Journal of Engineering, Project, and Production Management
2026, 16(4), 2025-219

Optimization and Visualization of Industrial Production Knowledge Graph Based on BERT and RGCN

Shucheng Li
Associate Professor, School of Law and School of Intellectual Property, Tiangong University, Tianjin, 300387, China, E-mail: tjgydxresearch2004@163.com

Production Management
Received October 10, 2025; revised December 22, 2025; accepted December 23, 2025
Available online May 29, 2026

Abstract: To improve the accuracy and efficiency of existing methods for optimizing and visualizing multivariate data knowledge graphs, a new model is proposed. The industrial production knowledge graph optimization and visualization model integrates a Bidirectional Encoder Representation (BERT) model, a Relational Graph Convolutional Network (RGCN), and Scalable Vector Graphics (SVG) technology. This model aims to enhance the accuracy of data fusion and semantic relationship mining by optimizing the bidirectional encoder representation through Conditional Random Fields (CRF). It also optimizes the knowledge graph using relational graph convolutional networks and Graph Attention Networks (GAN), and combines Scalable Vector Graphics (SVG) technology to complete its visualization. The main innovation of the research lies in overcoming the limitations of the existing method's insufficient adaptability to diverse industrial data by constructing a collaborative framework that features semantic enhancement, knowledge graph optimization, and efficient visualization. The study compared the proposed model with the optimization model based on deep learning. The results show that the proposed model's recall rate for equipment fault identification after optimizing the knowledge graph is 98.22%. The resource utilization rate and response delay during optimization are 27.81% and 12.8 minutes, respectively. Meanwhile, the average efficiency of this model in visualizing multiple-entity data is 96.69%, 93.66%, and 96.26%, respectively, with the structural clarity of visualizing the temperature data in the operation log at 97.26%. All experimental results outperform the comparison models, fully proving the feasibility and advantages of the proposed model. This study provides new ideas and methods for the further development of knowledge graphs.

Keywords: Multisource data, CRF-BERT, semantic relations, GAT-RGCN, knowledge graph visualization, scalable vector graphics.

DOI 10.32738/JEPPM-2025-219

1. Introduction

Applying knowledge graph technology to industrial production can improve efficiency and reduce risks, making its optimization necessary (Peng et al., 2023). The industrial knowledge graph optimization aims to improve knowledge graphs constructed in industrial production scenarios by using relevant algorithms or technologies, including but not limited to the integrity of the graph structure, the accuracy of semantic relations, and computational efficiency. Under the premise of ensuring the timeliness of industrial data and the consistency of entity relationships, produce high-quality knowledge graphs suitable for industrial tasks such as fault diagnosis and process optimization. However, existing knowledge graph optimization methods have limitations. For instance, the random forest algorithm is highly dependent on data and has poor recognition and processing capabilities for unknown data or cross-domain data. Although Bayesian networks are suitable for complex fault diagnosis involving multi-factor coupling, they have problems such as insufficient timeliness and insufficient accuracy in knowledge graph modeling when dealing with large-scale and diverse data. Therefore, a more comprehensive optimization method is needed (Choudhuri et al., 2023). The Bidirectional Encoder Representation from Transformers (BERT) model has powerful semantic representation and feature extraction capabilities. It can integrate and process diverse data in industrial production to construct high-quality knowledge graphs (Garrido-Merchan et al., 2023). The Relational Graph Convolution Network (RGCN), a variant of graph neural networks, endows it with powerful knowledge graph modeling capabilities through a relation-weight matrix and a direction-aware mechanism (Li et al., 2023). Based on this, the study adopts BERT and RGCN as the core frameworks for processing multisource semantic text data and optimizing the knowledge graph. It further introduces Conditional Random Fields (CRF) and Graph Attention Networks (GAT) to improve BERT and RGCN, respectively. Two hybrid algorithms were combined with Scalable Vector

Graphics (SVG) technology to construct a knowledge graph optimization and visualization model suitable for industrial production. It is expected to solve problems in current optimization methods, such as insufficient fusion of multiple data sources and low efficiency in both graph optimization and visualization. The innovation of this research lies in the development of a comprehensive solution for optimizing and visualizing industrial production knowledge graphs across three dimensions: theory, method, and application. Theoretically, by improving the loss function forms of CRF and BERT, supplementing the attention weight distribution constraints of GAT and RGCN, the gains of CRF for semantic error correction of BERT and GAT for parameter dimensionality reduction of RGCN are quantitatively demonstrated. In terms of methods, the research adopted an industrial entity group transfer matrix to reduce the parameter scale of CRF. It reduced the computational complexity of RGCN by dynamically adjusting the number of attention heads in GAT, and simultaneously shortened the deployment time of the SVG visual knowledge graph by incorporating an incremental rendering mechanism. In terms of application, the research integrates the semantic enhancement of multi-data, the optimization of knowledge graph structure, and lightweight visualization, forming an end-to-end closed loop that takes into account the resource constraints of industrial edge devices and solves the limitations of existing methods, such as complex deployment and insufficient adaptation to multi-data.

2. Related Works

The BERT model extracted contextual semantic information in both directions, enabling it to better understand various types of text data. Based on this, numerous scholars both at home and abroad have conducted related research. For example, Talaat (2023) proposed four deep learning models based on BERT to improve sentiment analysis accuracy and support strategic decision-making for businesses. The study used pretrained word embedding vectors to adjust the model parameters. The results showed that the BERT model with a BiGRU layer achieved the highest accuracy in sentiment classification. Bilal and Almazroi (2023) proposed a generalized classification method based on BERT to address the weak generalization ability of existing methods for predicting online platform customer reviews. They verified its effectiveness by comparing it with the traditional bag-of-words approach. The experiments showed that the proposed method demonstrated strong capabilities for handling long-sequence classification tasks. RGCN, with its unique relation-specific weight matrix, has been widely used in fields such as drug discovery and knowledge graph reasoning. For example, Wang et al. (2023) designed a prediction method based on RGCN to explore the internal relationship between microbes and diseases. They used principal component analysis to process data from multiple databases, and the resulting features as initial inputs for the prediction model. The results clearly and intuitively revealed the relationship between microbes and diseases. To reduce the impact of noise on predicting the remaining useful life of equipment, Zhu et al. (2023) proposed an RGCN with an uncertainty estimation. By mining the graph structure, the method learned spatial and temporal correlations. The study found that the model significantly improved the robustness of the prediction performance. Zhang et al. (2023) suggested a dual-attention graph convolutional network based on RGCN to improve the model's performance in handling text semantic relations. The method built bidirectional information flow to achieve multi-round interaction between contextual information and dependency relations. The results showed that this method effectively enhanced the performance of existing models.

Research on knowledge graph optimization methods have made significant progress, and many scholars have applied these methods to practical scenarios. For instance, to address the challenge of identifying hidden missing relations in knowledge graphs, Ren et al. (2023) built a knowledge graph framework and proposed an attention-based graph embedding model. They used the production process of an aerospace enterprise as the experimental case, and the results proved the model's effectiveness. Zhou et al. (2023) aimed to improve the processing efficiency of manufacturing systems for knowledge graphs. They proposed a semantic-aware event linking and reasoning method. By converting document information into industrial knowledge graphs, the method described dynamic semantic information in manufacturing environments. The results showed that it accurately captured the temporal characteristics of the system. To reduce the computational complexity and cost of traditional knowledge graph reasoning, Wang et al. (2024) introduced a two-stage algorithm based on TIGER. They decomposed complex problems into two subproblems to lower computational costs. The study found that this approach significantly reduced the running time of subgraph extraction. Xia et al. (2025) proposed a complex reasoning model for knowledge graphs based on large language models to improve the ability to answer complex queries over incomplete knowledge graphs. The model used a binary tree decomposition mechanism to enhance reasoning ability. The results showed that the model improved the average Mean Reciprocal Rank (MRR) score by 5.5% compared to traditional methods. To address the weak ability of existing knowledge graphs to capture semantic relations in text, Jin et al. (2023) developed an embedded model based on relational chains. They used explicit relational chains in natural language questions to enhance model performance. Experimental results showed that the model identified semantic data in multisource text.

In summary, although existing research on knowledge graphs has achieved meaningful results, most work on knowledge graph optimization focuses on general domains, lacking specific studies on the optimization of industrial production knowledge graphs. Moreover, there are problems such as poor quality of knowledge graph optimization and high computational complexity. Therefore, this study introduced CRF and GAT to optimize BERT and RGCN, respectively. Two hybrid algorithms were constructed and integrated with SVG technology to build an optimization and visualization model for industrial production knowledge graphs. This model aimed to optimize knowledge graphs by considering multisource data and semantic relations, and to present them clearly and intuitively.

3. Industrial Production Knowledge Graph Optimization and Visualization Model

3.1. CRF-BERT Hybrid Algorithm Design for Multisource Data Integration

Industrial production knowledge graphs refer to the comprehensive modeling of full-process knowledge and intelligent analysis using knowledge graph technology in the field of industrial production. They integrate multisource data to build a semantic network, enabling efficient use of the knowledge graph (Yahya et al., 2024). BERT, as a semantic pre-training model based on the transformer architecture, uses a bidirectional attention mechanism to accurately capture the polysemy of industrial terms and the contextual relevance in data. It extracts features from fragmented data (Talebi et al., 2024). The specific structure of BERT is shown in Fig. 1.

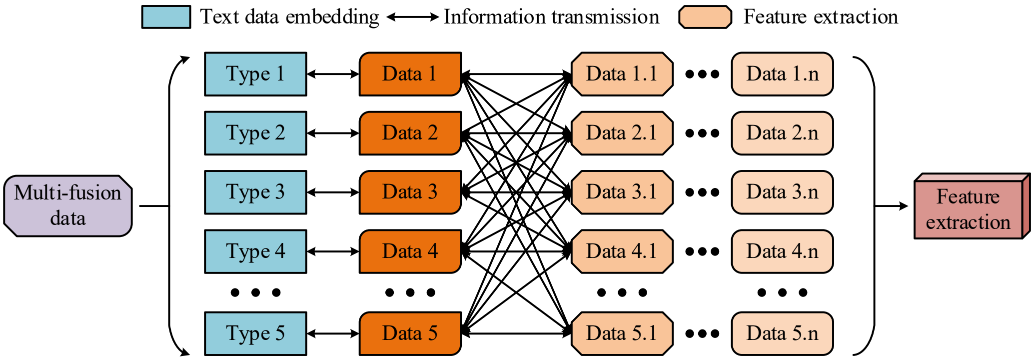


Fig. 1. Schematic diagram of BERT structure

As shown in Fig. 1, BERT classifies different types of semantic data and flexibly identifies data features by linking contextual information in both directions. This process lays the foundation for constructing knowledge graphs. BERT requires encoding multisource data into a format that matches its requirements for recognition. The process is expressed in Eq. (1).

$$X = \left\{ E_{token}, E_{value}, E_{pos} \right\} \tag{1}$$

In Eq. (1), X defines the encoding process of BERT, E_{token} refers to the embedding vector of the text data, E_{value} represents the linear transformation of data features, and E_{pos} denotes the embedding vector of positional or temporal information. BERT then extracts features through its encoder and outputs hidden states containing contextual information, as shown in Eq. (2).

$$H = \mathrm{BERT}(X) = \left\{ h_1, h_2, \ldots, h_n \right\} \tag{2}$$

In Eq. (2), H represents the feature extraction process, and h_n is the contextual information of the n-type data. When applied to industrial production knowledge graph optimization, a single BERT model has limitations, including low efficiency with long text and high computational complexity. Therefore, optimization is necessary. CRF predicts data sequence labels by analyzing contextual dependencies in the data and modeling medium- and long-distance dependencies within the sequences (Karthic and Kumar, 2023). For this reason, CRF is introduced to optimize BERT, forming the CRF-BERT hybrid algorithm. Its structure is shown in Fig. 2.

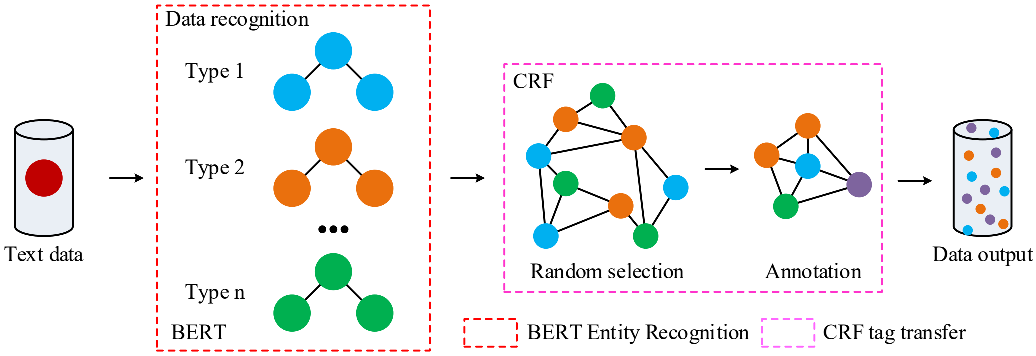


Fig. 2. Schematic diagram of CRF-BERT hybrid algorithm

As shown in Fig. 2, when processing multisource data with semantic relations, CRF-BERT first uses BERT to perform entity recognition on heterogeneous data and groups them based on contextual information. Then it applies the CRF label transition mechanism to the label data features, resolving ambiguities in the text. Finally, it generates results that integrate multiple data types with contextual semantic associations. Before CRF labels the data, it evaluates the features, as expressed in Eq. (3).

$$Score(X, y) = \sum_{i=1}^{i} s_i \left[y_i \right] + \sum_{i=1}^{i-1} A_{y_i, y_{i+1}} \tag{3}$$

In Eq. (3), $A_{y_i, y_{i+1}}$ denotes the probability of transitioning to the $i+1$-th data feature, $s_i\left[y_i\right]$ defines the evaluation result of CRF for the data feature, and $Score(X, y)$ refers to the total evaluation result of the y -th type of data feature. During the entire CRF-BERT data processing, a loss function is used to constrain the results and ensure output quality. This is shown in Eq. (4).

$$L = \frac{1}{N}\sum_{m=1}^{N} Score(X^{(m)}, y^{(m)}) - \log Z(X^{(m)}) \tag{4}$$

In Eq. (4), N is the total number of data processed by CRF-BERT, and $Z(X^{(m)})$ represents the normalization factor for all data sets m .

3.2. Knowledge Graph Optimization Method Based on RGCN-GAT

Although CRF-BERT processes multisource data effectively, it cannot optimize knowledge graphs. Therefore, it is necessary to introduce an algorithm that can optimize knowledge graphs. RGCN, with its relation-specific weight matrix and direction-aware mechanism, overcomes the limitations of traditional GCN in processing unstructured data. It is widely used in various knowledge graph optimization tasks (Zhang, 2025). The data transmission process in the graph convolutional layers of RGCN is expressed in Eq. (5).

$$M_i^{r(l)} = \sum_{r\in R} M_i^{r(l)} + W_0^l h_i'^l + b_0^l \tag{5}$$

In Eq. (5), $M_i^{r(l)}$ defines the message passing under a particular relation r , $h_i'^l$ is the feature vector of the knowledge graph i at the l -th convolutional layer, where W_0^l and b_0^l are both self-loop hyperparameters. Since knowledge graphs contain large amounts of data, RGCN often suffers from parameter explosion during optimization. GAT addresses this problem by introducing an attention mechanism that dynamically assigns attention weights to each node in the graph structure. This allows GAT to selectively capture important node dependencies, effectively avoiding parameter explosion (Zhang et al., 2023). Therefore, this study introduces GAT to improve RGCN. The process is shown in Fig. 3.

As shown in Fig. 3, the knowledge graph processed by RGCN is passed to GAT, which assigns attention weights and maps it to a linear representation to generate new feature representations. These features are then aggregated to update the initial knowledge graph. The core of GAT is its attention mechanism, which is expressed in Eq. (6).

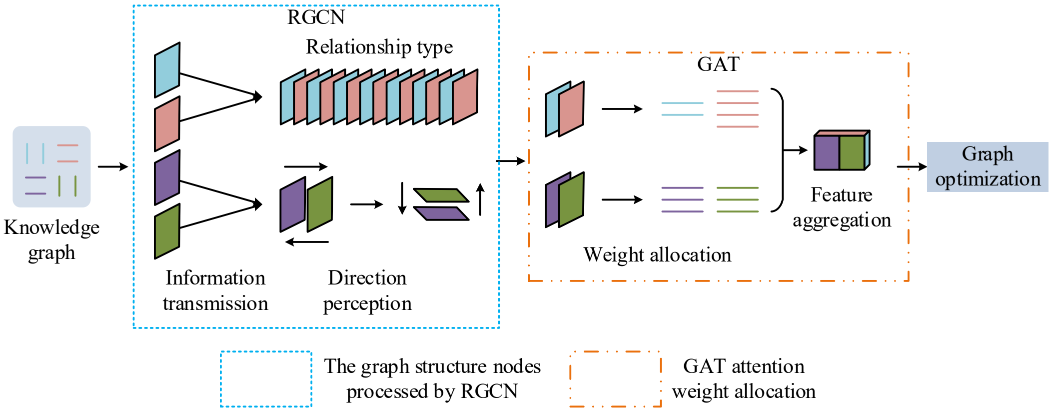


Fig. 3. Optimization mechanism of RGCN by GAT

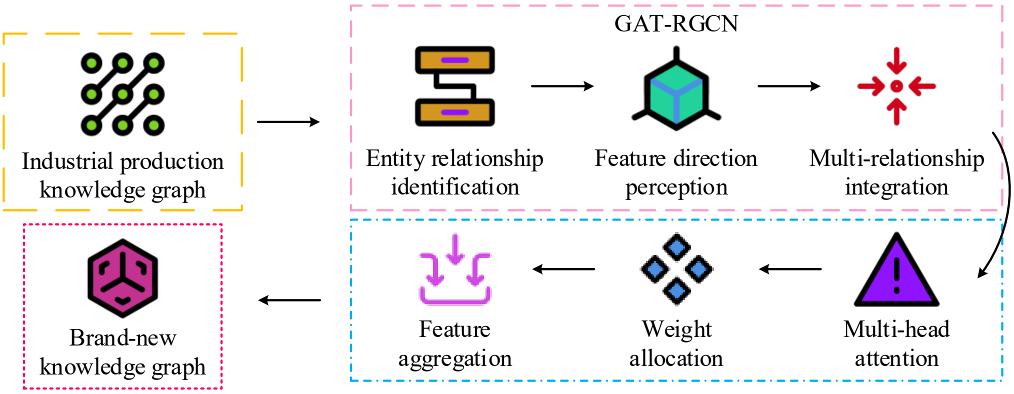


Fig. 4. Optimization process of industrial production knowledge graph based on GAT-RGCN (Icon source from: https://iconpark.oceanengine.com/home)

$$v_i^{l+1} = \Big\|_{k=1}^{K} \sigma\left(\sum\nolimits_{r\in R} \alpha_i^{r(k)} W_r^{lk} h_i^l\right) \tag{6}$$

In Eq. (6), v_i^{l+1} defines multi-head attention aggregation, K is the number of attention heads, $\alpha_{i,j}^{r(k)}$ represents the coefficient of the k-th attention head, and $\Big\|_{k=1}^{K}$ denotes the concatenated result of the attention weight distribution. The hybrid GAT-RGCN algorithm, obtained by applying GAT to optimize RGCN, performs well when handling knowledge graph data with complex relations and node features. Therefore, this study applies it to the optimization of industrial production knowledge graphs. The specific process is shown in Fig. 4.

As shown in Fig. 4, when applied to the optimization of industrial production knowledge graphs, GAT-RGCN first leverages RGCN's strong relational modeling to perform entity feature recognition and direction-aware analysis of the knowledge graph. Then, it aggregates features to complete knowledge graph optimization. GAT then assigns weights to the knowledge graph through its dynamic attention mechanism, prioritizing the aggregation of nodes with higher weight coefficients. Finally, the optimization of the industrial production knowledge graph is completed. The labeling process of entity data by GAT-RGCN is expressed in Eq. (7).

$$L' = -\sum\nolimits_{c\in C} y'_{i,c} \log\left(\frac{\exp(\omega_c h_i^l)}{\sum_{c'\in C}\exp(\omega_{c'} h_i^l)}\right) \tag{7}$$

In Eq. (7), L' is defined as the loss function of the entity knowledge graph v_i, C is the set of entity data, $y'_{i,c}$ represents the annotation of the entity knowledge graph under type c, and ω is the weight of the classification layer. The industrial entity knowledge graph optimized by the GAT-RGCN hybrid algorithm improves the accuracy of fault prediction and the efficiency of process optimization. Its mathematical representation is shown in Eq. (8).

$$h_i^{final} = \sigma(\alpha_{ij} W' h_{ij}^{RGCN}) \tag{8}$$

In Eq. (8), h_i^{final} defines the optimized industrial production knowledge graph, σ is the activation function, W' represents the updated self-loop hyperparameter, and h_{ij}^{RGCN} represents the knowledge graph that has been pre-processed by RGCN.

3.3. Construction of The Knowledge Graph Optimization and Visualization Model

CRF-BERT can integrate diverse data in industrial production and construct the relevant knowledge graph network. GAT-RGCN can optimize the knowledge graph through the attention mechanism while enhancing semantic relations. However, neither of them considers the visualization problem of the knowledge graph, and it is necessary to combine related technologies to achieve the visualization of the industrial production knowledge graph. SVG is a graphic technology suitable for high-resolution display and real-time interaction. Compared with traditional visualization technologies, SVG provides a more detailed visual experience (Goloboff and Morales, 2023). Therefore, this study combines CRF-BERT, GAT-RGCN, and SVG to develop an optimization and visualization fusion algorithm for industrial production knowledge graphs, with its specific structure illustrated in Fig. 5.

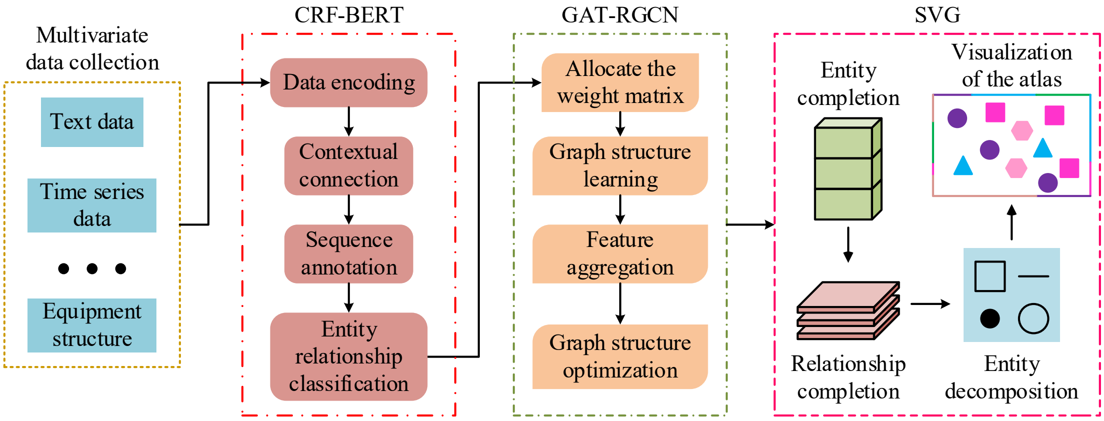


Fig. 5. Structure of the hybrid algorithm based on CRF-BERT, GAT-RGCN, and SVG

As shown in Fig. 5, when the fusion algorithm optimizes the knowledge graph, it first encodes diverse data from industrial production and then classifies relationships by connecting context-semantic information. Then, by virtue of the special mechanism of GAT-RGCN, the weight matrices of various types of knowledge graphs are calculated. Feature aggregation is carried out by learning the graph structure data of the knowledge graphs, thereby optimizing the knowledge graphs. Finally, through SVG for entity alignment and relationship completion, the complex and abstract knowledge graph is transformed into simple and precise edges, lines, and points, thereby achieving the visualization of the industrial production knowledge graph. When SVG performs entity alignment, it needs to determine the positions of each knowledge graph. The specific process is shown in Eq. (9).

$$\begin{cases} F_{attract}(i^*, j^*) = \dfrac{d(i^*, j^*)^2}{k'} \\ F_{repel}(i^*, j^*) = -\dfrac{k'^2}{d(i^*, j^*)} \end{cases} \tag{9}$$

In Eq. (9), $F_{attract}(i^*, j^*)$ represents the gravitational position of the knowledge graph, $F_{repel}(i^*, j^*)$ represents the repulsive position of the knowledge graph, and $d(i^*, j^*)$ is defined as the distance between nodes i^* and j^* within the knowledge graph, and k' is the layout area simulated by SVG. After that, the nodes positions are updated to complete the relation completion. This process is shown in Eq. (10).

$$x_i(t+1) = x_i(t) + \min(c', \max(c', \Delta x_i)) \tag{10}$$

In Eq. (10), $x_i(t+1)$ defines the updated position of node $t+1$ at the i-th iteration, and c' represents the specific type of entity data. In addition, SVG arranges entities in a concentric layout according to their importance. The process is expressed in Eq. (11).

$$x_i = r' \cdot \cos(\theta_i) + z_x \tag{11}$$

In Eq. (11), z_x is the center of the concentric circles, r' is the radius, and θ_i represents the level or importance of the entity data. The fusion algorithm can not only trim the redundant paths of similar relation groups through GAT-RGCN to alleviate problems such as node overlap and edge entanglement in the visualization of large-scale knowledge graphs, but also provide a more transparent topological structure for SVG. Moreover, it can precisely identify and categorize the knowledge graph using CRF-BERT, and encode and lay out node colors hierarchically, which is conducive to users quickly locating core knowledge graph entities and associated paths, and to improving the efficiency of information exploration and structural readability in industrial scenarios. Based on this, the study builds an optimization and visualization model for industrial production knowledge graphs using the hybrid algorithm. The process is shown in Fig. 6.

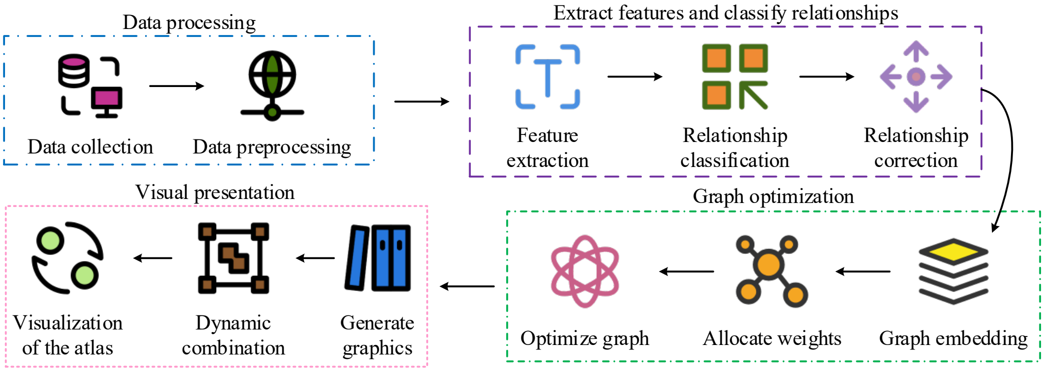


Fig. 6. Visualization process of the industrial production knowledge graph (Icon source from: https://iconpark.oceanengine.com/home)

As shown in Fig. 6, the model includes four main steps for optimizing and visualizing industrial production knowledge graphs. The first step is data collection and preprocessing. The second step is feature extraction and relation classification, which extracts features from the knowledge graph and identifies relations in the contextual information. The third step is graph evolution and optimization based on the weight matrix and attention mechanism. The final step converts the abstract graph into clear graphical representations and dynamically integrates them to complete the visualization. After transforming the knowledge graph into a graphical form, the weights of arrows, edges, and other elements must be mapped. This process is shown in Eq. (12).

$$stoke - width = \min(k^* \cdot \text{weight}, m^* \cdot \text{max} - width) \tag{12}$$

In Eq. (12), k^* and m^* are scaling factors, weight represents the relation weight, and $stoke - width$ defines the thickness of visual elements such as edges, nodes, and arrows. To enhance the operability of the model in practical applications, the study carried out lightweight optimization on the proposed model. In the semantic processing layer of BERT, the research reduces the 12-layer encoder to 6 layers by pre-training the weights using an industrial-domain knowledge graph, thereby reducing interference from irrelevant semantics. Meanwhile, in the knowledge graph optimization layer, the weight matrix of the shared RGCN is combined by merging similarity relations, and the number of attention heads in GAT is reduced according to node importance to lower the computational load. For the visualization operation of knowledge graphs, the research utilizes the incremental rendering mechanism to shorten the deployment time of the model. All the above lightweight optimizations are carried out to ensure the model's overall performance, thereby reducing its computational cost and implementation threshold.

4. Performance Verification and Application Evaluation of the Model

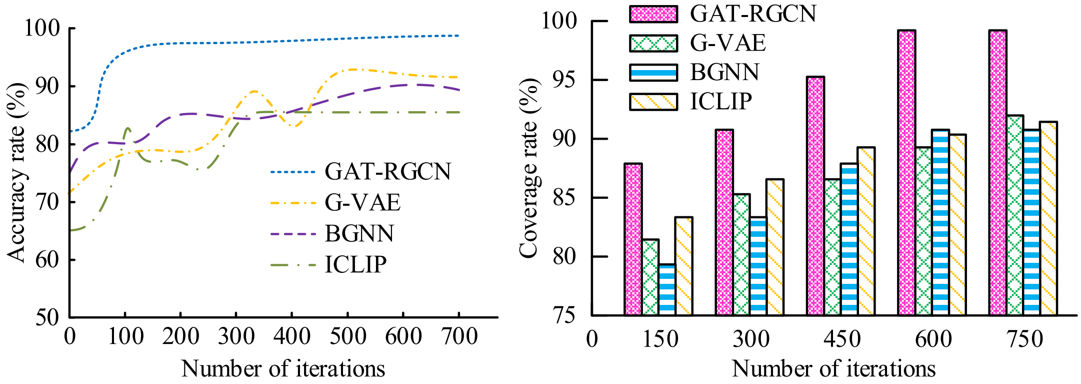
4.1. Performance Testing of GAT-RGCN

To evaluate the performance of GAT-RGCN in optimizing the knowledge graph, the study conducted comparative experiments with Graph Variational Autoencoder (G-VAE), Bayesian Graph Neural Network (BGNN), and Improved Contrastive Language-Image Pre-training (ICLIP). The specific experimental environment and parameter settings are shown in Table 1.

Table 1. Experimental environment and parameter settings

Computer configuration	Detailed parameter
CPU	Intel Core i7-14700KF
GPU	NVIDIA RTX 4070Ti
Internal memory	64GB DDR5, 6400MHz
store	2TB SSD and 2TB mechanical hard drive
Operating system	Windows 11 Professional
Programming language	Python 3.8
Dataset	SWAT data set

The SWAT dataset contained numerous data records, operation parameters, and equipment status, which supported the exploration of the industrial knowledge graph. The study highlights the superiority of GAT-RGCN by comparing the accuracy rate, node connection coverage rate, precision rate, and precision rate error of each algorithm. Among these, the accuracy rate, precision rate, and their errors are used to measure the overall performance of GAT-RGCN when optimizing the knowledge graph. In contrast, the node coverage rate is used to evaluate the integrity of entity data as GAT-RGCN processes the knowledge graph. The study first compared the accuracy of understanding entity semantics and node connection coverage during graph optimization using the four algorithms. The results are shown in Fig. 7.

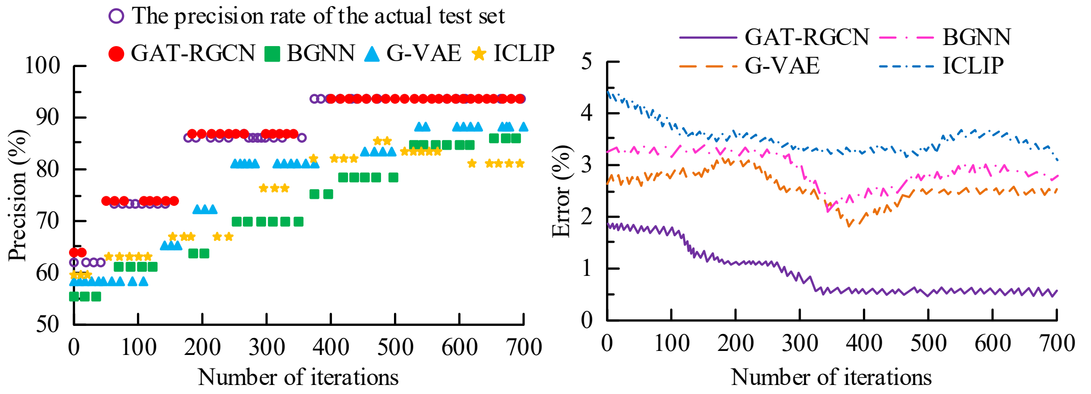


(a) Comparison of semantic understanding accuracy rates among the four algorithms

(b) Comparison of node connection coverage rates among the four algorithms

Fig. 7. Comparison of semantic understanding accuracy and node connection coverage

As shown in Fig. 7(a), GAT-RGCN achieved a semantic understanding accuracy of 97.94%, which was significantly higher than G-VAE (92.41%), BGNN (88.77%), and ICLIP (84.60%). Moreover, the accuracy of GAT-RGCN remained consistently higher than the three baseline models throughout the process, without any drop in performance as the number of iterations increased. However, after 300 iterations, G-VAE showed significant fluctuations in accuracy due to overfitting the local semantic features in the knowledge graph. BGNN and ICLIP were affected by the polysemy of terms in the knowledge graph, and their accuracy improvement rates gradually slowed as the number of iterations increased. From Fig. 7(b), the maximum coverage rate of GAT-RGCN node connections was 96.82%, which was significantly higher than that of the three comparison algorithms. This is because GAT-RGCN can preferentially aggregate the key entity pairs in the knowledge graph through the dynamic attention mechanism. At the same time, G-VAE is prone to losing low-frequency entity data during the knowledge graph reconstruction process. BGNN and ICLIP are limited by the knowledge graph's learning capacity and struggle to fully cover entity associations across modules. The study then compared the precision and error rate of fault diagnosis results using the graphs optimized by different algorithms. The results are presented in Fig. 8.



(a) Comparison of the fault diagnosis accuracy rates of the optimized graphs by different algorithms

(b) Comparison of fault diagnosis accuracy rate errors of the optimized spectra by different algorithms

Fig. 8. Results of precision and error in fault diagnosis

As shown in Fig. 8(a), after optimizing the knowledge graph with GAT-RGCN and applying it to fault diagnosis of industrial equipment, the accuracy is 94.73%, which is significantly higher than that of the comparison model. Moreover, after reaching the maximum value at the 400th iteration, it remains unchanged with the increase in the number of iterations without noticeable fluctuations. According to Fig. 8(b), the minimum precision error of GAT-RGCN was 0.51%, with an average error of only 0.96%, which was considerably lower than G-VAE (1.83%), BGNN (2.14%), and ICLIP (3.09%). These results indicated that GAT-RGCN provided highly reliable optimization for the industrial knowledge graph. Overall, optimizing the industrial knowledge graph with GAT-RGCN improved the accuracy of semantic knowledge representation and demonstrated strong capability in recognizing graph structural features. To further verify the universality, scalability, and resource consumption characteristics of GAT-RGCN, two datasets, namely ILPC22-Small and YAGO3-10, are newly introduced in this study. The knowledge graph used for training in the former contains 10,000 entities, 96 relationship types, and 78,000 triples, while the latter has over 12,000 entities, 37 relationship types, and 1 million triples. Meanwhile, two mainstream industrial graph databases, Neo4j and TigerGraph, are introduced as benchmarks. The comparison of resource consumption performance between GAT-RGCN and comparative algorithms under each dataset is shown in Table 2.

Table 2. Shows the comparison of the resource consumption of each algorithm under different datasets

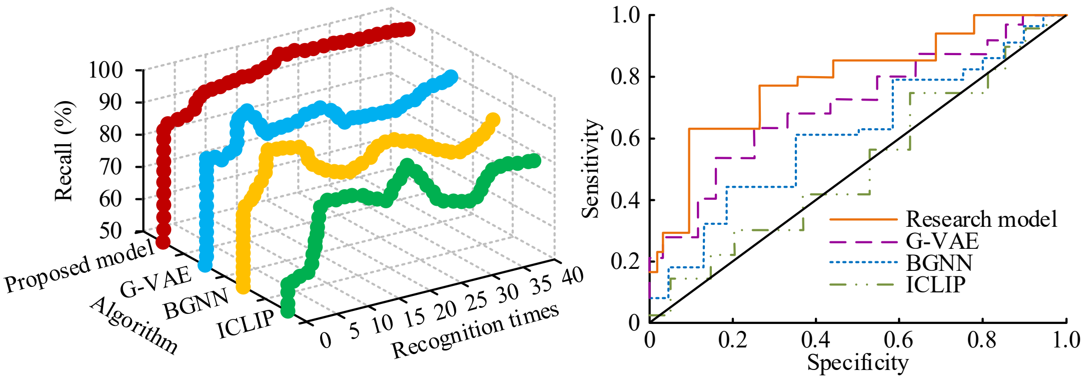
Indicator	SWAT			ILPC22-Small			YAGO3-10		
	GAT-RGCN	Neo4j	Tiger Graph	GAT-RGCN	Neo4j	Tiger Graph	GAT-RGCN	Neo4j	Tiger Graph
CPU usage rate/%	42.34	78.51	65.26	40.17	70.39	60.83	43.16	85.24	70.12
GPU memory usage/GB	3.38	/	5.82	3.52	/	6.24	3.19	/	6.04
Throughput/QBS	1865	973	1519	1804	1125	1328	1907	806	1573
Incremental update delay/ (s/ 10,000)	4.26	18.72	9.64	4.35	13.92	10.47	4.05	22.35	8.89
Write performance/ (Strips/s)	2437	1125	2032	2448	906	1863	2354	1074	1876

As can be seen from Table 2, the peak CPU usage rates of GAT-RGCN in the three datasets are 42.34%, 40.17%, and 43.16% respectively. The average GPU memory usage is 3.36GB, all of which are lower than those of Neo4j and TigerGraph. Moreover, the minimum incremental update delay of GAT-RGCN when facing different datasets is only 4.05 seconds, which is 81.88% faster than the maximum value of 22.35% of Neo4j. Meanwhile, GAT-RGCN simplifies entity association computation in the knowledge graph via a relation-weight matrix. Therefore, both its throughput and write performance are superior to those of the benchmark industrial mainstream graph databases.

Additionally, the research conducted experiments on the semantic understanding accuracy rate, node connection coverage rate, and fault diagnosis precision rate of GAT-RGCN in two specific scenarios: low-resource entities and long-tail relationships. It was found that, in recognizing entities related to special equipment maintenance components in the knowledge graph, the node connection coverage rate of GAT-RGCN decreased from the original 96.82% to 78.16%, and the fault diagnosis accuracy rate dropped from the initial 94.73% to 81.50%. The reason lies in the fact that the entity sample size of this knowledge graph is too small. When GAT allocates attention weights, it is easily masked by high-frequency entities, resulting in insufficient feature learning. Subsequent research plans to introduce an entity weight compensation mechanism, adding feature penalty terms to low-resource entities to enhance their weight in attention calculation. Meanwhile, in the mining of cross-domain relationships, the accuracy rate of GAT-RGCN in semantic understanding of the knowledge graph decreased from 97.94% to 88.35%, a decrease of 9.59%. The main reason is that the relationship weight matrix of RGCN struggles to capture sparse association features in long-tail relationships. Therefore, the research plans to construct a pre-trained dictionary for long-tail relationships in the future and to initialize cross-domain relationship weights based on the prior knowledge graph of the industrial domain, thereby reducing modeling error caused by sparse data.

4.2. Evaluation of the Optimization and Visualization Model Based on the Hybrid Algorithm

After validating GAT-RGCN's performance, the study further evaluated the proposed model's feasibility through field tests, comparing it with G-VAE, BGNN, and ICLIP. The study utilized the operation logs and maintenance records of large-scale working equipment in an industrial plant area in Sichuan Province as its data sources, and compared multiple indicators, including the recall rate, complexity, visualization efficiency, and results of four models. The results are shown in Fig. 9.



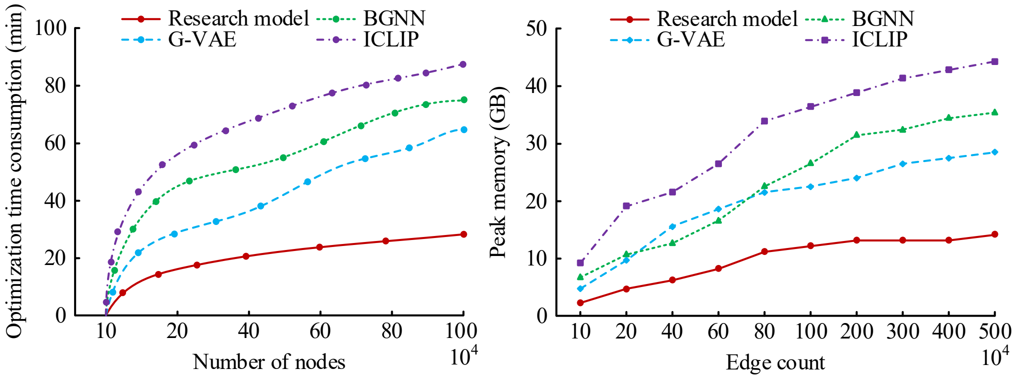
(a) Comparison of recall rates for fault identification in different models

(b) Comparison of AUC values for fault identification of different models

Fig. 9. Recall rate and AUC of fault recognition

As illustrated in Fig. 9(a), the maximum recall rate of the knowledge graph optimized by the proposed model for equipment failure type identification is 98.22%, which is significantly higher than that of the comparison model. However, G-VAE has insufficient semantic fusion and inadequate ability to identify latent faults, with a maximum recall rate of only 90.45%. BGNN and ICLIP are limited by the optimization accuracy of the knowledge graph, and their recall rates fluctuate within the range of 2.30% to 3.50% with an increase in the number of experiments. Fig. 9(b) shows that the AUC value of the proposed model reached 0.830, which was considerably higher than G-VAE (0.689), BGNN (0.625), and ICLIP (0.516). The research was then conducted to extend and generate multiple sets of experimental data from the SWAT dataset, comparing changes across different models with varying numbers of nodes and edges in the knowledge graph. The results are shown in Fig. 10.

It can be seen from Fig. 10(a) that the upward trend in time consumption for optimizing the knowledge graph of the proposed model, as the number of nodes increases, is gentler. The optimization time consumption for millions of nodes is only 29.1 minutes, which is significantly lower than that of the comparison model. This is because the GAT-RGCN in the proposed model can reduce redundant computations through the dynamic attention sharing mechanism and the relation weight matrix. As shown in Fig. 10(b), when the number of edges in the knowledge graph of the proposed model increased from 100,000 to 5 million, its peak memory usage rose from 2.42GB to 13.96GB, an increase of only 11.54GB. This is primarily attributed to the fact that the CRF-BERT module within the proposed model employs incremental storage of edge features, eliminating the need to cache all edge data, and retaining only the weights of the edges required for the current calculation. However, the contrast model needs to store the associated features of all edges, so when the number of edges increases, its memory usage rises significantly. In response to the large-scale application requirements in industrial scenarios where the number of knowledge graph nodes can reach hundreds of billions, the proposed model divides it into 4 to 6 hierarchical subgraphs, and further divides each subgraph into millions or hundreds of thousands of node fragments. Each fragment independently optimizes the knowledge graph. Meanwhile, data collaboration is achieved among the shards through the core entity association table, which helps avoid resource overload caused by loading the full knowledge graph. Industrial knowledge graphs mainly consist of three core entity data types: text data, graph-structured data, and device data. The visualization requirements for different types of data vary significantly. Therefore, it is of practical significance to evaluate the visualization efficiency of these three types of data. Thus, the study compared the effects of the knowledge graph visualization of each model. First, the visualization efficiency of four models was compared under three different types of entity data: text, graph-structured, and device data. Second, comparisons were made with two benchmark algorithms, ForceAtlas2 and Directed Acyclic Graph Layout Algorithm (DAGLAy). Text data visualization efficiency measures the speed and semantic integrity when presenting unstructured semantic information. Third, graph-structured data visualization efficiency assesses the speed of clear presentation and topological integrity of complex node and edge relationships within industrial knowledge graphs. Finally, the visualization efficiency of equipment data focuses on the operational data of equipment in industrial production, balancing the visualization speed of time-series numerical information with the retention of parameter correlations.



(a) Comparison of optimization time consumption of different models under each number of nodes

(b) Comparison of peak memory of different models under each number of edges

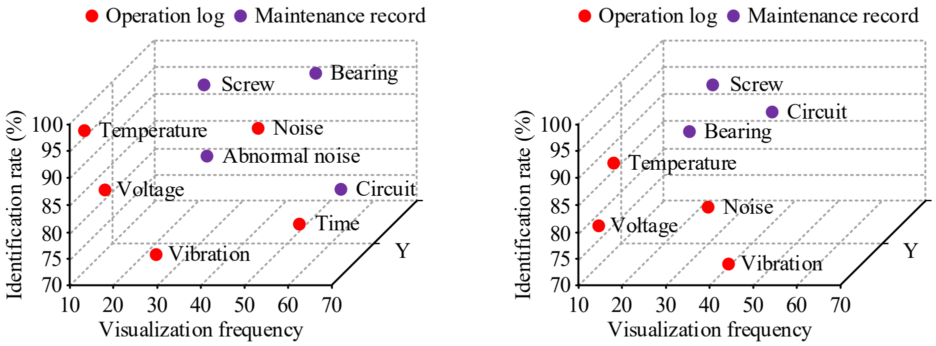
Fig. 10. The comparison results of the complexity of different models varying with the number of nodes and edges

Table 3. Comparison table of the visualization efficiency of each model under different entity data

Entity data	Number of experiments	Visualization efficiency/%					
		Research model	G-VAE	BGNN	ICLIP	ForceTars2	DAGLAy
Text data	10	97.34	81.42	71.83	80.57	85.24	79.51
	20	96.20	82.61	75.46	86.05	81.39	83.42
	30	96.53	82.56	73.79	88.40	87.51	85.86
Graph structure data	10	91.97	86.85	82.42	76.33	82.76	82.45
	20	93.25	81.97	83.63	72.96	85.83	84.90
	30	95.76	83.34	85.01	77.82	88.62	86.31
Equipment data	10	95.88	70.72	80.20	75.64	83.18	81.04
	20	96.21	73.25	75.37	73.16	86.47	83.57
	30	96.69	76.40	76.98	70.83	88.41	84.83

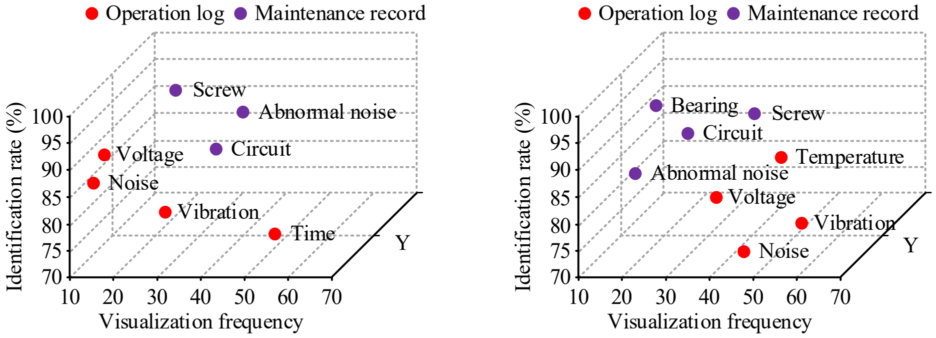
As shown in Table 3, the average visualization efficiency of the proposed model for the three different entity data sets is 96.69%, 93.66%, and 96.26%, respectively. The average overall visualization efficiency is 95.54%, which is higher than that of the three comparison models and exceeds the benchmark algorithm's 85.49% and 83.54% for ForceTars2 and DAGLAy, respectively. It indicates that the proposed model has strong adaptability when facing knowledge graphs with diverse data, and its visualization results are more in line with the usage requirements of industrial engineers. Among them, the proposed model performs exceptionally well in terms of text data visualization efficiency, thanks to the precise sorting of semantic relationships by CRF-BERT, which enables SVG to be visualized without requiring additional knowledge graph completion. Finally, the study conducted comparative experiments on the quality of knowledge graph visualization for the four models. Therefore, the clarity of the graphic structure is reflected in the cross-number optimization rate of 30.00%, the stress value optimization rate of 25.00%, the angle resolution compliance rate of 25.00%, and the user readability score of 20.00%, as determined by the recognized indicators in the graphic field. The crossover number is used to measure the degree of edge overlap, the stress value is used to evaluate the deviation between the node position and the ideal layout, the angular resolution indicates the discrimination of the edge, and the user readability score is the average score given by 10 engineers based on three aspects: fault location, relationship understanding, and information acquisition efficiency. The results are presented in Fig. 11.

As illustrated in Fig. 11(a), the proposed model achieved a structural clarity of 97.26% when visualizing temperature data from operation logs. And in 70 experiments, the knowledge graph data in the operation records and maintenance logs could all be visualized. The lowest structural clarity was observed in the visualization of runtime data, with a score of 82.64%. Figs. 11(b) to 11(d) show that the maximum structural clarity in G-VAE and BGNN was achieved in the visualization of screws in maintenance records, with peak values of 90.31% and 88.69%, respectively. However, both models failed to visualize the knowledge graphs after 50 iterations of the experiments. In addition, their visualization quantities were lower than those of the proposed model, demonstrating that the proposed model achieved the best performance in transforming entity relationships within knowledge graphs. In conclusion, the proposed model not only accurately and efficiently optimized industrial knowledge graphs containing diverse data types but also provided high-quality and intuitive visualization, while exhibiting a certain level of robustness.



(a) Comparison of the structural clarity of the visualization of the knowledge graph of the proposed model

(b) Comparison of the structural clarity of the visualization of the knowledge graph of the G-VAE



(c) Comparison of the structural clarity of the visualization of the knowledge graph of the BGNN

(d) Comparison of the structural clarity of the visualization of the knowledge graph of the ICLIP

Fig. 11. Comparison of the visualization quality of knowledge graphs

5. Conclusion

To address the problems of low quality and inefficiency in existing knowledge graph optimization methods, the study introduced CRF and GAT to improve BERT and RGCN, respectively, forming two hybrid algorithms: CRF-BERT and GAT-RGCN. These algorithms were then integrated with SVG technology to develop a fused approach. Based on this, the study developed a visualization model for industrial knowledge graphs that accounts for both multi-source data and semantic relationships. The feasibility of the proposed model was verified through comparative experiments. The results showed that the proposed model achieved a semantic understanding accuracy of 97.94% and a node connection coverage of 96.82% during the knowledge graph optimization process. In equipment fault diagnosis, a precision of 94.73% was achieved with an average error of 0.96%. Moreover, when the proposed model was used for equipment fault identification, the recall rate and AUC value were 98.22% and 0.830, respectively, with an average GPU memory usage of 3.36GB. In the evaluation of the visualization effect of the knowledge graph, the average visualization efficiency of the proposed model under three different entity datasets was 96.69%, 93.66%, and 96.26% respectively, and the clarity of the graphic structure based on recognized indicators such as the number of crosses and stress values was as high as 97.26%. All of these results outperformed the baseline models, fully demonstrating the proposed model's feasibility and superiority. Although the model performed remarkably well, the study did not distinguish between different types of knowledge graphs and focused only on industrial production knowledge graphs. Therefore, in the future, the proposed model in this research should be extended to various types of knowledge graphs, such as those in medical care and finance, to verify its universality. Targeted optimization and adjustment should be made based on the characteristics of different fields.

Funding

This research is supported by the 2021 Tianjin Educational Science Planning Project "Research on Promoting Xi Jinping Thought on the Rule of Law into Textbooks, Classrooms and Minds" (No.: CIE210158).

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author confirms that no AI tools were used in the preparation of this manuscript.

References

Bilal, M., and Almazroi, A., A. (2023). Effectiveness of fine-tuned BERT model in classification of helpful and unhelpful online customer reviews. *Electronic Commerce Research,* 23(4), 2737-2757. doi: 10.1007/s10660-022-09560-w.

Choudhuri, S., Adeniye, S., and Sen, A. (2023). Distribution Alignment Using Complement Entropy Objective and Adaptive Consensus-Based Label Refinement for Partial Domain Adaptation. *Artificial Intelligence and Applications,* 1(1), 43-51. doi: 10.47852/bonviewAIA2202524.

Garrido-Merchan, E., C., Gozalo-Brizuela, R., and Gonzalez-Carvajal, S. (2023). Comparing BERT against traditional machine learning models in text classification. *Journal of Computational and Cognitive Engineering,* 2(4), 352-356. doi: 10.47852/bonviewJCCE3202838.

Goloboff, P., A., and Morales, M., E. (2023). TNT version 1.6, with a graphical interface for MacOS and Linux, including new routines in parallel. *Cladistics,* 39(2), 144-153. doi: 10.1111/cla.12524.

Jin, W., Zhao, B., Yu, H., Tao, X., Yin, R., and Liu, G. (2023). Improving embedded knowledge graph multi-hop question answering by introducing relational chain reasoning. *Data Mining and Knowledge Discovery,* 37(1), 255-288. doi: 10.1007/s10618-022-00891-8.

Karthic, S., and Kumar, S., M. (2023). Hybrid optimized deep neural network with enhanced conditional random field based intrusion detection on wireless sensor network. *Neural Processing Letters,* 55(1), 459-479. doi: 10.1007/s11063-022-10892-9.

Li, W., Wang, M., Zhu, J., and Song, A. (2023). EEG-based emotion recognition using trainable adjacency relation driven graph convolutional network. *IEEE Transactions on Cognitive and Developmental Systems,* 15(4), 1656-1672. doi: 10.1109/TCDS.2023.3270170.

Peng, C., Xia, F., Naseriparsa, M., and Osborne, F. (2023). Knowledge graphs: Opportunities and challenges. *Artificial Intelligence Review,* 56(11), 13071-13102. doi: 10.1007/s10462-023-10465-9.

Ren, L., Li, Y., Wang, X., Cui, J., and Zhang, L. (2023). An ABGE-aided manufacturing knowledge graph construction approach for heterogeneous IoT data integration. *International Journal of Production Research,* 61(12), 4102-4116. doi: 10.1080/00207543.2022.2042416.

Talaat, A., S. (2023). Sentiment analysis classification system using hybrid BERT models. *Journal of Big Data,* 10(1), 110-127. doi: 10.1186/s40537-023-00781-w.

Talebi, S., Tong, E., Li, A., Yamin, G., Zaharchuk, G., and Mofrad, M., R. (2024). Exploring the performance and explainability of fine-tuned BERT models for neuroradiology protocol assignment. *BMC Medical Informatics and Decision Making,* 24(1), 40-51. doi: 10.1186/s12911-024-02444-z.

Wang, K., Xu, Y., and Luo, S. (2024). TIGER: Training Inductive Graph Neural Network for Large-scale Knowledge Graph Reasoning. *Proceedings of the VLDB Endowment,* 17(10), 2459-2472. doi: 10.14778/3675034.3675039.

Wang, Y., Lei, X., and Pan, Y. (2023). Microbe-disease association prediction using RGCN through microbe-drug-disease network. *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* 20(6), 3353-3362. doi: 10.1109/TCBB.2023.3247035.

Xia, T., Ding, L., Wan, G., Zhan, Y., Du, B., and Tao, D. (2025). Improving complex reasoning over knowledge graph with logic-aware curriculum tuning. *Proceedings of the AAAI Conference on Artificial Intelligence,* 39(12), 12881-12889. doi: 10.1609/aaai.v39i12.33405.

Yahya, M., Ali, A., Mehmood, Q., Yang, L., Breslin, J., G., and Ali, M., I. (2024). A benchmark dataset with Knowledge Graph generation for Industry 4.0 production lines. *Semantic Web,* 15(2), 461-479. doi: 10.3233/SW-233431.

Zhang, D., Liu, Z., Jia, W., Wu, F., Liu, H., and Tan, J. (2023). Dual attention graph convolutional network for relation extraction. *IEEE Transactions on Knowledge and Data Engineering,* 36(2), 530-543. doi: 10.1109/TKDE.2023.3289879.

Zhang, L. (2025). Implementing RGCN Model in Network Security Big Data Analysis. *Journal of Cyber Security and Mobility,* 14(2), 505-530. doi: 10.13052/jcsm2245-1439.14210.

Zhang, Q., Li, C., Su, F., and Li, Y. (2023). Spatiotemporal residual graph attention network for traffic flow forecasting. *IEEE Internet of Things Journal,* 10(13), 11518-11532. doi: 10.1109/JIOT.2023.3243122.

Zhou, B., Shen, X., Lu, Y., Li, X., Hua, B., Liu, T., and Bao, J. (2023). Semantic-aware event link reasoning over industrial knowledge graph embedding time series data. *International Journal of Production Research,* 61(12), 4117-4134. doi: 10.1080/00207543.2021.2022803.

Zhu, Q., Xiong, Q., Yang, Z., and Yu, Y. (2023). Rgcnu: Recurrent graph convolutional network with uncertainty estimation for remaining useful life prediction. *IEEE/CAA Journal of Automatica Sinica,* 10(7), 1640-1642. doi: 10.1109/JAS.2023.123369.

Shucheng Li earned his Ph.D. in International law (2015) from Nankai University in the PRC. Currently, he is an associate professor at the Intellectual Property School at Tiangong University. He was invited as a resourceful speaker to deliver talks on RAG, LLMs, KGs, Artificial Intelligence and Law. He is also serving and has served as a reviewer for national and international conferences and journals. He has published articles in more than 10 internationally reputed peer-reviewed journals and conference proceedings. His research interests include Artificial Intelligence and Law, LLM, AI agent, pattern recognition, intellectual property and information security.