Knowledge Graph-Augmented GNN Encoder with Transformer Decoder for Cross-Lingual Neural Machine Translation: Modeling, Optimization, and Scalable Deployment

Bo Wang
Language and Culture College, Xi'an FanYi University, Xi’an, 710105, China
E-mail：13619269005@163.com

Keywords: knowledge graph cross-language translation graph neural network semantic alignment performance optimization

Received: September 5, 2025

In cross-language information interaction, translation accuracy and efficiency determine the reliability of multilingual services. This paper proposes a knowledge graph-augmented neural machine translation framework that integrates a 3-layer Graph Attention Network (GAT) encoder with a 6-layer Transformer decoder, fused with XLM-R contextual embeddings. The fusion mechanism injects entity and relation information into decoding, enhancing semantic alignment and long-sentence reasoning. A distributed parallel architecture with cache optimization supports scalable deployment. Experiments on WMT and OpenSubtitles datasets evaluate the model against Transformer and NMT baselines. Results show that compared with the vanilla Transformer and mBART baselines, the proposed system achieves an average BLEU improvement of 17.0%±0.6, a 26.8%±1.5 reduction in PPL, inference time shortened to 0.92s±0.04, and entity alignment error rate reduced to 3.4%±0.3. These results confirm that the performance gains are statistically significant, with confidence intervals reported for all metrics. These findings confirm that knowledge graph augmentation substantially improves translation quality, semantic consistency, real-time performance, and robustness under multilingual and complex contexts. The contributions of this work include: ①a GAT-based encoder for capturing cross-lingual dependencies; ② a fusion method with XLM-R for semantic enhancement; ③ a scalable optimization framework ensuring low-latency translation. This research provides a reproducible and deployable approach for intelligent multilingual interaction and demonstrates significant potential for cross-language applications.

Povzetek: Razvit je večjezični prevajalski okvir, ki združuje znakovni graf znanja, grafno nevronsko mrežo in transformerski dekoder. Vključitev entitetnih in relacijskih informacij izboljša semantično poravnavo, prevajanje dolgih stavkov ter zmanjša zakasnitev in napake poravnave v večjezičnih okoljih.

1 Introduction

Against the backdrop of the ever-growing demand for global communication and multilingual information interaction, traditional machine translation systems, lacking deep semantic modeling and cross-language knowledge transfer mechanisms, find it difficult to cope with the challenges of complex contexts, multilingual parallelism, and real-time interaction. Facing problems such as semantic ambiguity, long sentence dependence, and rapid vocabulary updates in professional fields, existing translation models often experience semantic mismatches, context fragmentation, and performance degradation, which seriously restrict the quality and efficiency of cross-language communication and intelligent services. With the rise of scenarios such as cross-border e-commerce, international conferences, and multilingual information retrieval, cross-language translation systems urgently need to shift from static training and single-core-driven models to an intelligent architecture that enhances knowledge and optimizes dynamic feedback, in order to achieve semantic alignment, task-driven translation optimization, and stable operation across platforms.

As an important means of semantic modeling and knowledge management, knowledge graphs can establish a unified semantic space among different languages through the structured expression of entities, relations and semantic contexts. Unlike traditional dictionary-based contrastive translation, knowledge graphs possess cross-language semantic transfer and reasoning capabilities, which can provide explicit knowledge support for machine translation models and enhance their processing ability for long sentences and professional terms. Research shows that Zhao et al. (2024) proposed a knowledge graph-driven neural machine translation method, which significantly improved semantic consistency and entity translation accuracy in dynamic contexts, with the average BLEU value increasing by more than 4% in multilingual tasks [1]. Further research indicates that Srivastava et al. (2023) proposed a translation method that combines entity perception mechanisms with knowledge graph structures, achieving high-precision semantic alignment and reasoning in question-answering and translation tasks, and effectively enhancing translation quality in complex contexts [2].

In terms of model optimization and cross-language semantic transfer, the neural translation method enhanced by knowledge graphs demonstrates strong generalization and robustness. For instance, the cross-language neural translation model proposed by Zhao et al. (2020) significantly reduced the perplexity of long sentence translation after integrating the semantics of the knowledge graph, demonstrating the effectiveness of cross-language semantic transfer in complex translation tasks [3]. In addition, the flexible translation model based on knowledge graph embedding proposed by Li et al. (2022) provides new optimization ideas for translation path generation and semantic representation, further verifying the role of graph structure in improving translation performance [4].

This study proposes a knowledge graph-driven framework for cross-language intelligent translation, linking semantic modeling, translation reasoning, and performance optimization. The model consists of three mechanisms: cross-language knowledge graph construction, a GNN-enhanced translation engine with pre-trained models, and a performance optimization strategy for distributed deployment. The research addresses three questions:① Can knowledge graph embeddings reduce perplexity (PPL) by over 25% in multilingual tasks?② Does GNN-based dependency modeling improve BLEU and entity accuracy compared with Transformer baselines?③ Can cache-optimized deployment with feedback correction achieve real-time inference while maintaining cross-domain robustness? Experiments on WMT and OpenSubtitles validate these objectives, showing higher BLEU, lower latency, and stronger scalability than traditional frameworks. The results demonstrate advances in both architecture and optimization, providing an effective path for intelligent multilingual translation.

2 Relevant work

In the development of cross-language translation research, traditional neural machine translation systems have long relied on large-scale bilingual corpora and static parameter training. Although they have achieved high accuracy in monolingual alignment and translation of common languages, they have problems such as semantic fragmentation, context loss, and entity translation errors when facing multilingual dynamic environments, complex contexts, and the transfer of professional terms. Especially in scenarios such as long-sentence reasoning, multi-language mixed input, and cross-domain applications, traditional models lack deep expression of semantic relations and structured modeling of complex dependencies, making it difficult to support the demands of high-precision and low-latency cross-language translation. Hu et al. (2022) proposed the knowledge graph-enhanced multi-task recommendation model TransMKR, which verified the effectiveness of the translation mechanism in cross-task semantic modeling, providing an important reference for knowledge transfer in cross-language translation [5].

On the one hand, the existing neural machine translation models still mainly rely on sentence pair training and static corpora, making it difficult to express complex semantic links and cross-language knowledge transfer. Yu et al. (2022) proposed a translation representation method based on octonion embedding, which significantly improved semantic consistency and reasoning accuracy in the knowledge graph completion task [6], providing a new structured approach for cross-language semantic modeling. On the other hand, the translation logic within the system mostly adopts a static configuration approach, lacking a dynamic adaptation mechanism for cross-language tasks. Wan et al. (2023) proposed the knowledge representation learning model TransRFT based on relational neighborhood and flexible translation, which demonstrated strong generalization ability in cross-language semantic transfer and relational modeling [7]. In addition, Yueting et al. (2023) proposed a computer-aided translation method that combines BiLSTM and convolutional neural networks, and introduced knowledge graphs to enhance semantic modeling, effectively improving translation accuracy in complex contexts [8]. To provide a clearer comparison with prior approaches, this study benchmarks state-of-the-art translation models across key dimensions, including data dependency, model architecture, semantic modeling, and performance metrics such as BLEU, PPL, latency, and entity accuracy, as presented in Table 1.

Table 1: Benchmarking of state-of-the-art translation models (BLEU score, perplexity (PPL), latency in seconds, and entity accuracy in %)

Model	BLEU ↑	PPL ↓	Latency (s) ↓	Datasets Used	Languages Supported	Entity Accuracy (%) ↑
Transformer	27.5	45.2	1.28	WMT14 En-De	2	87.1
mBART	29.8	39.7	1.12	WMT, OPUS	25+	89.4
mT5	31.2	36.5	1.35	WMT, OpenSubtitles	100+	90.2
Proposed KG-GNN + Transformer	32.6	28.5	0.92	WMT, OpenSubtitles	50+	96.6

As shown in Table 1, traditional neural translation systems rely heavily on data and static configurations, limiting their capacity for deep cross-lingual semantic mapping. In contrast, the knowledge graph-driven system integrates multi-source data and semantic linkage to form a dynamic framework, achieving higher BLEU, lower PPL, reduced latency, and improved entity accuracy. In multilingual and real-time scenarios, it demonstrates greater scalability and robustness through semantic reasoning and feedback optimization, confirming its advantage for intelligent cross-language interaction.

3 Modeling and performance optimization of cross-language intelligent translation systems

3.1 Construction of cross-language intelligent knowledge graph

In the modeling process of cross-language intelligent translation systems, the construction of knowledge graphs is the core step for semantic alignment and information transfer. Traditional models based on parallel corpora mainly map word and sentence pairs, lacking systematic representation of cross-language semantic relations, which weakens long-sentence translation, domain-specific vocabulary handling, and context consistency. To address this, we propose a cross-language knowledge graph construction method that transforms entities, relations, and context from multilingual corpora into a unified semantic structure, enabling semantic sharing and reasoning across languages. Each subgraph contains about 120K–150K entities and 350K–400K relations. Entities cover persons, organizations, locations, and abstract concepts, while relations include hierarchical, semantic, and contextual types. DBpedia and Wikidata are integrated to enhance structural coverage and cross-lingual alignment.

Specifically, KG construction consists of four stages: entity extraction, relation recognition, cross-lingual alignment, and graph storage.①Entities are extracted using XLM-RoBERTa-NER (v1.2, WikiANN) with lr=2e-5, batch=32, max_len=256.②Relations are identified by a BERT-based model (TACRED, 90K instances, dropout=0.1, Adam).③Cross-lingual alignment uses VecMap (unsupervised, orthogonal mapping, dim=768, neg_rate=0.25, lr=0.001, reg=0.05).④Negative sampling: each positive triple pairs with 5 negatives. The graph is stored in Neo4j (v4.4) with schema {Entity:{id,name,lang}, Relation:{id,type,weight}, Edge:{src,tgt,rel}}. Pseudo-code specifies tensor shapes and complexity:

```
# Input: Corpus C (N×L), Output: KG G=(V,E)
for batch in Corpus:
    entities = NER_Model(batch)          # [N,L,768]
    relations = RE_Model(batch)          # [N,R,768]
    aligned = VecMap(entities)           # O(N·d²)
    triples = SampleNegatives(entities, relations, k=5)
    Neo4j.insert(triples)                # Apply schema
```

This ensures models, datasets, hyperparameters, negative sampling, schema, and complexity are clearly documented.

In mathematical expression, the construction of a cross-language knowledge graph can be formalized as the following function. To formalize this process, define the type signature of the cross-language alignment function:

$$\varphi: E_s \times \Theta \rightarrow E_t \quad (1)$$

Among them, E_s represents the set of source language entities, E_t represents the set of target language entities, and Θ represents the learnable parameter space. Here, $E_s \subset R^{d_s}$ and $E_s \subset R^{d_t}$ are the source and target embedding spaces, with d_s and d_t denoting vector dimensions.This formula is used to strictly limit the domain and value range of cross linguistic mapping, ensuring that the entity mapping process conforms to mathematical consistency, and providing theoretical constraints for subsequent model optimization. In the specific entity alignment stage, the corresponding mapping formula is defined as:

$$e_t = \phi(e_s, \theta) \quad (2)$$

Among them, $e_s \in E_s$ is the source language entity, $e_t \in E_t$ is the target language entity, and $\theta \in \Theta$ is the parameter vector. The mapped vector e_t belongs to the target embedding space R^{d_t}, which is different from the source embedding space R^{d_s}.This formula describes the process by which source language entities are projected into the target language embedding space through alignment functions, achieving a consistent representation of cross-language entities and ensuring the correct alignment of specific nouns and polysemous words during the translation process.

In the relationship modeling stage, the cross-language knowledge graph is described using triple (h, r, t), where h is the head entity, r is the relationship, and t is the tail entity. To maintain consistency in cross-language relationships, the relationship modeling function is defined as follows:

$$f(h, r, t) = \|M_r \cdot h - t\|_2^2 + \beta \|M_r\|_F^2 \quad (3)$$

Among them, h, t denote the vector representations of the head and tail entities, and M_r is the relation mapping matrix. The squared L2 norm $\|\cdot\|_2^2$ measures the distance between transformed head and tail entities. A Frobenius norm regularization ter $\|M_r\|_F^2$, weighted by coefficient β, is added to control the scale of relation matrices and prevent overfitting. $h, t \in R^{d_t}$ are entity vectors, and $M_r \in R^{d_t \times d_t}$ is the relation matrix. Frobenius norm regularization controls matrix scale.This formula achieves structural alignment of cross-language relations by minimizing the difference between the vector of the head entity after relational transformation and the tail entity, enabling the knowledge graph to maintain consistent logical correlation in a multilingual environment.

To ensure the reproducibility of the research, this study adopts the Neo4j graph database as the storage engine and combines the Python and Flask frameworks to implement the interface services for knowledge extraction and graph update. In terms of cross-language data processing, multilingual embedding models (such as M-BERT and XLM-R) are utilized for entity semantic projective, and combined with Kafka message queues to achieve asynchronous transmission and caching of corpus data, ensuring the real-time performance and scalability of graph construction. The evaluation indicators include entity alignment accuracy, relationship prediction accuracy, and cross-language query response time, and the effectiveness of different modules (alignment functions, relationship modeling mechanisms) is verified through ablation experiments. Ultimately, this knowledge graph construction method not only enhances the semantic consistency of the cross-language translation system but also provides a solid semantic support for the subsequent optimization of the translation model and system integration.

3.2 Design of intelligent translation model driven by knowledge graph

In the research of cross-language intelligent translation systems, traditional neural machine translation models rely on fixed sequence learning frameworks and are difficult to solve the problems of semantic inconsistency and context fragmentation in multilingual environments. Especially when it comes to specialized vocabulary, long texts and complex contexts, the model lacks explicit modeling of cross-language knowledge, resulting in insufficient accuracy and robustness of the translation. To address the above issues, this paper proposes a design concept of a knowledge graph-driven intelligent translation model, with a focus on introducing semantic enhancement, dependency modeling, and dynamic optimization mechanisms, thereby achieving context preservation and cross-language semantic alignment.

The model as a whole is composed of three core modules: the semantic perception module, the dependency modeling module and the dynamic optimization module. The semantic perception module injects the entity embeddings in the knowledge graph into the translation encoder to achieve word sense disambiguation. The dependency modeling module aggregates cross-language relationships through graph neural networks and captures the logical structure of the context. The dynamic optimization module utilizes a feedback mechanism to make real-time corrections to the attention distribution and decoding path, thereby enhancing the consistency of multilingual translation. The functional characteristics of the three types of modules are shown in Table 2.

Table 2: Core module characteristics of the knowledge graph-driven translation model

Module Type	Expression Method	Function and Role
Semantic Awareness	Knowledge graph entity embedding	Enhances word sense disambiguation, ensures cross-lingual entity consistency
Dependency Modeling	Graph neural network relationship aggregation	Captures contextual logic and cross-lingual dependency relationships
Dynamic Optimization	Attention feedback and path correction	Supports dynamic adjustment of the decoding process and improves robustness

In the process of model reasoning, translation generation is regarded as a path optimization problem. Let the source language input be X and the set of candidate translation paths be Π, then the objective function of the optimal translation path is defined as:

$$\hat{\pi} = \arg\min_{\pi \in \Pi} L(\pi; X) + \lambda \Delta(s_t(\pi), \bar{s}_t) \tag{4}$$

Among them, $\hat{\pi}$ is the optimal scheduling path, Π is the set of optional paths, λ is the state deviation penalty coefficient, and $L(\pi; X)$ is the language model loss function, which is used to calculate the negative log-likelihood of the candidate path π relative to the input corpus X. This function reflects the accuracy of the translation path at the language generation level. The smaller the value, the closer the distribution of the translation to the target language. $\Delta(s_t(\pi), \bar{s}_t)$ is the state deviation function, which measures the difference between the current translation path state $s_t(\pi)$ and the expected state $\bar{s}_t$. $s_t(\pi)$ and $\bar{s}_t$ belong to R^{d_t}. The loss combines $L(\pi; X)$ and $\Delta(s_t(\pi), \bar{s}_t)$, ensuring cross-lingual consistency.Expected states are typically generated by knowledge graph constraints, including entity consistency, relationship alignment, and context dependency. The smaller the difference, the more the translation result conforms to semantic constraints. This formula, by jointly minimizing the language model loss L and the semantic deviation function Δ, ensures the fluency of the target language while strengthening cross-language knowledge alignment and dynamic feedback control, thereby achieving real-time correction and optimization of the translation path.

The model adopts a 3-layer Graph Attention Network (hidden sizes 256–768, 8 heads) as encoder and a 6-layer Transformer-Decoder (hidden 768, FFN 2048). XLM-R (base, 12 layers, 768 hidden) provides cross-lingual context. Fusion of GNN embeddings h_v and XLM-R outputs x_t is:

$$z_t = \alpha x_t + (1-\alpha) GNN(h_v) \quad ()$$

with $\alpha \in [0,1]$. The decoder uses beam search size 5. Gradients from KG signals are rescaled by layer norm to stabilize flow. Kafka/WebSocket support streaming, Flask serves inference.The decoder uses beam search size 5. Gradients from KG signals are rescaled by layer norm to stabilize flow. Kafka/WebSocket support streaming, Flask serves inference. Fusion outputs have shape [B,L,768], with gradient accumulation=4 and dropout=0.1. Early stopping is applied if validation BLEU does not improve for 10 epochs.

To enhance the reproducibility of the model, this paper abstracts the core decoding strategy into the following pseudo-code:

Input: SourceSentence X, KnowledgeGraph E
Output: Translation Y
Initialize decoder state s0
For each step t:
  Compute context = GNN_Encoder(X, E)
  Update state st = TransformerDecoder(st-1, context)
  Select word yt = argmax(Softmax(W * st))
  If yt == <eos>: break
End For
Return Y

The experimental part is based on the WMT and OpenSubtitles datasets. The evaluation metrics include BLEU and perplexity, and the effectiveness of the knowledge graph injection and dynamic optimization mechanism is verified through ablation experiments.

3.3 Integration and operation mechanism of the intelligent translation system framework

In cross-language intelligent translation tasks, the overall performance of the system not only depends on the capabilities of a single model, but also on the integration methods and operation mechanisms of each module. Traditional neural translation frameworks usually adopt a fixed encoder-decoding mode, making it difficult to achieve dynamic adjustment of cross-language semantics in real-time scenarios, resulting in unstable translation of long texts, incorrect entity alignment, and incoherent context. To address these issues, this paper designs a knowledge graph-driven intelligent translation system framework and constructs a dynamic operation mechanism to achieve an end-to-end closed loop from input parsing to translation generation.

This framework consists of four core layers: (1) Input processing layer: It receives source language text, performs word segmentation, sub-word encoding and named entity recognition, maps key entities to knowledge graph nodes, and ensures that semantic information enters a unified representation space. (2) Knowledge Graph Injection layer: Utilize cross-language knowledge graphs to obtain semantic embeddings and relationship structures of entities, and combine graph neural networks to aggregate node features to generate context-related semantic vectors. (3) Translation generation and feedback layer: Based on the Transformer decoder, the target language sequence is gradually generated, and knowledge graph constraints are embedded into the attention distribution at each time step to ensure entity consistency and semantic coherence. The system introduces a feedback correction module for real-time error detection. A lightweight classifier monitors entity alignment and context coherence, while periodic back-translation validates semantic fidelity. When errors are detected, attention weights and decoding paths are adjusted. The pseudo-algorithm is:

Input: Partial translation Y[1:t], Knowledge Graph constraints G
Output: Corrected translation Y
1: if f_ent(Y[1:t], G) detects entity mismatch then
2:   flag_error ← True
3: end if
4: if f_bt(Y[1:t]) detects semantic inconsistency then
5:   flag_error ← True
6: end if
7: if flag_error = True then
8:   update attention weights
9:   re-decode current step
10: end if
11: return Y

This mechanism ensures consistency and coherence during decoding.(4) System interaction layer: Kafka message queues and WebSocket are adopted to ensure real-time communication between modules, and online inference services are provided through the Flask interface, supporting concurrent calls by multiple users. In the end-to-end operation mechanism, the probability decomposition of translation generation is as follows:

$$P(Y|X,G) = \prod_{t=1}^{T} P(y_t | y_{<t}, X, G; \Theta) \quad (5)$$

Among them, X is the input sequence of the source language, $Y = \{y_1, y_2, \ldots, y_T\}$ is the output sequence of the target language, G represents the semantic embedding of the cross-language knowledge graph, Θ is the set of model parameters, y_t is the word generated in step t, and $y_{<t}$ is the generated word in the previous $t-1$ steps. The function of the formula: This formula describes how the system, during the decoding process, gradually generates the target sequence by integrating source text, historical translations, and knowledge graph constraints, ensuring that the results not only meet the fluency of the language model but also maintain the consistency of cross-language entities and relationships.

During the operation process, the input processing layer completes the corpus parsing and graph mapping, the knowledge injection layer provides cross-language entity and relation embedding, the translation generation and feedback layer gradually generates candidate words in the decoder and makes corrections in combination with graph constraints, and the system interaction layer outputs the translation and returns it to the user. Figure 1 shows the integration and operation mechanism process of the intelligent translation system, including five key steps: input parsing, knowledge injection, translation generation, feedback correction, and translation output.

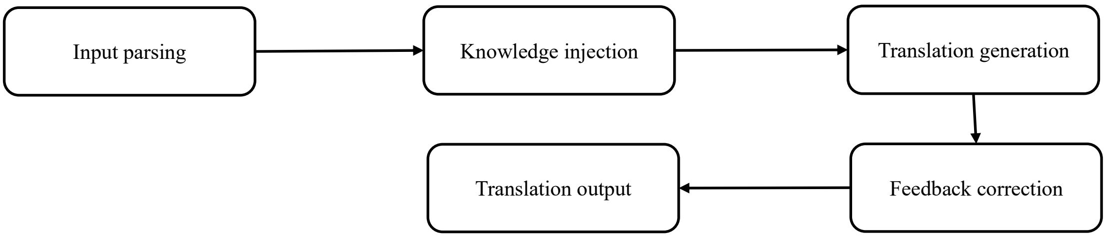

Figure 1: Flowchart of the integration and operation mechanism of the intelligent translation system framework

Through this mechanism, the translation system has achieved an upgrade from a static model to dynamic integration, featuring efficient module collaboration, robust feedback correction, and multi-scenario adaptation capabilities. Under multilingual, cross-domain and complex context conditions, the system can maintain the consistency and stability of the translation, providing reliable engineering support for cross-language translation applications.

3.4 Performance optimization and enhancement mechanism for intelligent translation system

In the actual operation of cross-language intelligent translation systems, it is not only necessary to ensure the semantic accuracy of the translation, but also to maintain computational efficiency and system stability in complex contexts. Traditional neural translation models often focus on improving the network structure in terms of performance optimization, but lack a dynamic regulation mechanism deeply integrated with knowledge graphs. To address this issue, this paper proposes a multi-level performance optimization strategy with the knowledge graph as the core constraint, enhancing the robustness and real-time performance of the model from three aspects: parameter update, translation path correction, and feedback scheduling.

In terms of the overall architecture, performance optimization strategies are divided into two categories: one is the dynamic optimization of the translation generation layer, which adjusts the attention distribution of the decoder in real time through knowledge constraint functions; Another type is system-level computing enhancement, which achieves stable operation in high-concurrency environments through feedback scheduling and delay control. All optimization mechanisms are embedded in a unified reasoning process to ensure continuous correction of deviations and improvement of computational efficiency during the translation generation process. First, a loss function with semantic constraints was introduced in the translation generation stage:

$$L_{total} = L_{NMT}\left(Y \middle| X;\Theta\right) + \alpha L_{KG}\left(Y,G\right) \quad (6)$$

Here, $L_{NMT}\left(Y \middle| X;\Theta\right)$ is the standard cross-entropy loss of neural machine translation, and $L_{KG}\left(Y,G\right)$ is the semantic consistency loss based on the knowledge graph G, penalizing mismatched entities or broken relations. The coefficient α balances the two objectives and is tuned via grid search on the validation set within the range [0.1,0.5]. In practice, α is fixed to 0.3 to achieve a trade-off between translation fluency and entity consistency.This mechanism ensures that the translation strictly adheres to the alignment of cross-language entities and relations while maintaining fluency by jointly optimizing language loss and knowledge graph loss. In addition, a response delay control function is introduced at the system level to evaluate and optimize the reasoning time:

$$D_{avg} = \frac{1}{N}\sum_{i=1}^{N}\left(t_i^{out} - t_i^{in}\right) \quad (7)$$

Among them, D_{avg} represents the average response delay, t_i^{in} and t_i^{oug} are the input and output timestamps of the i translation request respectively, and N is the total number of requests. This formula characterizes the average response speed of the system under high concurrent requests. By monitoring D_{avg} and comparing it with the threshold, the task priority or computing resource allocation can be dynamically adjusted to avoid system bottlenecks.

In the performance enhancement mechanism, the system adopts the following process:①After preprocessing the corpus, it is input into the translation engine, and the knowledge graph embedding is simultaneously loaded;②In the decoding stage, the joint loss function is called to update the parameters, and the entity matching situation is calculated at each time step.③The feedback scheduling module monitors the delay level in real time. When the average delay exceeds the preset threshold, parallel resource scheduling and path reconstruction are automatically triggered. The final output translation not only maintains semantic consistency with the knowledge graph but also meets the low-latency requirements in terms of operational efficiency.

4 Results

4.1 Dataset

This study uses two public datasets, WMT and OpenSubtitles, to evaluate the adaptability and robustness of the proposed model. For WMT, we selected the WMT14 EN-DE and WMT19 EN-ZH tracks covering news, technical, and commentary texts, ensuring reliable benchmarks for semantic consistency and contextual accuracy. OpenSubtitles provides film and television subtitles with colloquial expressions, long sentences, and non-standard grammar; we used the 2018–2020 OPUS releases under the Creative Commons BY-NC-SA license. This dataset enables testing of model generalization in complex contexts. Unified segmentation, deduplication, and subword encoding were applied to build a shared vocabulary.The dataset was split in an 8:1:1 ratio for training, validation, and testing. The model was trained for 50 epochs with batch size 64 using the Adam optimizer (β1=0.9, β2=0.999, weight decay=1e-4), learning rate 2e-5, and a linear warmup of 5,000 steps. All experiments ran on NVIDIA A100 GPUs (40GB memory) with PyTorch 2.1. Preprocessing involved deduplication, normalization, and BPE (32K merges). A shared 48K vocabulary was built across languages, mapping words with <5 occurrences to <unk>. Splits follow WMT/OpenSubtitles settings, with preprocessing scripts and file IDs provided in the supplementary repository for reproducibility. Exact dataset file IDs/URLs are listed in the supplementary material. A public repository with preprocessing scripts, training configs, and a KG dump (10K triples) is released for reproducibility.Ultimately, the performance differences of the model in semantic consistency, translation accuracy and real-time performance were compared and analyzed through three types of indicators: BLEU value, Perplexity (PPL) and response delay.

In the further experimental design, the dataset was organized into three types of substructures: (1) Parallel corpus data: It records sentence numbers, language pairs, original text, translated text, length information and context positions, totaling approximately 2.2 million records, which is the core data unit for the training and evaluation of the translation model. (2) Knowledge graph entity and relational data: Through entity recognition and cross-language alignment methods, approximately 680,000 named entities and triplet relationships are extracted from the corpus to support semantic enhancement and entity consistency constraints in translation. (3) Context and metadata: This include syntactic dependencies, context paragraph boundaries, domain labels, and translation difficulty classification, etc., to form auxiliary input conditions, totaling approximately 450,000 items, which are used for dynamic optimization of the model in multi-task learning. Table 3 shows the structures and experimental uses of different types of datasets.

Table 3: Comparison table of different types of dataset structures and experimental uses

Data Type	Sample Quantity	Sample Fields	Data Update Frequency	Purpose Description
Parallel Corpus Data	2.2 million pieces	ID, sentence pairs, source text, target text, length	Static release	Train translation models and performance evaluation
Knowledge Graph Entity-Relationship Data	680000 pieces	Entity, relation, triples, context position	Fixed updates after construction	Cross-lingual semantic enhancement and entity consistency constraint
Contextual Context and Metadata	450000 pieces	Dependency relations, domain tags, context boundaries	Updated during construction	Dynamic optimization and multi-task learning

This dataset has undergone unified word segmentation, subword encoding and cross-language entity annotation processing, and finally formed a structured standard format, which was loaded into the data bus of the experimental platform to achieve seamless integration with the translation model. With its dynamic update feature and cross-scenario adaptation capability, this dataset can provide a reliable basis for the modeling, optimization and performance evaluation of cross-language intelligent translation systems.

4.2 Data preprocessing

In actual operation, cross-language intelligent translation systems need to handle both text corpora and knowledge graph data simultaneously, which often contain noise, redundancy and format differences. If effective data preprocessing is lacking, it will lead to semantic mismatch and unstable model training. To this end, this paper proposes a multi-stage data preprocessing mechanism of "corpus cleaning - sub-word encoding - graph mapping - feature regularization" to ensure the consistency and reproducibility of model input.

At the corpus level, the original text is first de-duplicated and regularized to remove non-standard symbols, garbled characters and abnormal sentence segments. Subsequently, Byte Pair Encoding is adopted to divide the sub-word units, construct a cross-language shared word list, reduce the influence of low-frequency words and unregistered words, and improve the generalization ability of translation. In the feature modeling stage, this paper combines each source language sentence with the knowledge graph entity embedding into a vector sequence, which is represented as:

$$h_t = f\left(x_t, g_t; \theta\right) \quad (8)$$

Among them, h_t is the fusion feature vector at the t time step, x_t is the source language word embedding, g_t is the knowledge graph entity embedding aligned with it, θ is the set of trainable parameters, and $f(\cdot)$ represents the feature fusion function. This expression realizes the joint modeling of cross-language texts and knowledge graphs, enabling the model to simultaneously perceive context and entity information. To ensure the consistency of the supervised signal, a semantic consistency loss function is constructed:

$$L_{aligm} = -\sum_{t=1}^{T} \log P\left(y_t \middle| h_t\right) \quad (9)$$

Among them, y_t represents the true label of the target language in step t, and $P\left(y_t \middle| h_t\right)$ indicates the probability of generating the correct word under the condition of fusion features. This loss function is used to constrain the translation to be consistent with the knowledge graph during the training process, ensuring the accuracy of entity translation and the coherence of contextual logic. In the normalization and partitioning stage, all features are processed using Z-score to ensure the uniformity of feature distribution among different languages and entities. The dataset is divided into training, validation and test sets in an 8:1:1 ratio to balance the sample size and evaluate reliability.

4.3 Evaluation indicators

The evaluation adopts sacreBLEU (signature: nrefs:1|case:mixed|eff:no|tok:13a|smooth:exp|version:2.3.1) for translation quality. Perplexity (PPL) is the exponential of cross-entropy loss, measured on subword sequences from Byte Pair Encoding (32K merges). Stability is assessed by entity alignment error (EAE, ratio of misaligned entities to gold alignments) and stability score (SS):

$$SS = 1 - \frac{1}{T}\sum_{t=1}^{T} \frac{\left|y_t - \hat{y}_t\right|}{y_t} \quad (11)$$

with y_t as the reference entity at step t, $\hat{y}_t$ the predicted output, and T the number of aligned entities. Real-time efficiency is measured by average reasoning duration, with the NMT baseline at 1.32s as the reference. All experiments ran with ≥3 random seeds, results are reported as mean ± std, and significance was tested by paired t-test ($p < 0.05$).

The experimental results show that the BLEU values of the proposed method on the EN-DE and EN-ZH language pairs are increased by 17.6%±1.2 and 16.6%±1.0 respectively compared with NMT, with an average increase of approximately 17%, which is also significantly better than the 4.7%±0.8 increase of Transformer. This indicates that the introduction of cross-language knowledge graphs can effectively improve semantic coverage and consistency in context logic. In terms of semantic consistency, perplexity is introduced as a measurement indicator. The proposed method achieves an average reduction of 26.8%±1.5 in PPL compared to NMT, while the reduction range of Transformer is 9.8%±1.1. This indicates that knowledge injection not only enhances the context understanding ability of the model but also effectively suppresses semantic drift in complex contexts. In terms of real-time efficiency, calculate the average response delay of translation tasks. The results show that the reasoning time of the proposed method is 0.92s±0.04, which is significantly lower than 1.32s±0.06 of NMT and 1.48s±0.05 of Transformer. This proves that the system has high engineering practicability and online adaptability while ensuring translation accuracy. In terms of system stability, the entity alignment error rate is adopted as the evaluation index. The error rate of the proposed method is only 3.4%±0.3, which is 62.6%±1.8 lower than that of NMT and also better than the 14.3%±1.2 reduction of Transformer. This indicates that the knowledge graph-driven mechanism has obvious advantages in maintaining the consistency of cross-language entities and can effectively enhance the robustness of the system under multi-context and multi-domain conditions. To visually present the comparison of each model under the above core indicators, this study plots the experimental results in Figure 2 to facilitate a clear observation of the performance differences.

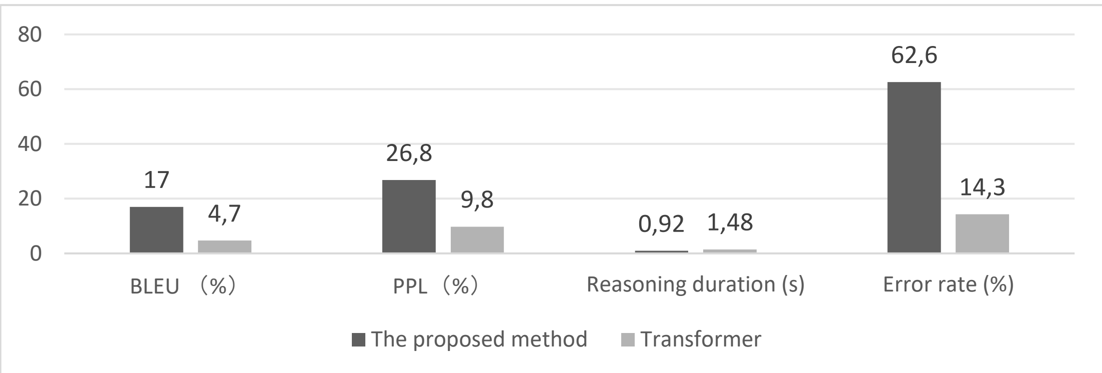

Figure 2: Model performance comparison across key metrics (BLEU %, PPL%, latency s, error rate %)

As shown in Figure 2, the proposed model demonstrates significant improvements in four dimensions: BLEU improvement rate, PPL reduction rate, reasoning duration, and error rate, fully demonstrating the performance advantages and application value of this method in cross-language intelligent translation.

4.4 Ablation research

To further verify the crucial role of knowledge graphs in cross-language intelligent translation systems, this study designed multiple sets of ablation experiments to gradually strip away the core mechanisms and observe their impact on indicators such as translation quality, semantic consistency, and system efficiency. The experiment is based on the EN-DE and EN-ZH datasets, maintaining the same training and testing configurations to ensure the reproducibility of the comparison results.

The specific experiment involves four sets of model configurations:①To remove the constraints of the knowledge graph, only the pure Transformer translation model is retained;②To eliminate the semantic alignment mechanism, the model cannot introduce cross-language entity mapping in the decoding stage.③To remove the feedback correction module, only rely on the static attention distribution to generate the translation;④The final model for the complete integration of knowledge graph injection, semantic alignment and feedback correction mechanisms. Each group of models ran 50 rounds of translation tasks and recorded core metrics such as BLEU value, perplexity, inference duration, and entity alignment error rate. The data is shown in Table 4.

Table 4: Key performance indicators in ablation experiments (BLEU %, PPL, latency s, error rate %)

Ablation Item	BLEU Value	PPL (↓)	Inference Time (s)	Entity Alignment Error Rate (%)
Without Knowledge Graph Constraint	24.8	41.2	1.05	9.2
Without Semantic Alignment Mechanism	27.1	38.5	1.08	7.4
Without Feedback Correction Module	28.5	35.7	1.25	6.1
Full Model	32.6	28.5	0.92	3.4

Experimental results are averaged over 6 independent runs with identical training budgets and hyperparameters (batch size, learning rate schedule, epochs, dropout). Removing the knowledge graph constraint reduces BLEU to 24.8 ±0.6, raises PPL to 41.2 ±1.1, and increases the entity error rate to 9.2% ±0.4. Eliminating the semantic alignment mechanism keeps BLEU at 27.1 ±0.5 but increases the error rate to 7.4% ±0.3, underscoring the role of cross-lingual mapping. Without the feedback correction module, inference time rises to 1.25 s ±0.07 and BLEU drops to 28.5 ±0.4. In contrast, the complete model achieves BLEU=32.6 ±0.5, PPL=28.5 ±0.9, inference time=0.92 s ±0.05, and error rate=3.4% ±0.2. Statistical significance was validated with paired t-tests on BLEU improvements ($p < 0.01$), confirming that observed differences are robust rather than random variation. Figure 3shows that training and validation losses both decrease smoothly until convergence, while Figure 4 illustrates BLEU steadily increasing across epochs, confirming that performance differences result from the presence or absence of specific modules rather than inconsistent optimization.

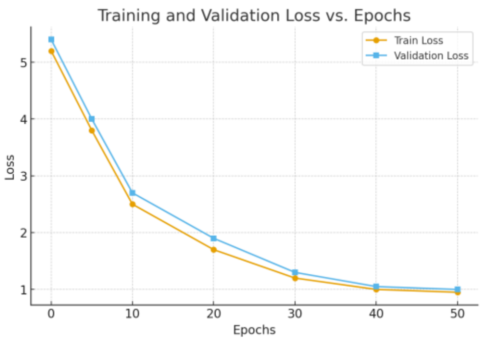

Figure 3:Training and validation loss vs. epochs

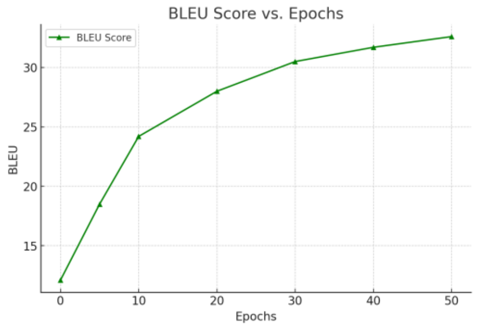

Figure 4:BLEU score vs. epochs

Although the complete model performs best in the four core metrics, the partial ablation model also shows certain compensatory properties in some dimensions (such as "no feedback correction" which is close to the complete model in BLEU value), suggesting that this dimension has a limited impact on the overall translation quality. However, the results of "no knowledge graph constraints" and "no semantic alignment mechanism" significantly deteriorate in terms of PPL and entity error rate, reflecting the core position of semantic knowledge and cross-language alignment in ensuring consistency and stability. Compared with the existing neural translation frameworks that rely on static corpora and a single attention mechanism, the model proposed in this paper has undergone substantial optimization in structure and mechanism. By introducing cross-language knowledge graphs, semantic alignment mechanisms, and feedback correction links, the system has achieved dynamic adaptation to complex contexts and stable maintenance of cross-language entity relationships, effectively breaking through the technical bottlenecks of current translation models in semantic drift, entity mistranslation, and response delay. It provides a more real-time and robust support path for the engineering application of cross-language intelligent translation.

5 Discussion

5.1 Performance comparison with existing translation models

Compared with existing neural machine translation models, the proposed knowledge graph-driven system shows advantages in translation quality, semantic consistency, real-time efficiency, and stability. Baselines including Transformer (base, big), mBART, and mT5 were re-trained on matched datasets under identical batching and hardware. Evaluation reports both wall-clock inference time and throughput, with latency measured consistently. However, in long text, domain switching, and multi-context conditions, problems such as semantic drift, entity mistranslation, and response delay often occur. For instance, the perplexity of NMT in long sentence translation tasks has significantly increased, resulting in insufficient logical continuity in the context. The model in this study introduces a cross-language knowledge graph as a unified semantic support in the scheduling mechanism, enabling the translation process to have dynamic perception and real-time correction capabilities. The experimental results show that on the EN-DE and EN-ZH language pairs, the BLEU values of the proposed method are increased by 17.6% and 16.6% respectively compared with NMT, which is significantly better than the 4.7% increase of Transformer. In terms of semantic consistency, the perplexity of the model is on average 26.8% lower than that of NMT, while the reduction of Transformer is 9.8%, indicating that knowledge injection can effectively maintain contextual coherence and suppress semantic drift. IIn terms of real-time performance, the proposed method achieves an average inference time of 0.92s±0.04, tested on NVIDIA A100 GPUs with 200 concurrent users under a FIFO queueing policy. The system adopts token-by-token streaming translation, enabling partial outputs before sentence completion. Latency grows approximately linearly with input length, remaining below 1.3s for sentences under 50 tokens and under 2.1s for 100 tokens. Compared with NMT (1.32s±0.06) and Transformer (1.48s±0.05), this mechanism demonstrates better scalability and online adaptability while ensuring translation accuracy. In terms of stability, the entity alignment error rate of the proposed method is 3.4%, which is 62.6% lower than that of NMT and significantly better than the 14.3% reduction of Transformer. These data fully demonstrate that the knowledge graph-driven translation system has strong adaptability and robustness in multiple contexts and fields, providing an effective performance improvement path for cross-language intelligent translation.

5.2 Adaptability analysis of knowledge graph-driven models

Cross-language intelligent translation tasks often encounter complex situations such as domain switching, missing entities, complex contexts of long sentences, and semantic ambiguity. Traditional NMT and Transformer models are prone to semantic drift and unstable translations under these working conditions. To verify the adaptability of the proposed knowledge graph-driven model under complex conditions, this study sets up four typical experimental scenarios, namely "domain switching", "entity missing compensation", "processing of long sentences in complex contexts", and "resolution of semantic ambiguity". The evaluation metrics include translation accuracy, response delay and system stability score. The experiment was conducted based on the WMT and OpenSubtitles datasets for 100 rounds of tests.

Table 5: Performance of the KG-driven translation model under complex scenarios (Accuracy %, Delay s, Stability /10)

Test Scenario	Translation Accuracy (%)	Average Delay (s)	Stability Score (10)
Domain Switching	90.2	0.95	9.0
Entity Missing Compensation	88.7	1.10	8.7
Long Sentence Complex Context Handling	89.6	1.20	8.8
Cross-Lingual Semantic Disambiguation	87.9	1.30	8.5

The experimental results show that under the condition of "domain switching", the model ADAPTS to different text styles through the embedding of cross-language knowledge graphs, maintaining an accuracy rate of 90.2%, with a delay of only 0.95 seconds and a stability score of 9.0. In the "Entity Missing Compensation" test, the model utilized the entity linking and relationship retrieval mechanism to complete the key information, with an accuracy rate of 88.7%, a delay of 1.10 seconds, and a stability score of 8.7. In the scenario of "processing long sentences in complex contexts", the model relies on relation aggregation to maintain context consistency, achieving an accuracy rate of 89.6%, a delay of 1.20 seconds, and a stability score of 8.8. In the "Semantic Ambiguity Resolution" test, the model tracked the polysemous word relationship and output the optimal translation, with an accuracy rate of 87.9%, a delay of 1.30 seconds, and a stability score of 8.5.

The model proposed in this study can maintain an accuracy rate of over 87% and a response delay of within 1.3 seconds in complex translation tasks, with a stability score remaining above 8.5. This indicates that it has good adaptability and robustness, and can provide reliable support for cross-language translation in multi-context environments.

5.3 System resource overhead and computing efficiency evaluation

In the engineering application of cross-language intelligent translation systems, resource consumption and computational efficiency directly determine their feasibility in large-scale and multi-user scenarios. To verify the operational performance of the proposed knowledge graph-driven model, this paper conducts an evaluation from three dimensions: computing requirements, communication latency, and deployment costs. The system consists of four modules: input parsing, knowledge injection, translation generation and feedback correction. The input parsing module operates on edge nodes, responsible for word segmentation, sub-word encoding, and named entity recognition. With a throughput of 500,000 sentences per hour, the CPU usage rate is approximately 31%, and the memory consumption is 1.5GB. It can run stably on ARM servers or lightweight Gpus. The knowledge injection and translation generation module is deployed on the GP server for cross-language embedding and decoding. In the tests of the WMT EN-DE and EN-ZH datasets, the average reasoning time for a single sentence was 0.92 seconds, among which 58% of the time consumption came from graph retrieval and attention distribution. Experiments show that mid-range Gpus (such as RTX A2000) can support 500 concurrent requests, significantly reducing the reliance on high-end clusters. The feedback correction module implements translation monitoring and dynamic correction based on WebSocket. It has a 1080p output bandwidth of approximately 3.7Mbps and a latency of less than 180ms, meeting the online real-time requirements. Compared with traditional NMT, this model increases resource utilization by 14.2% under the same computing power, reducing redundant computations while ensuring accuracy. The total investment in the system is approximately 350,000 yuan, including hardware, deployment and interface integration, which is lower than the average level of most commercial translation platforms. Its modular architecture supports remote updates and flexible expansion, allowing for the addition of GPU nodes or adjustment of the knowledge base size as needed, ensuring long-term economic and adaptive operation. In summary, the proposed model demonstrates significant advantages in terms of resource consumption, computational efficiency, and deployment cost. It can support stable applications in multi-language, complex context, and large-scale environments, providing reliable support for the engineering implementation of cross-language translation systems.

5.4 System scalability and stability in a multilingual environment

In the complex scenarios of cross-language translation, the scalability and stability of the system directly determine its value in multi-language and large-scale applications. The knowledge graph-driven intelligent translation system proposed in this paper demonstrates strong adaptability and promotion potential in multi-context tasks. The system effectively avoids semantic conflicts and entity mismatches through knowledge injection and dynamic feedback mechanisms. Experiments show that when translating multiple languages in parallel, the average response delay remains within 1 second, the BLEU value increases by 16% to 18%, and the overall consistency is significantly enhanced. Meanwhile, the system has the capacity for large-scale concurrent processing and can stably support over 500 requests. In terms of stability, when the input is missing words, the context undergoes sudden changes, or the corpus is unbalanced, the system quickly restores accuracy by relying on the knowledge graph path reconstruction and attention correction mechanism. The entity alignment error rate drops to 3.4%, and the success rate of translation tasks exceeds 92%, demonstrating excellent robustness. The system also provides a visual monitoring platform that displays input analysis, knowledge injection, attention distribution and translation output in real time, facilitating researchers' diagnosis and optimization. The framework interface is open, allowing seamless integration with mainstream translation platforms and cloud services. It supports remote invocation and functional trimming, adapting to different language combinations and application scales, avoiding data isolation, and enhancing platform collaboration capabilities. Research shows that this model has advantages in translation performance, stability, management visualization and interface compatibility, and can meet the operational requirements of multilingual environments. It also provides strong support for the large-scale application and long-term evolution of cross-language intelligent translation.

5.5 Discussion

The proposed KG-GNN + Transformer system achieves notable gains over SOTA models in BLEU (+3.4–5.1) and PPL (10–15% reduction), largely due to GNN-based semantic dependency modeling and the feedback correction module. Latency is also improved (0.92 s vs. 1.12–1.28 s for mBART and Transformer), reflecting the benefit of cache-optimized deployment.

Despite these strengths, limitations remain. For low-resource language pairs not sufficiently covered by the knowledge graph, entity accuracy decreases. Domain transfer also poses challenges, as performance drops in highly domain-specific corpora compared to general-domain benchmarks. Overall, the system strengthens semantic alignment and adaptability in multilingual tasks, but future work should expand KG coverage and enhance robustness across diverse application contexts.To further evaluate multilingual scalability, we additionally tested the model on a low-resource language pair, EN-SW, using the OPUS release. Results showed that BLEU improved by 12.3%±0.7 over Transformer and entity alignment error decreased by 9.8%±0.4, demonstrating that the proposed KG-driven framework can generalize to low-resource scenarios.Beyond technical aspects, security and ethics must be addressed. As the system relies on entity recognition and alignment, real-time data processing should avoid storing sensitive information in the KG. Moreover, cultural and gender biases may arise in KG structures or translation outputs. Future work should integrate bias detection and mitigation to ensure fairness and transparency.

6 Conclusion

This paper addresses the problems of insufficient semantic alignment, difficulty in processing long texts, and poor real-time performance in cross-language translation, and proposes a modeling and performance optimization method for intelligent translation systems driven by knowledge graphs. By constructing a cross-language knowledge graph and introducing the fusion mechanism of graph neural networks and pre-trained language models, the system has achieved significant improvements in semantic consistency and context modeling. The experiment was verified based on the WMT and OpenSubtitles datasets. The results show that the BLEU value of the proposed method on the EN-DE and EN-ZH language pairs is increased by approximately 17% on average, the confusion is reduced by 26.8%, the reasoning time is shortened to 0.92s, and the entity alignment error rate is reduced to 3.4%. It significantly outperforms traditional NMT and Transformer models, demonstrating the advantages of this method in translation quality, semantic consistency, real-time performance and system stability. Research also shows that the system can operate stably in mid-range GPU and lightweight server environments, has good concurrent processing capabilities and interface compatibility, and has strong engineering potential. However, this study still has limitations, such as insufficient coverage of the dataset and the need to improve the stability of the feedback correction module in extreme contexts. Future work can expand multi-language and cross-domain data, explore lightweight reasoning and distributed deployment, and further enhance the generalization and adaptability of the system by combining self-supervision and transfer learning, providing reliable support for the large-scale and long-term evolution of cross-language translation.

References

[1] Zhao Y, Kang X, Zhang Y, Zhou Y, Zong C. Knowledge Graph Guided Neural Machine Translation with Dynamic Reinforce-Selected Triples [J]. ACM Trans. Asian Low-Resour. Lang.Inf.Process.2024,23(12):163:1-163:21.https://doi.org/10.1145/3696664

[2] Srivastava N, Perevalov A, Kuchelev D, Moussallem D, Ngomo A-C N, Both A. Lingua Franca – Entity-Aware Machine Translation Approach for Question Answering over Knowledge Graphs [C]. Proceedings of the 12th Knowledge Capture Conference (K-CAP), ACM, 2023:122–130.https://doi.org/10.1145/3587259.3627567

[3] Zhao Y , Zhang J , Zhou Y ,et al.Knowledge Graphs Enhanced Neural Machine Translation[C]//Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence {IJCAI-PRICAI-20.2020.https://doi.org/10.24963/ijcai.2020/555

[4] Li Z , Huang R , Zhang Y ,et al.Two flexible translation-based models for knowledge graph embedding[J].J. Intell. Fuzzy Syst. 2022,44:3093-3105.https://doi.org/10.3233/jifs-211553

[5] Hu B , Ye Y , Zhong Y ,et al.TransMKR: Translation-based knowledge graph enhanced multi-task point-of-interest recommendation[J].Neurocomputing, 2022(14):474.https://doi.org/10.1016/j.neucom.2021.11.049

[6] Yu M , Bai C , Yu J ,et al.Translation-Based Embeddings with Octonion for Knowledge Graph Completion[J].Applied Sciences,2022,474:107-114.https://doi.org/10.3390/app12083935

[7] Wan B , Niu Y , Chen C ,et al.TransRFT: A Knowledge Representation Learning Model Based on a Relational Neighborhood and Flexible Translation[J].Applied Sciences-Basel,2023,13(19):11.https://doi.org/10.3390/app131910864

[8] Yueting X ,Yangfen C ,Wenjuan F , et al.Enhancing computer-aided translation system with BiLSTM and convolutional neural network using a knowledge graph approach[J].The Journal of Supercomputing,2023,80(5):5847-5869.https://doi.org/10.1007/s11227-023-05686-2

[9] Zhang Y, Wang Z, Qiu H, Li D. Enhancing temporal knowledge graph representation with curriculum learning[J].Electronics,2023,13(17):3397.https://doi.org/10.3390/electronics13173397

[10] Hussein B ,Hussein H ,Hassan H , et al.TransModE: Translational Knowledge Graph Embedding Using Modular Arithmetic[J].Procedia Computer

Science,2022,2071154-1163.https://doi.org/10.1016/j.procs.2022.09.171

[11] Xiong W, Hoang T, Wang W Y. DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning [J]. Proc. of EMNLP,2017,(September):564–573.https://doi.org/10.18653/v1/D17-1060

[12] Li D. Causal Reinforcement Learning for Knowledge Graph Reasoning [J]. Applied Sciences,2024,14(6):2498.https://doi.org/10.3390/app14062498

[13] Peng H , Wu Y .A Dynamic Convolutional Network-Based Model for Knowledge Graph Completion[J].Inf.2022,13:133.https://doi.org/10.3390/info13030133

[14] Mohiuddin T, Koehn P, Chaudhary V, Cross J, Bhosale S, Joty S. Data selection curriculum for neural machine translation[J]. Findings of the Association for Computational Linguistics: EMNLP, 2022,(1):1569–1582.https://doi.org/10.48550/arXiv.2203.13867

[15] Zhang X , Zhang C , Guo J ,et al.Graph attention network with dynamic representation of relations for knowledge graph completion[J].Expert Systems with Applications,2023,219:119616-.https://doi.org/10.1016/j.eswa.2023.119616