Constructing Industrial Knowledge Graph through Ontology and Link Prediction

Yuwei Wan, Zheyuan Chen, Ying Liu *Member, IEEE*, Michael Packianather and Rui Wang

Abstract— **Given that the numerous data embedded in manufacturing processes and products are separated, it is challenging to tackle and integrate heterogeneous data in industrial scenarios. In this context, an industrial knowledge graph (iKG) has been developed as a promising semantic organisation to leverage the rich information from multiple resources. However, relations are usually missing and hidden in original iKGs, which results in the necessity for iKG completion. Given these two perspectives, a framework of iKG construction is proposed based on ontology and link prediction in this study. Firstly, an ontology design framework is deployed to generate domain-centric ontologies after extracting numerous data (e.g., entities and relations). Secondly, the missing relations between each couple of entities are discovered over existing knowledge to increase the number of edges that complete and refine iKGs. Thirdly, iKG visualisation is conducted by importing data into the generated ontology. The feasibility and effectiveness of the proposed framework are substantiated and demonstrated in a case study using real-world data.**

I. Introduction

As the industry develops, enormous data embedded in manufacturing processes and products are accumulated and stored in isolated silos [1]. A complex industrial phenomenon usually involves abundant data from multiple different sources, which provide comprehensive and versatile information for follow-up tasks [2]. However, it is challenging to integrate the data derived since there are varying distance metrics across facet boundaries [3]. Specifically, there is still a lack of semantic organisation of numerous heterogeneous manufacturing resources, which enables the free flow of ever-evolving knowledge among processing modules, information systems, and users, thereby impeding the effective exploitation of knowledge in industrial scenarios [4].

Knowledge graph (KG) is proposed as a semantic network for knowledge representation, which displays a powerful expressive ability and a high degree of modelling flexibility, making it a promising content-retrieval approach [5]. Concisely, KG adopts practical and straightforward representation approaches based on triplets (represented by <head, relation, tail>), containing entities and relations. An entity refers to an individual (e.g., organisation, person, event, location etc.), and a relation represents a specific relational connection between an entity pair [6]. Presently, KGs have been utilised energetically in many scenarios, such as recommendation systems [7], question answering [8], and knowledge visualisation [9] etc. In this context, it has been conducted that the KG has a crucial role in breaking the semantic gap in industrial scenarios as it provides a promising mechanism to fuse more sophisticated knowledge and structure from multiple sources [10].

Compared with other domains, since relations usually have physical significance in industrial applications, knowledge demands in industrial products and services require higher synthesis and creation. Even though knowledge extraction can be performed accurately for the development of iKGs, it is common that the information on relations is insufficient [11, 12]. Concisely, the relation is often missing in iKGs constructed from raw data, which means that the potential relations are needed to be predicted and discovered. In other words, iKGs need to completed by link prediction before applying in application scenarios. According to the essential physical significances, potential relations are discovered through more sophisticated relation structures in industrial fields [4, 13-15]. Presently, most research on link prediction in iKGs relies on shallow models and methods based on path sorting. Despite their simplicity, intuitiveness, and computational lightness, conventional methods cannot learn more complex patterns posed by in-depth iKGs. A further challenge in the construction of iKGs is that the scale is large. Therefore, the iKG must have good scalability, whereas these models have poor scalability. In this context, with the successful utilisation of deep learning in natural language processing, methods based on deep learning are going to be more advantageous, such as graph neural networks (GNNs) and their variants [16, 17].

To address the above-mentioned issues, a framework of iKG construction has been presented based on ontology and link prediction in this study. The proposed iKG construction framework serves as a model that aims to design a reliable ontology and complete relations in constructing iKG for information management and knowledge sharing. In summary, the following are the major contributions of this study:

- A framework of iKG construction is presented to solve the problem of 'data island' by associating multi-source data in industrial scenarios.
- A link prediction model based on GraphSAGE is built to identify the potential relations over the existing knowledge, enhancing the reliability of iKGs.

Yuwei Wan, Zheyuan Chen, Ying Liu and Michael Packianather are with the Institute of Mechanical and Manufacturing Engineering, School of Engineering, Cardiff University, Cardiff, UK (corresponding author e-mail: wanY8@cardiff.ac.uk).

Rui Wang is with the UNICMICRO, Science Town Commercial Plaza, Kexue Road, Guangzhou 510000, China

- A real-life iKG of the strip-breakage phenomenon in the cold rolling process is constructed, which serves as an example to validate the proposed framework.

The remainder of this study is given as follows. Section 2 introduces related work. In section 3, a detailed technical roadmap has been demonstrated and explained. A real-world case study from the steel industry is shown in section 4. Section 5 concludes this study and gives future insights.

I. Related Work

A. Link Prediction

Link prediction has been proposed as a process of completing KG by discovering and adding the missing and implicit relations. Existing knowledge is utilised to infer potential relations between entity pairs for KG completion. Essentially, link prediction augments KG in some respects by increasing the edge number to enhance reliability.

The existing methods of link prediction are divided into three types: decomposition-based methods, path-based methods, and embedding-based methods [18]. Since many parameters need to be adjusted, decomposition-based methods suffer two major limitations: low efficiency and poor scalability [19]. A path from one entity to another entity has been regarded as a sequence of edges in path-based models, and relation-structure information has been considered as input features regarding a stronger representation of new relations [20]. Additionally, these approaches have been easily extended to large-scale graphs. However, these models have more computational complexity, resulting in high computational costs. Unlike the two former methods, embedding-based models have been proposed to learn the semantic representations of entities and relations simultaneously. One major advantage of KG embedding algorithms is that the structural and underlying semantic information has been preserved in the process of mapping entities and relations into a low-dimensional vector space to a great extent. As a powerful technique, embedding-based approaches play an important role predict the potential links in KGs. Especially, the link prediction models based on deep learning (e.g., GraphSAGE) have proven to be more beneficial accompanied by the success of deep learning.

B. Industrial KG Construction

Generally, iKG construction is composed of two major parts: knowledge extraction and iKG completion. The earliest research focused on knowledge extraction, including named entity recognition and relation extraction. These models have continued to be popular and dominated most research in constructing iKGs. For instance, a hybrid approach was built to gain an effective representation of sentence semantics and output the maximum probability sequence for chemical-named entity recognition [4]. A reinforcement learning approach was deployed to collect knowledge from the ternary space of humans, cyberspace and the Internet of Things (IoT) [21]. The interdependencies between different time series and the before-and-after relations of time series were mined using a CNN and LSTM, respectively [22].

Subsequently, as one of the promising research topics, iKG completion is arousing great attention. For instance, a GCN and tensor factorisation model was used as an end-to-end learning model to predict the missing entities in knowledge completion [23]. A pre-training model, namely LP-BERT, was proposed to leverage the linkage prediction strategy using the semantic matching representation [24].

As mentioned previously, KG has been referred to as the advanced technology of knowledge representation, which provides an important way to integrate multi-source information in industrial fields. However, the construction of iKG has certain limitations. Relation information of iKGs constructed from raw data is inadequate, where the missing relationships are not included in the iKG.

II. Methodology

In this section, a framework of iKG construction is proposed based on ontology and link prediction. In Figure 1, the overall framework consists of three main stages. In the first stage, a two-step section comprising ontology design and knowledge extraction is conducted. After clarifying the request for the iKG construction, the class hierarchy structure is designed to conform to the requirements thoroughly. Based on that, the property hierarchy is confirmed by the class hierarchy, including the object property hierarchy and the data property hierarchy. For the second stage, a two-layer GNN model is built to predict potential relations for iKG completion. Specifically, the edges are divided into two types: existing edges and non-existing edges. The relations represented by the embeddings of two entities are classified in the pre-trained GNN model. Finally, the third stage contains iKG construction and iKG visualisation. The overall triplets are described by the RDF format based on the triplet integration. Subsequently, the open-source platform is used to fuse the triplets for iKG construction, and the visualisation tool is deployed to present the generated iKG in a graphic format.

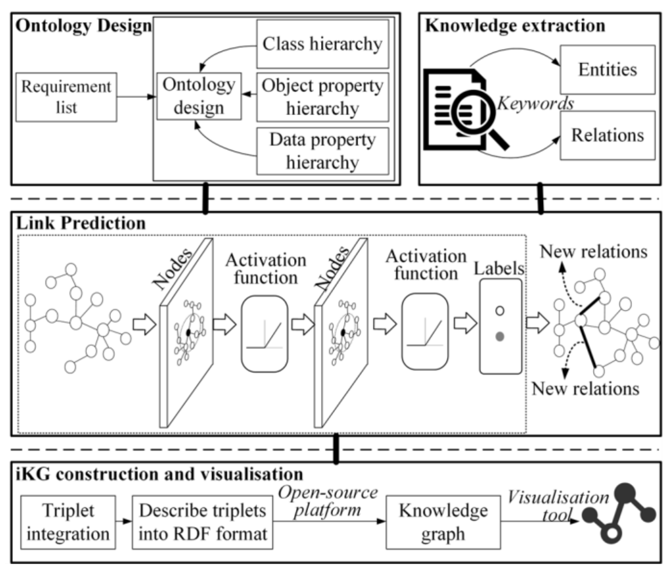

Figure 1. The overall framework of iKG construction.

A. *Industrial Ontology Design*

For iKG construction, the first key step is to determine the applicable ontologies. As semantic data models, industrial ontologies are mainly utilised to describe the relationships between concepts in a given domain and provide standardised, clear and unambiguous definitions that can be shared. Specifically, although the general types of things that share certain properties are modelled in industrial ontologies, these models do not contain information about specific individuals from domains.

In this section, a flowchart of ontology design has been proposed in Figure 2. Two minor revisions are incorporated to enhance the effectiveness of the ontology design process. Firstly, 5M1E (including man, machine, material, method, measurement, and environment) was introduced to determine whether or not the data collected was comprehensive. Secondly, to aid in the process of the design process, two tools were introduced (SmartKG and *OOPS!*) to be used to examine and test temporary ontologies. It avoids the need to evaluate and revise the final ontology after it has been constructed. In this case, it is possible to modify and improve ontologies in design procedures in an effective and time-saving manner.

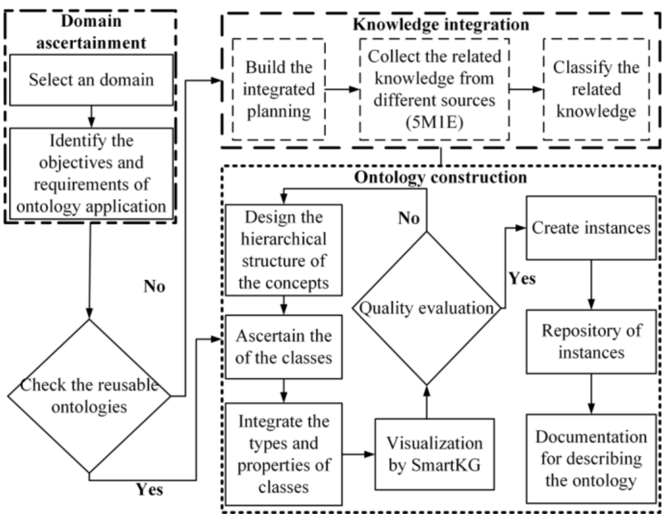

Figure 2. The detailed flowchart of industrial ontology design.

Firstly, the specific industrial scenario is determined, allowing us to identify the objectives and requirements of ontology applications. Subsequently, it is necessary to check if there are any reusable ontologies for the selected domain, providing participants to gain insight into both its opportunities and challenges. In the case of reusable ontologies, the ontology construction is carried out based on the previous ontologies in the next step immediately. Otherwise, industrial knowledge integration should be accomplished before constructing ontology. Secondly, three regular procedures have been applied to integrate the domain knowledge in the stage of industrial knowledge integration. The fragmented domain knowledge is collected from six different sources (5M1E) comprehensively after building the integrated planning. After that, the collected information should be further classified into different and incompatible subsets for the next step. In the last step, the hierarchical structure of concepts should be designed first in the context of a given domain. The top-level classes have been depicted as the root of concepts in the hierarchical structure. It has been ascertained that classes have properties and interrelations, and constraints. Then, an implementation of the SmartKG framework, developed by Microsoft, is made to visualise the developed hierarchy structure quickly. It will be possible to quickly determine if the hierarchy structure is appropriate and complete. In the case of high-quality classes, the instances are created and collected into a repository for ontology construction. Following that, it is to document the ontologies into a file, as it is essential to achieve iKG. Otherwise, the properties and interrelationship constraints of the classes are needed to be ascertained and fused again.

B. *Link Prediction*

As an inductive representation learning for node embedding, the GraphSAGE algorithm is especially advantageous and useful for large graphs with rich node attribute information [25-27]. The main idea of GraphSAGE is to adhere to GNN and aggregate the neighbours' information by embedding them into each node. Generally, the GraphSAGE-based link prediction method involves propagating GraphSAGE networks forwards and propagating GraphSAGE networks backwards.

A graph is represented by $=(\mathbf{V},\mathbf{E})$, where $\mathbf{V}$ denotes entities, and $\mathbf{E}$ is relations between entity pairs. Specifically, the i_{th} node is indicated by $v_i \in \mathbf{V}$, and the features of all nodes are defined as $\mathbf{X}_v, \forall v \in \mathbf{V}$. Meanwhile, the adjacency matrix $\mathbf{A} \in \mathbf{R}^{|n|\times|n|}, A_{\boldsymbol{ij}} \in \{0,1\}$ is usually used to describe $\mathbf{E}$, which is a $|n| \times |n|$ square matrix. If an edge exists between node v_i and node v_j, then $A_{\boldsymbol{ij}}=1$, otherwise $A_{\boldsymbol{ij}}=0$.

In the embedding process, the node's current representation h_u^{k-1} has been concatenated with the aggregated neighbours' vector $\mathrm{h}_{N(v)}^{k-1}$. Then, this combined vector is fed into a fully connected layer with nonlinear activation function σ, which updates the representations for the final representation. The mean aggregator has been applied in this study:

$$\mathrm{h}_v^k \leftarrow \sigma(\mathbf{W}\cdot MEAN(\{\mathrm{h}_u^{k-1}\} \cup \{\mathrm{h}_u^{k-1}, \forall u \in N(v)\})) \qquad (1)$$

For learning the weights of aggregators and embeddings, the cross-entropy function is usually utilised as a loss function in this study, as shown as follows:

$$\mathcal{L}=-\frac{1}{N}\sum_i \mathcal{L}_i=\frac{1}{N}\sum_i -[y_i \cdot \log(p_i)+(1-y_i)\cdot \log(1-p_i)] \quad (2)$$

where: $y_i \in \{0,1\}$ is label, $y_i=1$ for the positive sample, and $y_i=0$ for the negative sample; p_i is the predicted probability that sample i is a positive sample.

Figure 3 illustrates a link prediction model based on GraphSAGE. Firstly, iKGs are used to derive and compute the node-feature matrices and adjacency matrices. In this context, iKGs are expressed by the attributes of the entities (defined as node features) and the relationship features (regarded as adjacency matrices). Then, the two types of matrices are fed into a two-layer GraphSAGE model, which is a supervised training model. The features of each node have been updated by its neighbour within the share parameters. After propagating two GraphSAGE layers, the updated representations of each node have been obtained. Compared with the common GraphSAGE model, the sample labels are different from the node classification. Since the goal task is regarded as the link classification, the edges are divided into two different types: positive edges and negative edges. In terms of positive edges, the real-existing edges are defined as positive edges. Similarly, the non-existing edges are considered negative edges. In this context, the link prediction task can be considered a binary classification. For the next step, the embeddings of each edge have been represented by two linked entities in the proposed model. Accordingly, the proposed model employs the rule that minimises the classification error to achieve the final embeddings of nodes. In this context, the proposed model has already been trained to estimate the possible edges.

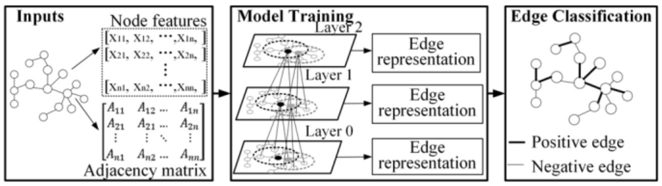

Figure 3. The architecture of the two-layer GraphSAGE model.

C. iKG Construction and Visualization

The key steps of iKG construction and visualisation are present in this section. The potential triplets have been discovered by the GraphSAGE-based link prediction model. After that, all triplets are integrated into a united repository. Through using open-source platforms, the RDF language is used to describe each triplet for the industrial knowledge representation. In other words, the iKG is constructed and stored on these platforms. Meanwhile, graph visualisation tools are often embedded in these platforms. In this context, iKG visualisation is usually accompanied by iKG construction.

Several independent open-source packages can directly achieve iKG visualisation. Presently, several open-source graph databases are available, including Neo4j, Gephi, Grakn etc. The iKG has been visualised using Gephi in this study because of its user-friendliness and high visual performance.

III. CASE STUDY

A number of important industries (i.e. military and defence, and manufacturing) rely on the production of iron and steel by providing raw materials, making it one of the largest industries in the world [28]. Meanwhile, cold rolling in the steel-making industry is recognised as an important process in the production of electrical steel strips because of its advantages with regard to accuracy, efficiency, and output rate. Presently, cold rolling contributes to the improvement of the properties of steel strips on changes both in the microstructure and thickness of the steel. Since the properties that have been improved include surface smoothness, tensile strength, yield strength and hardness, cold-rolled products usually have superior mechanical properties, small dimensional tolerances and high-quality surfaces [29]. As science and technology continue to advance, the quality requirements for steel strip products from cold rolling processes are becoming more detailed and demanding. Therefore, it is imperative increasingly to analyse and monitor the quality of cold-rolled products. This section conducted a real-world experimental study on the cold rolling process to validate the proposed framework.

A. Cold Rolling in the Steel Industry

Regarding the modern steel industry, steel strips are produced by cold rolling in a high-speed, high-precision, and continuous process. It is not uncommon for cold rolling to encounter certain defects in the same manner as the process of metal forming. As the most serious defect, strip breakage needs to be paid special attention to as it causes huge financial losses. Specifically, strip breakage has damaged rolls and mill accessories badly, not only the increase of production costs. Hence, the information integration of different resources around the cold rolling process of the steel industry contributes significantly to the prediction of this failure. The relevant procedures contribute to the influence of different degrees of the quality of cold-rolled products, including hot-rolling process, annealing process, pickling process, cold-rolling process, and quality inspection. The dataset was collected and stored from these resources, which covers a production period of six months in this study. Specifically, this historical dataset contains 1254 samples, including 94 variables.

B. Ontology for Cold Rolling

In this section, a strip-breakage ontology for cold rolling was designed. The strip breakage knowledge involves steel-coil material, chemical reagents, transportation, processing, treatment, etc. The class hierarchy considers the concepts of processes, facilities, products, operations, parameters, chemicals, fault diagnosis, etc. Meanwhile, these seven parts are regarded as the top classes. Based on the provided concepts, the subclasses of each top class are determined subsequently.

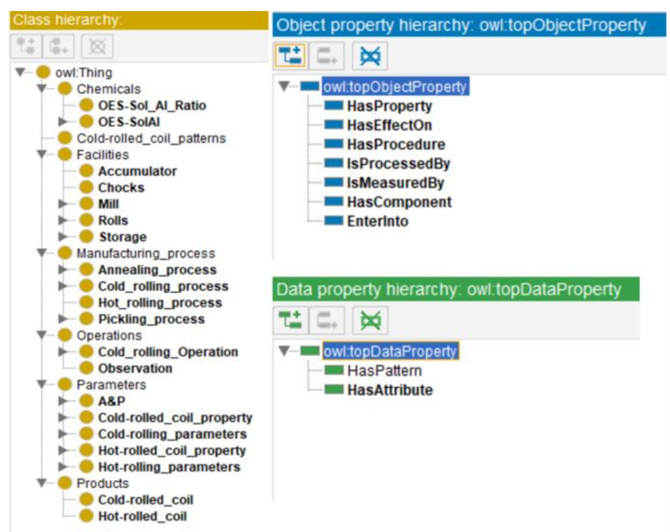

Figure 4. The classes, object properties and data properties of the strip-breakage ontology.

The strip-breakage ontology was finished in open-source software (Protégé5.5), which supports OWL. The concepts, properties, and associated relationships of the strip-breakage ontology were defined in this software. Figure 4 shows an illustrative example of the classes, object properties and data properties of the strip-breakage ontology. There are seven main concepts in this ontology, including 'chemicals', 'cold-rolled coil patterns', 'facilities', 'manufacturing processes', 'operations', 'parameters', and 'products'. In sum, the second level of ontology comprises 20 subclasses, such as 'OES-

SoIAI', 'Mill', 'Observation', 'accumulator', 'pickling process', etc. Meanwhile, the object properties and the data properties are defined as well for a better understanding of the concepts in Figure 4.

C. Link Prediction

In this section, the experiments were conducted to predict the potential edges for relation completion of the strip-breaking KG. A brief overview of the dataset is provided in the following. The dataset contains 13113 samples, which are divided into two types (existing edges and non-existing edges). Regarding training and testing data split, the 5-fold cross-validation tests were conducted to evaluate the performances, and the mean values of the five folds were outputted and labelled. Specifically, 80 per cent of the overall data were selected for training, and the remaining 20 per cent was reserved for testing.

As mentioned above, the experiments were conducted using our proposed two-layer GraphSAGE model. Meanwhile, three common machine learning algorithms were built to compare the performances, including back propagation neural network (BPNN), SVM, and random forest (RF). For the BPNN model, the hidden unit number and learning rate were set to 50 and 0.71, respectively, and the training epoch was set to 500. In terms of the SVM model, the kernel function was set as a radial basis function. Lastly, for the RF model, the number of estimators was set to 600. Based on that, the edges were fed respectively into three different models (BPNN, SVM, and RF) after calculating the representations of entities. Meanwhile, as shown in Figure 3, the representations of graphs (entity matrices and adjacency matrices) were fed into the two-layer GraphSAGE mode to predict the potential for relation completion.

Moreover, in the task of link prediction, the goal is to output the potential edges by predicting the edge types. Since the edges are classified into two types (existing edge and non-existing edge), the link prediction is considered a binary classification. In this context, five following metrics were introduced to evaluate the performances of the proposed model, including accuracy, precision, recall, F1-Score, and False Alarm Rate (FAR):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

$$F1-Score = 2 \times \frac{Recall \times Precision}{Recall+Precision} \tag{6}$$

$$FAR = \frac{FP}{FP+TN} \tag{7}$$

where: TP, TN, FP, FN mean true positive, true negative, false positive, and false negative, respectively.

TABLE I. THE BINARY CLASSIFICATION RESULT OF FOUR MODELS.

Models	Accuracy	Precision	Recall	F1-Score	FAR
BPNN	76.99	70.45	73.17	71.78	20.46
SVM	77.34	71.13	72.96	72.04	19.74
RF	78.06	72.17	73.48	72.82	18.89
GraphSAGE	**82.02**	**76.6**	**79.26**	**77.9**	**16.14**

Table 1 gives the detailed binary classification result by five different metrics of BPNN, SVM, RF, and GraphSAGE models. It is obvious that the GraphSAGE model shows advantages over the other three machine learning models. Specifically, the GraphSAGE model surpasses the other machine learning baselines in terms of accuracy (82.02%), precision (76.6%), recall (79.26%), and F1-Score (77.9%). Additionally, the experimental results indicate that the GraphSAGE model had a smaller value on the FAR metric (16.14%) than the other three models. In this context, the GraphSAGE classifier performs well across all different performance metrics, showing that incorporating the neighbours' information is better for the link prediction task in this study. Moreover, for the considered three benchmark models, it can be observed that the RF model achieves much better performances on all different metrics than SVM and BPNN models. Despite the fact that the BPNN classifier had a better recall than the SVM classifier, both classifiers showed about the same level of performance during the comparison experiment.

After training the GraphSAGE classifier, the existing graphs were fed into the proposed model to learn the edge representations and discover the potential edges. As shown in Figure 5, the missing relations are found to complete the strip-breakage KG, for example, the relation between 'DSP width' and 'Z6Temp std dev' and the relation between 'Gauge average (microns)' and 'Crown min (microns)'.

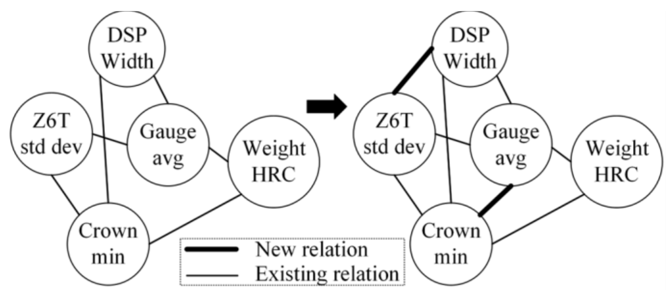

Figure 5. An illustrative example of link prediction.

D. Construction of Knowledge Graph

Based on ontology and link prediction, all triplets were integrated to construct the strip-breakage-centric KG of the cold rolling process. In this section, the free software Gephi was implemented to store and visualise the generated iKG due to its easy operation and great function. Specifically, 2295 triplets were integrated and imported into Gephi to construct and visualise the strip-breakage KG in the cold rolling process of the steel industry. In Figure 6, an entire strip-breakage-centric KG composed of seven subclasses is present. The entire strip-breakage KG contains 230 entities and 2295 relations.

IV. CONCLUSIONS AND FUTURE WORK

In this study, a framework of iKG construction is proposed, containing two parts: ontology design and link prediction. Firstly, the relevant knowledge is extracted and integrated to construct a strip-breakage ontology in the cold rolling process.

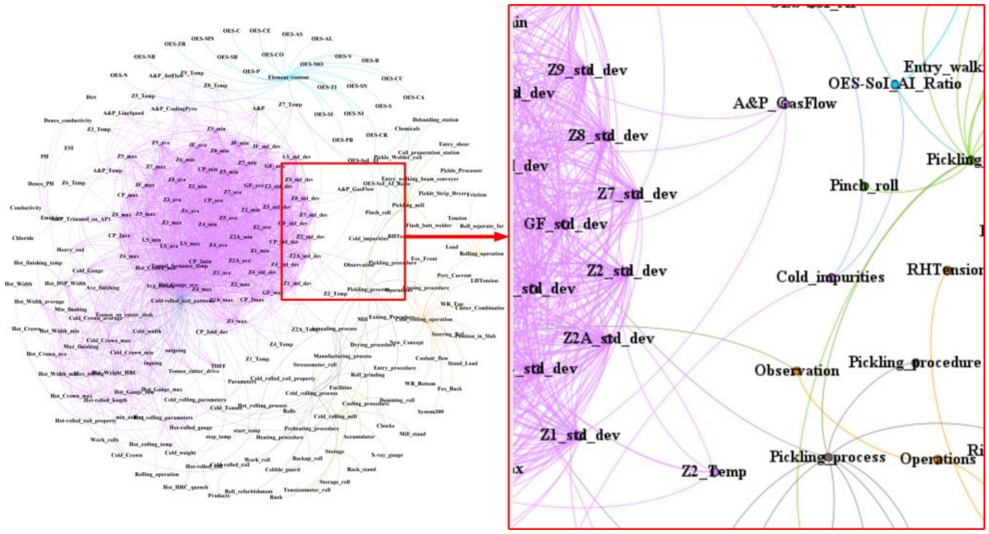

Figure 6. The entire strip-breakage KG.

Secondly, the potential relations are predicted by a two-layer GraphSAGE model from the existing knowledge. The experimental results show that the proposed model performs better on five different metrics than other baseline models. Methodologically, this framework improves the reliability, efficiency and effectiveness of constructing iKGs, and enables the effective exploitation of knowledge in industrial scenarios. In the future, the following research work will focus on developing and refining iKGs automatically.

REFERENCES

[1] K. Zhang et al., "Review on the Application of Knowledge Graph in Cyber Security Assessment," in IOP Conference Series: Materials Science and Engineering, 2020, vol. 768, no. 5, p. 052103.

[2] M. Cocchi, "Introduction: Ways and means to deal with data from multiple sources," in Data handling in science and technology, vol. 31: Elsevier, 2019, pp. 1-26.

[3] W. Zheng et al., "Pay attention to doctor-patient dialogues: Multi-modal knowledge graph attention image-text embedding for COVID-19 diagnosis," Information Fusion, 2021.

[4] P. Zheng et al., "Towards Self-X cognitive manufacturing network: An industrial knowledge graph-based multi-agent reinforcement learning approach," Journal of Manufacturing Systems, vol. 61, pp. 16-26, 2021.

[5] H. Paulheim, "Knowledge graph refinement: A survey of approaches & evaluation methods," Semantic web, vol. 8, no. 3, pp. 489-508, 2017.

[6] T. Al-Moslmi et al., "Named entity extraction for knowledge graphs: A literature overview," IEEE Access, vol. 8, pp. 32862-32881, 2020.

[7] Q. Guo et al., "A survey on knowledge graph-based recommender systems," IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 8, pp. 3549-3568, 2020.

[8] M. Yasunaga et al., "QA-GNN: Reasoning with language models and knowledge graphs for question answering," arXiv preprint arXiv:2104.06378, 2021.

[9] X. Zheng et al., "A knowledge graph method for hazardous chemical management: Ontology design and entity identification," Neurocomputing, vol. 430, pp. 104-111, 2021.

[10] M. Lyu et al., "Achieving Knowledge-as-a-Service in IIoT-driven smart manufacturing: A crowdsourcing-based continuous enrichment method for Industrial Knowledge Graph," Advanced Engineering Informatics, vol. 51, p. 101494, 2022.

[11] L. Getoor et al, "Entity resolution: theory, practice & open challenges," Proceedings of the VLDB Endowment, vol. 5, no. 12, pp. 2018-2019, 2012.

[12] A. McCallum, "Information extraction: Distilling structured data from unstructured text," Queue, vol. 3, no. 9, pp. 48-57, 2005.

[13] B. Abu-Salih, "Domain-specific knowledge graphs: A survey," Journal of Network and Computer Applications, vol. 185, p. 103076, 2021.

[14] X. Li et al., "Exploiting knowledge graphs in industrial products and services: a survey of key aspects, challenges, and future perspectives," Computers in Industry, vol. 129, p. 103449, 2021.

[15] X. Li et al, "Achieving Cognitive Mass Personalization via the Self-X Cognitive Manufacturing Network: An Industrial Knowledge Graph- and Graph Embedding-Enabled Pathway," Engineering, 2021.

[16] E. E. Kosasih et al., "A machine learning approach for predicting hidden links in supply chain with graph neural networks," International Journal of Production Research, vol. 60, no. 17, pp. 5380-5393, 2022.

[17] S. Yan et al., "Semi-supervised fault diagnosis of machinery using LPS-DGAT under speed fluctuation and extremely low labeled rates," Advanced Engineering Informatics, vol. 53, p. 101648, 2022.

[18] J. Ma et al., "ELPKG: A high-accuracy link prediction approach for knowledge graph completion," Symmetry, vol. 11, no. 9, p. 1096, 2019.

[19] K.-W. Chang et al., "Typed tensor decomposition of knowledge bases for relation extraction," in Proceedings of the 2014 conference on empirical methods in NLP (EMNLP), 2014, pp. 1568-1579.

[20] N. Lao et al., "Random walk inference and learning in a large scale knowledge base," in Proceedings of the 2011 conference on empirical methods in natural language processing, 2011, pp. 529-539.

[21] S. Chen et al., "Top-down human-cyber-physical data fusion based on reinforcement learning," IEEE Access, vol. 8, pp. 134233-134245, 2020.

[22] B. Zhou et al., "Semantic-aware event link reasoning over industrial knowledge graph embedding time series data," International Journal of Production Research, pp. 1-18, 2022.

[23] A. Zeb et al. , "KGEL: A novel end-to-end embedding learning framework for knowledge graph completion," Expert Systems with Applications, vol. 167, p. 114164, 2021.

[24] D. Li et al., "LP-BERT: Multi-task Pre-training Knowledge Graph BERT for Link Prediction," arXiv preprint arXiv:2201.04843, 2022.

[25] W. Hamilton et al., "Inductive representation learning on large graphs," Advances in neural information processing systems, vol. 30, 2017.

[26] H. Yu et al., "Automatic segmentation of parallel drainage patterns supported by a graph convolution neural network," Expert Systems with Applications, vol. 211, p. 118639, 2023.

[27] Z. Chen et al., "A new energy consumption prediction method for chillers based on GraphSAGE by combining empirical knowledge and operating data," Applied Energy, vol. 310, p. 118410, 2022.

[28] Z. Fan et al., "Low-carbon production of iron and steel: Technology options, economic assessment, and policy," Joule, vol. 5, no. 4, pp. 829-862, 2021.

[29] C. Wu et al., "Multiscale Interface Stress Characterisation in Cold Rolling," Metals and Materials International, vol. 27, no. 7, pp. 1997-2013, 2021.