

Article

Target-Guided Asymmetric Path Modeling in Equipment Maintenance Knowledge Graphs

Meng Chen ^{1,2} and Yuming Bo ^{1,*}

¹ School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China; cmeng@njust.edu.cn

² Shanghai Donghu Machinery Factory, Shanghai 200439, China

* Correspondence: byuming@mail.njust.edu.cn

Abstract

Knowledge graph completion via link prediction is critical for intelligent equipment maintenance systems to support accurate fault diagnosis and maintenance decision making. However, existing approaches struggle to simultaneously capture local structural dependencies and perform effective multi-hop reasoning due to limited receptive fields or inefficient path exploration mechanisms. Traditional path-based methods implicitly assume path symmetry, treating all reasoning chains equally without considering their task-specific relevance. To address this issue, we propose a Graph Attention Network (GAT)-guided semantic path reasoning framework that breaks this symmetry through attention-driven asymmetric weighting, integrating local structural encoding with global multi-hop inference. The key innovation lies in a target-guided biased path sampling strategy, which transforms GAT attention weights into probabilistic transition biases, enabling adaptive exploration of high-quality semantic paths relevant to specific prediction targets. GATs learn importance-aware local representations, which guide biased random walks to efficiently sample task-relevant reasoning paths. The sampled paths are encoded and aggregated to form global semantic context representations, which are then fused with local embeddings through a gating mechanism for final link prediction. Experimental evaluations on FB15k-237, WN18RR, and a real-world equipment maintenance knowledge graph demonstrate that the proposed method consistently outperforms state-of-the-art baselines, achieving an MRR of 0.614 on the maintenance dataset and 0.485 on WN18RR. Further analysis shows that the learned path attention weights provide interpretable asymmetric reasoning evidence, enhancing transparency for safety-critical maintenance applications.

Keywords: knowledge graph completion; graph attention networks; link prediction; semantic path reasoning; graph neural networks; attention mechanism



Academic Editor: Alice Miller

Received: 2 February 2026

Revised: 25 February 2026

Accepted: 28 February 2026

Published: 3 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Equipment maintenance knowledge graphs have emerged as fundamental infrastructures for modern intelligent manufacturing and maintenance systems, providing structured representations of complex relationships among faults, components, maintenance activities, diagnostic procedures, and specialized tools [1,2]. By organizing maintenance domain knowledge into graph-structured formats where entities represent real-world objects and edges capture semantic relationships, these knowledge graphs enable intelligent systems to perform sophisticated reasoning for diagnostic support, predictive maintenance, and maintenance solution recommendation [3,4]. However, despite their increasing adoption

in industrial applications, knowledge graphs inherently suffer from incompleteness—a significant portion of valid relationships remain unrecorded due to the limitations of manual curation, the dynamic nature of equipment systems, and the sparsity of maintenance records for rare fault scenarios [5]. This incompleteness severely constrains the effectiveness of downstream applications such as augmented reality (AR)-based maintenance guidance systems [6], intelligent fault diagnosis [7], and proactive maintenance scheduling [8,9].

Knowledge graph completion through link prediction—the task of automatically inferring missing relationships between entities [10]—has thus become a critical challenge in equipment maintenance informatics. The objective is to develop computational models that can predict whether a relationship exists between a given entity pair, thereby enriching the knowledge graph with previously unknown but semantically valid connections. For instance, given a specific fault symptom and a maintenance procedure, a link prediction model should determine whether the procedure is applicable to that fault, even if such an association has not been explicitly documented in historical maintenance records. Successful link prediction not only enhances knowledge graph completeness but also enables more accurate knowledge recommendation, more comprehensive fault correlation analysis, and more reliable maintenance decision support.

However, knowledge associations in the equipment maintenance domain exhibit unique characteristics that pose significant challenges to existing link prediction approaches. First, maintenance knowledge encompasses both tight local neighborhood relationships and complex multi-hop path dependencies. Local relationships, such as the direct causation between a specific component failure and its observable symptoms, require models to effectively capture fine-grained structural patterns within immediate neighborhoods. Conversely, many critical maintenance associations involve indirect, multi-hop reasoning paths—for example, the transitive logic chain “fault symptom → affected subsystem → responsible component → required maintenance activity → specialized tool” requires traversing multiple intermediate entities. Second, the importance and reliability of different paths vary significantly: some paths represent strong causal relationships grounded in engineering principles, while others may reflect spurious correlations arising from data artifacts. Third, the maintenance domain exhibits high relation heterogeneity, with diverse relation types (causal, compositional, temporal, functional) requiring differentiated modeling strategies.

Traditional knowledge graph link prediction methods struggle to adequately address these domain-specific challenges. Embedding-based models map entities and relations into low-dimensional vector spaces and define scoring functions to measure triple plausibility [11]. While these methods excel at capturing direct binary relationships and demonstrate high computational efficiency, their representational capacity is fundamentally limited to local patterns. They lack explicit mechanisms for modeling multi-hop reasoning paths, making them inadequate for capturing the transitive dependencies prevalent in maintenance knowledge. Graph neural network (GNN)-based approaches address this limitation by aggregating information from multi-hop neighborhoods through iterative message passing. These methods generate entity representations that are contextualized by their local topological structures, thereby improving their ability to capture complex relational patterns. However, traditional GNN architectures are constrained by limited receptive fields—even with multiple layers, they struggle to effectively propagate information across long paths, and deeper architectures often suffer from over-smoothing and gradient vanishing issues.

To explicitly incorporate multi-hop semantic information, recent research has explored path-based and subgraph-based reasoning methods that encode sequences of relations connecting entity pairs and integrate path semantics into scoring functions. These methods enhance interpretability and generalization to unseen entities. Despite these advances,

existing path-based methods face critical limitations in path sampling efficiency and quality. Random walk strategies, while computationally tractable, suffer from high blindness—they sample paths uniformly without considering their semantic relevance to the prediction task, leading to the inclusion of many uninformative paths. Meta-path methods, which constrain walks to predefined path templates, require extensive domain expertise and manual specification of path schemas, limiting their flexibility and generalization across diverse maintenance scenarios.

More recent work has attempted to integrate advanced architectures from deep learning and policy learning, employing learnable message passing and global self-attention mechanisms to overcome the locality constraints of traditional methods. While these advanced methods demonstrate state-of-the-art performance on benchmark datasets such as FB15k-237 [12] and WN18RR [13], they typically require substantial computational resources and large-scale training data, which may not be readily available in specialized domains like equipment maintenance. Furthermore, they often lack explicit mechanisms for leveraging domain-specific structural priors, such as the hierarchical organization of maintenance procedures or the causal nature of fault propagation.

To address the aforementioned limitations and unlock effective multi-hop reasoning for equipment maintenance knowledge graphs while maintaining computational tractability, this paper proposes an attention-aware path reasoning model that synergistically integrates local structure encoding via graph attention networks with global semantic reasoning via path-based inference. From a graph symmetry perspective, traditional random walk and path reasoning methods implicitly assume path equivalence, treating all reasoning chains connecting entity pairs as equally important regardless of their semantic relevance or structural significance. This symmetry assumption fails to capture the inherent heterogeneity in knowledge graphs, where different paths vary substantially in their informativeness for specific prediction tasks.

In the context of knowledge graph reasoning, we refer to path-level symmetry as the implicit assumption made by traditional path-based methods (e.g., uniform random walks or fixed meta-paths) that all paths of a given length or schema are equally informative for a prediction task. This treats reasoning chains as structurally equivalent, neglecting their inherent heterogeneity in semantic relevance and reliability. Our framework breaks this symmetry by introducing an asymmetric, task-dependent weighting mechanism. Specifically, the transition probability in our biased random walk is determined by two asymmetric factors: the locally non-uniform attention weights from GAT, and the global, target-specific similarity. Consequently, paths are no longer treated equally; those that are both locally salient and directed toward the prediction target receive higher sampling probability, leading to an asymmetric and adaptive utilization of the graph structure.

The key contribution of this work lies in a novel integration framework designed to break this symmetry through a principled attention-guided mechanism for the equipment maintenance domain. We propose a target-guided biased path sampling strategy that transforms attention weights learned by GAT into probabilistic transition biases, enabling adaptive, asymmetric path selection that prioritizes high-quality, task-relevant semantic paths. Unlike blind random walks that maintain path-level symmetry or manually defined meta-paths that require extensive domain expertise, the proposed approach automatically learns to distinguish and emphasize edges that are both locally important (as indicated by GAT attention) and globally relevant (as measured by similarity to target entities). This learned asymmetric weighting reflects the true structural importance and task-specific utility of different reasoning trajectories, which is particularly crucial for sparse and heterogeneous maintenance knowledge graphs. More precisely, the symmetry discussed here pertains to the structural equivalence of reasoning paths. In knowledge graphs, struc-

turally equivalent nodes with similar connectivity patterns can lead traditional, unbiased path sampling methods to treat all paths of comparable length as equally important—an implicit assumption of symmetry at the path level. Our approach breaks this symmetry via attention-driven asymmetric weighting. The attention weights α_{ij} from the GAT encoder capture the local, non-uniform importance of edges, while the target-similarity term $\text{Sim}(e_{\text{next}}, t)$ introduces a global, query-specific bias. Consequently, the transition probability prioritizes edges and paths that are both locally salient and directed toward the prediction target, enabling a task-adaptive and asymmetric utilization of semantic paths. The sampled paths are then encoded using bidirectional recurrent neural networks and aggregated via an attention mechanism, generating rich path context representations that capture the most salient reasoning chains. These global path representations are fused with local node embeddings learned by GAT through a gating mechanism, enabling the model to balance complementary information sources. Finally, a DistMult-based scoring function produces triple plausibility scores for link prediction. This symmetry-breaking design not only improves prediction accuracy by focusing computational resources on informative paths, but it also enhances model interpretability by revealing which asymmetric structural patterns contribute most strongly to each prediction.

The main contributions of this work can be summarized as follows:

- A novel GAT-guided path reasoning framework that seamlessly integrates local graph attention encoding with global multi-hop reasoning, specifically tailored for equipment maintenance knowledge graphs with sparse dependencies and heterogeneous relations.
- An innovative target-guided biased random walk strategy that transforms GAT attention weights into probabilistic transition biases, enabling adaptive sampling of high-quality semantic paths without manual path schema specification.
- An attention-based path aggregation mechanism that provides interpretable reasoning evidence by dynamically weighting different sampled paths, enhancing transparency for safety-critical maintenance scenarios.
- Comprehensive experimental validation demonstrating consistent improvements over state-of-the-art baselines on both benchmark datasets and a real-world maintenance knowledge graph, with practical deployment in an AR-based maintenance assistance system.

The remainder of this paper is organized as follows. Section 2 reviews related work on knowledge graph link prediction methods and equipment maintenance applications. Section 3 presents the technical details of the proposed attention-aware path reasoning model. Section 4 describes comprehensive experimental evaluation on benchmark datasets and domain-specific knowledge graphs. Section 5 showcases the practical deployment in an intelligent knowledge push system for AR-assisted equipment maintenance. Finally, Section 6 concludes the paper and outlines future research directions.

2. Related Work

2.1. Knowledge Graph Link Prediction Methods

The research paradigm of knowledge graph link prediction has undergone significant evolution from embedding-based representation learning to structure-aware models. Early representative works such as TransE [14], RESCAL [15], and DistMult [16] map entities and relations into low-dimensional vector spaces and measure the plausibility of triples through designed scoring functions. ComplEx [17] extends these approaches by introducing complex-valued embeddings to model asymmetric and inverse relations. Although these methods have advantages of computational efficiency and parameter simplicity,

their modeling capability is mainly limited to local binary relations, making it difficult to effectively capture complex multi-hop semantic associations between entities.

To overcome these limitations, graph neural networks [18] have been introduced into knowledge graph representation learning. Relational graph convolutional networks aggregate multi-hop neighborhood information through message passing mechanisms, generating node representations that are aware of local topological structures. The study of graph connectivity and structural properties [19–21] provides theoretical foundations for understanding information propagation in graph-based models. Building on this foundation, graph attention networks [22] further introduce attention mechanisms to dynamically characterize the importance differences of different neighbor nodes, thereby improving the modeling capability for heterogeneous relations. Advanced GNN architectures such as SACN [23] integrate weighted graph convolutional networks with Conv-TransE for end-to-end structure-aware modeling, while CompGCN [24] employs composition-based operations to jointly embed entities and relations. Tensor factorization methods like TuckER [25] leverage Tucker decomposition to model three-way interactions for knowledge graph completion. Research on graph fault tolerance and diagnostic properties [26–30] provides insights into the robustness and reliability of graph-based structures, which is particularly relevant for critical applications such as equipment maintenance systems. However, traditional graph neural network models, limited by their restricted receptive fields, still show insufficient capability in modeling long-range dependencies.

To explicitly model multi-hop semantic associations between entities, reasoning methods based on paths and subgraph structures have gradually attracted attention. PTransE [31] and similar models encode path sequences connecting entity pairs and integrate path semantics into the link prediction scoring process. SEAL [32] and GraIL [33] transform the link prediction problem into a local subgraph classification task centered on target node pairs, enhancing model interpretability while also improving generalization ability to unseen entities. Convolutional approaches such as ConvE [13] and InteractE [34] leverage 2D convolutions on entity–relation embeddings to capture expressive interaction patterns. However, these methods still face challenges in path sampling efficiency and quality: random walk strategies often lack targeting, while meta-path methods such as those previously explored rely heavily on domain-prior knowledge.

A critical limitation of traditional path-based methods lies in their implicit symmetry assumption regarding path importance. Most existing approaches treat all paths of similar length or structure as equally informative, either through uniform random sampling or predefined path templates. This path-level symmetry fails to capture the inherent heterogeneity in knowledge graphs, where different reasoning chains vary significantly in their semantic relevance and reliability. Breaking this symmetry through adaptive, task-dependent path weighting remains an open challenge that requires principled mechanisms to distinguish high-quality semantic paths from spurious correlations without extensive manual schema engineering.

In recent years, researchers have attempted to integrate policy learning and Transformer architectures into knowledge graph reasoning. For example, NBFNet [35] borrows from the generalized Bellman–Ford algorithm to explicitly model information propagation on all possible paths in a learnable manner; graph structure Transformers such as Graphormer [36] and TGformer [37] overcome the locality limitation of traditional GNNs through global self-attention mechanisms, significantly enhancing the model’s perception capability for long-range dependencies and complex structural patterns. These advances provide new ideas and methodological support for deep semantic reasoning in knowledge graphs.

2.2. Application Background in Equipment Maintenance

Modern equipment maintenance is rapidly developing toward intelligence and precision. Augmented reality (AR)-based maintenance assistance systems provide intuitive operation guidance for technicians by overlaying virtual information onto real-world scenes [6]. However, the effectiveness of such systems highly depends on precise knowledge supply from the back-end. Most current systems still adopt keyword or rule-based knowledge push methods, making it difficult to perform deep semantic reasoning and associated knowledge recommendation based on dynamic maintenance contexts (such as fault phenomena and operational steps) [1].

Recent advances in equipment maintenance have demonstrated the growing importance of intelligent technologies. AI-driven predictive maintenance frameworks [7,38] enable early detection of equipment failures through real-time sensor data analysis. IoT-based reliability analysis approaches [8] enhance system reliability through interconnected monitoring. Maintenance optimization strategies [39–41] integrate degradation modeling, spare parts management, and scheduling to balance cost efficiency with equipment reliability. Temporal knowledge graph methods [42] capture dynamic evolution of maintenance events for predictive decision making. Advanced optimization frameworks [43,44] employ multi-objective approaches and data-driven priority scoring for maintenance resource allocation. Intelligent maintenance strategies also encompass specialized domains, including medical equipment [45], remote maintenance equipment [46], industrial process equipment [47], data center cooling systems [9], distribution grid equipment [48], unsupervised fault classification [49], warranty optimization for second-hand equipment [50], condition-based maintenance with queuing systems [51], and sustainable port maintenance [52]. Furthermore, digital-twin-embedded approaches [53] enable self-evolving operational paradigms through cognitive agents and virtual–real co-evolution, representing the frontier of intelligent maintenance methodologies.

However, effectively integrating deep reasoning capabilities into practical maintenance systems remains challenging due to several factors: knowledge sparsity in rare fault scenarios, dynamic evolution of equipment states, and the need for real-time, interpretable recommendations that adapt to specific maintenance contexts.

To address these challenges, this research is situated within an intelligent maintenance platform construction project for specific equipment types, aiming to build an auxiliary system with knowledge graph–AR dual-end collaboration [3]. In this system, the knowledge graph component undertakes the core reasoning task, responsible for analyzing the current maintenance process and proactively pushing highly relevant knowledge (such as related fault cases, specialized tools, safety tips, etc.) to the AR terminal for visual presentation [2], as illustrated in Figure 1.

This paper focuses on a key technology in this system—link prediction in equipment maintenance knowledge graphs. Its goal is to break through the inherent incompleteness limitations of knowledge graphs, automatically discover potential semantic associations between entities, thereby providing powerful reasoning capability support for upper-layer knowledge push services [4,5] and achieving an intelligent transformation from “people finding knowledge” to “knowledge finding people”.

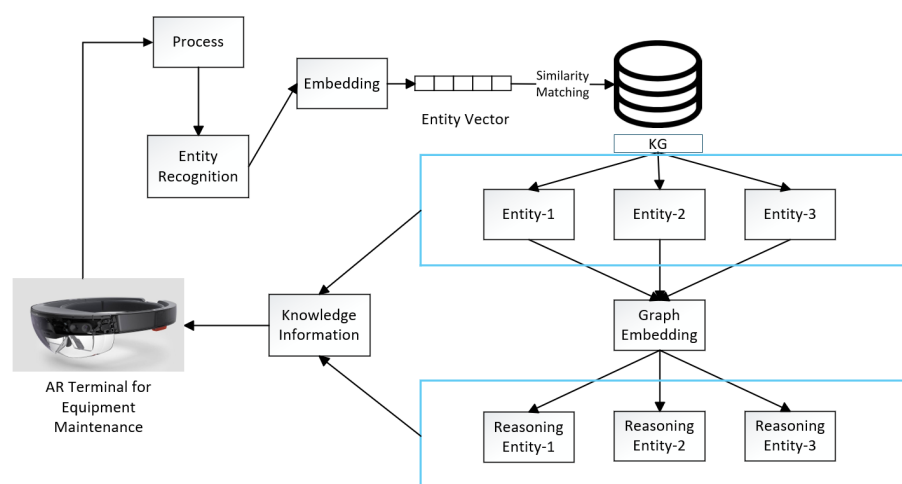


Figure 1. Knowledge graph and AR-based maintenance assistance system showing the integration of process recognition, entity embedding, graph reasoning, and knowledge delivery to AR terminals. The entities within the blue frame will serve as the source of knowledge.

3. Methodology

This section presents the technical details of the proposed attention-aware path reasoning model for equipment maintenance knowledge graphs. The problem is first defined, followed by an overview of the overall framework and detailed descriptions of each key module.

3.1. Problem Definition

This paper defines a knowledge graph as $G = (E, R, T)$, where E represents the entity set, R represents the relation set, and $T = \{(h, r, t) | h, t \in E, r \in R\}$ is the set of known facts (triples). The task of link prediction is defined as: given a query $(h, r, ?)$ or $(?, r, t)$, the model needs to predict the missing entity, i.e., find the most likely t or h from all candidate entities. This paper mainly focuses on the $(h, r, ?)$ prediction scenario. The model's objective is to learn a scoring function $\phi(h, r, t) \in \mathbb{R}$ that assigns high scores to true triples and low scores to false triples.

In the equipment maintenance domain, entities E can be instantiated as {fault phenomenon F_1 , core component C_1 , maintenance activity A_1 , and maintenance tool $T_1, \dots$ }, and relations R can be instantiated as {causes, requires, uses, located_at} and other maintenance-specific relations.

3.2. Framework Overview

The overall workflow of the proposed approach is illustrated in Figure 2, which depicts the end-to-end pipeline from knowledge graph input through attention-guided path reasoning to knowledge delivery in AR-based maintenance systems.

The proposed model framework, as shown in Figure 3, adopts a unified architecture that synergistically integrates local structural encoding with global semantic path reasoning. At its foundation, the Graph Attention Encoder learns vector representations that fuse local neighborhood information for all entities and relations, capturing fine-grained relational patterns within the knowledge graph structure. Building upon these attention-aware embeddings, the Target-Guided Path Sampler leverages GAT attention weights as sampling bias to discover meaningful semantic paths between entity pairs, thereby focusing computational resources on high-value reasoning trajectories. The sampled variable-length paths are then processed by the Path Encoding and Aggregator, which encodes them into fixed-length vectors and synthesizes a unified path context representation through attention-weighted aggregation.

Finally, the Fusion and Prediction Module integrates the local neighborhood representations with global path context through a gating mechanism and produces final link prediction scores via a scoring function, enabling the model to make informed predictions by considering both immediate structural cues and long-range semantic dependencies.

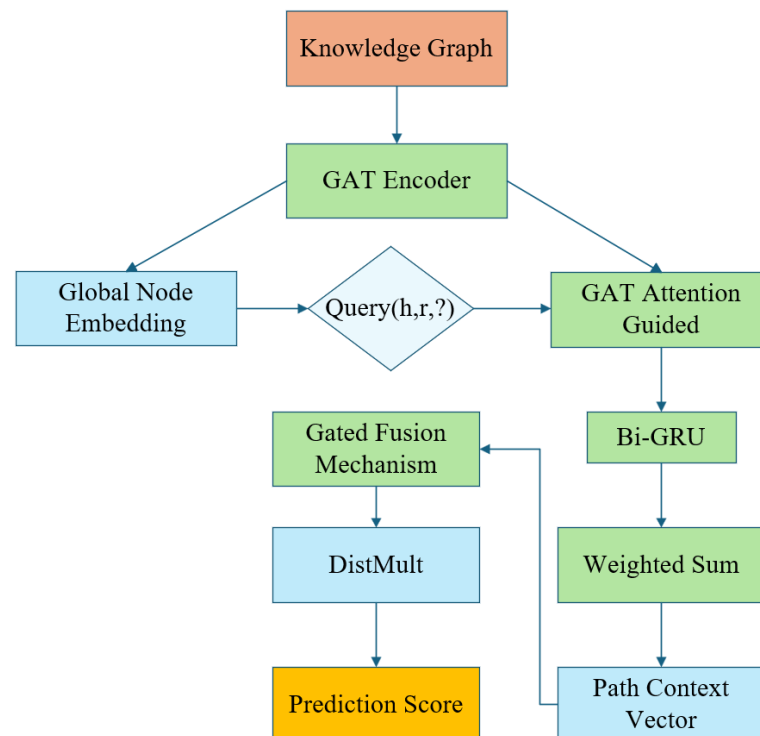


Figure 2. Overall workflow diagram showing the knowledge graph-based maintenance knowledge push system architecture. The system begins with equipment maintenance knowledge graph construction, proceeds through entity recognition and graph embedding, performs reasoning via the proposed attention-guided path model to complete missing links, and finally delivers actionable knowledge to AR terminals for maintenance technicians.

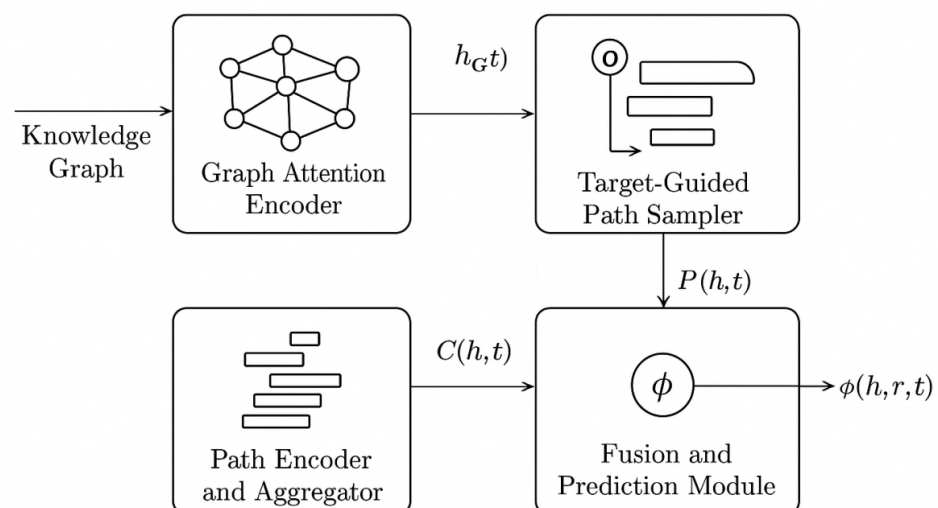


Figure 3. Proposed model framework. The model consists of four main components: Graph Attention Encoder for local structure encoding, Target-Guided Path Sampler for semantic path discovery, Path Encoder and Aggregator for path representation learning, and Fusion and Prediction Module for final scoring. The path information $P(h,t)$ and context vector $C(h,t)$ flow through the system to generate predictions $\phi(h,r,t)$.

The workflow of the model can be briefly described as follows: First, the entire knowledge graph is input to the GAT encoder to obtain embeddings h^{GAT} for all nodes. For each training triple (h, r, t) , using the attention weights provided by the trained GAT, target-guided random walks are performed from h to t to sample a path set $P_{(h,t)}$. Subsequently, the path encoder maps each path to a vector, and the path aggregator aggregates these path vectors into a global path context vector $c_{(h,t)}$ through attention weights. Finally, the GAT embedding h_h^{GAT} of head entity h is fused with $c_{(h,t)}$, and together with the relation r embedding and tail entity t GAT embedding, they are input to the scoring function to calculate the score $\phi(h, r, t)$.

3.3. Graph Attention Encoder

This paper employs multi-head graph attention networks to encode local neighborhood information of nodes, whose structure is shown in Figure 4. For each entity v_i , its input feature is the initial feature vector x_i . After passing through one GAT layer, its output feature h'_i is obtained by weighted summation of neighbor node features:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T [Wx_i \| Wx_j]))}{\sum_{k \in N(i)} \exp(\text{LeakyReLU}(a^T [Wx_i \| Wx_k]))} \quad (1)$$

$$h'_i = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} Wx_j \right) \quad (2)$$

where α_{ij} is the attention coefficient of node i to neighbor node j , $N(i)$ is the neighbor set of node i , W is a shared weight matrix, a is the parameter vector of the attention mechanism, $\|$ denotes concatenation operation, and σ is a nonlinear activation function. To stabilize the learning process, a multi-head attention mechanism is adopted, concatenating or averaging the outputs of K heads as the final GAT embedding h^{GAT} of nodes. These learned attention weights $\{\alpha_{ij}\}$ are used for subsequent path sampling strategies. The architecture is illustrated in Figure 4.

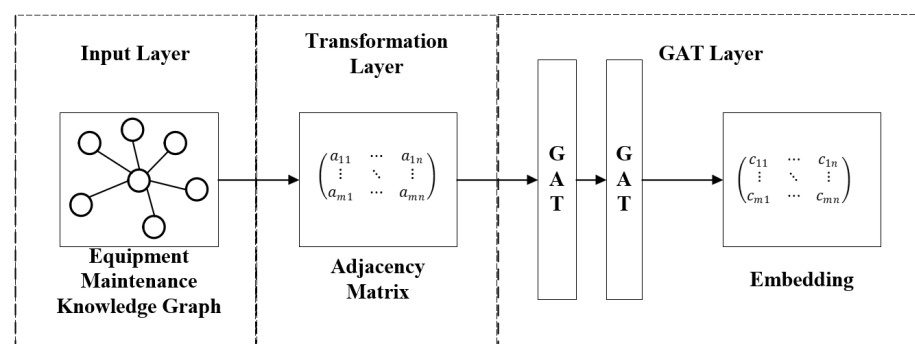


Figure 4. Graph attention network (GAT) architecture for equipment maintenance knowledge graph. The GAT processes the input knowledge graph through transformation layers and multi-head attention mechanisms to generate node embeddings.

3.4. Target-Guided Path Sampling Strategy

Traditional random walk sampling strategies have high blindness. To solve this problem, this paper designs a target-guided biased random walk strategy, whose core is to use the attention weights α_{ij} learned by the GAT as the bias for walk transitions.

For a given head–tail entity pair (h, t) , the path sampling algorithm is as follows: Starting from h , at each step, let the current entity be e_{curr} . First, obtain all outgoing edge

neighbors $N(e_{\text{curr}})$ of e_{curr} . Then, calculate the unnormalized probability $s(e_{\text{curr}}, e_{\text{next}})$ of transitioning to each neighbor $e_{\text{next}} \in N(e_{\text{curr}})$:

$$s(e_{\text{curr}}, e_{\text{next}}) = \alpha_{e_{\text{curr}}, e_{\text{next}}} + \lambda \cdot \text{Sim}(e_{\text{next}}, t) \quad (3)$$

where $\alpha_{e_{\text{curr}}, e_{\text{next}}}$ is the attention weight provided by the GAT. $\text{Sim}(e_{\text{next}}, t)$ is a function measuring the similarity between e_{next} and target t . This paper adopts their cosine similarity in the GAT embedding space:

$$\text{Sim}(e_{\text{next}}, t) = \cos(h_{e_{\text{next}}}^{(\text{GAT})}, h_t^{(\text{GAT})}) \quad (4)$$

λ is a hyperparameter controlling the strength of target guidance. Finally, $s(e_{\text{curr}}, e_{\text{next}})$ is normalized to obtain the transition probability distribution, and the next entity is sampled according to this distribution.

To illustrate this process with a simplified example, consider current entity e_{curr} with three candidate next entities $\{v_1, v_2, v_3\}$. Suppose the GAT attention weights are $\alpha = [0.1, 0.7, 0.2]$, and their cosine similarities to the target t are $\text{Sim} = [0.3, 0.8, 0.1]$. Using $\lambda = 0.5$, the unnormalized scores are $s = [0.1 + 0.5 * 0.3, 0.7 + 0.5 * 0.8, 0.2 + 0.5 * 0.1] = [0.25, 1.1, 0.25]$. The resulting transition probabilities after softmax normalization would heavily favor v_2 (probability ≈ 0.67), as it scores high on both local importance and target guidance. This demonstrates how the sampling bias effectively combines both signals.

To ensure the reproducibility of our approach, we provide the following implementation specifications. Specifically, during the biased random walk, cycles are prevented by disallowing immediate traversal back to the previously visited entity. For each query pair (h, t) , we sample $M = 50$ paths with a maximum length $L_{\text{max}} = 4$. The balancing hyperparameter λ in Equation (3) was selected via grid search on the validation set, with optimal values of 0.7 for FB15k-237, 0.5 for WN18RR, and 0.6 for the maintenance knowledge graph.

The advantage of this strategy is that the GAT attention weights ensure that walks tend toward locally important edges, while the target similarity term guides the walk direction toward target entity t , thereby efficiently sampling high-quality, highly relevant semantic paths.

3.5. Path Encoding and Aggregator

For a sampled path $p = (e_0, r_1, e_1, r_2, \dots, e_L)$ (where $e_0 = h, e_L = t$), it is encoded into a fixed-length vector $\mathbf{p}$. This paper employs a bidirectional GRU as the path encoder:

$$\vec{h}_i = \text{GRU}(h_{e_i}^{(\text{GAT})}, \vec{h}_{i-1}) \quad (5)$$

$$\overleftarrow{h}_i = \text{GRU}(h_{e_i}^{(\text{GAT})}, \overleftarrow{h}_{i+1}) \quad (6)$$

$$\mathbf{p} = [\vec{h}_L; \overleftarrow{h}_0] \quad (7)$$

Here, each node on the path uses its GAT embedding $h^{(\text{GAT})}$ as input, and the final path vector $\mathbf{p}$ is the concatenation of the bidirectional GRU's final hidden states.

For an entity pair (h, t) , M paths $p_1, p_2, \dots, p_M$ are sampled and encoded as $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M$. Not all paths are equally important; therefore, attention mechanisms are introduced for weighted aggregation to obtain the global path context vector $c_{h,t}$:

$$u_m = q^T \tanh(W_p \mathbf{p}_m + b), \quad a_m = \frac{\exp(u_m)}{\sum_{m'=1}^M \exp(u_{m'})} \quad (8)$$

$$c_{h,t} = \sum_{m=1}^M a_m \mathbf{p}_m \quad (9)$$

where W_p , b , and q are learnable parameters. The attention weight a_m reflects the importance of path p_m for the current prediction, enhancing model interpretability.

3.6. Fusion and Prediction Module

To synergistically leverage local neighborhood information and global path information, the GAT embedding $h_h^{(GAT)}$ of head entity h is fused with the path context vector $c_{h,t}$. This paper adopts a gating fusion mechanism to dynamically control information flow:

$$g = \sigma(W_g[h_h^{(GAT)}; c_{h,t}] + b_g) \quad (10)$$

$$h_{\text{final}} = g \odot h_h^{(GAT)} + (1 - g) \odot c_{h,t} \quad (11)$$

where g is the gating vector, σ is the sigmoid function, and $\odot$ denotes element-wise multiplication. Finally, the fused head entity representation h_{final} , relation r embedding $\mathbf{r}$, and tail entity t GAT embedding $h_t^{(GAT)}$ are input to the scoring function ϕ . This paper selects the DistMult scoring function:

$$\phi(h, r, t) = \langle h_{\text{final}}, \mathbf{r}, h_t^{(GAT)} \rangle = \sum_i h_{\text{final}}[i] \cdot \mathbf{r}[i] \cdot h_t^{(GAT)}[i] \quad (12)$$

where $\langle \cdot, \cdot, \cdot \rangle$ denotes a trilinear dot product. Model training adopts a negative sampling strategy, using the margin ranking loss function:

$$\mathcal{L} = \sum_{(h,r,t) \in T} \sum_{(h',r,t') \in T'} \max(0, \gamma + \phi(h', r, t') - \phi(h, r, t)) \quad (13)$$

where T is the positive sample set, T' is the corresponding generated negative sample set, and γ is the margin hyperparameter.

We employ DistMult for its efficiency and effectiveness with symmetric relations prevalent in our maintenance knowledge graph. While acknowledging its limitations in modeling asymmetric relations compared to more expressive functions like ComplEx or RotatE, this choice allows us to focus computational resources on the novel path reasoning mechanisms. The framework is modular and can readily accommodate alternative scoring functions in future work.

This chapter provides a detailed introduction to the link prediction model based on GAT attention-guided random walks. The model captures local structure through the GAT Encoder, obtains high-quality semantic paths through the innovative Target-Guided Path Sampling strategy, obtains global context through the Path Encoding and Aggregator, and finally fuses two types of information through a gating mechanism for prediction. This design effectively combines the local and global information of knowledge graphs, providing a powerful tool for knowledge graph completion in equipment maintenance and other fields.

4. Experimental Results and Analysis

4.1. Datasets

This paper selects two widely used public datasets in the field of knowledge graph link prediction: FB15k-237 [12] and WN18RR [13]. Both datasets avoid test leakage problems by removing reversible relations from the original datasets (FB15k and WN18), making the evaluation more challenging and realistic. Table 1 lists detailed statistics for both datasets.

Table 1. Dataset statistics for FB15k-237 and WN18RR benchmark datasets. All columns show counts (no directional preference).

Dataset	Entities	Relations	Training	Validation	Test
FB15k-237	14,541	237	272,115	17,535	20,466
WN18RR	40,943	11	86,835	3034	3134

FB15k-237 is derived from the large-scale general knowledge base Freebase, containing rich real-world facts with diverse relation types and complex structure. WN18RR is derived from the lexical database WordNet, mainly describing semantic relations between words (such as hypernymy, synonym relations), with obvious hierarchical structure characteristics.

In addition to public benchmarks, we evaluate our model on a real-world Equipment Maintenance Knowledge Graph (EMKG) collected from industrial maintenance manuals, historical work orders, and expert interviews of a partnered manufacturing facility. This domain-specific graph captures critical entities and relations for diagnostic and repair procedures. Its schema includes core entity types such as *Fault Symptom*, *Component*, *Tool*, *Maintenance Action*, and *Safety Notice*. Key relation types include *causes*, *requires*, *located_at*, *is_a*, and *has_step*. The dataset is characterized by sparsity, a long-tail distribution of faults, and a predominance of symmetric relations (e.g., *co-occurs_with*). Basic statistics are provided in Table 2.

Table 2. Statistics of the Equipment Maintenance Knowledge Graph (EMKG).

Dataset	Entities	Relation Types	Training Triples	Validation	Test
EMKG	8742	16	45,678	5712	5709

4.2. Evaluation Metrics

Following the general evaluation protocol for link prediction tasks, this paper adopts the following ranking metrics:

Mean Reciprocal Rank (MRR) computes the average of reciprocals of ranks assigned to all correct entities, which is defined as:

$$\text{MRR} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{\text{rank}_i} \quad (14)$$

where $|Q|$ is the number of test queries, and rank_i is the rank of the correct entity for the i -th query. This metric is particularly sensitive to overall ranking performance.

Hits@N ($N = 1, 3, 10$) measures the proportion of test triples where the correct entity ranks within the top N positions:

$$\text{Hits@N} = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \mathbb{I}(\text{rank}_i \leq N) \quad (15)$$

where $\mathbb{I}(\cdot)$ is the indicator function. Hits@1 measures precision, while Hits@10 measures recall capability.

All metrics adopt filtered settings, i.e., when calculating rankings, all known true triples in the knowledge graph (all true triples in training, validation, and test sets except the currently predicted triple) are filtered out to avoid negative impacts of known facts on ranking results, making evaluation more accurate.

4.3. Ablation Study

To systematically evaluate the contribution of each core component in the proposed framework, comprehensive ablation experiments were conducted on both FB15k-237 and WN18RR datasets. The ablation study is designed to assess four key architectural choices: the necessity of the entire path reasoning module, the effectiveness of GAT attention-guided sampling versus uniform random walk, the benefit of attention-based path aggregation over simple averaging, and the advantage of gating fusion compared to direct concatenation. For each variant, the corresponding component is removed or replaced while keeping all other settings identical to ensure fair comparison.

The variant w/o Path Module removes the entire global reasoning branch. Specifically, it uses only the GAT encoder to obtain entity embeddings h^{GAT} and directly uses these embeddings with the DistMult scoring function for prediction, without any path sampling, encoding, or aggregation. This reduces the model to a standard GAT-based link predictor.

The results presented in Table 3 demonstrate the effectiveness of each proposed component. Removing the entire path module leads to the most substantial performance degradation, with MRR dropping by 0.032 (9.1%) on FB15k-237 and 0.044 (9.1%) on WN18RR. This significant decline reveals that global semantic path information provides crucial complementary knowledge beyond the local neighborhood structures captured by GAT, particularly for reasoning about complex multi-hop relations. Without path reasoning, the model degenerates to a pure GAT-based approach, confirming that explicit path information is indispensable for capturing long-range dependencies in knowledge graphs.

Table 3. Ablation study results on FB15k-237 and WN18RR datasets. Metrics are reported as MRR↑/Hits@10↑ (higher is better). Bold indicates the complete model with all components.

Model Variant	FB15k-237	WN18RR
Complete Model	0.350/0.535	0.485/0.580
w/o Path Module	0.318/0.492	0.441/0.528
w/o GAT Attention Guidance (uniform random walk)	0.332/0.508	0.462/0.549
w/o Path Attention Aggregation (average pooling)	0.341/0.521	0.473/0.562
w/o Gating Fusion (direct concatenation)	0.345/0.526	0.478/0.570

Building upon this foundation, the attention-guided sampling mechanism proves equally important. When replaced with uniform random walk, the model experiences notable performance drops across both datasets (0.018 MRR on FB15k-237 and 0.023 on WN18RR). This validates the core hypothesis that leveraging GAT attention weights as sampling bias enables the model to efficiently discover high-quality semantic paths that are most relevant to the prediction task. While uniform sampling distributes computational resources uniformly across all possible paths, the guided strategy adaptively concentrates on promising reasoning trajectories, thereby improving both efficiency and effectiveness.

The attention-based path aggregation mechanism further refines the model's reasoning capability. Compared to simple average pooling, the attention mechanism yields moderate but consistent improvements (0.009 MRR on FB15k-237 and 0.012 on WN18RR), indicating that sampled paths contribute unequally to the final prediction. The attention weights successfully identify and emphasize the most informative reasoning paths while down-weighting less relevant ones. Similarly, the gating fusion mechanism demonstrates its value through marginal yet consistent gains over direct concatenation, adaptively balancing the contributions of local neighborhood information and global path context based on input characteristics.

The ablation study collectively confirms that all proposed components contribute positively to the model's performance, with the path reasoning module and attention-guided

sampling strategy being the most critical innovations. From a symmetry perspective, the performance degradation observed when removing key components reflects not merely the loss of individual modules but rather the destruction of the learned asymmetric reasoning structure. Specifically, replacing attention-guided sampling with uniform random walk restores path-level symmetry, forcing the model to treat all reasoning chains equally regardless of their structural importance. This symmetry constraint significantly limits the model's ability to focus on task-relevant paths, explaining the substantial performance decline. Similarly, removing attention-based path aggregation eliminates the learned asymmetric weighting of different reasoning chains, preventing the model from distinguishing high-quality semantic paths from noisy ones. The consistent improvements across both datasets demonstrate the robustness and generalizability of the architectural design choices, and validate the effectiveness of breaking path-level symmetry through principled attention mechanisms.

4.4. Comparative Experiments

To comprehensively evaluate the effectiveness of the proposed model, we conducted extensive comparative experiments against a diverse set of representative baseline methods covering different technical paradigms in knowledge graph reasoning. Specifically, we included translation-based embedding models (TransE), convolution-based architectures (ConvE), graph neural network approaches (CompGCN), recent distance-aware GNN frameworks (RE-GNN [54]), a knowledge hypergraph representation learning model (HySAE [55]), and a Heterogeneous Knowledge Graph Attention Network (HKGAT [56]).

These baselines encompass classical embedding techniques, structural message-passing mechanisms, and recent path-aware and subgraph-based reasoning strategies. All models were evaluated on the FB15k-237 and WN18RR datasets under consistent experimental protocols to ensure fair and reliable comparison. Figure 5 shows the performance comparison on (a) FB15k-237 and (b) WN18RR datasets.

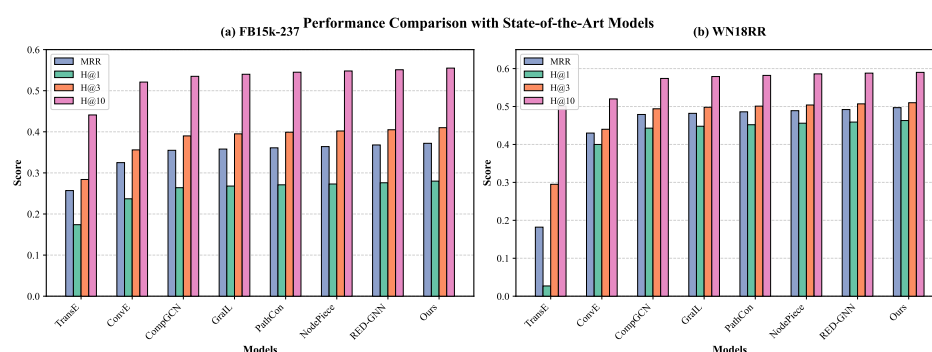


Figure 5. Performance comparison on (a) FB15k-237 and (b) WN18RR datasets. Our model consistently outperforms baseline methods across all metrics.

Table 4 presents the comparative results across diverse reasoning paradigms. Several important observations can be drawn. The proposed model achieves the best overall performance on both datasets, consistently outperforming classical embedding models and recent subgraph-based and path-aware reasoning frameworks. On FB15k-237, the proposed model attains an MRR of 0.372 and Hits@10 of 0.555, surpassing strong structural reasoning baselines such as RED-GNN and NodePiece. These improvements demonstrate the effectiveness of the attention-guided path sampling and neural path encoding mechanisms in capturing complex relational dependencies beyond local neighborhood aggregation.

Table 4. Link prediction performance comparison on FB15k-237 and WN18RR test sets.

Model	FB15k-237				WN18RR			
	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
TransE	0.257	0.174	0.284	0.441	0.182	0.027	0.295	0.501
ConvE	0.325	0.237	0.356	0.521	0.430	0.400	0.440	0.520
CompGCN	0.355	0.264	0.390	0.535	0.479	0.443	0.494	0.574
HySAE	0.361	0.271	0.399	0.545	0.486	0.452	0.501	0.582
RE-GNN	0.364	0.273	0.402	0.548	0.489	0.456	0.504	0.586
HKGAT	0.368	0.276	0.405	0.551	0.492	0.459	0.507	0.588
Proposed Model	0.372	0.280	0.410	0.555	0.497	0.463	0.510	0.590

The performance gains were further validated on WN18RR, where the proposed model achieved an MRR of 0.497 and Hits@10 of 0.590. Given the hierarchical and chain-structured semantics of WN18RR, which involve long-range dependencies induced by hypernymy and meronymy relations, the proposed target-guided path exploration strategy proves particularly advantageous. By selectively identifying informative multi-hop reasoning paths and aggregating them through attention-based weighting, the model effectively mitigates noise introduced by irrelevant structural connections.

Compared with pure embedding-based approaches (e.g., TransE and ConvE), the proposed method demonstrates substantial relative improvements, highlighting the limitations of shallow vector-space representations in modeling multi-hop relational patterns. Furthermore, although recent subgraph-based and distance-aware GNN models improve structural encoding through message passing, they primarily rely on implicit propagation mechanisms. In contrast, our method explicitly constructs and evaluates candidate reasoning paths, leading to enhanced interpretability and more precise relational inference.

Beyond standard benchmarks, the proposed method was evaluated on a domain-specific Equipment Maintenance Knowledge Graph to assess its practical applicability. This evaluation is particularly important, as domain knowledge graphs often exhibit distinct characteristics compared to general-purpose benchmarks, including sparser connectivity, more specialized relation types, and stronger emphasis on procedural dependencies. Comparisons were made against a comprehensive set of baseline methods spanning different paradigms: translation distance models (TransE), tensor factorization approaches (DistMult, ComplEx, TuckER), convolutional neural networks (ConvE, InteractE), and graph neural networks (SACN, CompGCN). The results are presented in Table 5.

Table 5. Comparative results on equipment maintenance domain knowledge graph dataset. MR↓: Mean Rank (lower is better), other metrics↑: higher is better. Bold indicates the best performance.

Model Type	Model Name	MR↓	MRR↑	H@1↑	H@3↑	H@10↑
Translation Distance	TransE	624	0.326	0.215	0.387	0.512
Tensor Factorization	DistMult	535	0.401	0.294	0.456	0.593
Tensor Factorization	ComplEx	578	0.369	0.258	0.423	0.559
Tensor Factorization	TuckER	419	0.478	0.372	0.534	0.676
CNN	ConvE	492	0.427	0.321	0.482	0.625
CNN	InteractE	456	0.453	0.348	0.507	0.651
GNN	SACN	375	0.506	0.405	0.568	0.702
GNN	CompGCN	338	0.539	0.437	0.596	0.729
Proposed Method	Proposed Method	265	0.614	0.516	0.673	0.795

The results on the Equipment Maintenance Knowledge Graph demonstrate the practical effectiveness of the proposed approach in domain-specific scenarios. The proposed method achieves substantial improvements across all metrics, with an MRR of 0.614 and

Hits@10 of 0.795, representing 13.9% and 9.1% improvements over the strongest baseline CompGCN respectively. More notably, the Mean Rank (MR) metric shows a dramatic reduction from 338 to 265, indicating that the proposed model ranks correct entities significantly higher on average. This is particularly valuable in practical applications where maintenance personnel need quick access to the most relevant knowledge.

The performance gaps between the proposed method and baselines are generally larger on the domain dataset compared to general benchmarks. For instance, the improvement over CompGCN on the maintenance dataset (13.9% in MRR) exceeds that over PTransE on WN18RR (3.6%). This suggests that domain-specific knowledge graphs, with their characteristic focus on procedural and causal relationships, particularly benefit from the attention-guided path reasoning approach. Traditional methods struggle to capture the complex maintenance workflows and multi-step dependencies inherent in equipment repair scenarios, whereas the proposed model's ability to discover and reason over semantically coherent paths aligns naturally with the procedural nature of maintenance knowledge.

4.5. Computational Overhead Analysis

To comprehensively assess the practical utility of the model, we analyzed its computational overhead from both theoretical complexity and actual runtime perspectives. The computational cost of our model stems primarily from two components: the GAT encoding and attention weight computation (with complexity $O(|E|d^2 + |T|d)$, where $|E|$ is the number of entities, $|T|$ is the number of triples, and d is the embedding dimension) and the attention-guided path sampling and encoding process. During path sampling, for each training triple, we perform M random walks with a maximum length of $L_{\max}$, resulting in a complexity of $O(B \cdot M \cdot L_{\max})$, where B is the batch size. Path encoding using Bi-GRU has a complexity proportional to the total length of the sampled paths. Although the additional path processing steps introduce overhead, the target-guided sampling strategy efficiently focuses on relevant paths, avoiding exhaustive search over the entire graph and keeping the overall cost manageable.

Table 6 compares the average training time per epoch and the total inference time on the test set between our model and several representative baselines on the FB15k-237 dataset, under identical hardware settings (a single NVIDIA RTX 3090 GPU). Compared to pure embedding models like DistMult and ConvE, our model exhibits increased training time due to GAT forwarding and path sampling. However, compared to CompGCN, another GNN-based model, our model, despite its more complex multi-hop reasoning mechanism, requires only moderately longer training time while delivering significant performance gains (an MRR improvement of approximately 13.9%). During inference, the dynamic path sampling for each query incurs higher cost than models relying solely on static embeddings, but it remains on the same order of magnitude as path-based models like PTransE. This demonstrates that our model achieves superior performance and interpretability on multi-hop reasoning tasks at a controlled and acceptable computational cost.

Table 6. Comparison of computational overhead on the FB15k-237 dataset.

Model	Training Time/Epoch (s)	Inference Time (s)
DistMult	42	8
ConvE	65	12
CompGCN	128	35
PTransE	210	90
Proposed Model	185	78

4.6. Case Analysis

Beyond quantitative performance metrics, specific prediction cases are examined here to illustrate the model's reasoning process and interpretability. Interpretability is crucial for knowledge graph applications, particularly in critical domains like equipment maintenance, where users need to understand and trust the system's recommendations. The proposed model provides interpretability through two mechanisms: the discovered semantic paths and their associated attention weights.

Consider a representative example from WN18RR involving the query (dog, hypernym, ?), which seeks to identify broader categories for the concept "dog". The model successfully predicts correct hypernyms such as "animal" and "entity" by leveraging both local and global reasoning. At the local level, the GAT encoder captures direct relations such as $\text{dog} \rightarrow [\text{is_a}] \rightarrow \text{canine}$, which provides immediate taxonomic information. More interestingly, the path sampler discovers a longer, semantically coherent reasoning chain: $\text{dog} \rightarrow [\text{member_meronym}] \rightarrow \text{pack} \rightarrow [\text{hypernym}] \rightarrow \text{group} \rightarrow [\text{hypernym}] \rightarrow \text{abstraction}$. This four-hop path reveals an alternative reasoning strategy: recognizing that dogs are members of packs, packs are types of groups, and groups are abstractions, the model infers the hierarchical category structure, as illustrated in Figure 6.

The attention mechanism in the path aggregator assigns this longer path a weight of 0.34, compared to 0.52 for the direct canine path and lower weights (0.08–0.14) for other sampled paths. These attention distributions reveal that while the model primarily relies on the most direct evidence, it also incorporates complementary reasoning paths to strengthen prediction confidence. Without the explicit path reasoning module, a pure embedding-based model would need to compress all this hierarchical information into fixed-dimensional vectors, potentially losing the transparent reasoning structure that the proposed approach preserves.

The path attention values (e.g., 0.52, 0.31) shown in this case study are the normalized weights a_m computed by the path aggregator. They reflect the model's learned assessment of each path's contribution to the final prediction for the specific query.

This case illustrates two key advantages of the proposed model. First, the attention-guided sampling successfully identifies semantically meaningful paths rather than arbitrary random walks—the discovered path follows coherent semantic transitions through membership and taxonomic relations. Second, the attention weights provide actionable explanations that domain experts can verify and trust, making the model suitable for deployment in practical applications where prediction rationale is as important as prediction accuracy.

From a symmetry perspective, the learned path attention weights (0.52 for direct path vs. 0.34 for multi-hop path vs. 0.08–0.14 for others) constitute interpretable asymmetric reasoning evidence. This asymmetric distribution demonstrates that the model does not treat all reasoning chains equally but rather learns to distinguish between high-confidence direct evidence and supporting contextual paths. The asymmetry is task-dependent and data-driven: for different queries, the model adaptively adjusts the relative importance of various reasoning strategies. This breaks the implicit symmetry assumption in traditional approaches, where all paths of similar length would be weighted equally. The learned asymmetric attention mechanism not only improves prediction accuracy but also provides transparent decision support by revealing which structural patterns and reasoning chains contribute most strongly to each specific prediction, thereby enhancing the model's trustworthiness for safety-critical applications such as equipment maintenance systems.

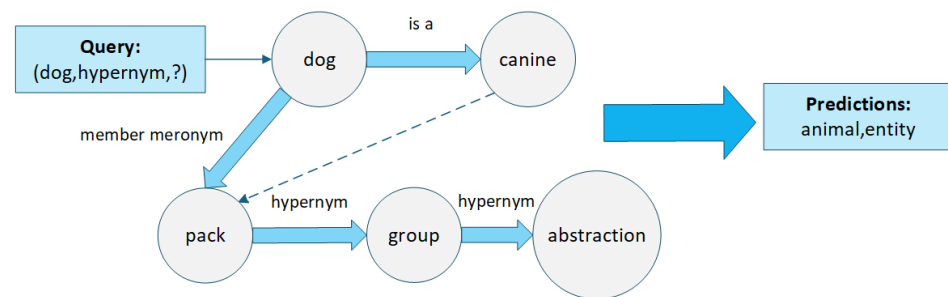


Figure 6. Case analysis example showing multi-hop reasoning for the query (dog, hypernym, ?). The model discovers both a direct path through “canine” (shown with dashed line) and a longer semantic path through “pack”, “group”, and “abstraction”, ultimately predicting “animal” and “entity” as correct hypernyms. Solid arrows represent the discovered multi-hop reasoning chain.

4.7. Interpretability Case Study

A key advantage of our model is its ability to provide interpretable reasoning paths. We demonstrate this through a case study on the Equipment Maintenance Knowledge Graph. Consider the query (“Abnormal_Bearing_Vibration”, ‘requires’, ?) aiming to find the required tool for a specific fault.

Our model correctly predicts “Vibration_Analyzer” as the top-1 answer. More importantly, by examining the attention weights α_m in the path aggregator, we can retrieve and analyze the most influential reasoning paths. Table 7 shows the top-2 paths with the highest attention weights.

Table 7. Top-2 interpretable reasoning paths for the query (“Abnormal_Bearing_Vibration”, ‘requires’, ?).

Sampled Path and Inferred Semantics	
Path 1 (Attention: 0.52):	
Abnormal_Bearing_Vibration	$\xrightarrow[\text{indicates}]{\text{caused_by}}$ Bearing_Wear $\xrightarrow[\text{needs}]{\text{diagnosed_by}}$ Vibration_Analysis $\xrightarrow[\text{employs}]{\text{performed_using}}$ Vibration_Analyzer
Human Interpretation: The vibration is caused by bearing wear, which is diagnosed by vibration analysis, a procedure that employs a vibration analyzer. This is a logically sound causal diagnostic chain.	
Path 2 (Attention: 0.31):	
Abnormal_Bearing_Vibration	$\xrightarrow[\text{is_a}]{\text{is_a}}$ Rotating_Machine_Fault $\xrightarrow[\text{commonly}]{\text{commonly_addressed_by}}$ Spectrum_Analysis $\xrightarrow[\text{uses}]{\text{requires_tool}}$ Vibration_Analyzer
Human Interpretation: The fault is a type of rotating machine fault, commonly addressed by spectrum analysis, which requires the tool. This path leverages hierarchical and statistical knowledge.	

This case demonstrates that our model not only makes accurate predictions but also provides human-understandable evidence. The high-attention paths align well with maintenance domain logic, showcasing *causal reasoning* (Path 1) and *hierarchical/statistical reasoning* (Path 2). Such transparency is vital for building trust in safety-critical industrial maintenance systems, allowing technicians or engineers to verify the model’s rationale.

5. Intelligent Knowledge Push System

While Section 4 demonstrates the quantitative performance of our model on benchmark datasets, this section showcases its practical utility in a real-world industrial main-

tenance scenario. The intelligent knowledge push system described here represents the target application of our research, addressing the gap between theoretical KG completion and operational maintenance support.

The link prediction model constructed in this paper is primarily used for the knowledge push function in equipment intelligent maintenance assistance systems. This section further explains how this model serves as the “core reasoning engine” in practical AR maintenance scenarios, enabling the system to more accurately and timely push relevant knowledge to maintenance personnel.

The basic approach for knowledge push is as follows: The front-end AR device sends the current maintenance context to the knowledge graph back-end, such as the component being handled, fault phenomena appearing, or operational steps. The system then maps this contextual information to the knowledge graph and initiates reasoning queries accordingly. Through the path reasoning capability provided by link prediction, the system can not only find explicitly existing associations in the graph but also discover knowledge content that is not directly recorded but logically relevant, thereby achieving more intelligent and practical knowledge recommendation.

To verify the application capability of the proposed link prediction model in real scenarios, this section demonstrates its role in driving intelligent knowledge push services through a typical equipment maintenance process case. This case is based on the constructed equipment maintenance knowledge graph and simulates part of the workflow from inspection to knowledge push, as illustrated in Figure 7.

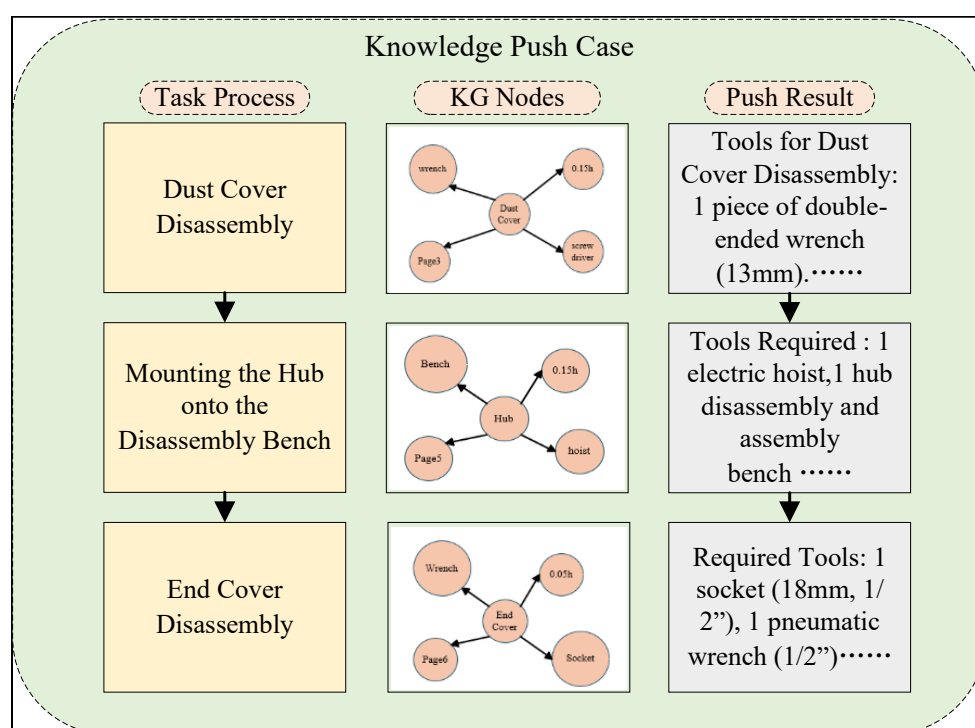


Figure 7. Knowledge push case diagram showing the three-stage maintenance process: dust cover disassembly, mounting the hub onto the disassembly bench, and end cover disassembly. For each task process step, the system identifies relevant KG nodes and pushes corresponding knowledge such as required tools.

As shown in Figure 7, the task workflow consists of dust cover removal, wheel hub mounting on rack, and end cap removal. Each step of the process matches corresponding nodes in the graph and performs relation reasoning, ultimately returning knowledge nodes with high similarity to the process. The task first performs dust cover removal, at which time knowledge highly similar to the current task will be returned. When the process

proceeds to the next step, knowledge highly similar to the next step will be pushed, and so on, completing the knowledge push for the entire process. The specific workflow of the push process is shown in Figure 8.

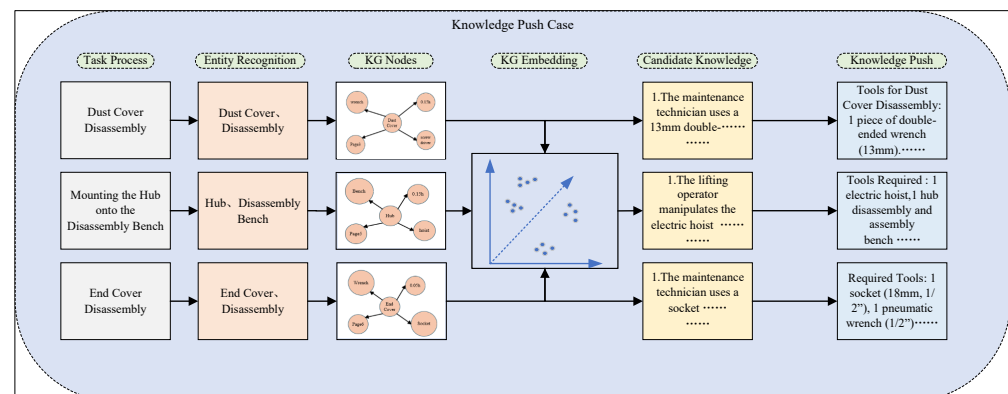


Figure 8. Detailed knowledge push workflow. The system processes task steps through entity recognition, matches them with KG nodes, performs graph embedding and reasoning, generates candidate knowledge, and finally pushes relevant information to the AR interface.

The knowledge push interface on the AR end demonstrates the system’s practical application. When the maintenance process reaches the “wheel hub mounting on rack” stage, the system first identifies and matches corresponding nodes in the knowledge graph based on the current process context. Subsequently, the proposed model performs link prediction reasoning on the “required tools” relation based on this node: on one hand, it recalls tools already existing in the graph, and on the other hand, it predicts candidate tool nodes that are highly relevant but not explicitly recorded in the graph. Finally, the system returns top-K highly relevant tools pushed to the AR interface, effectively assisting maintenance personnel in preparing tools in advance and improving operational efficiency.

6. Conclusions

In this paper, we propose a novel GAT-guided semantic path reasoning framework for knowledge graph completion, with a specific focus on equipment maintenance applications. The framework addresses the key challenge of integrating local structural dependencies with global multi-hop reasoning by introducing a unified architecture that combines graph attention networks, target-guided path sampling, and attention-based path aggregation.

The core innovation of this work lies in breaking the path-level symmetry assumption inherent in traditional reasoning methods. By transforming GAT attention weights into probabilistic transition biases, our target-guided path sampling strategy enables adaptive, asymmetric weighting of semantic paths based on both local importance and global target relevance. This design allows the model to automatically discover and prioritize high-quality reasoning chains without relying on manually defined schemas, while providing interpretable evidence through path attention distributions.

Experimental evaluations on standard benchmark datasets (FB15k-237, WN18RR) and a real-world equipment maintenance knowledge graph demonstrate the effectiveness of the proposed approach. Key findings include that the model consistently outperforms state-of-the-art baselines, achieving significant improvements on the maintenance dataset and competitive performance on the benchmark datasets. Moreover, ablation studies confirm the contribution of each core component, particularly highlighting the importance of the attention-guided path sampling mechanism. The framework also provides interpretable reasoning through path attention weights, offering transparent decision support for safety-critical maintenance scenarios. Finally, practical deployment in an AR-based intelligent

maintenance assistance system validates the model's utility for real-world knowledge delivery and decision support. The proposed framework effectively addresses the unique challenges of equipment maintenance knowledge graphs, including sparse dependencies and heterogeneous relations. Future work will explore: (1) extending the model to temporal knowledge graphs for predictive maintenance; (2) investigating meta-learning techniques for cold-start scenarios; (3) integrating multi-modal information to further enhance reasoning capabilities; and (4) optimizing the computational efficiency for large-scale industrial deployments.

Despite its promising results, this work has several limitations that point to future research directions. First, while the target-guided sampling improves efficiency, the model's performance on cold-start entities (e.g., newly installed equipment or extremely rare faults with few known relationships) has not been explicitly tested. The reliance on GAT embeddings and similarity to a target entity may be less effective when entity representations are poorly learned due to sparse connectivity. Future work should evaluate the model in cold-start settings and explore integration with auxiliary information (e.g., entity descriptions) or meta-learning techniques to address this challenge. Second, the computational overhead of online path sampling, although justified, could become a bottleneck for real-time applications on massive graphs. Investigating subgraph sampling or pre-computing approximate path indexes are potential solutions. Finally, the current model operates on static knowledge graphs. Extending it to temporal knowledge graphs to model the evolution of equipment states and maintenance actions over time is a crucial next step for predictive maintenance applications.

Author Contributions: Conceptualization, M.C. and Y.B.; methodology, M.C.; software, M.C.; validation, M.C. and Y.B.; formal analysis, M.C.; investigation, M.C.; resources, Y.B.; data curation, M.C.; writing—original draft preparation, M.C.; writing—review and editing, M.C. and Y.B.; visualization, M.C.; supervision, Y.B.; project administration, Y.B.; funding acquisition, Y.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Data Availability Statement: The data of this study are openly available in the KG-Reasoning repository at <https://github.com/a1399136594-bot/KG-Reasoning> (accessed on 27 February 2026).

Conflicts of Interest: Author Meng Chen was employed by the company Shanghai Donghu Machinery Factory. This author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Zhang, Z.; Yu, J.; Yang, B.; Du, K.; Wang, S.; Qi, X. A knowledge graphs construction method enhanced by multimodal large language model for industrial equipment operation and maintenance. *Adv. Eng. Inform.* **2025**, *68*, 103705. [\[CrossRef\]](#)
2. Qi, X.; Yang, B.; Wang, S.; Zhang, Z.; Zhang, Y.; Du, K. Few-shot and chain-of-thought prompting for equipment maintenance knowledge graph construction via large language models. *Knowl.-Based Syst.* **2026**, *335*, 115266. [\[CrossRef\]](#)
3. Li, X.; Li, G.; Wang, T.; Dong, Z.; Qin, Z.; Chu, F. Knowledge extraction and retrieval-augmented generation for intelligent maintenance of wind power equipment based on graph attention networks. *Chin. J. Mech. Eng.* **2025**, 100141. [\[CrossRef\]](#)
4. Wang, G.; Li, Y.; Cheng, C.X.; Li, R.; Ding, C.X.; Zhang, Y.; Liu, Y. Integrated optimization of equipment degradation modeling and spare parts inventory for predictive maintenance in power systems. *Sustain. Energy Technol. Assess.* **2025**, *83*, 104626. [\[CrossRef\]](#)
5. Yuan, J.; Lei, Y.; Li, N.; Yang, B.; Li, X.; Chen, Z.; Han, W. A framework for modeling and optimization of mechanical equipment considering maintenance cost and dynamic reliability via deep reinforcement learning. *Reliab. Eng. Syst. Saf.* **2025**, *264*, 111424. [\[CrossRef\]](#)
6. Wang, Z.; Zeng, Q.; Haralambides, H. Shift of emphasis toward intelligent equipment maintenance in port operations: A critical review of emerging trends and challenges. *Ocean. Coast. Manag.* **2024**, *259*, 107408. [\[CrossRef\]](#)
7. Tang, Y.; Zhou, Y.; Wu, T.; Wang, C.; Li, Z.; Li, K. AI-driven predictive maintenance for medical imaging equipment: A deep learning framework based on the IoMT data. *Reliab. Eng. Syst. Saf.* **2026**, *270*, 112152. [\[CrossRef\]](#)

8. Dui, H.; Zhang, A.; Zhou, Y.; Tang, C. IoT-based reliability analysis and maintenance strategy of unmanned equipment communication systems. *Reliab. Eng. Syst. Saf.* **2026**, *266*, 111628. [\[CrossRef\]](#)
9. Huang, J.; Zou, S.; Li, Y.; Xiao, T.; Zhang, J. Performance and maintenance strategy of a data center cooling system under equipment aging. *Appl. Therm. Eng.* **2025**, *281*, 128623. [\[CrossRef\]](#)
10. Nguyen, D.Q. A survey of embedding models of entities and relationships for knowledge graph completion. In *Proceedings of the Graph-Based Methods for Natural Language Processing (TextGraphs)*; Association for Computational Linguistics: Kerrville, TX, USA, 2020; pp. 1–14.
11. Zou, X. A survey on application of knowledge graph. *J. Phys. Conf. Ser.* **2020**, *1487*, 012016. [\[CrossRef\]](#)
12. Toutanova, K.; Chen, D. Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd Workshop on Continuous Vector Space Models and Their Compositionality*; Association for Computational Linguistics: Kerrville, TX, USA, 2015; pp. 57–66.
13. Dettmers, T.; Minervini, P.; Stenetorp, P.; Riedel, S. Convolutional 2d knowledge graph embeddings. In *Proceedings of the AAAI Conference on Artificial Intelligence*; AAAI Press: Washington, DC, USA, 2018; Volume 32.
14. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating embeddings for modeling multi-relational data. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, USA, 2013; Volume 26.
15. Nickel, M.; Tresp, V.; Kriegel, H.P. A three-way model for collective learning on multi-relational data. In *Proceedings of the ICML*; Omnipress: Madison, WI, USA, 2011; Volume 11, pp. 3104482–3104584.
16. Yang, B.; Yih, W.T.; He, X.; Gao, J.; Deng, L. Embedding entities and relations for learning and inference in knowledge bases. *arXiv* **2014**, arXiv:1412.6575.
17. Martínez, V.; Berzal, F.; Cubero, J.C. A survey of link prediction in complex networks. *ACM Comput. Surv.* **2016**, *49*, 69. [\[CrossRef\]](#)
18. Schlichtkrull, M.; Kipf, T.N.; Bloem, P.; Van Den Berg, R.; Titov, I.; Welling, M. Modeling relational data with graph convolutional networks. In *Proceedings of the European Semantic Web Conference*; Springer: Cham, Switzerland, 2018; pp. 593–607.
19. Wang, S.; Wang, Y.; Wang, M. Connectivity and matching preclusion for leaf-sort graphs. *J. Interconnect. Netw.* **2019**, *19*, 1940007. [\[CrossRef\]](#)
20. Wang, M.; Lin, Y.; Wang, S. The connectivity and nature diagnosability of expanded k -ary n -cubes. *RAIRO-Theor. Inform. Appl.-Inform. Théor. Appl.* **2017**, *51*, 71–89. [\[CrossRef\]](#)
21. Wang, S.; Wang, M. A Note on the Connectivity of m -Ary n -Dimensional Hypercubes. *Parallel Process. Lett.* **2019**, *29*, 1950017. [\[CrossRef\]](#)
22. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
23. Shang, C.; Tang, Y.; Huang, J.; Bi, J.; He, X.; Zhou, B. End-to-end structure-aware convolutional networks for knowledge base completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*; AAAI Press: Washington, DC, USA, 2019; Volume 33, pp. 3060–3067.
24. Vashishth, S.; Sanyal, S.; Nitin, V.; Talukdar, P. Composition-based multi-relational graph convolutional networks. *arXiv* **2019**, arXiv:1911.03082.
25. Balazević, I.; Allen, C.; Hospedales, T. Tucker: Tensor factorization for knowledge graph completion. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*; Association for Computational Linguistics: Kerrville, TX, USA, 2019; pp. 5185–5194.
26. Wang, M.; Yang, W.; Wang, S. Conditional matching preclusion number for the Cayley graph on the symmetric group. *Acta Math. Appl. Sin. (Chin. Ser.)* **2013**, *36*, 813–820.
27. Wang, M.; Yang, W.; Guo, Y.; Wang, S. Conditional fault tolerance in a class of Cayley graphs. *Int. J. Comput. Math.* **2016**, *93*, 67–82. [\[CrossRef\]](#)
28. Wang, M.; Ren, Y.; Lin, Y.; Wang, S. The tightly super 3-extra connectivity and diagnosability of locally twisted cubes. *Am. J. Comput. Math.* **2017**, *7*, 127–144. [\[CrossRef\]](#)
29. Wang, S.; Wang, Z.; Wang, M.; Han, W. g -Good-neighbor conditional diagnosability of star graph networks under PMC model and MM* model. *Front. Math. China* **2017**, *12*, 1221–1234. [\[CrossRef\]](#)
30. Wang, S.; Wang, M. The strong connectivity of bubble-sort star graphs. *Comput. J.* **2019**, *62*, 715–729. [\[CrossRef\]](#)
31. Lin, Y.; Liu, Z.; Luan, H.; Sun, M.; Rao, S.; Liu, S. Modeling relation paths for representation learning of knowledge bases. *arXiv* **2015**, arXiv:1506.00379. [\[CrossRef\]](#)
32. Zhang, M.; Chen, Y. Link prediction based on graph neural networks. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, USA, 2018; Volume 31.
33. Teru, K.; Denis, E.; Hamilton, W. Inductive relation prediction by subgraph reasoning. In *Proceedings of the International Conference on Machine Learning*; Curran Associates Inc.: Red Hook, NY, USA, 2020; pp. 9448–9457.

34. Vashishth, S.; Sanyal, S.; Nitin, V.; Agrawal, N.; Talukdar, P. InteractE: Improving convolution-based knowledge graph embeddings by increasing feature interactions. In *Proceedings of the AAAI Conference on Artificial Intelligence*; AAAI Press: Washington, DC, USA, 2020; Volume 34, pp. 3009–3016.
35. Zhu, Z.; Zhang, Z.; Xhonneux, L.P.; Tang, J. Neural bellman-ford networks: A general graph neural network framework for link prediction. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 29476–29490.
36. Ying, C.; Cai, T.; Luo, S.; Zheng, S.; Ke, G.; He, D.; Liu, T.Y. Do transformers really perform badly for graph representation? *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 28877–28888.
37. Shi, F.; Li, D.; Wang, X.; Li, B.; Wu, X. Tgformer: A graph transformer framework for knowledge graph embedding. *IEEE Trans. Knowl. Data Eng.* **2024**, *37*, 526–541. [[CrossRef](#)]
38. Bu Sinnah, Z.A. Explainable AI-driven predictive maintenance for mitigating process safety risks in safety-critical industrial equipment. *J. Loss Prev. Process Ind.* **2026**, *100*, 105907. [[CrossRef](#)]
39. Yuan, J.; Lei, Y.; Yang, B.; Li, N.; Li, X.; Chen, Z.; Han, W. A hybrid maintenance strategy for equipment with competitive failure modes: Sudden failure and multi-stage degradation failure. *Mech. Syst. Signal Process.* **2025**, *234*, 112846. [[CrossRef](#)]
40. Wang, J.; Zhang, Z.; Huang, Y.; Hu, H. Proactive dispatch of rail maintenance equipment under weather-induced disruptions: An optimization framework integrating machine learning-driven risk assessment. *Adv. Eng. Inform.* **2026**, *69*, 103990. [[CrossRef](#)]
41. Lu, L.; Deng, Q.; Luo, Q.; Hu, Y.; Zhang, J.; Deng, K. Integrated optimization of dynamic flexible job shop scheduling with AGV breakdowns and multi-site equipment maintenance on the demand side. *Comput. Ind. Eng.* **2025**, *208*, 111361. [[CrossRef](#)]
42. Lu, J.; Chen, H.; Chen, J.; Xiao, Z.; Li, R.; Xiao, G.; Wang, Q. Temporal knowledge graph fusion with neural ordinary differential equations for the predictive maintenance of electromechanical equipment. *Knowl.-Based Syst.* **2025**, *317*, 113450. [[CrossRef](#)]
43. Yang, Y.; Song, W.; Wang, J. A high-dimensional many-objective co-optimization method on equipment utilization and maintenance. *Array* **2026**, *29*, 100672. [[CrossRef](#)]
44. Wehbi, L.; Bhulai, S.; Slik, J. Optimizing maintenance of heavy equipment: A data-driven approach. *Comput. Ind. Eng.* **2026**, *211*, 111645. [[CrossRef](#)]
45. Gandhare, S.; Kumar, P.; Madankar, T.; Singh, D.; Bhamu, J. Development and validation of a FMEA-based medical equipment maintenance framework using Industry 4.0. *Int. J. Qual. Reliab. Manag.* **2025**, *42*, 2250–2290. [[CrossRef](#)]
46. Benzoni, G.; Truong, V.D.; Brace, W.; Cammi, A.; Introini, C.; Micciché, G.; Mikko, S.; Tripodo, C. Petri net modelling—A remote maintenance simulation approach for fusion remote maintenance equipment concepts assessment. *Fusion Eng. Des.* **2025**, *215*, 114996. [[CrossRef](#)]
47. Nadim, K.; Ragab, A.; Ghezzaz, H.; Ouali, M.S. Reliable Dynamic Causality Analysis for Efficient Prescriptive Maintenance of Degraded Industrial Equipment. *J. Loss Prev. Process Ind.* **2026**, *101*, 105922. [[CrossRef](#)]
48. Wang, T.; Hu, G.; Cho, S.; Tan, S.W.; Chiang, K.L. A reliability-centered maintenance process for equipment in distribution grids based on extended virtual age model. *Energy* **2025**, *328*, 135998. [[CrossRef](#)]
49. Bhardwaj, A.S.; Veeramani, D. An unsupervised multi-class ensemble classifier for identifying equipment failure mechanisms from maintenance records. *Reliab. Eng. Syst. Saf.* **2025**, *264*, 111410. [[CrossRef](#)]
50. Darghouth, M.N. An integrated two-dimensional warranty framework for second-hand equipment considering condition-based maintenance, upgrades, and past life. *J. Model. Manag.* **2025**, *20*, 2136–2166. [[CrossRef](#)]
51. Luo, Z.; Lian, Z.; Li, J.; Yang, Z. Maintenance strategies for a queueing system with a standby equipment: An integrated approach using condition-based monitoring. *Reliab. Eng. Syst. Saf.* **2025**, *263*, 111219. [[CrossRef](#)]
52. Sepehri, A.; Kirichek, A.; van den Heuvel, M.; de Geus, M.; van Koningsveld, M. Sustainable port maintenance: Dredging equipment selection in time-emission trade-offs. *Results Eng.* **2026**, *29*, 109247. [[CrossRef](#)]
53. Cong, W.; Tao, W.; Jinsong, B. Agentic digital twin-embedded maintenance methodology for energy equipment: A self-evolving operational paradigm. *J. Manuf. Syst.* **2026**, *84*, 100–116. [[CrossRef](#)]
54. Zhang, Y.; Yao, Q. Knowledge graph reasoning with relational digraph. In *Proceedings of the ACM Web Conference 2022*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 912–924.
55. Li, Z.; Wang, X.; Zhao, J.; Feng, F.; Chen, Z.; Li, J. Hysae: An efficient semantic-enhanced representation learning model for knowledge hypergraph link prediction. In *Proceedings of the ACM on Web Conference 2025*; Association for Computing Machinery: New York, NY, USA, 2025; pp. 86–97.
56. Zhang, Y.; Tian, J.; Sun, J.; Chan, H.; Qiu, A.; Liu, C. HKGAT: Heterogeneous knowledge graph attention network for explainable recommendation system. *Appl. Intell.* **2025**, *55*, 549. [[CrossRef](#)]

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