independent static vectors become context-dependent dynamic vectors. Xiao et al. [43] proposed the GPT model, which differs from the ELMo model in that GPT uses Transformer instead of RNN as the feature extractor, but GPT only uses one-way encoding. BERT [44] is innovative over GPT in that it uses the Self-attention mechanism of Transformer to implement bidirectional encoding and constructs a more general input and output layer. Simply modifying the downstream tasks of the model can be used in a variety of tasks, enriching the original tasks of GPT, including sentence-to-sentence relationship judgment, sentence classification tasks, reading comprehension tasks, and sequence labeling tasks [45,46].

3 Construction of Ontology

3.1 Framework

The construction of domain ontology needs to determine the domain and scope of the ontology first. The domain-oriented in this paper is the field of equipment defects, which is to integrate and manage the data related to data defects of various structures and provide basic data services for the specific application of upper-level equipment maintenance. Fig. 1 shows the construction framework of the equipment defect field.

As shown in Fig. 1, the overall construction framework of equipment defect domain ontology is shown. The construction framework is based on the seven-step method proposed by Stanford University and combined with the idea of ontology evaluation in the methodology method. After determining the domain and scope of the ontology, the next step is to construct an ontology for the field of equipment defects, which is used for the unified management of equipment defect-related data with multiple structures. The main work is to define the classes. This work includes defining concept cores and constructing the concept hierarchy based on the core classes. After this work, we should define the relationship between concepts formally. The reuse of existing related ontologies can be considered to define the concepts and the relationship between them, which can significantly improve the efficiency of ontology construction. After defining the relationship between the domain-related concepts, we should define the properties. This part of the work needs to combine the characteristics of existing data sources. The sets of properties are defined according to the existing structure of various

types of data. After completing the construction of ontology, the constructed ontology needs to be evaluated. The main content of this paper is to investigate whether the constructed equipment defect domain ontology can effectively organize the current equipment defect data. If not, continue to adjust the sets of concepts and modify the relationship between them until the ontology is evaluated.

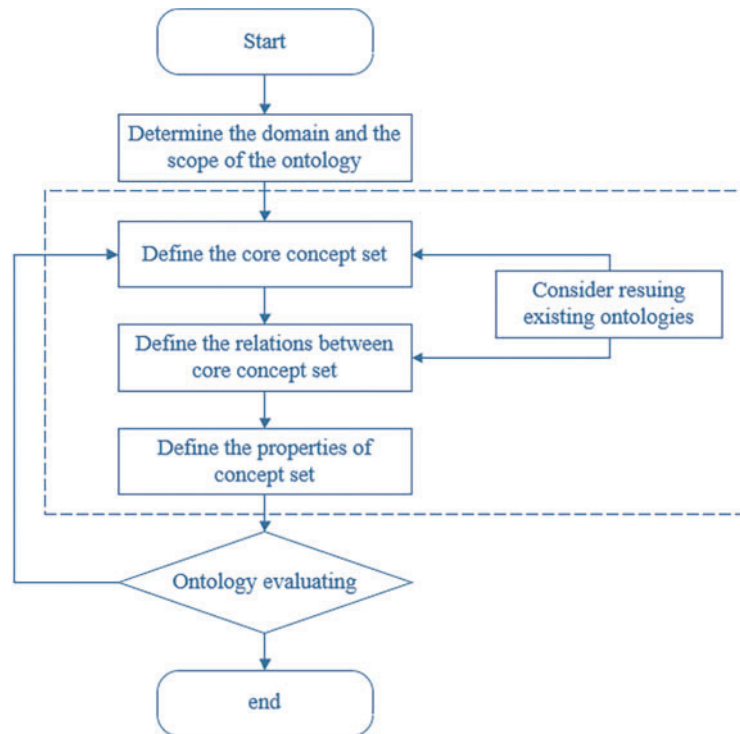


Figure 1: Framework of the domain ontology

3.2 Construction of Core Concept Sets

This section is mainly to define the set of concepts involved in the equipment defect domain ontology. First, list the important terms in the equipment defect domain and collect domain concepts that may be used. Second, construct the set of equipment defect domain concepts after sorting and analyzing. If it is time-consuming and laborious to build concepts from scratch, we can consider reusing the concepts involved in the equipment defect management process. [Table 1](#) lists the collection of concepts extracted from the existing equipment defect business process.

The defect business process can be divided into four parts: defect discovering, defect reporting, defect handling, and defect elimination. Defect discovery includes defective equipment information (such as device type, parts, manufacturer, etc.), power network topology information (such as voltage level, power line, etc.), discoverer information (unit, team, etc.), and defect-specific description information (defect description, content, cause, etc.).

Defect reporting includes the registrant information (work group, registration time, etc.), the confirmation information of defects (such as classification basis, defect nature, etc.), and System generated defect information (such as defect ID, corresponding defect standard library, etc.).

Defect handling includes Outage Management System (OMS) information on the defects (such as OMS ID, OMS defect ID), information on the person in charge of the work (such as ID of the

person in charge of the work, receiving unit, etc.), and information of the inspection and construction proposals (such as the inspection and repair proposals, repair categories, etc.).

Defect elimination mainly includes defect elimination information (such as elimination date, etc.), acceptance information (such as acceptance team, acceptance unit, acceptor, etc.), and defect handling result information (such as legacy problems, reasons for the delay, whether missing or not).

Table 1: Sets of core concepts based on equipment defect business process

Equipment defect business process	Sets of core concepts
Defect discovering	Defective equipment information Power network topology information Discoverer information
Defect reporting	Defect specific description information Registrant information Confirmation information of defects System generated defect information
Defect handling	OMS information of the defects Information of the person in charge of the work Inspection and construction proposals
Defect elimination	Defect elimination information Acceptance information Defect handling results information

Based on the relevant core concepts proposed in [Table 1](#), we combined the characteristics of existing data sources and introduced some new concepts to build a set of core classes. Taking the equipment defect as the core, referring to the basic ontology design concepts such as “people, events, things, and rules”, combined with the equipment defect business and model, the equipment defect ontology is constructed by selecting the team, equipment, equipment defect, defect description, defect treatment, defect elimination and acceptance, family defect and defect management documents as the main vertex concepts. At the same time, we divide the core concepts in the field of equipment defects into four dimensions. They are equipment information, defect information, organization, and defect management documents, which is showed in [Table 2](#).

Equipment information: equipment information includes equipment, manufacturer, and topology information of the defective equipment.

Defect information: it covers the classification and treatment of equipment defects and other defect businesses, and defines equipment defect treatment, defect elimination, family defects, and other contents.

Organization: it covers the management and implementation of equipment defects, and defines the organization and personnel involved in defect management.

Defect documents: management rule documents covering equipment defects. This part defines documents, specifications, guidelines, and other contents.

Table 2: Sets of core concepts based on equipment defect dimension

Dimension	Concept set
Organization	Team, acceptor, registrant, discoverer, defect eliminator
Equipment information	Power grid, manufacturers, parts, suppliers
Defect information	Defect description, defect treatment, defect elimination acceptance, family defects
Defect document	Standard specification, documents, guideline system

3.3 Construction of Relations between Concept Sets

In the previous section, the core concept set in the field of equipment defects is constructed based on four dimensions. The main content of this section is to determine the relationship between the concepts in the core concept set. In the constructed concept set, equipment defect is a very important concept, and other concepts will be directly or indirectly related to the concept of vulnerability. Therefore, this section starts from the equipment defect to construct the relationship.

Family defect is the sub-concept of equipment defect, and family defect and equipment defect are the sub-concept relationships. The equipment defect occurs on a certain part of the equipment, so the relationship from the equipment defect to the equipment and parts is expressed by occurring. Defect description is a specific description of specific defects, so the relationship from equipment defects to defect description is represented by description. For equipment defects, the defect management document plays a guiding role in defect handling operation and defect elimination and acceptance operation. Therefore, the relationship between the defect management document and equipment defects is expressed by the guide, and the relationship between defect handling and defect elimination and acceptance is operating. The treatment of equipment defects is carried out by specific teams, so the relationship between equipment defects and teams is represented by handling.

Then make a specific analysis for the specific concept set of each dimension. In the organizational dimension, the acceptor, registrant, discoverer, and defect eliminator belong to their corresponding teams, so their relationship is represented by a “subclass of”.

In the dimension of equipment information, each equipment has its own power grid topology, so its relationship is represented by “belong to”. The manufacturer produces the equipment, and the supplier supplies the equipment. Therefore, manufacture represents the relationship between the manufacturer and the equipment, and supply represents the relationship between the supplier and the equipment. Each part of the equipment constitutes the components of the equipment, and each part of the equipment constitutes the specific equipment. Therefore, the relationship between parts and components and between components and equipment is represented by composition.

In the dimension of defect documents, standard specifications, guideline systems, and documents are three important components of defect management documents. “Subclass of” is used to express the relationship between them and the concept of defect documents. Under the standard norms, there are defect classification standards and defect terminology standards. In terms of guideline system, technical super-vision is an important part of the guideline system, and the relationship between

them is expressed by “subclass of”. The relationship between technical supervision and supervision basis is represented by the guide. For the same reason, the relation between technical supervision and supervision requirements is also a “guide”.

In the dimension of defect information, defect description includes two key elements: defect location and defect causes. Therefore, the relationship between defect location, defect cause, and defect description is represented by “subclass of”. The classification basis is an important subconcept in defect description, and the subclass is used to represent the relationship between defect description and classification basis. The concept of defect elimination acceptance includes the evaluation of the state, specific maintenance operation, and specific test operation. We use operation as the relationship between defect elimination acceptance and equipment. Evaluate is used as the relationship between condition evaluation and defect elimination acceptance. The analysis report after each equipment defect contains three parts: rectification measures, experimental data, and analysis summary. We use “subclass of” to express the relationship between the analysis report and these three parts. The specific relationship between related concepts is shown in Fig. 2.

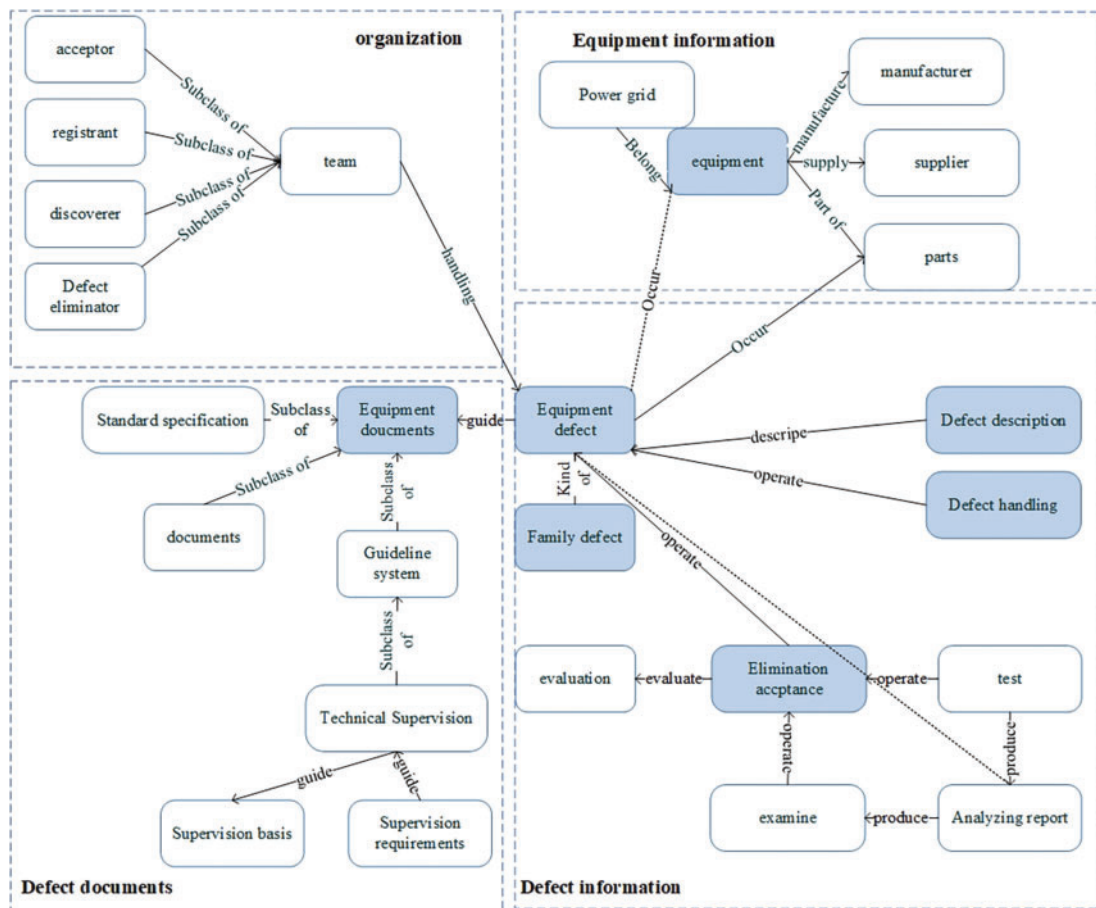


Figure 2: Relation graph of core concept set

3.4 Construction of Properties

The core concept set of equipment defect domain has been constructed in the previous two sections. If only the core concept set and the relationship between concepts are defined, it is impossible to describe the equipment defect domain in detail. Therefore, in addition to the work done in the first two sections, we also need to describe the internal structure of the concept in detail. In this paper, the internal structure is described by properties. Therefore, the main content of this section is to construct the property set corresponding to each concept in combination with the structural characteristics of existing equipment defect data sources. The following mainly introduces the construction of defect management documents and property sets of equipment defects.

3.4.1 Construction of Defect Information Property Set

The construction of a defect information set mainly refers to the equipment defect Standard Specification, quality event report, and specific equipment defect data. The specific property set is integrated. Through the integration of each property set, the redundant properties are removed, and the similar property is uniformly named.

3.4.2 Construction of Property Set of Defect Management Document

The construction of a defect management document property set mainly refers to the standard specification, guideline system, and detailed supervision rules. The specific management document property set is integrated, the redundant properties are removed, and the similar properties are uniformly named.

3.5 Evaluation of Ontology

Ontology evaluation is an essential link in the process of continuous iteration and improvement of ontology. The following is a qualitative analysis and evaluation of the constructed equipment defect domain ontology from four aspects: conceptual integrity, conceptual correctness, conceptual scalability, and relationship scalability between concepts.

Conceptual integrity: the concept set involved in the ontology should contain important concepts in the field of equipment defects. The core concepts of ontologies constructed in this chapter are extracted from the existing formed ontologies in the field of equipment defects. These ontologies cover all levels and aspects of the field of equipment defects, with wide coverage, and meet the requirements of integrity.

Conceptual Correctness: the concepts in the ontology should be clear and accurate without ambiguity. We analyze the core concepts one by one, and there is no ambiguity in each concept.

Concept Extensibility: the concept extensibility of ontology is mainly reflected in the continuous addition of new concepts according to the actual situation in the application process of ontology so that the ontology can be continuously improved. We divide the core concept set of equipment defect domain ontology into four dimensions. With the enrichment of data sources, we can add subclasses from four dimensions and establish new relationships, to continuously expand the ontology, Make the ontology more perfect.

Integrity between concepts: it mainly measures the integrity of the relationship category between concepts. The equipment defect domain ontology constructed by us includes various relationship categories such as cause and effect, use, dependency, and so on. The relationship categories are rich, and the equipment defect processing process can be modeled through these rich relationships.

4 Named Entity Recognition

4.1 Framework

Named entities in the domain of equipment defect have the characteristics of diverse entity names and text structures. Therefore, the problems of sparse semantics, long sequence semantics learning ability, and standardization constraints in the task of named entity recognition need to be solved. The BERT layer can understand the context information better and improve the accuracy of feature extraction. The BiLSTM layer can process sequence information well. The attention layer can obtain the word structure information. The CRF layer needs to be added to the named entity recognition model to ensure that the final output tag meets the BIO standard. By the combination of these layer the problems of sparse semantics, long sequence semantics learning ability, and standardization constraints in the task of named entity recognition can be solved well.

The model presented in this paper consists of four modules, BERT, BiLSTM, attention layer, and CRF, and the model framework is shown in Fig. 3. First, the text to be recognized is input into the BERT layer for pre-training, then the output vector sequence is input into the BiLSTM module for semantic encoding. After the semantic encoding, the output of the BiLSTM layer is sent into the CRF layer to calculate the optimal tag sequence. Because we introduced the BERT pre-training language model in the pre-training language model module, we can get the word vectors with more dimension information and improve the word ambiguity. Enhance the ability to represent sentences and get a better solution to the problem of sparse semantics. BiLSTM layer can learn both long-term and short-term memory. It solves the problem of long-sequence semantic feature learning through long-term and short-term memory learning. Due to the different weights of long-term memory and short-term memory, it may cause the problem of semantic dilution of long sequences. Using the attention layer to deal with the output of BiLSTM model can improve the weight of keywords in sentences so that the model can focus on the target entity and reduce the role of other irrelevant words CRF layer calculates the results of BiLSTM layer output and converts them into sequence tags to get the final prediction results.

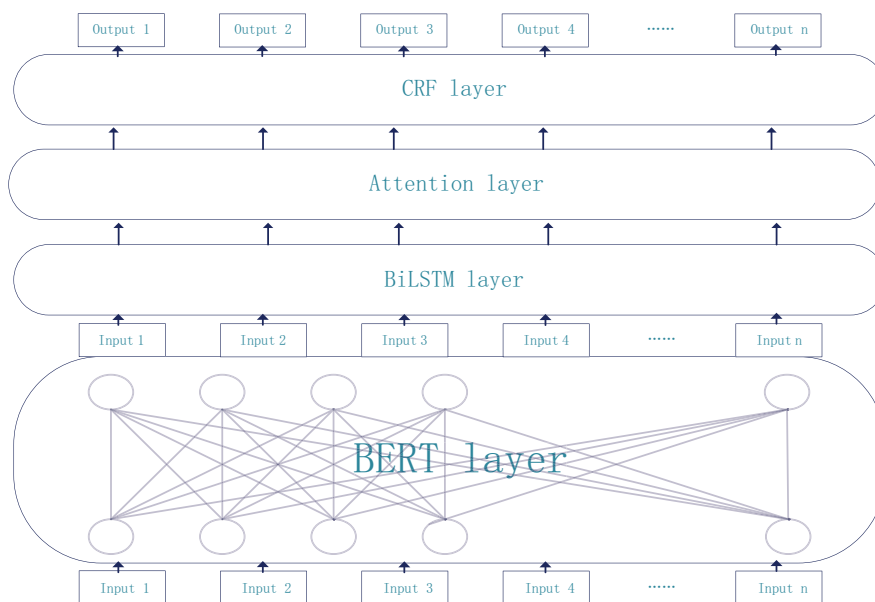


Figure 3: Framework of BERT-BiLSTM-CRF model

4.2 BERT

The BERT [27] model was proposed in 2018 and has achieved good results in the field of natural language processing. It obtains the characteristics of the entire input sequence through bidirectional parallel computation of the self-attention network. This method is independent of the time series. Compared with other pre-training models, it can understand the context information better and improve the accuracy of feature extraction. The BERT model inputs the sum of the word and position vectors. Word vectors can represent the main information about words in the model because the self-attention mechanism cannot remember time series information, so adding position vectors can improve the accuracy of recognition. An example of the BERT model input is shown in Fig. 4. The input vector consists of the sum of the word vector and the position vector.

BERT model training is divided into two steps, pre-training BERT and model fine-tuning BERT. BERT's pre-training uses the Masked LM strategy. Masked LM randomly covers a certain proportion of the input Token during training, and after covering, the encoder predicts the original words of the obscured Token. The training sample consisted of 15% of the words in the random masking sentence. The masking replacement rules were as follows: (1) 80% probability of replacing Token with [MASK]; (2) The probability of replacing a word with random sampling is 10%; (3) The probability of no replacement is 10%. In this paper, the BERT model only adjusts the weight through the multi-head self-attention mechanism.

The core of the BERT model is Encoder and Decoder, which uses a bidirectional transformer to model the text based on the attention mechanism. This model abandons the circular network structure of RNN (Recurrent Neural Network), whose encoding unit is shown in Fig. 4.

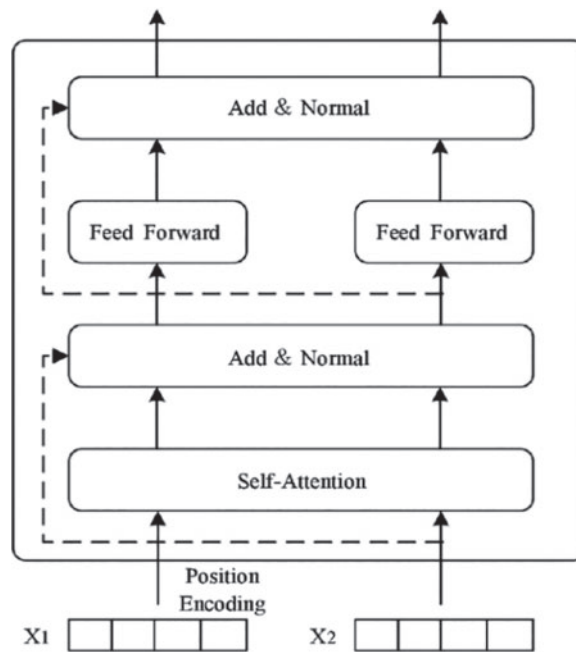


Figure 4: Workflow of transformer encoder

The attention mechanism considers the characteristics of each Token in the sequence in parallel. In the Encoder model, each module incorporates a fully connected feed-forward network and a multi-headed self-attention mechanism, which splices each attention to obtain multi-headed attention

information so that each attention mechanism can consider the information in the sequence. The process of the Transformer module is as follows:

(1) Get the query matrix Q, the value matrix K, and the index matrix V through the linear layer.

$$Q = \text{Liner}_Q(I) = W_Q^T I \quad (1)$$

$$K = \text{Liner}_K(I) = W_K^T I \quad (2)$$

$$V = \text{Liner}_V(I) = W_V^T I \quad (3)$$

(2) The value of attention is obtained by calculating V after the Softmax calculation. The formula is

$$\text{att}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V \quad (4)$$

The self-attention mechanism in the Transformer model uses the multi-head self-attention mechanisms that traditional self-attention mechanisms do not have to expand the “representation subspace” of the attention unit. This improvement improves the model’s ability to focus on different locations, the formula is as follows:

$$M_h = \text{Concat}(h_1, h_2, \dots, h_h) W^O \quad (5)$$

$$h_i = \text{att}(QW_i^O, KW_i^K, VW_i^V) \quad (6)$$

W_i^O, W_i^K, W_i^V in the formula is the weight of the corresponding matrix.

BERT’s input is represented as the addition of word vectors, segment vectors, and position embeddings for each word. Compared with other language models, BERT’s pre-training language model can make full use of the information on both sides of a word to get a better representation of word distribution.

4.3 BiLSTM

The BiLSTM model is derived from the LSTM model. LSTM and BiLSTM are good at processing sequence information. LSTM is an improved RNN. The LSTM model effectively solves the problem of gradient disappearance in traditional RNNs and realizes the effective use of long text information. LSTM’s ability to process sequence information comes from three control units called gates and one memory unit. Input gate, output gate and forget gate are the three gates in the LSTM network structure. The input gate selectively records the new information into the new memory unit, the forgotten gate determines what information will be discarded in the memory unit, and the output gate decides to transfer the input data and the state of the memory unit to the next hidden layer. The key to the ability of LSTM to remember long-term dependencies is the input gate and the forget gate. Memory cells can manage and store information. In contrast, BiLSTM is composed of two LSTM models that are forward and backward. This method enables both forward and backward sequences of data to be used for analysis, thus providing better robustness. Fig. 5 shows the cell structure of LSTM.

As shown in Fig. 5, the input of an LSTM cell at time t consists of three parts: memory cell c_{t-1} , previously hidden layer h_{t-1} , and input layer x_t . Implied layer h_t and memory cell form the output of the unit at time t. The computation detail of the LSTM is as follows:

(1) The forget gate selectively forget the messages passed in from the last unit.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (7)$$

where f_t is the forget gate, sigmoid is chosen to be the activation function.

(2) The unit will selectively remember the input

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{8}$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \tag{9}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{10}$$

where \tilde{C}_t is the updated value of the unit, i_t is the input gate. i_t is used to determine which features of \tilde{C}_t are used to update C_t .

(3) The output gate will determine which will be used as the output:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t * \tanh(C_t) \tag{12}$$

where o_t is the output gate, and h_t will input into the next unit.

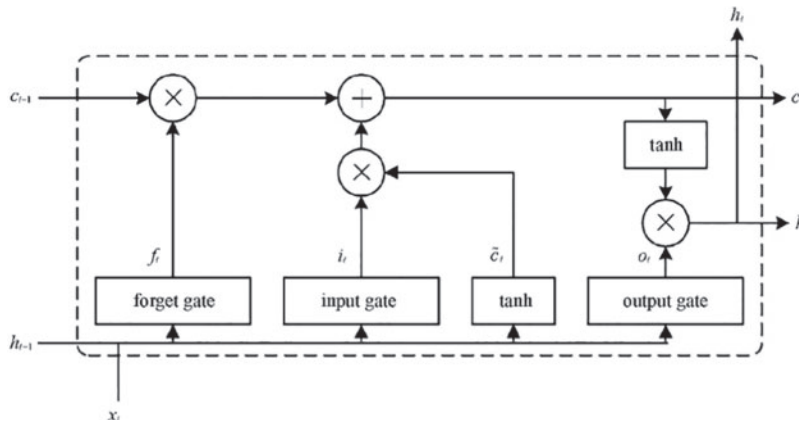


Figure 5: Unit structure of BiLSTM

LSTM only gets the information above in the text to be recognized, and the information below is also very important for NER tasks. BiLSTM is the optimized method based on LSTM and has a prominent performance in sequence marking tasks in the field of NLP. Therefore, to obtain context information at the same time, this paper uses two-way LSTM for model training, and its structure is shown in Fig. 6.

4.4 Attention Layer

The essence of the attention mechanism is the weight distribution, which obtains word structure information by calculating the degree of association between words. The attention mechanism can effectively solve the problem of semantic dilution in the front end of a long sequence caused by the fixed semantic vector generated by BiLSTM network. The attention layer is used to obtain sentence-level information, which can extract long-distance dependency information and integrate the coding layer information. The attention layer mainly calculates the attention weight vector r_i , which is defined as shown in formula (13).

$$r_i = \sum_{j=1}^N b_{ij}, y_j \tag{13}$$

where y_j represents the feature sequence output by the BiLSTM layer, and b_{ij} represents the correlation probability distribution between words in the document. The definition of b_{ij} is as follows:

$$b_{ij} = \frac{\exp(f(w_i, w_j))}{\sum_{k=1}^N \exp(f(w_i, w_k))} \quad (14)$$

where w_i and w_j represent the i th and j th word in the text. N is the number of the text. $f(w_i, w_k)$ indicates the similarity score of the word w_i and w_j .

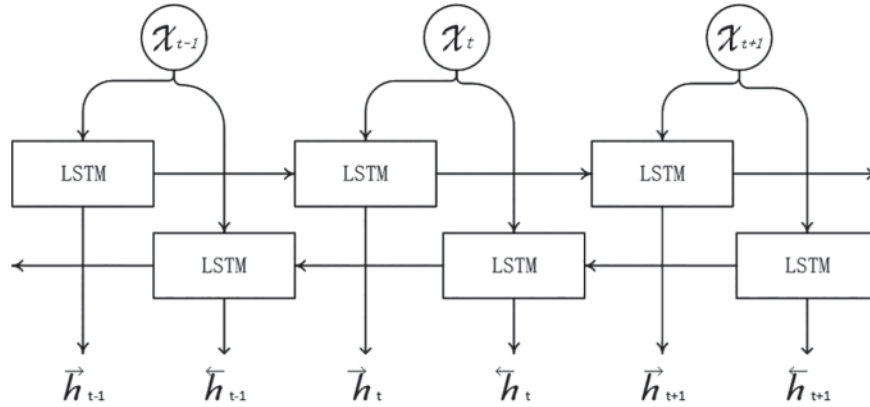


Figure 6: The model of BiLSTM

By processing the feature sequence output from BiLSTM layer through the attention layer, the correlation between the current word and other words in the sequence is obtained, and the global feature representation is obtained. Finally, the feature sequences obtained by the attention layer are fed into the CRF for prediction.

4.5 CRF Layer

Named entity recognition can be defined as a sequence labelling task, and the output sequence labels in the previous layer are only affected by the word itself and context characteristics, regardless of the label dependency. However, in BIO labelling mode, labels depend on each other, and label B must be preceded by the label I. Therefore, a CRF layer needs to be added to the named entity recognition model to ensure that the final output tag meets the BIO standard. The input of CRF module is a word vector trained by BERT and BiLSTM layers, and the sequence markers at the sentence level are obtained from the input of CRF. Corresponding weight values λ_j for each word vector eigenvalue f_j has been trained in the previous two models, using $score(l|s) = \sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})$, calculate the value of each tag sequence l to mark the entire sentence s , and get the score of the tag sequence. Where s is the sentence to be marked, i is the position of the word in the sentence, l_i is the label of the current word, l_{i-1} is the label of the previous work, m is the corresponding feature number, n is the corresponding sentence length. Normalize fractions into probability values of 0–1 as follows:

$$p(l|s) = \frac{\exp[score(l|s)]}{\sum_{l'} \exp[score(l'|s)]} \quad (15)$$

The $p(l|s)$ of each tag sequence is calculated, and the maximum is the final tag sequence.

5 Experiment

5.1 Experiment Design

Experimental Environment: The hardware environment for the case was AMD Ryzen 5 5600X 6-Core Processor, 16 GB memory, Windows10 operating system, and GTX3060Ti GPU. Set up a dependency environment of python3+tensorflow2.

Experiment data: The equipment defect knowledge graph is based on all kinds of data such as defect data, defect reports, standard specifications, rules, and regulations of each business system in the data desk. The project group collates data about equipment defects in PMS, power transmission and substation equipment status monitoring, intelligent transportation inspection, control, and other systems, involving 212 data tables and 4699 fields. At present, the construction of the defect graph is mainly around 22 tables and related tables of substation accounts. The sorting of unstructured documents is shown in Table 3. The document data includes standard specification and quality event report. These data are also the guidelines about device defect management. The existing device defect data includes equipment ledger data, Statistics on defects of equipment in Hubei and China. These data are highly complex, heterogeneous, and valuable.

Table 3: Sorting of unstructured documents

Data type	Data content	Number
Document data	Standard specification	20
	Quality event report	20
Device defect data	Equipment ledger data	233
	Statistics on defects of equipment in China	49000
	Statistics on defects of equipment in Hubei	2684

Comparison of different model experiments: In order to verify the performance of BERT-BiLSTM-CRF model on named entity recognition in device defect corpus, experiments were performed using RNN-CRF, LSTM-CRF, BiLSTM-CRF, BERT-CRF and BERT-BiLSTM-IDCNN-CRF models compared with the BERT-BiLSTM-CRF (without attention layer) models proposed in this paper.

Assessment Indicators: The aim of the named entity recognition task in the field of equipment defect is to automatically identify named entities in text. Therefore, the accuracy and recall rate of named entity recognition will directly affect the quality of knowledge graph construction. Therefore, Precision, recall rate and F1 value are selected as performance evaluation indicators for named entity recognition tasks in the domain of power equipment. The Precision P, recall R, and F1 values used in this paper are defined as the evaluation indexes for named entity recognition as follows:

$$R = \frac{\text{Number of correct entities identified}}{\text{Number of entities for all labels}} \times 100\% \quad (16)$$

$$R = \frac{\text{Number of correct entities identified}}{\text{Number of entities for all labels}} \times 100\% \quad (17)$$

$$F_1 = \frac{2PR}{P + R} \times 100\% \quad (18)$$

Experiment parameters: The training parameters used in the experiment are as follows: the number of network layers is 12, the hidden layer is 768, the learning rate is $5e-5$, the number of iterations is 40, the dropout rate is 0.5, the vector dimension is 300, and the optimizer is Adam.

5.2 Experiment Result and Analyzation

The RNN-CRF, LSTM-CRF, BiLSTM-CRF [41], BERT-CRF, BERT-BiLSTM-IDCNN-CRF [45] and BERT-BiLSTM-CRF (with and without attention layer) [46] models were tested. The training data and test data of all model experiments are in the same dataset. The test results are shown in Table 4.

Table 4: Results of different models (%)

Model	Precision	Recall	F1
RNN-CRF	73.60	75.20	74.39
LSTM-CRF	76.25	79.82	77.99
BiLSTM-CRF	79.91	81.22	80.56
BERT-CRF	83.21	82.94	83.07
BERT-BiLSTM-CRF	85.19	86.27	85.73
BERT-BiLSTM-IDCNN-CRF	86.21	85.83	86.02
BERT-BiLSTM-Att-CRF	87.92	87.55	87.73

Table 4 shows that the RNN-CRF had the worst precision, recall, and F1 value. Because this model cannot handle long-distance dependence due to the gradient disappearance problem of the traditional RNN model. Compared with the traditional RNN-CRF model, LSTM-CRF can solve the problem of gradient disappearance in RNN method. LSTM adds the input gates, forgotten gates, and output gates. These gates enable the model to obtain the relationship between sentences so that the precision, recall, and F1 are improved by about 3%. BiLSTM-CRF transforms the LSTM model to bidirectional so that the model can get the relationship between forward and backward sentences with a precision of 79.91%. The BERT pre-training model can obtain global information in parallel, and its precision, recall, and F1 value are all increased to more than 80%. When BiLSTM is added to the BERT-CRF model, better experimental results are obtained, which are 1.98, 3.33, and 2.66 percentage points higher than the BERT-CRF model in precision, recall rate, and F1 value, respectively. If we add an attention layer between BiLSTM layer and the CRF layer, the experiment gets the best performance the precision is 87.92, the recall is 87.55, and the F1 value is 87.83. The experiment shows that the performance of the attention layer is better than the IDCNN layer.

To further evaluate the recognition effect of various indicator entities, four models were compared to identify four entities: person, rule, equipment, and event. The precision results are shown in Figs. 7–9.

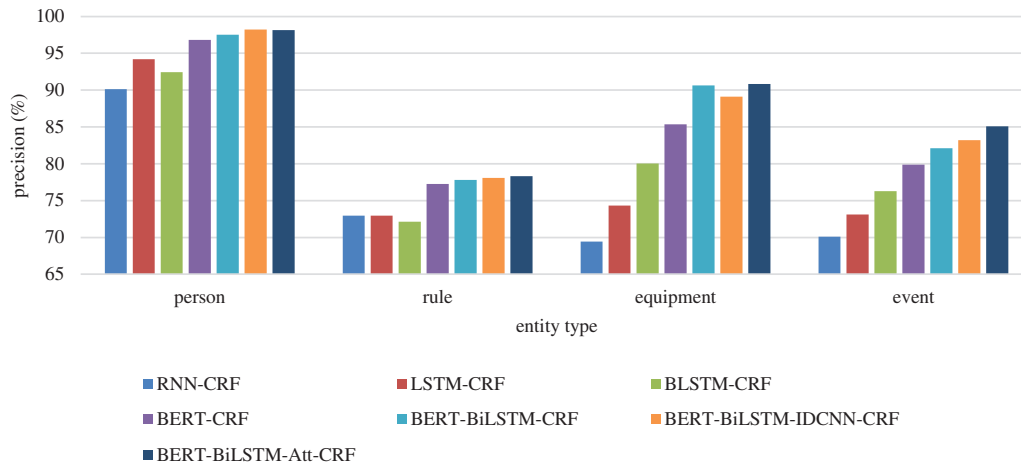


Figure 7: Precision results of different entities

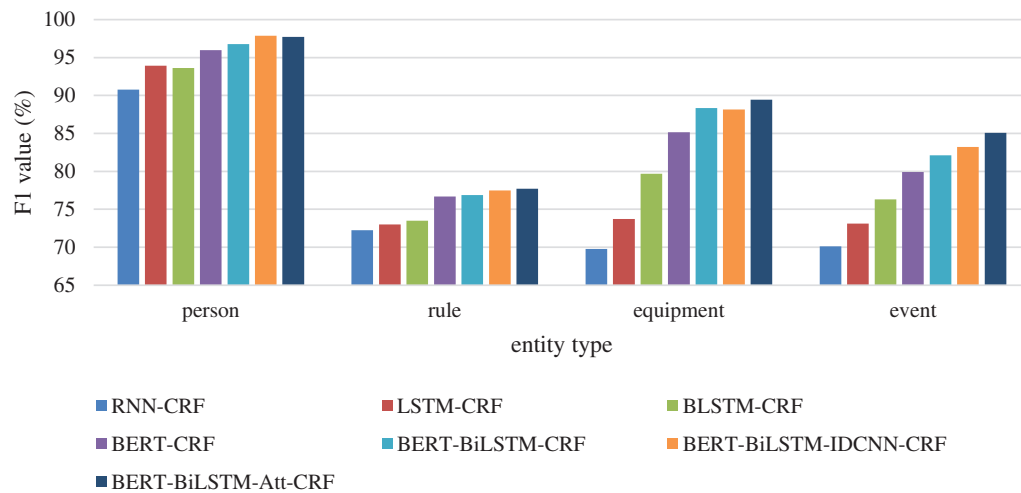


Figure 8: F1 value results of different entities

It can be seen from the precision results that among the recognition results of four different types of entities, the precision result of Person is the best. Compared with other types of entity, the entity features of a Person are more obvious. At the same time, the result of the rule is the worst because the text of the rule has many proper nouns. It is difficult to recognize them when the document is not enough. However, the event and equipment are composed of words, which is highly like the context text. The boundary between the indicator name and the context needs to be divided first, and then the entity recognition. The errors caused by the boundary division will be transferred to the entity recognition task, thus reducing the accuracy of the recognition. Therefore, the recognition effect of equipment and event is not good.

Based on the above experiment, the recognition result of the combined BERT-BiLSTM-Att-CRF model is better than the other models. The recognition accuracy is close to 90%. Compared with the BERT-BiLSTM-CRF model, one more attention mechanism can efficiently allocate attention

resources, selectively focus on relevant information, and learn feature information more accurately. The experiment also proved that the result of one more attention layer is better than the IDCNN layer.

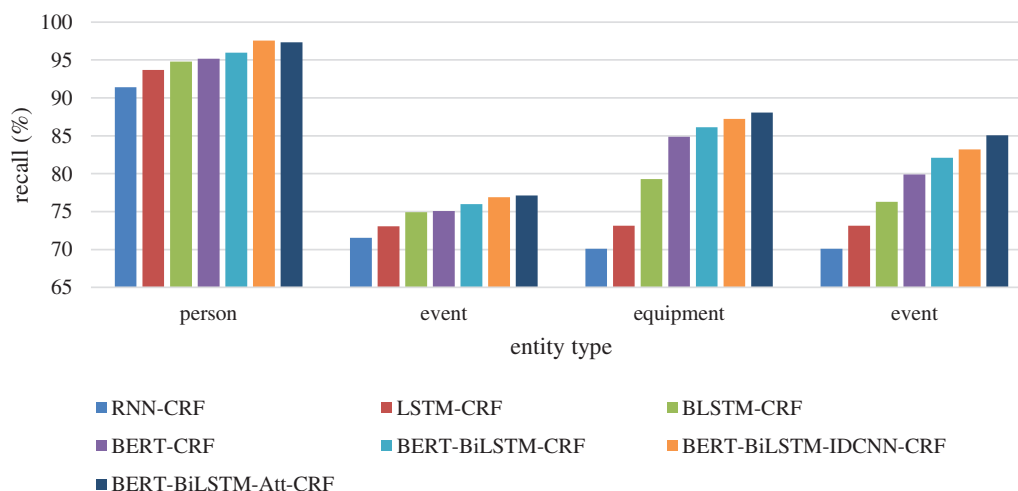


Figure 9: Recall results of different entities

6 Conclusion

In view of the current situation that there is much unstructured intelligent equipment defect information in the equipment defect domain in the era of the industrial Internet of things, this paper research how to use the existing information to build the equipment defect domain ontology. And use the BERT-BiLSTM-Att-CRF model to extract the knowledge from the existing text. This research has great significance to the construction of the knowledge graph of the equipment defect domain and further data processing. The specific research work is summarized as follows:

1. The construction process of domain ontology is studied, and the construction process of domain ontology is determined according to the research and data status of the equipment defect domain.
2. According to the existing structured data and semi-structured data, the core concept set is abstracted, and the relationship between concept sets and property sets is constructed. The qualitative analysis and evaluation of the constructed equipment defect ontology prove that the ontology concept we constructed is complete and correct, and the relationship between concepts has strong scalability.
3. Based on ontology, this paper proposed the BERT-BiLSTM-Att-CRF model to recognize the entities. This method solves the problem of semantic sparseness, gradient explosion, and insufficient feature information extraction in the field of equipment defect field.

Although the proposed method reaches great performance, there are still some avenues for future work. (1) The construction of ontology did not consider the more complicated actual situation. (2) There are a lot of other pretraining models that can be used in sequential processing. It means the NER task can be replaced by a more effective self-attention model.

Acknowledgement: Thank you for the assistance provided by the editor and the staff of NARI during the paper writing process.

Funding Statement: This work is supported by the fund project: Research on Basic Capability of Multi-modal Cognitive Graph (Granted No. 524608210192).

Author Contributions: Huafei Yang performed the methodology; Wenqing Yang performed the formal analysis; Nan Zhang performed the validation; Liu Sun wrote the manuscript, Yingnan Shang performed the data curation.

Availability of Data and Materials: Data available on request.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] L. Cattaneo, P. Adalberto and M. Marco, "A framework to integrate novelty detection and remaining useful life prediction in Industry 4.0-based manufacturing systems," *International Journal of Computer Integrated Manufacturing*, vol. 35, no. 4–5, pp. 388–408, 2022.
- [2] A. K. Sandhu, "Big data with cloud computing: Discussions and challenges," *Big Data Mining and Analytics*, vol. 5, no. 1, pp. 32–40, 2021.
- [3] W. Fan, L. Guangshun, W. Yilei, R. Wajid, R. K. Mohammad *et al.*, "Privacy-aware traffic flow prediction based on multi-party sensor data with zero trust in smart city," *ACM Transactions on Internet Technology*, 2022. <https://doi.org/10.1145/3511904>
- [4] H. Niu, Z. Chu, Z. Zhu and F. Zhou, "Aerial intelligent reflecting surface for secure wireless networks: Secrecy capacity and optimal trajectory strategy," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 119–133, 2022.
- [5] Q. Lianyong, L. Wenmin, Z. Xuyun, D. Wanchun, X. Xialong *et al.*, "A correlation graph based approach for personalized and compatible web APIs recommendation in mobile APP development," *IEEE Transactions on Knowledge and Data Engineering*, 2022. <https://doi.org/10.1109/TKDE.2022.3168611>
- [6] Y. W. Liu, Z. L. Song, X. L. Xu, W. Rafique, X. Y. Zhang *et al.*, "Bidirectional GRU networks-based next POI category prediction for healthcare," *International Journal of Intelligent Systems*, vol. 37, no. 7, pp. 4020–4040, 2021.
- [7] W. Gu, F. Gao, R. Li and J. Zhang, "Learning universal network representation via link prediction by graph convolutional neural network," *Journal of Social Computing*, vol. 2, no. 1, pp. 43–51, 2021.
- [8] Y. Dongxiao, Z. Lifang, L. Qi, C. Xiuzhen, Y. Jiguo *et al.*, "Fast skyline community search in multi-valued networks," *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 171–180, 2020.
- [9] Y. W. Liu, D. J. Li, S. H. Wan, F. Wang, W. Dou *et al.*, "A long short-term memory-based model for greenhouse climate prediction," *International Journal of Intelligent Systems*, vol. 37, no. 1, pp. 135–151, 2021.
- [10] L. Kong, L. Wang, W. Gong, C. Yan, Y. Duan *et al.*, "LSH-aware multitype health data prediction with privacy preservation in edge environment," *World Wide Web Journal*, vol. 25, no. 5, pp. 1783–1808, 2021.
- [11] C. Catlett, P. Beckman, N. Ferrier, H. Nusbaum, M. E. Papka *et al.*, "Measuring cities with software-defined sensors," *Journal of Social Computing*, vol. 1, no. 1, pp. 14–27, 2020.
- [12] F. Wang, H. B. Zhu, G. Srivastava, S. C. Li, M. R. Khosravi *et al.*, "Robust collaborative filtering recommendation with user-item-trust records," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 986–996, 2021.
- [13] T. P. Zhou, W. C. Liu, N. B. Li, X. Y. Yang, Y. L. Han *et al.*, "Secure scheme for locating disease-causing genes based on multi-key homomorphic encryption," *Tsinghua Science and Technology*, vol. 27, no. 2, pp. 333–343, 2022.
- [14] G. Zhou, C. Pan, H. Ren, K. Wang, K. K. Chai *et al.*, "User cooperation for IRS-aided secure MIMO systems," *Intelligent and Converged Networks*, vol. 3, no. 1, pp. 86–102, 2022.

- [15] M. Azrour, J. Mabrouki, A. Guezzaz and Y. Farhaoui, "New enhanced authentication protocol for Internet of Things," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 1–9, 2021.
- [16] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?," *International Journal of Human-Computer Studies*, vol. 43, no. 6, pp. 907–928, 1995.
- [17] Y. Yihong, Y. Xuan, H. Mohsen, M. A. Khan, G. Srivastava *et al.*, "ASTREAM: Data-stream-driven scalable anomaly detection with accuracy guarantee in IIoT environment," *IEEE Transactions on Network Science and Engineering*, 2022. <https://doi.org/10.1109/TNSE.2022.3157730>
- [18] L. Y. Qi, Y. W. Liu, Y. L. Zhang, X. X. Xu, M. Bilal *et al.*, "Privacy-aware Point-of-Interest category recommendation in Internet of Things," *IEEE Internet of Things Journal*, 2022. <https://doi.org/10.1109/JIOT.2022.3181136>
- [19] X. Zhou, W. Liang, K. Yan, W. Lin, K. I. Wang *et al.*, "Edge enabled two-stage scheduling based on deep reinforcement learning for Internet of Everything," *IEEE Internet of Things Journal*, pp. 1, 2022. <https://doi.org/10.1109/JIOT.2022.3179231>
- [20] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker *et al.*, "Dbpedia-a crystallization point for the web of data," *Journal of Web Semantics*, vol. 7, no. 3, pp. 154–165, 2009.
- [21] K. Bollacker, C. Evans, P. Paritosh, T. Sturge and J. Taylor, "Freebase: A collaboratively created graph database for structuring human knowledge," in *Proc. of the 2008 ACM SIGMOD Int. Conf. on Management of Data*, New York, NY, USA, pp. 1247–1250, 2008.
- [22] B. Kamsu-Foguem and D. Noyes, "Graph-based reasoning in collaborative knowledge management for industrial maintenance," *Computers in Industry*, vol. 64, no. 8, pp. 998–1013, 2013.
- [23] D. Le-Phuoc, H. N. M. Quoc, H. N. Quoc, T. T. Nhat and M. Hauswirth, "The graph of things: A step towards the live knowledge graph of connected things," *Journal of Web Semantics*, vol. 37, pp. 25–35, 2016.
- [24] S. A. Ahmed, D. P. Dogra, S. Kar and P. P. Roy, "Surveillance scene representation and trajectory abnormality detection using aggregation of multiple concepts," *Expert Systems with Applications*, vol. 101, no. 1, pp. 43–55, 2018.
- [25] H. F. Wang, J. Cao and D.Y. Lin, "Deep analysis of power equipment defect based on semantic framework text mining technology," *CSEE Journal of Power and Energy Systems*, vol. 8, no. 4, pp. 1157–1164, 2019.
- [26] L. Y. Qi, Y. H. Yang, X. K. Zhou, W. Rafique and J. H. Ma, "Fast anomaly identification based on multi-aspect data streams for intelligent intrusion detection toward secure Industry 4.0," *IEEE Transactions on Industrial Informatics*, 2022. <https://doi.org/10.1109/TII.2021.3139363>
- [27] M. Gaha, A. Zinflou, C. Langheit, A. Bouffard, M. Viau *et al.*, "An ontology-based reasoning approach for electric power utilities," *International Conference on Web Reasoning and Rule Systems*, vol. 7994, pp. 95–108, 2013.
- [28] X. Zhou, W. Liang, W. Li, K. Yan, S. Shimizu *et al.*, "Hierarchical adversarial attacks against graph neural network based IoT network intrusion detection system," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9310–9319, 2022.
- [29] R. Junyu, L. Jinze, L. Huaxing and Q. Tuanfa, "Task offloading strategy with emergency handling and blockchain security in SDN-empowered and fog-assisted healthcare IoT," *Tsinghua Science and Technology*, vol. 27, no. 4, pp. 760–776, 2022.
- [30] H. S. Pinto and J. P. Martins, "Ontologies: How can they be built?," *Knowledge and Information Systems*, vol. 6, no. 4, pp. 441–464, 2004.
- [31] L. Dandan, D. Jianwei and Y. Shuzhen, "Research on computer science domain ontology construction and information retrieval," *Knowledge Engineering and Management*, vol. 123, pp. 603–608, 2011.
- [32] M. S. Fox, B. Mihai and G. Michael, "An organisation ontology for enterprise modeling: Preliminary concepts for linking structure and behaviour," *Computers in Industry*, vol. 29, no. 2, pp. 123–134, 1996.
- [33] H. Panetto, B. Iung, D. Ivanov, G. Weichhart and X. Wang, "Challenges for the cyber physical manufacturing enterprises of the future," *Annual Reviews in Control*, vol. 47, no. 4, pp. 200–213, 2019.
- [34] C. P. Menzel and R. J. Mayer, *IDEF5 Ontology Description Capture Method: Concept Paper*, Houston, USA: NASA Johnson Space Center, pp. 190285, 1990.

- [35] M. Uschold and M. Gruninger, "Ontologies: Principles, methods and applications," *The Knowledge Engineering Review*, vol. 11, no. 2, pp. 93–136, 1996.
- [36] M. Damerchiloo and M. S. Hosseini Beheshti, "Converting thesaurus to ontology (a systematic review)," *Library and Information Science Research*, vol. 11, no. 2, pp. 105–127, 2022.
- [37] M. Nowroozi, M. Mirzabeigi and H. Sotudeh, "Constructing an ontology based on a thesaurus: A case of ASIS&TOnto based on the ASIS&T web-based thesaurus," *The Electronic Library*, vol. 36, no. 4, pp. 750–764, 2018.
- [38] M. Nowroozi, M. Mirzabeigi and H. Sotudeh, "The comparison of thesaurus and ontology: Case of ASIS&T web-based thesaurus and designed ontology," *Library Hi Tech*, vol. 36, no. 4, pp. 665–684, 2018.
- [39] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu *et al.*, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, no. 11, pp. 2493–2537, 2011.
- [40] R. Alzaidy, C. Caragea and C. L. Giles, "Bi-LSTM-CRF sequence labeling for keyphrase extraction from scholarly documents," in *The World Wide Web Conf.*, New York, NY, USA, pp. 2551–2557, 2019.
- [41] P. C. Jason and E. Nichols, "Named entity recognition with bidirectional LSTM-CNNs," *Transactions of the Association for Computational Linguistics*, vol. 4, no. 26, pp. 357–370, 2016.
- [42] M. Peters, M. Neumann, M. Iyyer, M. Gardner and C. Clark, "Deep contextualized word representations," in *Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, Louisiana, USA, vol. 1, pp. 2227–2237, 2018.
- [43] Y. Xiao and J. Zhezhi, "Summary of research methods on pre-training models of natural language processing," *Open Access Library Journal*, vol. 8, no. 7, pp. 1–7, 2021.
- [44] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.
- [45] Y. Chang, L. Kong, K. Jia and Q. Meng, "Chinese named entity recognition method based on BERT," in *2021 IEEE Int. Conf. on Data Science and Computer Application (ICDSCA)*, Dalian, China, pp. 294–299, 2021.
- [46] W. Zhang, S. Jiang, S. Zhao, K. Hou, Y. Liu *et al.*, "A BERT-BiLSTM-CRF model for Chinese electronic medical records named entity recognition," in *2019 12th Int. Conf. on Intelligent Computation Technology and Automation (ICICTA)*, Xiangtan, China, pp. 166–169, 2019.