A Survey of Link Prediction in N-ary Knowledge Graphs

Jiyao Wei[1,2], **Saiping Guan**[1,2]*, **Da Li**[1,2], **Zhongni Hou**[3],
Miao Su[1,2], **Yucan Guo**[1,2], **Xiaolong Jin**[1,2]*, **Jiafeng Guo**[1,2], **Xueqi Cheng**[1,2]
[1]School of Computer Science and Technology, University of Chinese Academy of Sciences;
[2]State Key Laboratory of AI Safety, Institute of Computing Technology,
Chinese Academy of Sciences; [3]Meituan.
{weijiyao20z, guansaiping, jinxiaolong}@ict.ac.cn

Abstract

N-ary Knowledge Graphs (NKGs) are a specialized type of knowledge graph designed to efficiently represent complex real-world facts. Unlike traditional knowledge graphs, where a fact typically involves two entities, NKGs can capture n-ary facts containing more than two entities. Link prediction in NKGs aims to predict missing elements within these n-ary facts, which is essential for completing NKGs and improving the performance of downstream applications. This task has recently gained significant attention. In this paper, we present the first comprehensive survey of link prediction in NKGs, providing an overview of the field, systematically categorizing existing methods, and analyzing their performance and application scenarios. We also outline promising directions for future research.

1 Introduction

Since Google introduced Knowledge Graph (KG) to enhance its search services, KGs have attracted growing attention from both academia and industry (Lehmann et al., 2015). A traditional KG stores numerous facts, typically represented in the form of triples (h, r, t), indicating a specific relation r between a head entity h and a tail entity t, such as ($Biden, the\ President\ of, the\ USA$).

However, many real-world facts involve more than two entities, requiring a more expressive representation. N-ary Knowledge Graphs (NKGs) address this need by enabling the representation of complex facts involving multiple entities, commonly referred to as n-ary facts. For instance, the fact *"Einstein studied physics at the University of Zurich and received his PhD"* can be represented as $\{person : Einstein, institution : Uni.\ Zurich, major : Physics, degree : PhD\}$ in NKGs.

N-ary facts are prevalent in real-world scenarios (Fatemi et al., 2021). Statistically, in Freebase, over a third of entities are involved in n-ary facts (Wen et al., 2016), and more than 61% facts are n-ary facts (Fatemi et al., 2021). Like traditional KGs, NKGs are inevitably incomplete, due to the complex process of their construction (Li et al., 2024d). The incompletion of NKGs hinders the performance of downstream applications, including information retrieval (Zhao et al., 2020) and recommendation systems (Liang et al., 2023). To address this, link prediction in NKGs is proposed to predict missing elements in facts therein, helping populate and enrich NKGs (Wen et al., 2016).

Traditional link prediction methods for KGs encode facts as triples. To handle n-ary facts, they decompose each n-ary fact into multiple triples, such as introducing Compound Value Type (CVT) entities in Freebase (Bollacker et al., 2008). However, this decomposition complicates inference, leads to structural information loss, increases model parameters, and risks incorrect reasoning (Wen et al., 2016). More details on the disadvantages of the decomposition are shown in Appendix A.

As shown in Figure 1, recent efforts have increasingly focused on directly modeling n-ary facts without decomposition. Methods for link prediction in NKGs fall into three categories: spatial mapping-based (Wen et al., 2016), tensor decomposition-based, and neural network-based approaches (Guan et al., 2019). These methods address both general scenarios (Wen et al., 2016) and specialized scenarios, including temporal (Hou et al., 2023), few-shot (Zhang et al., 2022b), and inductive settings (Ali et al., 2021).

Although there exist numerous surveys on link prediction in KGs, covering general (Wang et al., 2017; Guan et al., 2022), temporal (Cai et al., 2023; Wang et al., 2023c), multi-modal KGs (Zhu et al., 2022; Peng et al., 2023), and sparse KGs (Chen

* Corresponding Authors.

Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, pages 28534–28556
November 4-9, 2025

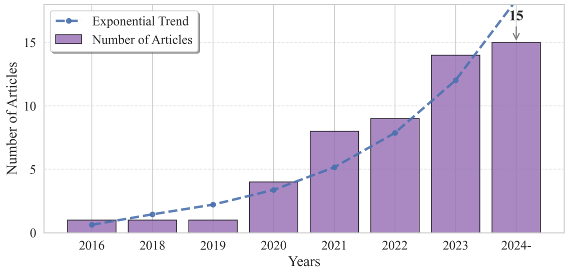

Figure 1: Number of articles published each year (2016-) on link prediction in NKGs.

et al., 2023c), none specifically focus on NKGs. More statistics of existing surveys are shown in Appendix B. Despite the rapid development of link prediction methods for NKGs, with nearly 50 methods proposed, existing surveys only briefly mention it, lacking a comprehensive and in-depth analysis. A dedicated survey on link prediction in NKGs is crucial for understanding the progress, challenges, and future directions of this field. This paper aims to fill this gap by providing a detailed and timely survey on link prediction in NKGs, facilitating further research in this area.

The rest of this paper is organized as follows: Section 2 introduces the related definitions of link prediction in NKGs and categorizes existing methods. Section 3 analyzes existing methods. Section 4 reports their performance on benchmarks. Section 5 highlights representative applications. Finally, Section 6 suggests future research directions. Compilation and details of papers used for the survey can be found via our repository[1].

2 Preliminary

In this section, we introduce the related definitions of link prediction in NKGs, the formalizations of n-ary facts, the classifications of existing methods for link prediction in NKGs, and the applicability of link prediction in KGs and NKGs.

2.1 Definition

Definition 1. ***KG****: a set of facts, each of which is represented as a triple (h, r, t), where h and t denote its head entity and tail entity, respectively, and r denotes the relation between them.*

For example, in a KG, the fact *"Biden is the president of the USA"* is represented as ($Biden$, $the\ President\ of$, $the\ USA$).

[1] https://github.com/JiyaoWei/LP_NKGs

Definition 2. ***NKG****: a set of facts, each of which may contain more than two entities, which is also referred to as an n-ary fact.*

For example, in an NKG, the n-ary fact *"Einstein studied physics at the University of Zurich and received his PhD."* contains four entities, including *Einstein, the University of Zurich, Physics*, and *PhD*. Therefore, it is called a 4-ary fact.

Definition 3. ***Link prediction in NKGs****: predict missing elements in facts in NKGs based on the existing facts.*

For example, predict the missing entity *the University of Zurich* in the n-ary fact *"Einstein studied physics at ? and received his PhD"*.

2.2 The formalizations of N-ary Facts

Typical n-ary fact formalizations include hyperedge, role-value pair, and hyper-relation formalizations, as illustrated in Figure 2.

2.2.1 Hyperedge Formalization

A hyperedge connects all entities in an n-ary fact (Wen et al., 2016), e.g., $(H, e_1, ..., e_n)$, where e_* is the *-th entity and hyperedge H indicates the role of each entity in the fact. For example, the facts in Figure 2 can be represented as follows:

Fact 1: ($educated\_with\_degree\_major$, $Einstein$, $Uni.\ Zurich$, PhD, $Physics$).

Fact 2: ($awarded\_with\_time\_place$, $Einstein$, $Nobel\ Prize\ in\ Physics$, 1921, $Switzerland$).

Note that under the hyperedge formalization, the entities in the formalized fact are ordered with a fixed number of entities. Each position in the hyperedge represents a fixed role. The hyperedge formalization directly builds connections between multiple entities in an n-ary fact. When dealing with binary facts, this formalization is simplified to the triple formalization in traditional KGs.

2.2.2 Role-value Pair Formalization

An n-ary fact is formulated as multiple role-value pairs (Guan et al., 2019), such as $\{r_i : v_i\}_{i=1}^n$, where value v_i is an entity and plays the role r_i in the fact; n is the total number of entities within the fact. For example, the two facts about Einstein introduced above can be represented as follows:

$$\begin{aligned}\text{Fact 1: } \{&person : Einstein,\\ &institution : Uni.\ Zurich,\\ &degree : PhD,\\ &major : Physics\}.\end{aligned}$$

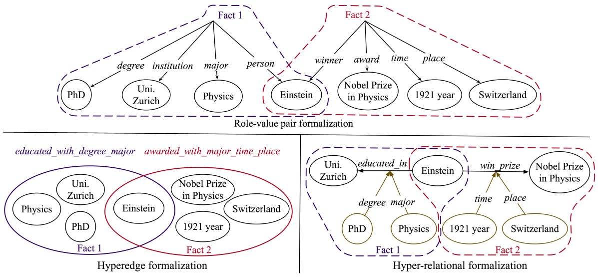

Figure 2: Examples of different formalizations of n-ary facts.

Fact 2: $\{winner : Einstein,$
$award : Nobel\ Prize\ in\ Physics,$
$place : Switzerland,$
$time : 1921\}.$

Note that in the role-value pair formalization, the role-value pairs within a fact are unordered and may involve an arbitrary number of entities. This representation offers flexibility in specifying the roles of entities in n-ary facts. However, it fails to account for the varying importance or prominence of different entities within the same fact.

2.2.3 Hyper-relational Formalization

An n-ary fact is formulated as a primary triple coupled with a set of qualifier role-value pairs (Rosso et al., 2020; Guan et al., 2020), i.e., $((h, r, t), \{r_i : v_i\}_{i=1}^{n-2})$, where h and t are the head entity and tail entity of the fact and r denotes the relation between them; role-value pairs $\{r_i : v_i\}_{i=1}^{n-2}$ qualify the triple (h, r, t). For example, the two facts about Einstein introduced above can be represented as follows:

Fact 1: $((Einstein, educated, Uni.\ Zurich),$
$-|\{major : Physics,$
$-|degree : PhD\}).$

Fact 2: $((Einstein, won, Nobel\ Prize\ in\ Physics),$
$-|\{time : 1921,$
$-|place : Switzerland\}).$

Note that when there is no clear subject (i.e., head entity) or object (i.e., tail entity) in the facts, it is not appropriate to use the hyper-relational formalization. Additionally, no matter which formalization is used, link prediction in NKGs aims to predict missing elements in facts.

2.3 Classification of Methods for Link Prediction in NKGs

From a technical perspective, link prediction methods for NKGs fall into three main categories: spatial mapping-based, tensor decomposition-based, and neural network-based. Spatial mapping-based methods project entities into semantic space (e.g., Euclidean, hyperbolic, or complex) and then assess fact plausibility via entity positions. Tensor decomposition-based methods model n-ary facts as higher-order tensors, indicating fact validity. Neural network-based methods use Fully Connected Network (FCN), Convolutional Neural Network (CNN), Transformer, or Graph Neural Network (GNN) to encode element associations in n-ary facts.

Most of the above methods are designed for the general scenario, while some GNN-based methods address special scenarios, such as the temporal, few-shot, and inductive settings. In addition, different methods use different formalizations of n-ary facts. The correspondence between fact formalizations and methods is shown in Appendix D.

2.4 Applicability of NKGs and Their Link Prediction

This subsection discusses the scenarios where NKGs are particularly beneficial and highlights the fundamental differences between link prediction in NKGs and traditional KGs.

2.4.1 When to Use NKGs: Suitable Scenarios and Motivations

NKGs extend traditional KGs by effectively representing facts that involve three or more entities. Their use is particularly advantageous in the following scenarios: (1) Multi-party participation: a fact involves multiple semantically related entities; (2) Semantic coupling: entities in facts form an inseparable semantic unit that cannot be split into multiple triples without loss of meaning; (3) Context-dependent facts: facts are context-dependent, requiring time, location, or other conditions to be fully understood. Assessing these aspects can help determine whether NKGs provide advantages over traditional KGs for more accurate downstream reasoning. Further details are provided in Appendix C.

2.4.2 Link Prediction in NKGs vs. Traditional KGs: Key Differences

While traditional KGs and NKGs share basic components (entities, relations, facts) and adopt similar techniques (e.g., spatial mapping, tensor decomposition, neural network methods), key differences in fact structure and task definition necessitate tailored modeling strategies for NKGs.

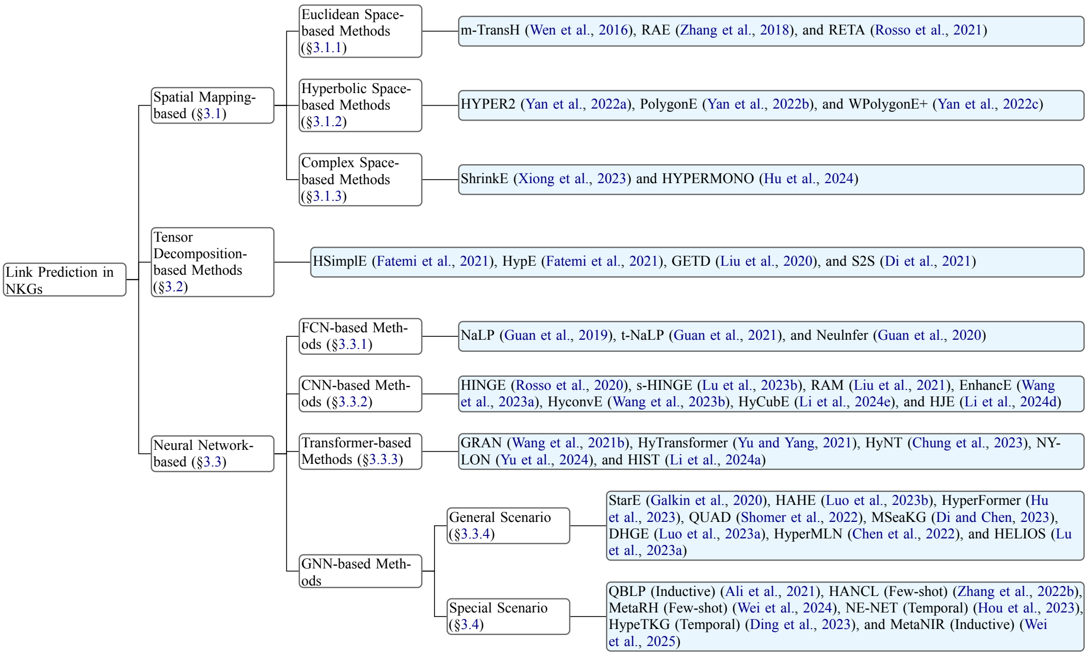

Figure 3: Classification of Methods for Link Prediction in NKGs. "()" marks specific special settings.

Modeling Structure: Traditional KG models handle simple triples (h, r, t), whereas NKGs must represent complex n-ary facts, such as hyper-relational facts involving main triples and multiple qualifier role-value pairs, demanding higher expressive power.

Prediction Tasks: Traditional methods focus on completing missing elements in triples (e.g., $(h, r, ?)$), while link prediction in NKGs often involves multiple missing roles, values, or entities, requiring more flexible and robust reasoning.

These distinctions drive specific adaptations in NKG methods. See Appendix E for a detailed comparison.

3 Methods for Link Prediction in NKGs

This section begins with introducing link prediction approaches for NKGs in general scenarios, after which it examines methods tailored to specific scenarios such as temporal, inductive, and few-shot settings. For the methods in general scenarios, we introduce each category of methods one by one, first introducing their general ideas and then delving into specific methods.

3.1 Spatial Mapping-based Methods

These methods map entities into a shared embedding space, enforcing geometric constraints to ensure meaningful spatial relationships among them. They can be further divided into three types based on the embedding space: Euclidean, hyperbolic, and complex space-based methods.

Euclidean Space-based Methods. m-TransH (Wen et al., 2016) projects entities in a fact onto a hyperplane according to their corresponding roles, and then evaluates the fact with spatial positions of the entities. However, its complexity grows with the number of missing entities in the n-ary fact. RAE (Zhang et al., 2018) reduces this complexity by assuming high similarity among entities within a fact and only calculating entities with high similarity.

Hyperbolic Space-based Methods. The number of related entities grows exponentially along the NKG hierarchy, which aligns with the superlinear growth in hyperbolic space. To capture such hierarchical structures, HYPER2 (Yan et al., 2022a) and PolygonE (Yan et al., 2022b) embed entities in hyperbolic space. HYPER2 projects entities to the tangent space to integrate qualifier values, then maps them back for scoring. PolygonE treats n-ary facts as gyro-polygons and evaluates entity compatibility via vertex-gyrocentroid distances, preserving both structure and semantics. To address the assumption of equal entity importance, WPolygonE+ (Yan et al., 2022c) introduces learned entity weights and enhances fact representation by linking gyro-midpoints and centroids.

Complex Space-based Methods. These methods effectively capture inference patterns in complex space, particularly monotonicity: if two role-value pairs q_i and q_j satisfy q_i implies q_j, then forattaching either to a fact should yield the same truth value. ShrinkE (Xiong et al., 2023) models a primary triple as a spatial-functional transformation specific to its relation, mapping the head entity to a query box in complex space that contains potential answer entities. Each qualifier constrains this box by shrinking it, ensuring that the contracted box remains inside the original—providing a geometric view of monotonicity through box containment. HYPERMONO (Hu et al., 2024) first aggregates neighbor information to enhance entity representation, and then to achieve qualifier monotonicity HYPERMONO adopts cone embedding. Each time a qualifier is added, the angle of the cone is reduced, thereby reducing the answer set.

3.2 Tensor Decomposition-based Methods

Such methods represent the set of facts in an NKG as a high-order tensor, where each tensor entry indicates the truth value of a particular fact. By reconstructing, decomposing, and optimizing this tensor, the model uncovers latent pattern features among the elements of n-ary facts, thereby enhancing link prediction accuracy. HSimplE (Fatemi et al., 2021) shifts entity embeddings by position and combines them with hyperedge embeddings for scoring. HypE (Fatemi et al., 2021) enhances this by using convolutional filters to generate position-specific embeddings.

GETD (Liu et al., 2020) generalizes Tucker decomposition by reshaping it into a higher-order tensor and applying tensor ring decomposition (Wang et al., 2018) to reduce parameters. However, GETD cannot handle multiple facts of different number of entities at the same time, which easily leads to data sparsity problems. S2S (Di et al., 2021) addresses this issue by partitioning embeddings to enable parameter sharing across facts of varying sizes, thereby improving efficiency.

3.3 Neural Network-based Methods

These methods leverage neural networks to encode NKGs and perform link prediction with learned element representations. They can be categorized into four types: FCN-based, CNN-based, Transformer-based, and GNN-based methods. Each type employs a corresponding neural network architecture to encode n-ary facts.

3.3.1 FCN-based Methods

NaLP (Guan et al., 2019) models n-ary facts as sets of role-value pairs, using FCNs to extract and aggregate features for truth value prediction, but treats all pairs equally. To address this, NeuInfer (Guan et al., 2020) decomposes a fact into a primary triple and qualifiers, scoring both the triple and its compatibility with qualifiers via FCNs. t-NaLP (Guan et al., 2021) further enhances NaLP by incorporating entity types and improved negative sampling.

3.3.2 CNN-based Methods

HINGE extends NeuInfer by modeling n-ary facts as a primary triple with role-value pairs, using CNNs and min-pooling for feature aggregation and an FCN for scoring. s-HINGE (Lu et al., 2023b) further incorporates entity type information to enhance performance, similar to t-NaLP. RAM (Liu et al., 2021) introduces a latent space to model role semantics, where role embeddings are generated via linear combinations of basis vectors, and role-specific pattern matrices evaluate entity-role compatibility using a multilinear scoring function.

To better exploit entity context, EnhancE (Wang et al., 2023a) enriches entity representations with position and neighbor information, and integrates semantics into relation embeddings. To leverage CNNs' representational power, HyconvE (Wang et al., 2023b) uses 3D convolution with role-aware and position-aware filters to capture intricate intra-fact interactions. HJE (Li et al., 2024d) enhances HyconvE with learnable position embeddings, while HyCubE (Li et al., 2024e) improves efficiency by introducing 3D circular convolutions and a masked stacking strategy.

3.3.3 Transformer-based Methods

HyTransformer (Yu and Yang, 2021) utilizes a Transformer (Vaswani et al., 2017) with a regularization layer to encode n-ary facts. It initializes position embeddings to represent type information of elements within n-ary facts, but does not explicitly model the relationship type between two elements in an n-ary fact. GRAN (Wang et al., 2021b) addresses this by representing each n-ary fact as a heterogeneous graph and introducing multiple edge types to encode relationship types between elements. It processes these heterogeneous graphs using multiple fully connected attention layers with edge-aware biases, improving the performance of the model.

The above methods can only handle discrete entities, however, real-world NKGs often contain numeric entities. For example, a number-related qualifier role-value pair (starting time, 1911) is associated with a triple (J.R.R., educated at, Oxford University). Recognizing the importance of numeric values, HyNT (Chung et al., 2023) encodes numeric literals within both primary triples and qualifier role-value pairs with a context transformer and a prediction transformer.

Different from the above methods, HIST (Li et al., 2024a) integrates text information and structural information in NKG to enhance the representation of elements in NKGs. This method uses GNNs to extract structural information and effectively integrates text and structural information through structural soft prompt tuning (Chen et al., 2023a). Recently, NYLON (Yu et al., 2024) extends GRAN to handle noisy NKGs, exploring robust link prediction in noisy NKGs. NYLON employs a Transformer with learnable edge biases to compute fact confidence and element confidence. Based on these confidences, efficient selective annotation is performed for annotation groups.

3.3.4 GNN-based Methods

GNN-based methods for link prediction in NKGs in general scenarios fall into two categories: fact modeling and schema modeling.

Fact Modeling Methods focus on encoding the elements within a fact and use GNNs to enhance their semantic representations. StarE (Galkin et al., 2020) aggregates qualifier information into relations to update entity embeddings, but lacks reverse information flow. QUAD (Shomer et al., 2022) improves this by enabling bidirectional aggregation between primary triples and qualifier pairs. HAHE (Luo et al., 2023b) models both global hypergraph structures and local semantic sequences using dual attention modules. HyperFormer (Hu et al., 2023) reduces noise from multi-hop neighbors via a bidirectional interaction mechanism and Mixture-of-Experts for parameter efficiency. To address the rigidity of fixed architectures, MSeaKG (Di and Chen, 2023) introduces a neural architecture search framework with diverse message functions adaptable to various NKG formats.

Schema Modeling Methods, on the other hand, not only encode these intra-fact elements but also incorporate schema information related to entities. By encoding schema information, Schema Modeling Methods further enrich the semantic representations of entities. DHKG (Luo et al., 2023a) uses a dual-view encoder to model instance and ontology views, while HELIOS (Lu et al., 2023a) enhances type representation using GATs and self-attention. HyperCL (Lu et al., 2024) further introduces hierarchical ontologies and concept-aware contrastive learning to balance fact and schema influences, achieving improved prediction performance.

Schema Modeling Methods can be seen as an extension of Fact Modeling Methods. They not only encode the elements within a fact but also consider schema information (e.g., type hierarchies and ontologies), providing inductive biases and constraints that factual modeling alone cannot capture. They can capture fine-grained intra-fact semantics and often achieve better performance when comprehensive ontology information is available. However, this comes with trade-offs: acquiring and maintaining high-quality schema data can be costly and labor-intensive. In scenarios where schema information is sparse or noisy, these methods may be more vulnerable to errors, reducing their effectiveness compared to factual modeling approaches that rely less on schema completeness.

3.4 Methods for Link Prediction in NKGs in Special Scenario

Recent advances have extended GNN-based methods to specialized settings, including temporal, inductive, and few-shot scenarios.

3.4.1 Temporal Setting

N-ary facts often include temporal information, yet many methods either ignore it or treat time as a generic role, blurring the distinction between relational and temporal semantics. Without explicit temporal modeling, models fail to capture ordering, duration, and the influence of historical patterns on future facts. NE-Net (Hou et al., 2023) addresses this by leveraging an entity-role encoder based on a GNN to capture precise entity evolution representations. HypeTKG (Ding et al., 2023) further considers the influence of time-invariant relations on temporal reasoning.

3.4.2 Inductive Setting

In real-world scenarios, new elements often emerge after the training phase, presenting a challenge for methods that struggle to handle unseen elements. To handle unseen elements emerging post-training, QBLP (Ali et al., 2021) generates embed-

Table 1: Comparison of link prediction methods for NKGs.

Types	Introduction	Advantages	Drawbacks
Spatial Mapping-based	They use spatial transformations to model relationships between entities and roles in an n-ary fact, ensuring that their embedding vectors maintain specific geometric constraints in the embedding space.	They are characterized by low time complexity, minimal model parameters, and fast training. This efficiency allows these methods to handle extensive datasets without significant computational overhead.	They have limited expressive power, cannot capture the complex interactions between entities and roles, and usually have poor prediction results.
Tensor Decomposition-based	They represent an NKG as a high-order tensor where each cell indicates the validity of a corresponding fact.	They possess strong expressive power, enabling them to capture complex relationships and interactions within NKGs. This results in relatively better model performance.	They have high time complexity, especially for n-ary facts with high arity.
Neural Network-based	They use neural networks to extract features from n-ary facts and then score them to predict missing elements in NKGs.	They effectively extract complex features from n-ary facts; they usually have high prediction accuracy.	The training process of these methods usually requires a large amount of data and a long training time; they have poor interpretability.

Table 2: Complexity analysis of NKG link prediction methods. The space and time complexities for m-TransH, RAE, HypE, HINGE, and NeuInfer are sourced from Liu et al. (Liu et al., 2021), while those for GETD and S2S are based on Di et al. (Di et al., 2021). Here, n_e, n_r, n_a, and d denote the number of entities, the number of roles, the maximum arity in NKGs, and the dimension of the embedding, respectively.

Methods	Years	Space Complexity	Time Complexity
m-TransH	2016	$O(n_e d + 2n_r d)$	$O(d)$
HypE	2016	$O(n_e d + n_r d)$	$O(d^2)$
RAE	2018	$O(n_e d + 2n_r d)$	$O(d^2)$
HINGE	2020	$O(n_e d + n_r n_a d)$	$O(d^2)$
NeuInfer	2020	$O(n_e d + n_r n_a d)$	$O(d^2)$
S2S	2021	$O(n_e d + n_r d)$	$O(d)$
GRAN	2021	$O(n_e d + n_r n_a d)$	$O(d^2)$
HAHE	2023	$O(n_e d + n_r n_a d)$	$O(d^2)$

dings from auxiliary facts and text. MetaNIR (Wei et al., 2025) adopts meta-learning (Finn et al., 2017) to simulate inductive tasks and generate adaptive embeddings. HART (Yin et al., 2025) combines hypergraph GNNs and Transformers with a role-aware mechanism to mine complex subgraph semantics for inductive prediction.

3.4.3 Few-shot Setting

Both few-shot and inductive settings address unseen elements, but few-shot learning focuses on scenarios with very limited examples rather than none. To predict links involving sparse relations, HANCL (Zhang et al., 2022b) leverages GNNs and attention mechanisms to enhance entity representations and match queries to limited support instances. MetaRH (Wei et al., 2024) applies meta-learning to refine relation representations and improve generalization in few-shot settings.

3.5 Comparison of link prediction methods for NKGs

To intuitively compare the characteristics of the three types of methods, Table 1 shows the modeling idea, advantages, and drawbacks of each type of method. The spatial mapping-based methods have high computational efficiency, a small number of parameters, and are good at encoding large NKGs. However, due to the simple modeling idea, it has a great disadvantage in dealing with complex relationship types, and the model effect is usually poor. The tensor decomposition-based methods have a strong model expression ability in theory because their basic idea is to fully model the information contained in the NKG. However, it usually has more parameters, which is not conducive to application in large-scale NKGs. The neural network-based methods have strong feature learning ability, but the model is weak in interpretability. To intuitively compare the computational efficiency of the methods, Table 2 shows the time and space complexity of several representative link prediction methods.

3.6 Model Selection Guidance

In practical scenarios, model selection should balance predictive accuracy, scalability, and domain requirements. For high accuracy and complex reasoning, GNN-based models like HAHE are recommended, especially when relational structure and role semantics are critical. When computational resources are limited, simpler neural models (e.g., NeuInfer or HINGE) offer a trade-off between performance and efficiency. Tensor decomposition-based models such as S2S are suitable for scenarios that require high structural interpretability. Spatial mapping-based embedding methods are generally not recommended unless extreme efficiency is needed. Due to space limitations, the comparison of each type of method is shown in Appendix 3.5.

4 Performance of Existing Methods

This section presents the benchmarks, evaluation metrics, and the performance of existing methods. Due to space constraints, we briefly report results in general scenarios here; results in special scenarios are provided in Appendix F.

4.1 Benchmarks

JF17K (Wen et al., 2016), WikiPeople (Guan et al., 2019), and WD50K (Galkin et al., 2020) are the most commonly used benchmarks for evaluating link prediction methods for NKGs in general settings. Specifically, JF17K is derived from Freebase (Bollacker et al., 2008), while WikiPeople and WD50K are based on Wikidata (Vrandečić and Krötzsch, 2014). Table 3 shows the basic statistics of these datasets, where #X is the number of X in the dataset, E and R represent entities and roles, respectively, Arity represents the number of entities in an n-ary fact, that is, N is the proportion of n-ary facts, respectively.

Table 3: Benchmarks of link prediction in NKGs in general scenarios. Statistics are based on the original paper.

Dataset	#E	#R	Arity	N	#Facts
JF17K	28,645	501	2-6	45.9%	100,947
WikiPeople	47,765	193	2-9	11.6%	382,229
WD50K	47,155	531	2-67	13.6%	236,507

4.2 Metrics

To evaluate link prediction, the model ranks candidate answers by score, aiming to rank the correct answer as high as possible. The two most widely used metrics are Mean Reciprocal Rank (MRR) and Hits@K. MRR is the average inverse rank of correct answers, while Hits@K is the proportion of correct answers ranked within the top K positions (e.g., $K \in \{1, 3, 5, 10\}$). Both range from 0 to 1, with higher values indicating better performance.

4.3 Results

Table 4 reports the results of several representative link prediction methods on JF17K, WikiPeople, and WD50K, with best scores in bold. Neural network-based methods consistently outperform others, with HAHE (GNN-based) achieving the best result across all datasets, demonstrating the strength of GNNs in capturing complex entity interactions. In contrast, spatial mapping-based methods (m-TransH, RAE) perform worst, suggesting that simple projections are inadequate for n-ary fact modeling. Tensor decomposition-based methods (HypE, S2S) show moderate performance, but still fall short compared to neural network-based methods.

5 Applications of Link Prediction in NKGs

Due to its ability to represent complex semantic relationships among multiple entities, link prediction in NKGs has shown great potential in various domains, including biomedicine, recommender systems, and financial technology. In biomedicine, it enables the modeling of intricate relationships such as drug-target-disease interactions, supporting applications like drug repositioning and personalized treatment planning. In recommender systems, NKG captures rich contextual signals (e.g., time, location, device), offering fine-grained user behavior modeling and improving recommendation accuracy. In financial technology, NKG supports the structured representation of multi-role financial facts, facilitating tasks like risk inference and high-risk case detection. Detailed examples and case studies are provided in Appendix G.

6 Future Prospects

Link prediction in NKGs has made significant progress but remains a nascent field. Further research is needed in the following directions.

6.1 Link Prediction in NKGs with LLMs

LLMs have shown strong capabilities in both natural language understanding and structured data processing(Zhang et al., 2025; Liu et al., 2024; Zeng et al., 2025; Zhou et al., 2024; Yi et al., 2025; Wang et al., 2024a; Liang et al., 2024b; Yang et al., 2025). Recent work leverages this by converting KG elements into text and employing LLMs for link prediction. For example, KG-LLaMA (Yao et al., 2023) frames triples as textual sequences

Table 4: Link prediction results in NKGs in general scenarios. "-" indicates that the method was not evaluated on the corresponding dataset in the original paper. Data for HypE and S2S are from (Di et al., 2021), others from (Luo et al., 2023b).

Type	Method	JF17K		WikiPeople		WD50K	
		MRR	Hits@1	MRR	Hits@1	MRR	Hits@1
spatial mapping-based	m-TransH	0.102	0.069	-	-	-	-
	RAE	0.310	0.219	0.172	0.102	-	-
Tensor Decomposition-based	HypE	0.494	0.408	0.292	0.162	-	-
	S2S	0.528	0.457	0.372	0.277	-	-
Neural Network-based	HINGE	0.473	0.397	0.333	0.259	-	-
	NeuInfer	0.517	0.436	0.350	0.282	0.232	0.164
	GRAN	0.656	0.582	0.479	0.410	0.309	0.240
	HAHE	**0.668**	**0.597**	**0.495**	**0.420**	**0.402**	**0.327**

and fine-tunes LLaMA (Touvron et al., 2023) to learn KG knowledge directly. KICGPT (Wei et al., 2023) addresses the hallucination issue in LLMs via a hybrid reranking mechanism that uses LLMs to refine candidates from traditional models. Sehwag et al.(2024) treat link prediction as prompt-based question answering, incorporating structural features like entity neighbors. CTLP(Li et al., 2024b) employs structural paths between head and tail entities to conduct zero-shot link prediction. Similarly, KoPA (Zhang et al., 2024b) converts KG structures, such as local subgraphs and relational context, into extended prompts, allowing LLMs to implicitly learn structural patterns through language. MuKDC (Li et al., 2024c) addresses few-shot learning challenges by prompting LLMs to generate synthetic structured knowledge, effectively augmenting training with high-quality pseudo-samples.

Despite these advances, to the best of our knowledge, LLMs have not been applied to link prediction in NKGs, likely due to two main challenges: (1) converting structured n-ary facts into formats compatible with LLMs, and (2) overcoming input length limitations that hinder simultaneous processing of all candidate entities. Addressing these challenges could open new avenues for LLM-based link prediction in NKGs.

6.2 Link Prediction in NKGs in Special Scenarios

Most link prediction methods for NKGs focus on general scenarios. Research on special scenarios—temporal, inductive, and few-shot—remains in its early stages, offering ample opportunities for further exploration. For example, existing methods in temporal scenarios (Hou et al., 2023) overlook the local structure of n-ary facts, and existing methods in few-shot scenarios (Zhang et al., 2022b; Wei et al., 2024) require extensive few-shot tasks for training, which are difficult to construct in real-world applications. Additionally, real-world NKGs are dynamic, frequently incorporating new facts. It is crucial to develop methods for growing NKGs that can adaptively learn from new facts while retaining previously acquired knowledge.

6.3 Expainable Link Prediction in NKGs

To our knowledge, HyperMLN (Chen et al., 2022) is the only method that explicitly addresses explainability in link prediction in NKGs. It employs a random field model to capture dependencies among n-ary facts and improves interpretability by extracting predefined first-order logic rules (e.g., self-inverse, symmetric, subrelation). However, its focus on predefined rule types limits its ability to explain more complex relational combinations frequently observed in n-ary facts. Explainability should move beyond rule extraction to encompass broader interpretable reasoning, such as causal attribution (Jaimini et al., 2024) and counterfactual analysis (Zhao et al., 2022). Future work should pursue more flexible and comprehensive approaches to improve the transparency and reliability of link prediction in NKGs.

7 Conclusion

Link prediction in NKGs has emerged as a significant research area. In this survey, we provided the first comprehensive overview of existing work in this field. We began by introducing the definitions of NKGs and link prediction tasks within them, followed by a classification of current methods based on underlying techniques and application scenarios. Subsequently, we reported the performance and applications of existing methods. Finally, we outlined several promising future directions for advancing link prediction in NKGs.

Limitations

This study presents a comprehensive overview of recent advances in link prediction in NKGs, covering a wide range of modeling paradigms and application scenarios. However, the current version primarily focuses on high-level comparisons of different approaches, with limited discussion on practical aspects such as computational efficiency, scalability to large-scale NKGs, and robustness to noise. In addition, while the survey categorizes methods based on modeling techniques and special settings, it does not deeply analyze cross-scenario concerns such as interpretability and generalization to out-of-distribution data. Moreover, the evaluation is based on a few widely used benchmarks, which may not fully reflect the challenges present in real-world applications. Future work could incorporate more detailed empirical analyses and consider broader deployment factors to offer a more holistic assessment of link prediction methods for NKGs.

8 Acknowledgments

The work is supported by the Strategic Priority Research Program of the CAS under Grant No. XDB0680102, the National Natural Science Foundation of China under Grant No. 62441229, and the Innovation Funding of ICT, CAS under Grant No. E561010. We thank anonymous reviewers for their insightful comments and suggestions.

9 Ethics Statement

This paper presents a comprehensive review of existing methods and research progress on link prediction in NKGs. It does not involve any experiments with human subjects, personal data, or potentially harmful content. All datasets and methods discussed are publicly available and widely used in academic research. Therefore, this work does not pose any ethical concerns.

References

Mehdi Ali, Max Berrendorf, Mikhail Galkin, Veronika Thost, Tengfei Ma, Volker Tresp, and Jens Lehmann. 2021. Improving inductive link prediction using hyper-relational facts. In *ISWC*.

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: a collaboratively created graph database for structuring human knowledge. In *SIGMOD*.

Rebecca Braken, Alexander Paulus, André Pomp, and Tobias Meisen. 2023. An evaluation of link prediction approaches in few-shot scenarios. *Electronics*.

Dan Brickley, Ramanathan V Guha, and Brian McBride. 2014. Rdf schema 1.1. *W3C recommendation*, page 10.

Borui Cai, Yong Xiang, Longxiang Gao, He Zhang, Yunfeng Li, and Jianxin Li. 2023. Temporal knowledge graph completion: a survey. In *IJCAI*.

Jiahang Cao, Jinyuan Fang, Zaiqiao Meng, and Shangsong Liang. 2024. Knowledge graph embedding: A survey from the perspective of representation spaces. *ACM Computing Surveys*, 56(6):1–42.

Debyani Chakravarty, Jianjiong Gao, Sarah Phillips, Ritika Kundra, Hongxin Zhang, Jiaojiao Wang, Julia E Rudolph, Rona Yaeger, Tara Soumerai, Moriah H Nissan, et al. 2017. Oncokb: a precision oncology knowledge base. *JCO precision oncology*, 1:1–16.

Chen Chen, Yufei Wang, Aixin Sun, Bing Li, and Kwok-Yan Lam. 2023a. Dipping plms sauce: Bridging structure and text for effective knowledge graph completion via conditional soft prompting. *arXiv preprint arXiv:2307.01709*.

Jiaoyan Chen, Yuxia Geng, Zhuo Chen, Ian Horrocks, Jeff Z Pan, and Huajun Chen. 2021. Knowledge-aware zero-shot learning: Survey and perspective. *arXiv:2103.00070*.

Jiaoyan Chen, Yuxia Geng, Zhuo Chen, Jeff Z Pan, Yuan He, Wen Zhang, Ian Horrocks, and Huajun Chen. 2023b. Zero-shot and few-shot learning with knowledge graphs: A comprehensive survey. *Proceedings of the IEEE*.

Mingyang Chen, Wen Zhang, Yuxia Geng, Zezhong Xu, Jeff Z Pan, and Huajun Chen. 2023c. Generalizing to unseen elements: a survey on knowledge extrapolation for knowledge graphs. In *IJCAI*.

Mingyang Chen, Wen Zhang, Wei Zhang, Qiang Chen, and Huajun Chen. 2019. Meta relational learning for few-shot link prediction in knowledge graphs. 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)*, pages 4217–4226.

Sulin Chen and Jingbin Wang. 2022. A survey on temporal knowledge graphs-extrapolation and interpolation tasks. In *The International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*.

Xiaojun Chen, Shengbin Jia, and Yang Xiang. 2020a. A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications*.

Yong Chen, Xinkai Ge, Shengli Yang, Linmei Hu, Jie Li, and Jinwen Zhang. 2023d. A survey on multimodal knowledge graphs: Construction, completion and applications. *Mathematics*.

Zhe Chen, Yuehan Wang, Bin Zhao, Jing Cheng, Xin Zhao, and Zongtao Duan. 2020b. Knowledge graph completion: A review. *Ieee Access*.

Zirui Chen, Xin Wang, Chenxu Wang, and Jianxin Li. 2022. Explainable link prediction in knowledge hypergraphs. In *CIKM*.

Chanyoung Chung, Jaejun Lee, and Joyce Jiyoung Whang. 2023. Representation learning on hyper-relational and numeric knowledge graphs with transformers. In *SIGKDD*.

UniProt Consortium. 2019. Uniprot: a worldwide hub of protein knowledge. *Nucleic acids research*, 47(D1):D506–D515.

Yuanfei Dai, Shiping Wang, Neal N Xiong, and Wenzhong Guo. 2020. A survey on knowledge graph embedding: Approaches, applications and benchmarks. *Electronics*.

Shimin Di and Lei Chen. 2023. Message function search for knowledge graph embedding. In *WWW*.

Shimin Di, Quanming Yao, and Lei Chen. 2021. Searching to sparsify tensor decomposition for n-ary relational data. In *WWW*.

Zifeng Ding, Jingcheng Wu, Jingpei Wu, Yan Xia, and Volker Tresp. 2023. Exploring link prediction over hyper-relational temporal knowledge graphs enhanced with time-invariant relational knowledge. *arXiv:2307.10219*.

Felix Engel, Mark Vanin, and Nenad Krdzavac. 2024. Anticipate risk with the value and trade flows knowledge graph. *ESWC*.

Bahare Fatemi, Perouz Taslakian, David Vazquez, and David Poole. 2021. Knowledge hypergraphs: prediction beyond binary relations. In *IJCAI*.

Ilaria Ferrari, Giacomo Frisoni, Paolo Italiani, Gianluca Moro, and Claudio Sartori. 2022. Comprehensive analysis of knowledge graph embedding techniques benchmarked on link prediction. *Electronics*.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *ICML*.

Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck, and Jens Lehmann. 2020. Message passing for hyper-relational knowledge graphs. *arXiv:2009.10847*.

Xiou Ge, Yun Cheng Wang, Bin Wang, C-C Jay Kuo, et al. 2024. Knowledge graph embedding: An overview. *APSIPA Transactions on Signal and Information Processing*, 13(1).

Rishab Goel, Seyed Mehran Kazemi, Marcus Brubaker, and Pascal Poupart. 2020. Diachronic embedding for temporal knowledge graph completion. In *Proceedings of the AAAI conference on artificial intelligence*, pages 3988–3995.

Malachi Griffith, Nicholas C Spies, Kilannin Krysiak, Joshua F McMichael, Adam C Coffman, Arpad M Danos, Benjamin J Ainscough, Cody A Ramirez, Damian T Rieke, Lynzey Kujan, et al. 2017. Civic is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nature genetics*, 49(2):170–174.

Saiping Guan, Xueqi Cheng, Long Bai, Fujun Zhang, Zixuan Li, Yutao Zeng, Xiaolong Jin, and Jiafeng Guo. 2022. What is event knowledge graph: A survey. *TKDE*.

Saiping Guan, Xiaolong Jin, Jiafeng Guo, Yuanzhuo Wang, and Xueqi Cheng. 2020. Neuinfer: Knowledge inference on n-ary facts. In *ACL*.

Saiping Guan, Xiaolong Jin, Jiafeng Guo, Yuanzhuo Wang, and Xueqi Cheng. 2021. Link prediction on n-ary relational data based on relatedness evaluation. *TKDE*.

Saiping Guan, Xiaolong Jin, Yuanzhuo Wang, and Xueqi Cheng. 2019. Link prediction on n-ary relational data. In *WWW*.

Zhen Han, Peng Chen, Yunpu Ma, and Volker Tresp. Explainable subgraph reasoning for forecasting on temporal knowledge graphs. In *International conference on learning representations*.

Zhongni Hou, Xiaolong Jin, Zixuan Li, Long Bai, Saiping Guan, Yutao Zeng, Jiafeng Guo, and Xueqi Cheng. 2023. Temporal knowledge graph reasoning based on n-tuple modeling. In *EMNLP*.

Zhiwei Hu, Víctor Gutiérrez-Basulto, Zhiliang Xiang, Ru Li, and Jeff Z Pan. 2023. Hyperformer: Enhancing entity and relation interaction for hyper-relational knowledge graph completion. In *CIKM*.

Zhiwei Hu, Víctor Gutiérrez-Basulto, Zhiliang Xiang, Ru Li, and Jeff Z Pan. 2024. Hypermono: A monotonicity-aware approach to hyper-relational knowledge representation. *arXiv:2404.09848*.

Nicolas Hubert, Pierre Monnin, and Heiko Paulheim. 2023. Beyond transduction: A survey on inductive, few shot, and zero shot link prediction in knowledge graphs. *arXiv:2312.04997*.

Utkarshani Jaimini, Cory Henson, and Amit P Sheth. 2024. Causallp: Learning causal relations with weighted knowledge graph link prediction. *arXiv preprint arXiv:2405.02327*.

Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and S Yu Philip. 2021. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems*.

Robin Jia, Cliff Wong, and Hoifung Poon. 2019. Document-level n-ary relation extraction with multi-scale representation learning. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3693–3704.

Xuhui Jiang, Chengjin Xu, Yinghan Shen, Xun Sun, Lumingyuan Tang, Saizhuo Wang, Zhongwu Chen, Yuanzhuo Wang, and Jian Guo. 2023. On the evolution of knowledge graphs: A survey and perspective. *arXiv preprint arXiv:2310.04835*.

Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia. *Semantic web*.

Jake Lever, Martin R Jones, Arpad M Danos, Kilannin Krysiak, Melika Bonakdar, Jasleen K Grewal, Luka Culibrk, Obi L Griffith, Malachi Griffith, and Steven JM Jones. 2019. Text-mining clinically relevant cancer biomarkers for curation into the civic database. *Genome medicine*, 11(1):78.

Lijie Li, Hui Wang, Jiahang Li, Xiaodi Xu, Ye Wang, and Tao Ren. 2024a. Integrating structure and text for enhancing hyper-relational knowledge graph representation via structure soft prompt tuning. In *CIKM*.

Mingchen Li, Chen Ling, Rui Zhang, and Liang Zhao. 2024b. Zero-shot link prediction in knowledge graphs with large language models. In *2024 IEEE International Conference on Data Mining (ICDM)*, pages 753–760. IEEE.

Qian Li, Zhuo Chen, Cheng Ji, Shiqi Jiang, and Jianxin Li. 2024c. Llm-based multi-level knowledge generation for few-shot knowledge graph completion. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, volume 271494703.

Yujia Li, Shiliang Sun, and Jing Zhao. 2022a. Tirgn: Time-guided recurrent graph network with local-global historical patterns for temporal knowledge graph reasoning. In *IJCAI*, pages 2152–2158.

Zhao Li, Chenxu Wang, Xin Wang, Zirui Chen, and Jianxin Li. 2024d. Hje: joint convolutional representation learning for knowledge hypergraph completion. *TKDE*.

Zhao Li, Xin Wang, Jianxin Li, Wenbin Guo, and Jun Zhao. 2024e. Hycube: Efficient knowledge hypergraph 3d circular convolutional embedding. *arXiv:2402.08961*.

Zixuan Li, Saiping Guan, Xiaolong Jin, Weihua Peng, Yajuan Lyu, Yong Zhu, Long Bai, Wei Li, Jiafeng Guo, and Xueqi Cheng. 2022b. Complex evolutional pattern learning for temporal knowledge graph reasoning. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 290–296.

Ke Liang, Lingyuan Meng, Meng Liu, Yue Liu, Wenxuan Tu, Siwei Wang, Sihang Zhou, Xinwang Liu, Fuchun Sun, and Kunlun He. 2024a. A survey of knowledge graph reasoning on graph types: Static, dynamic, and multi-modal. *TPAMI*.

Shuang Liang, Jie Shao, Jiasheng Zhang, and Bin Cui. 2023. Graph-based non-sampling for knowledge graph enhanced recommendation. *TKDE*.

Xuechen Liang, Yangfan He, Meiling Tao, Yinghui Xia, Jianhui Wang, Tianyu Shi, Jun Wang, and Jing-Song Yang. 2024b. Cmat: A multi-agent collaboration tuning framework for enhancing small language models. *arXiv preprint arXiv:2404.01663*.

Wanlong Liu, Junying Chen, Ke Ji, Li Zhou, Wenyu Chen, and Benyou Wang. 2024. Rag-instruct: Boosting llms with diverse retrieval-augmented instructions. *arXiv preprint arXiv:2501.00353*.

Yi Liu, Hongrui Xuan, Bohan Li, Meng Wang, Tong Chen, and Hongzhi Yin. 2023. Self-supervised dynamic hypergraph recommendation based on hyper-relational knowledge graph. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 1617–1626.

Yu Liu, Quanming Yao, and Yong Li. 2020. Generalizing tensor decomposition for n-ary relational knowledge bases. In *WWW*.

Yu Liu, Quanming Yao, and Yong Li. 2021. Role-aware modeling for n-ary relational knowledge bases. In *WWW*.

Yuhuan Lu, Bangchao Deng, Weijian Yu, and Dingqi Yang. 2023a. Helios: Hyper-relational schema modeling from knowledge graphs. In *MM*.

Yuhuan Lu, Dingqi Yang, Pengyang Wang, Paolo Rosso, and Philippe Cudre-Mauroux. 2023b. Schema-aware hyper-relational knowledge graph embeddings for link prediction. *TKDE*.

Yuhuan Lu, Weijian Yu, Xin Jing, and Dingqi Yang. 2024. Hypercl: A contrastive learning framework for hyper-relational knowledge graph embedding with hierarchical ontology. In *ACL, findings*.

Haoran Luo, E Haihong, Ling Tan, Gengxian Zhou, Tianyu Yao, and Kaiyang Wan. 2023a. Dhge: dual-view hyper-relational knowledge graph embedding for link prediction and entity typing. In *AAAI*.

Haoran Luo, Yuhao Yang, Yikai Guo, Mingzhi Sun, Tianyu Yao, Zichen Tang, Kaiyang Wan, Meina Song, Wei Lin, et al. 2023b. Hahe: Hierarchical attention for hyper-relational knowledge graphs in global and local level. *arXiv:2305.06588*.

Haodi Ma and Daisy Zhe Wang. 2023. A survey on few-shot knowledge graph completion with structural and commonsense knowledge. *arXiv:2301.01172*.

Fanshen Meng, Zhenhua Meng, Ru Jin, Rongheng Lin, and Budan Wu. 2025. Doge: Llms-enhanced hyper-knowledge graph recommender for multimodal recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 12399–12407.

Siyuan Meng, Jie Zhou, XuXin Chen, Yufei Liu, Fengyuan Lu, and Xinli Huang. Structure-information-based reasoning over the knowledge graph: A survey of methods and applications. *TKDD*.

Lars-Peter Meyer, Claus Stadler, Johannes Frey, Norman Radtke, Kurt Junghanns, Roy Meissner, Gordian Dziwis, Kirill Bulert, and Michael Martin. 2023. Llm-assisted knowledge graph engineering: Experiments with chatgpt. In *Working conference on artificial intelligence development for a resilient and sustainable tomorrow*, pages 103–115. Springer Fachmedien Wiesbaden Wiesbaden.

Dat Quoc Nguyen. 2020. 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)*.

Maria José Nobre. 1986. *Inner speech as the basis for artistic conceptualization: Soviet psycholinguistics and semiotics of art*. The Ohio State University.

Jinghui Peng, Xinyu Hu, Wenbo Huang, and Jian Yang. 2023. What is a multi-modal knowledge graph: A survey. *Big Data Research*.

Manita Pote. 2024. Survey on embedding models for knowledge graph and its applications. *arXiv preprint arXiv:2404.09167*.

Andrea Rossi, Denilson Barbosa, Donatella Firmani, Antonio Matinata, and Paolo Merialdo. 2021. Knowledge graph embedding for link prediction: A comparative analysis. *TKDD*.

Paolo Rosso, Dingqi Yang, and Philippe Cudré-Mauroux. 2020. Beyond triplets: hyper-relational knowledge graph embedding for link prediction. In *WWW*.

Paolo Rosso, Dingqi Yang, Natalia Ostapuk, and Philippe Cudré-Mauroux. 2021. Reta: A schema-aware, end-to-end solution for instance completion in knowledge graphs. In *WWW*.

Udari Madhushani Sehwag, Kassiani Papasotiriou, Jared Vann, and Sumitra Ganesh. 2024. In-context learning with topological information for llm-based knowledge graph completion. In *ICML 2024 Workshop on Structured Probabilistic Inference {\&} Generative Modeling*.

Tong Shen, Fu Zhang, and Jingwei Cheng. 2022. A comprehensive overview of knowledge graph completion. *KBS*.

Jiawei Sheng, Shu Guo, Zhenyu Chen, Juwei Yue, Lihong Wang, Tingwen Liu, and Hongbo Xu. 2020. Adaptive attentional network for few-shot knowledge graph completion. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1681–1691.

Harry Shomer, Wei Jin, Juanhui Li, Yao Ma, and Jiliang Tang. 2022. Learning representations for hyper-relational knowledge graphs. *arXiv*.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv:2302.13971*.

Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, and Partha Talukdar. 2019. Composition-based multi-relational graph convolutional networks. *arXiv:1911.03082*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *NIPS*.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Communications of the ACM*.

Chenxu Wang, Zhao Li, Xin Wang, and Zirui Chen. 2023a. Enhance: Enhanced entity and relation embedding for knowledge hypergraph link prediction. In *WWW*.

Chenxu Wang, Xin Wang, Zhao Li, Zirui Chen, and Jianxin Li. 2023b. Hyconve: A novel embedding model for knowledge hypergraph link prediction with convolutional neural networks. In *WWW*.

Jiapu Wang, Boyue Wang, Meikang Qiu, Shirui Pan, Bo Xiong, Heng Liu, Linhao Luo, Tengfei Liu, Yongli Hu, Baocai Yin, et al. 2023c. A survey on temporal knowledge graph completion: Taxonomy, progress, and prospects. *arXiv:2308.02457*.

Junqiao Wang, Zeng Zhang, Yangfan He, Zihao Zhang, Yuyang Song, Tianyu Shi, Yuchen Li, Hengyuan Xu, Kunyu Wu, Xin Yi, et al. 2024a. Enhancing code llms with reinforcement learning in code generation: A survey. *arXiv preprint arXiv:2412.20367*.

Meihong Wang, Linling Qiu, and Xiaoli Wang. 2021a. A survey on knowledge graph embeddings for link prediction. *Symmetry*.

Peijie Wang, Jianrui Chen, Lide Su, and Zhihui Wang. 2023d. N-ary relation prediction based on knowledge graphs with important entity detection. *Expert Systems with Applications*, 221:119755.

Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. 2017. Knowledge graph embedding: A survey of approaches and applications. *TKDE*.

Quan Wang, Haifeng Wang, Yajuan Lyu, and Yong Zhu. 2021b. Link prediction on n-ary relational facts: A graph-based approach. *arXiv:2105.08476*.

Weiguang Wang, Xuanyi Zhang, Juan Zhang, Wei Cai, Haiyan Zhao, and Xia Zhang. 2024b. Mhre: Multivariate link prediction method for medical hyper-relational facts. *Applied Intelligence*.

Wenqi Wang, Yifan Sun, Brian Eriksson, Wenlin Wang, and Vaneet Aggarwal. 2018. Wide compression: Tensor ring nets. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9329–9338.

Jiyao Wei, Saiping Guan, Xiaolong Jin, Jiafeng Guo, and Xueqi Cheng. 2024. Few-shot link prediction on n-ary facts. *COLING*.

Jiyao Wei, Saiping Guan, Xiaolong Jin, Jiafeng Guo, and Xueqi Cheng. 2025. Inductive link prediction in n-ary knowledge graphs. In *Proceedings of the 31st International Conference on Computational Linguistics*.

Yanbin Wei, Qiushi Huang, Yu Zhang, and James Kwok. 2023. Kicgpt: Large language model with knowledge in context for knowledge graph completion. In *EMNLP*.

Jianfeng Wen, Jianxin Li, Yongyi Mao, Shini Chen, and Richong Zhang. 2016. On the representation and embedding of knowledge bases beyond binary relations. In *IJCAI*.

Bo Xiong, Mojtaba Nayyer, Shirui Pan, and Steffen Staab. 2023. Shrinking embeddings for hyper-relational knowledge graphs. *arXiv:2306.02199*.

Shiyao Yan, Zequn Zhang, Xian Sun, Guangluan Xu, Li Jin, and Shuchao Li. 2022a. Hyper2: Hyperbolic embedding for hyper-relational link prediction. *Neurocomputing*.

Shiyao Yan, Zequn Zhang, Xian Sun, Guangluan Xu, Shuchao Li, Qing Liu, Nayu Liu, and Shensi Wang. 2022b. Polygone: Modeling n-ary relational data as gyro-polygons in hyperbolic space. In *AAAI*.

Shiyao Yan, Zequn Zhang, Guangluan Xu, Xian Sun, Shuchao Li, and Shensi Wang. 2022c. Modeling n-ary relational data as gyro-polygons with learnable gyro-centroid. *KBS*.

Minglai Yang, Ethan Huang, Liang Zhang, Mihai Surdeanu, William Wang, and Liangming Pan. 2025. How is llm reasoning distracted by irrelevant context? an analysis using a controlled benchmark. *arXiv preprint arXiv:2505.18761*.

Liang Yao, Jiazhen Peng, Chengsheng Mao, and Yuan Luo. 2023. Exploring large language models for knowledge graph completion. *arXiv:2308.13916*.

Zi Ye, Yogan Jaya Kumar, Goh Ong Sing, Fengyan Song, and Junsong Wang. 2022. A comprehensive survey of graph neural networks for knowledge graphs. *IEEE Access*.

Qiang Yi, Yangfan He, Jianhui Wang, Xinyuan Song, Shiyao Qian, Xinhang Yuan, Li Sun, Yi Xin, Jingqun Tang, Keqin Li, et al. 2025. Score: Story coherence and retrieval enhancement for ai narratives. *arXiv preprint arXiv:2503.23512*.

Gongzhu Yin, Hongli Zhang, Yuchen Yang, and Yi Luo. 2025. Inductive link prediction on n-ary relational facts via semantic hypergraph reasoning. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pages 1821–1832.

Donghan Yu and Yiming Yang. 2021. Improving hyper-relational knowledge graph completion. *arXiv:2104.08167*.

Weijian Yu, Jie Yang, and Dingqi Yang. 2024. Robust link prediction over noisy hyper-relational knowledge graphs via active learning. In *WWW*.

Mohamad Zamini, Hassan Reza, and Minou Rabiei. 2022. A review of knowledge graph completion. *Information*.

Shuang Zeng, Xinyuan Chang, Mengwei Xie, Xinran Liu, Yifan Bai, Zheng Pan, Mu Xu, and Xing Wei. 2025. Futuresightdrive: Thinking visually with spatio-temporal cot for autonomous driving. *arXiv preprint arXiv:2505.17685*.

Chuxu Zhang, Kaize Ding, Jundong Li, Xiangliang Zhang, Yanfang Ye, Nitesh V Chawla, and Huan Liu. 2022a. Few-shot learning on graphs. *arXiv:2203.09308*.

Chuxu Zhang, Huaxiu Yao, Chao Huang, Meng Jiang, Zhenhui Li, and Nitesh V Chawla. 2020. Few-shot knowledge graph completion. In *Proceedings of the AAAI conference on artificial intelligence*, pages 3041–3048.

Jixiao Zhang, Yongkang Li, Ruotong Zou, Jingyuan Zhang, Renhe Jiang, Zipei Fan, and Xuan Song. 2024a. Hyper-relational knowledge graph neural network for next poi recommendation. *World Wide Web*, 27(4):46.

Le Zhang, Bo Wang, Xipeng Qiu, Siva Reddy, and Aishwarya Agrawal. 2025. Rearank: Reasoning re-ranking agent via reinforcement learning. *arXiv preprint arXiv:2505.20046*.

Richong Zhang, Junpeng Li, Jiajie Mei, and Yongyi Mao. 2018. Scalable instance reconstruction in knowledge bases via relatedness affiliated embedding. In *WWW*.

Xuan Zhang, Xun Liang, Xiangping Zheng, Bo Wu, and Yuhui Guo. 2022b. When true becomes false: Few-shot link prediction beyond binary relations through mining false positive entities. In *MM*.

Yichi Zhang, Zhuo Chen, Lingbing Guo, Yajing Xu, Wen Zhang, and Huajun Chen. 2024b. Making large language models perform better in knowledge graph completion. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pages 233–242.

Huaxuan Zhao, Yueling Pan, and Feng Yang. 2020. Research on information extraction of technical documents and construction of domain knowledge graph. *Ieee Access*.

Tong Zhao, Gang Liu, Daheng Wang, Wenhao Yu, and Meng Jiang. 2022. Learning from counterfactual links for link prediction. In *International Conference on Machine Learning*, pages 26911–26926. PMLR.

Ziqi Zhou, Jingyue Zhang, Jingyuan Zhang, Yangfan He, Boyue Wang, Tianyu Shi, and Alaa Khamis. 2024. Human-centric reward optimization for reinforcement learning-based automated driving using large language models. *arXiv preprint arXiv:2405.04135*.

Xiangru Zhu, Zhixu Li, Xiaodan Wang, Xueyao Jiang, Penglei Sun, Xuwu Wang, Yanghua Xiao, and Nicholas Jing Yuan. 2022. Multi-modal knowledge graph construction and application: A survey. *TKDE*.

Xiaohan Zou. 2020. A survey on application of knowledge graph. In *Journal of Physics: Conference Series*.

A Limitations of Decomposition-Based Representations for N-ary Facts

To represent n-ary facts using binary relational formats, link prediction methods for traditional KGs often rely on decomposition strategies such as reification (Brickley et al., 2014) and star-to-clique (S2C) (Wen et al., 2016) transformation. These techniques convert each n-ary fact into multiple triples, enabling the use of standard KG embedding models originally designed for binary relations.

Reification introduces a virtual auxiliary entity to represent the original n-ary fact, and connects this auxiliary node to each participating entity through distinct role-specific relations. For instance, the 3-ary fact *meeting(UN, Geneva, 2023)* can be decomposed into the binary facts (m1, *organizer*, UN), (m1, *location*, Geneva), and (m1, *year*, 2023), where `m1` is the reified fact node. This approach preserves the full semantic structure of the original fact but increases graph complexity by introducing additional nodes and relations.

In contrast, the S2C approach avoids introducing auxiliary entities by directly connecting all participating entities in a fully connected subgraph (clique), assigning a specific relation to each entity pair. Using the same example, S2C may generate triples like (UN, *met_in*, Geneva), (UN, *met_on*, 2023), and (Geneva, *hosted_in*, 2023). This transformation simplifies the graph by keeping only the original entities, but it may obscure the unified semantic context of the original fact, making it harder for models to infer higher-level relations across multiple roles.

Overall, both decomposition strategies involve trade-offs: reification preserves relational integrity at the cost of increased structural complexity, while S2C maintains simplicity but risks losing contextual semantics essential for accurate reasoning.

B Positioning Our Work: A Survey Comparison on Link Prediction in NKGs

Table 5 provides a comprehensive comparison between our survey and existing surveys on link prediction in KGs. As shown, most existing surveys primarily focus on general KGs, temporal KGs, sparse KGs, or multi-modal KGs, and few of them systematically address link prediction in NKGs. Notably, while Shen et al. (2022) and Guan et al.

(2022) partially cover NKGs by introducing task definitions and a limited number of methods, they lack a comprehensive exploration of NKG-specific models, benchmarks, and applications. In contrast, our survey specifically targets NKGs, providing a thorough task definition, a systematic summary of nearly 50 methods, over 10 benchmarks, as well as discussions on applications and future directions. To the best of our knowledge, our survey is the first to comprehensively address the link prediction task in NKGs.

C Evaluation Criteria for Modeling Facts with NKGs

NKGs are a generalization and extension of traditional KGs, providing the advantage of more accurately representing complex facts involving multiple entities. When a fact involves three or more core entities, NKGs can be prioritized to better capture such complex structures. To further assess whether the knowledge in a given domain is particularly suitable for modeling with NKG, three evaluation dimensions can be considered:

1. **Multi-party Participation:** When a fact involves three or more entities that are semantically related, NKG is recommended to be used. The more participating entities involved in a fact, the stronger its multi-party participation. For example, the fact "A, B, C, and D are university classmates" has a stronger multi-party participation compared to "A, B, and C are university classmates," thus making NKG modeling more preferable for maintaining the completeness of the fact.

2. **Semantic Coupling:** If the entities within a fact are tightly semantically coupled and cannot be reasonably decomposed into independent binary relations without losing essential semantics, NKGs should be used. For example, in the fact "Student A received scholarship D at school C in year B", all elements collectively form an inseparable semantic whole. Decomposing it into multiple binary facts, such as (Student A, studies at, School C) and (Student A, received, Scholarship D), would fail to accurately capture the original semantics.

3. **Context Dependence:** This refers to facts whose validity depends on contextual conditions such as time, location, or state. These contextual elements are integral parts of the fact's semantics. In such cases, NKG is recommended, ensuring the completeness of contextual information. For instance, in the fact "Einstein received the Nobel Prize in Physics in 1921 in Switzerland," both "1921" and "Switzerland" are essential contextual components. Ignoring them during modeling would compromise the accuracy of the fact. Therefore, such context-dependent facts are better modeled uniformly using NKG.

Analyzing the above evaluation dimensions can support making choices between NKGs and traditional KGs, thereby improving the accuracy of downstream reasoning tasks.

D Fact Formalization

From the perspective of fact formalization methods, as previously discussed, current link prediction methods for NKGs can be broadly categorized into three types: hyperedge-based, role-value pair-based, and hyper-relational-based. The correspondence between these methods and their adopted fact representations is summarized in Table 6. Different formalization methods directly affect the design of link prediction models in NKGs.

For instance, hyperedge-based methods emphasize modeling the overall structural relationships among multiple entities within a fact and are adept at capturing complex interactions among entities. Role-value pair-based methods focus on role-centered modeling, which is effective in capturing the semantic influence of different roles on entities. Hyper-relational-based methods introduce modeling of entity importance, enabling a more precise reflection of the varying contributions and roles of entities within a fact.

Among these, hyper-relational formalization is the most widely applied in existing research. This approach not only possesses strong capability in modeling n-ary facts but also maintains compatibility with traditional KG triple facts, providing good flexibility and generalization ability. Consequently, it demonstrates strong adaptability and transferability across various real-world tasks.

Overall, when applying link prediction methods for NKGs to specific scenarios, researchers can flexibly select the most suitable fact formalization approach by referring to the comparative analysis in Figure 3 and Table 6, considering factors such

Table 5: Comparison of Existing Surveys on Link Prediction in KGs with Our Survey. "-" indicates that the survey does not introduce methods for link prediction in NKGs.

Surveys	Years	KG types	Contents related to link prediction in NKGs
Wang et al. (2017), Nguyen (2020), Dai et al. (2020), Ji et al. (2021), Zou (2020), Chen et al. (2020b), Chen et al. (2020a), Chen et al. (2021), Rossi et al. (2021), and Wang et al. (2021a)	Before 2021	General KG	-
Zamini et al. (2022) and Ye et al. (2022)	2022	General KG	-
Chen and Wang (2022)	2022	Temporal KG	-
Zhu et al. (2022)	2022	Multi-modal KG	-
Shen et al. (2022)	2022	General KG, temporal KG, and NKG	Contain task definition, 4 methods, and 2 benchmarks
Guan et al. (2022)	2022	General KG and NKG	Contain task definition and 4 methods
Liang et al. (2024a)	2022	General, sparse, and multi-modal KG	-
Ferrari et al. (2022) and Hubert et al. (2023)	2023	General KG	-
Zhang et al. (2022a), Braken et al. (2023), Ma and Wang (2023), Chen et al. (2023b), and Chen et al. (2023c)	2023	Sparse KG	-
Jiang et al. (2023)	2023	General, sparse, and temporal KG	-
Cai et al. (2023) and Wang et al. (2023c)	2023	Temporal KG	-
Peng et al. (2023) and Chen et al. (2023d)	2023	Multi-modal KG	-
Cao et al. (2024), Ge et al. (2024), Pote (2024), and Meng et al.	2024	General KG	-
Ours	2024	NKG	Contain task definition, nearly 50 methods, more than 10 benchmarks, applications, and future prospect

as task types, data characteristics, and reasoning requirements.

E Detailed Comparison between Link Prediction in NKGs and Link Prediction in Traditional KGs

Traditional KGs and NKGs share some commonalities in their basic components, such as entities, relations, and facts. Consequently, they both adopt similar technical approaches for link prediction, including spatial, tensor decomposition, and neural networks. However, significant differences exist in the structural characteristics of their modeling targets and the definitions of their prediction tasks. These differences necessitate specific extensions and optimizations in the structural design and modeling strategies of link prediction methods for NKGs. Specifically, the key differences between link prediction in traditional KGs and link prediction in NKGs are reflected in the following two aspects.

E.1 Modeling Object Structure

Traditional KG focuses on modeling triple facts (h, r, t), while NKG deals with more flexible and complex n-ary facts. For example, for hyper-relational facts in the form of $(h, r, t), \{r_i : v_i\}_{i=1}^{n}$, link prediction in NKGs requires the model to not only capture the relation r between the head entity h and the tail entity t, but also handle the correspondence between qualifier roles r_i and qualifier values v_i, as well as the interactions between these role-value pairs $\{r_i : v_i\}_{i=1}^{n}$ and the main triple (h, r, t). Such complex structures within n-ary facts impose higher requirements on the model's representation capability.

Table 6: Classification of link prediction methods for NKGs by Fact Formalization Approach

Fact Formalization	Methods	Advantages
Hyperedge-based	m-TransH (Wen et al., 2016), RAE (Zhang et al., 2018), m-SimplE (Fatemi et al., 2021), HypE (Fatemi et al., 2021), GETD (Liu et al., 2020), S2S (Di et al., 2021), RAM (Liu et al., 2021), EnhancE (Wang et al., 2023a), HyconvE (Wang et al., 2023b), HyCubE (Li et al., 2024e), and HJE (Li et al., 2024d)	Good at modeling the overall structural relationships among multiple entities in n-ary facts.
Role-value pair-based	PolygonE (Yan et al., 2022b), NaLP (Guan et al., 2019), t-NaLP (Guan et al., 2021), NE-NET (Hou et al., 2023), and HypeTKG (Hou et al., 2023)	Effectively captures the semantic roles of each entity within n-ary facts.
Hyper-relational-based	HYPER2 (Yan et al., 2022a), WPolygonE+(Yan et al., 2022c), HYPERMONO(Hu et al., 2024), NeuInfer (Guan et al., 2020), HINGE (Rosso et al., 2020), s-HINGE (Lu et al., 2023b), GRAN (Wang et al., 2021b), HyTransformer (Yu and Yang, 2021), HyNT (Chung et al., 2023), HIST (Wang et al., 2023d), NYLON (Wang et al., 2023d), StarE (Galkin et al., 2020), HAHE (Luo et al., 2023b), HyperFormer (Hu et al., 2023), QUAD (Shomer et al., 2022), DHGE (Luo et al., 2023a), HELIOS (Lu et al., 2023a), HyperCL (Chen et al., 2022), HANCL (Zhang et al., 2022b), and MetaRH (Wei et al., 2024)	Effectively distinguishes the importance differences among entities within n-ary facts.

E.2 Prediction Task Definition

In traditional KGs, link prediction mainly focuses on completing missing entities within triples, such as $(?, r, t)$, $(h, r, ?)$, or $(h, ?, t)$. In contrast, link prediction in NKGs is broader and more flexible, involving not only missing entities and relations but also missing qualifier roles and values, and often requires the simultaneous completion of multiple missing elements. This task setting demands that the model possesses stronger structural modeling capabilities and flexible reasoning mechanisms to handle various types of missing elements.

Based on these differences, link prediction methods for NKGs have been specifically improved to address the modeling challenges of NKGs. For instance:

- **Spatial Mapping-based Methods:** Introduce role-specific spatial transformation functions to capture semantic differences of entities under different role contexts.
- **Tensor Decomposition-based Methods:** Utilize tools such as Tucker decomposition to handle higher-dimensional and structurally complex tensors, and address challenges such as nested structures and variable numbers of entities.
- **Neural Network-based Methods:** Emphasize the modeling of unique structures in NKG, such as hypergraphs formed by inter-fact relations or fully connected graphs composed of elements within a fact.

Furthermore, in response to more complex prediction tasks, most current link prediction methods for NKGs support the completion of arbitrary missing elements, and some methods can even predict multiple missing elements simultaneously. For example, the HAHE method based on neural networks employs an autoregressive encoder combined with a MASK mechanism to effectively support the prediction of multiple missing elements within n-ary facts.

F More Details of Performance of Existing Methods

F.1 Benchmarks

F.1.1 Temporal Scenario

Hou et al. (Hou et al., 2023) constructed two datasets specifically for link prediction in NKGs in temporal scenarios: NWIKI and NICE. The NWIKI dataset is derived from Wikidata and contains a large number of n-ary facts, providing a rich foundation for link prediction in temporal NKGs. To ensure effective model training, they filtered out low-frequency entities and retained only high-frequency ones during the data construction process. In contrast, the NICE dataset is based on ICEWS, where the temporal information is more prominent, making it particularly suitable for tasks that require modeling dynamically evolving facts over time. These two datasets provide an important experimental foundation in the temporal NKG domain and have promoted research on n-ary fact modeling methods in temporal scenarios.

At the same time, Di et al. (Ding et al., 2023) extended the traditional temporal KGs into temporal NKGs by identifying qualifying role-value pairs from Wikidata within the existing Wikidata11k (Nobre, 1986) and YAGO1830 (Han et al.) datasets. The resulting datasets are named Wiki-hy and YAGO-hy, respectively. Table 7 summarizes the statistics of NWIKI, NICE, Wiki-hy, and YAGO-hy datasets, where Timestamps indicates the recorded time points of facts, and Time Interval refers to the minimum time interval between facts.

F.1.2 Few-shot Scenario

Zhang et al. (Zhang et al., 2022b) constructed three datasets for link prediction in NKGs in few-shot scenarios: WikiAnimals, WikiCompanies, and WD50K-Few. These datasets aim to simulate the learning challenges of rare relations in real-world settings and evaluate models' generalization abilities under limited data conditions. Specifically, WikiAnimals and WikiCompanies were derived from Wikidata by extracting facts related to animals and companies, respectively, while WD50K-Few was curated from a subset of the WD50K dataset. These datasets provide important experimental benchmarks for research on few-shot link prediction in NKGs.

Concurrently, Wei et al. (Wei et al., 2024) further developed the F-WikiPeople, F-JF17K, and F-WD50K datasets by extending WikiPeople, JF17K, and WD50K, respectively, into few-shot scenarios. These datasets cover various knowledge domains and exhibit distinctive data distributions and n-ary fact structures, further enriching the experimental foundation for few-shot link prediction in NKGs. Table 8 summarizes the key statistics of these benchmarks for few-shot link prediction in NKGs, where "E-q" and "R-q" respectively represent the number of entities and roles involved in qualifying role-value pairs, reflecting the complexity of n-ary facts in each dataset, and "Tasks" indicates the number of few-shot relation link prediction tasks.

F.1.3 Inductive Scenario

Ali et al. (Ali et al., 2021) constructed a series of datasets for inductive link prediction in NKGs based on WD50K, including multiple datasets under different settings to evaluate model generalization in inductive scenarios. This subsection focuses on the representative datasets WD20K(25), WD20K(100) V1, and WD20K(100) V2; for more details on other datasets, please refer to the original paper. In the inductive setting, these datasets typically include entities with textual descriptions or additional inference graphs (containing supporting instances related to unseen entities) to assist in generating representations for unseen entities during the testing phase. Specifically, WD20K(25) only provides textual descriptions without inference graphs containing unseen entities, requiring the model to rely solely on text features to complete inductive link prediction tasks. In contrast, WD20K(100) V1 and WD20K(100) V2 provide both textual descriptions and inference graphs, enabling models to leverage structural information to infer representations of unseen entities. Moreover, WD20K(100) V1 offers larger training data compared to WD20K(100) V2, allowing models to learn richer features during training. The number in parentheses in the dataset names indicates the proportion of hyper-relational facts, reflecting the diversity of facts across datasets. Furthermore, Wei et al. (Wei et al., 2024) constructed the JF-Ext, WIKI-Ext, and WD-Ext datasets, based on extensions of JF17K, WikiPeople, and WD50K, respectively, providing additional high-quality benchmarks for inductive link prediction in NKGs. These datasets cover different knowledge domains and vary in data scale, proportion of hyper-relational facts, and richness of entity information, thus offering a more comprehensive experimental benchmark for future research. Table 9 summarizes the statistics of these benchmarks for inductive link prediction in NKGs. The "Inference" column indicates the inference graph used during testing.

F.2 Metrics

During evaluation, the model assigns scores to all candidate answers and ranks them accordingly. A higher rank for the correct answer indicates better performance. Mean Reciprocal Rank (MRR) and Hits@K are the most commonly used evaluation metrics, which assess the model's link prediction capabilities from different perspectives.

Table 7: Baseline Datasets for Link Prediction in NKGs in Temporal Scenarios. All statistics are reported from the original papers. "Timestamps" indicates the number of recorded time points, and "Time Interval" refers to the minimum temporal resolution of the dataset. "N" denotes the proportion of n-ary facts in the dataset.

Dataset	#Entities	#Timestamps	Time Interval	N (%)	#Train	#Valid	#Test
NWIKI	17,481	205	1 year	81.9%	108,397	14,370	15,591
NICE	10,860	4,017	24 hours	97.5%	368,868	5,268	46,159
Wiki-hy	11,140	507	1 year	9.5%	111,252	13,900	13,926
YAGO-hy	10,026	188	1 year	6.9%	51,193	10,973	10,977

Table 8: Baseline Datasets for Few-shot Link Prediction in NKGs. All statistics are reported from the original papers. "E-q" and "R-q" indicate the number of entities and roles involved in qualifying role-value pairs, reflecting the complexity of n-ary facts in the dataset. "Tasks" refers to the number of few-shot relation link prediction tasks.

Dataset	#Entities	#Relations	#E-q	#R-q	N (%)	#Tasks	#Facts
WikiAnimals	2,925,278	167	396,739	48	49.7%	49	5,964,839
WikiCompanies	30,781	164	27,511	127	18.1%	125	1,128,040
WD50K-Few	47,156	532	5,460	45	13.6%	126	236,507
F-WikiPeople	40,529	237	4,663	75	9.0%	30	319,140
F-JF17K	19,721	480	4,928	127	47.6%	52	91,572
F-WD50K	43,802	697	10,242	85	13.1%	118	379,653

F.2.1 Mean Reciprocal Rank (MRR)

Mean Reciprocal Rank (MRR) is primarily used to evaluate the ranking quality of the correct answer among the prediction results, reflecting the overall link prediction performance of the model. The calculation formula is as follows:

$$\mathrm{MRR} = \frac{1}{|Q|} \sum_{q \in Q} \frac{1}{\mathrm{rank}_q}, \tag{1}$$

where Q denotes the set of queries, and rank_q is the rank of the correct answer in the sorted list for query q. The value of MRR ranges from 0 to 1, where a higher value indicates better prediction performance.

F.2.2 Hits@K

Hits@K calculates the proportion of queries where the correct answer is ranked within the top K, measuring the model's performance at different precision levels to meet specific application needs. The formula is defined as:

$$\mathrm{Hits@K} = \frac{|\{q \in Q : \mathrm{rank}_q \leq K\}|}{|Q|}. \tag{2}$$

Similar to MRR, Hits@K ranges from 0 to 1, with higher values indicating that the model is more capable of ranking the correct answer in the top positions. Common values of K include 1, 3, 5, and 10, corresponding to different requirements of prediction accuracy in various application scenarios.

F.3 Results

F.3.1 Temporal Scenario

NE-Net and HypeTKG are specifically designed for link prediction in NKGs in temporal scenarios. Table 10 presents their experimental results on the temporal datasets NWIKI and Wiki-hy, along with several representative baseline models. CEN (Li et al., 2022b), TiGRN (Li et al., 2022a), and DE-SimplE (Goel et al., 2020) are typical temporal link prediction methods for binary facts, capable of effectively modeling temporal information in KGs but without considering the qualifier role-value pairs in multi-fact settings. HINGE, RAM, and HyTransformer are representative methods for link prediction in NKGs but do not leverage temporal information in temporal NKGs. The results demonstrate that NE-Net and HypeTKG achieve the best performance on the two datasets, significantly outperforming both temporal binary fact modeling methods and non-temporal NKG modeling methods. For example, NE-Net achieves an MRR of 0.720 on the NWIKI dataset, considerably surpassing other methods, further verifying its effectiveness in temporal link prediction in NKGs. These results indicate that jointly modeling qualifier role-value pairs and temporal evolution can significantly enhance predictive capability in temporal link prediction in NKGs.

F.3.2 Few-shot Scenario

HANCL and MetaRH are specifically designed for few-shot link prediction in NKGs. Ta-

Table 9: Benchmark datasets for inductive link prediction in NKGs. Statistics are from the original papers.

Dataset	Train		Validation		Test		Inference	
	#Facts	**N**	**#Facts**	**N**	**#Facts**	**N**	**#Facts**	**N**
WD20K(25)	39,819	30.0%	4,252	25.0%	3,453	22.0%	0	0
WD20K(100) V1	7,785	100.0%	295	100.0%	364	100.0%	2,667	100.0%
WD20K(100) V2	4,146	100.0%	538	100.0%	678	100.0%	4,274	100.0%
JF-Ext	3,305	54.0%	1,061	30.0%	1,283	21.0%	5,012	28.1%
WIKI-Ext	3,905	2.1%	6,480	2.5%	4,733	2.9%	4,880	6.6%
WD-Ext	5,112	1.0%	2,610	3.6%	3,053	2.0%	3,382	6.1%

Table 10: Link prediction results in temporal scenarios. The results on NWIKI and Wiki-hy are from (Hou et al., 2023), (Ding et al., 2023).

Method	NWIKI			Wiki-hy		
	MRR	**Hits@1**	**Hits@10**	**MRR**	**Hits@1**	**Hits@10**
HINGE	0.217	0.191	0.259	0.543	0.497	0.694
HypE	0.252	0.249	0.257	0.624	0.604	0.658
CEN	0.406	0.302	0.610	-	-	-
DE-SimplE	0.138	0.108	0.191	0.351	0.218	0.640
TiGRN	0.611	0.506	0.811	-	-	-
NE-Net	**0.720**	**0.668**	**0.802**	-	-	-
HypeTKG	-	-	-	**0.693**	**0.642**	**0.792**

ble 11 presents their experimental results on the few-shot datasets WikiAnimals and F-WD50K. FSRL (Zhang et al., 2020), FAAN (Sheng et al., 2020), and MetaR (Chen et al., 2019) are representative few-shot link prediction methods for binary facts, capable of learning relational representations from limited data but neglecting qualifier role-value pairs in multi-fact settings. The results show that HANCL and MetaRH achieve the best performance on both datasets, significantly outperforming other methods. For instance, HANCL achieves an MRR of 0.318 on the WikiAnimals dataset, showing a substantial improvement over other approaches. These results suggest that traditional few-shot learning methods are insufficient for capturing knowledge in multi-fact settings, while incorporating qualifier role-value pairs can effectively enhance reasoning capability and significantly boost prediction performance.

F.3.3 Inductive Scenario

QBLP, MetaNIR, and HART are specifically designed for inductive link prediction in NKGs. Table 12 shows their experimental results on the WD20K(100) V1, WD20K(100) V2, and WD-Ext datasets. BLP (Chen et al., 2019) is a representative inductive link prediction method in the KG domain, which generates embeddings for unseen entities by encoding textual descriptions, while StarE and CompGCN (Vashishth et al., 2019) generate embeddings for unseen entities based on neighborhood information in reasoning graphs. Both BLP and CompGCN overlook qualifier role-value pairs in multi-fact settings. The results show that HART and MetaNIR achieve competitive results, demonstrating that leveraging qualifier role-value pairs in multi-fact settings is essential for inductive link prediction in NKGs.

G Details Applications of Link Prediction in NKGs

Due to their ability to represent complex semantic relationships among multiple entities, NKGs have been widely adopted in knowledge modeling and reasoning tasks across various domains, including biomedicine, recommender systems, and financial technology. This section provides a detailed discussion of how link prediction in NKGs is applied in these representative scenarios and highlights its practical value.

G.1 Biomedicine

In the biomedical domain, knowledge often involves complex multi-entity relationships, such as “a drug treats a disease by targeting a specific biomarker” or “a gene mutation causes a disease within a certain population.” These facts require modeling of tightly connected semantic entities.

NKGs enable more expressive and accurate knowledge representation by explicitly specifying the roles of each entity (e.g., drug, target, disease, population). On this basis, link prediction

Table 11: Link prediction results in few-shot scenarios. The results on WikiAnimals and F-WD50K are from (Zhang et al., 2022b), (Wei et al., 2024).

Method	WikiAnimals			F-WD50K		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
StarE	0.265	0.233	0.215	0.102	0.057	0.177
GRAN	0.253	0.199	0.221	0.126	0.077	0.222
FSRL	0.236	0.201	0.230	-	-	-
FAAN	0.270	0.225	0.246	0.116	0.059	0.226
MetaR	-	-	-	0.108	0.064	0.183
HANCL	**0.318**	**0.288**	**0.258**	-	-	-
MetaRH	-	-	-	**0.192**	**0.109**	**0.340**

Table 12: Link prediction results in inductive scenarios. The results on WD20K(100) V1 and V2 are from (Yin et al., 2025), and WD-Ext results are from (Wei et al., 2025).

Method	WD20K(100) V1			WD20K(100) V2			WD-Ext		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
BLP	0.057	0.019	0.123	0.039	0.014	0.092	-	-	-
CompGCN	0.104	0.057	0.183	0.025	0.007	0.053	-	-	-
StarE	0.112	0.061	0.212	0.049	0.019	0.110	0.079	0.021	0.131
QBLP	0.107	0.039	0.245	0.066	0.034	0.120	-	-	-
HART	**0.385**	**0.294**	**0.522**	**0.258**	**0.176**	**0.468**	-	-	-
MetaNIR	-	-	-	-	-	-	**0.582**	**0.433**	**0.901**

in NKGs facilitates the discovery of novel associations, such as between drug combinations and indications, thus supporting tasks like drug repositioning and the identification of combination therapies. For instance, Wang et al. (Wang et al., 2024b) constructed the MedCKG dataset from clinical data of China Medical University and applied link prediction in NKGs to assist in generating personalized treatment plans. Peng et al. (Lever et al., 2019) leveraged graph structures to detect potential n-ary facts in oncology knowledge bases. Moreover, MULTISCALE (Jia et al., 2019) has been applied to literature retrieval and fact completion tasks across oncology and protein translation datasets (Chakravarty et al., 2017; Griffith et al., 2017; Consortium, 2019), underscoring the broad potential of NKGs in real-world scientific domains.

G.2 Recommender Systems

User behaviors in recommender systems are influenced by various contextual factors, such as time, location, device, and behavior type. A typical user action can be described as "a user browses a product at a certain time, in a specific location, using a certain device." However, traditional KGs often model such behaviors as simple triples (e.g., (user, buy, item)), which fail to capture these contextual dependencies.

NKGs offer a more comprehensive representation by encoding multi-faceted facts (e.g., user, item, behavior type, time, location), enabling a finer-grained understanding of user behavior. Link prediction on NKGs can identify users' potential interests under specific contexts, thereby improving recommendation performance. For example, SDK (Liu et al., 2023) models such multi-entity interactions holistically and significantly improves the representation quality of users and items. Moreover, SDK demonstrates enhanced generalization in cold-start and sparse-data settings through multidimensional reasoning. DOGE (Meng et al., 2025) integrates LLM-based textual semantics with NKG structural semantics to enable multimodal recommendation. HKGNN (Zhang et al., 2024a) leverages self-attention mechanisms to capture access sequences. Collectively, these approaches highlight the advantages of NKGs in capturing user behavior semantics, alleviating data sparsity, and enhancing personalized recommendations.

G.3 Financial Technology

In financial technology, real-world facts often involve multiple components, such as "a bank issues a loan to a customer at a specific time for a particular project," or "an institution invests in an asset in a certain market while facing specific risks." These complex relationships are difficult to model using traditional KGs.

NKGs allow for accurate representation of each component and its role (e.g., lender, borrower, purpose, risk level), enabling comprehensive modeling of financial interactions. This supports advanced analysis tasks such as risk assessment and retrieval of similar historical cases, greatly enhancing automation in financial data analytics. For example, Hou et al. (Hou et al., 2023) constructed a financial NKG based on real-world transaction data and applied link prediction in NKGs to identify potentially high-risk loan cases, significantly improving the effectiveness of risk warning systems. Similarly, Engel et al.(Engel et al., 2024) modeled international trade data with n-ary facts to support complex analyses, demonstrating the value of NKGs in enhancing supply chain resilience and risk management.