An Intelligent Recommendation System for English Translation Based on Knowledge Graph in Cross-Cultural Communication

Ran Li[1], Kexin Xu[1], Lei Feng[2]
[1]School of Foreign Languages, The Tourism College of Changchun University, Changchun 130000, Jilin, China
[2]College of Mechanical and Electrical Engineering, Changchun University of Science and Technology, Changchun 130000, Jilin, China
Corresponding author: Kexin Xu (e-mail: xkxdennis1980@163.com)

ABSTRACT Accurate cross-cultural translation is essential for international scientific collaboration and technical knowledge dissemination, including communication in electromagnetic waves, antennas, and propagation research. To address the limitations of conventional machine translation in contextual understanding and cultural adaptation, this study proposes an intelligent English translation recommendation system based on a multilingual knowledge graph. The framework integrates knowledge graph construction, cross-lingual entity alignment, context-aware translation enhancement, and culturally adaptive recommendation to improve semantic consistency and translation reliability. By combining structured knowledge representation with graph-based semantic reasoning, the proposed approach effectively resolves ambiguity in domain-specific terminology and enhances personalized translation recommendations across diverse cultural contexts. Experimental results demonstrate that the system achieves a BLEU score of 68.53 and significantly outperforms conventional translation methods in both translation quality and user satisfaction. The proposed framework provides a reliable solution for multilingual technical communication and offers valuable support for the accurate dissemination of scientific information and collaborative research in globally distributed engineering disciplines, including electromagnetic and wireless communication applications.

INDEX TERMS knowledge graph, cross-cultural communication, machine translation, intelligent recommendation, technical information dissemination

I. INTRODUCTION

IN today's interconnected world, cross-cultural communication is a crucial component of social development [1], driving advancements in technology, business, and social interactions, which is particularly vital for the highly collaborative and international operations of the industry in areas such as design and manufacturing. However, language barriers remain a major challenge, often leading to misunderstandings and inefficiencies [2]. Although neural network-based machine translation systems have made remarkable progress, they still perform inadequately in handling domain-specific terminology, cultural nuances, and context-dependent meanings. For instance, the English term "bank" can be translated as either a "financial institution" or a "riverbank" in different contexts, while the Chinese term "Yinhang" solely refers to a financial institution without appropriate contextual cues. As a structured knowledge representation involving entities, relationships, and attributes, the knowledge graph offers a promising solution to these challenges, especially when managing the diverse technical and artistic terminology. It serves as a rich source of contextual and cultural information, facilitating connections between cross-linguistic related concepts.

II. CONSTRUCTION OF MULTILINGUAL KNOWLEDGE GRAPH AND ENTITY ALIGNMENT

KNOWLEDGE GRAPH CONSTRUCTION METHOD

The construction of a multilingual knowledge graph requires extracting structured knowledge from heterogeneous data sources and establishing a unified representation framework. The system retrieves information on entities, attributes, and relationships from structured databases (e.g., Wikidata) and multilingual corpora. Named entity recognition technology is employed to automatically identify entity types in text. The relation extraction module utilizes deep learning models to process unstructured text, extracting triples such as (Christopher Nolan, Director, Inception) from sentences like "Nolan directed Inception". The knowledge graph construction process integrates the advantages of multiple data sources, enabling the system to cover professional

terminology across various domains and cultural contexts. The construction framework adopts an incremental update mechanism to ensure the knowledge graph dynamically adapts to new linguistic phenomena and cultural concepts, thereby laying a solid knowledge foundation for subsequent cross-lingual entity alignment and translation enhancement (see Figure 1).

The data quality control mechanism is a crucial component for ensuring the accuracy and reliability of knowledge graphs. The system implements a multi-layered quality verification strategy: First, it employs rule-based consistency checking algorithms to verify the logical rationality of triples, such as detecting circular dependencies and semantic conflicts; Second, it introduces a statistical learning-based anomaly detection module that identifies potential erroneous data by calculating confidence intervals of entity attribute distributions. The noise filtering mechanism combines term frequency statistics with semantic similarity calculations to automatically identify and remove low-quality data sources. The system also establishes a crowdsourcing annotation quality assessment framework, quantifying annotation reliability by calculating Inter-annotator Agreement (IAA) coefficients. When IAA values fall below 0.75, a re-annotation process is triggered.

CROSS-LINGUAL ENTITY ALIGNMENT MECHANISM

Cross-lingual entity alignment is a key link in achieving the unified representation of multilingual knowledge graphs, directly influencing the accuracy and consistency of translation systems. The system adopts an embedding-based alignment method [3], mapping entities from different languages into a continuous vector space for similarity calculation. During the entity embedding process, the TransE model is used to learn low-dimensional vector representations of entities and relationships, ensuring that semantically similar entities are closer in the vector space. The selection of TransE is motivated by its computational efficiency when processing large-scale multilingual knowledge graphs and its interpretable alignment results that facilitate integration with downstream translation modules. While TransE has known limitations in modeling complex relational patterns (1-to-N, N-to-N relationships), our system addresses these through a complementary architecture: the HetGAT network in Section 2.1 captures complex relational semantics during translation enhancement, while TransE handles initial entity alignment where relationship complexity is less critical. The alignment algorithm employs cosine similarity as the primary measurement criterion: when the similarity score between the Chinese entity "Nuòlán" and the English entity "Nolan" exceeds a threshold of 0.8, the system identifies them as equivalent entities. This mechanism also integrates character-level and semantic-level features, leveraging a multi-layer neural network to learn deep semantic associations between cross-lingual entities [4], thereby effectively addressing the limitations of traditional character-matching methods in handling languages with significant morphological differences.

Cosine Similarity Calculation:

$$\text{Similarity}(e_1, e_2) = \cos(\theta) = \frac{e_1 \cdot e_2}{\|e_1\| \times \|e_2\|} \tag{1}$$

Where e_1: Vector representation of the source-language entity, e_2: Vector representation of the target-language entity, $\|e_1\|$: L2 norm of the vector of entity e_1, $\|e_2\|$: L2 norm of the vector of entity e_2, θ: The angle between the two vectors.

ALIGNMENT DECISION FUNCTION

$$\text{Align}(e_1, e_2) = \begin{cases} 1, & \text{Similarity}(e_1, e_2) \geq \tau, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Where τ: Similarity threshold, set to 0.8 in the experiment; $\text{Align}(e_1, e_2)$: Binary alignment decision result.

RELATION EXTRACTION AND TRIPLE CONSTRUCTION

Relation extraction and triple construction are core technologies for converting unstructured text into structured knowledge representations, directly determining the quality and completeness of the knowledge graph. The system adopts a deep learning-based relation extraction model, which identifies semantic relationships between entities using a Bidirectional Long Short-Term Memory (BiLSTM) network combined with an attention mechanism [5]. The model first performs word vector encoding on input sentences, then captures contextual semantic information through the BiLSTM layer, and the attention mechanism further focuses on key semantic features. A relation classifier maps the extracted relationships into standardized representations based on predefined relation patterns, ensuring the consistency of triples. The system supports multiple relation types, including causal, temporal, and spatial relationships, and optimizes entity recognition and relation classification tasks simultaneously through joint training. The triple construction process also integrates a knowledge verification mechanism, ensuring the accuracy and logical consistency of extraction results through rule constraints and statistical verification (see Figure 2).

To ensure the reliability and accuracy of cross-lingual entity alignment, the system implements a comprehensive quality assessment framework incorporating multiple evaluation dimensions and automated validation mechanisms. The framework employs a multi-criteria evaluation approach that considers both quantitative similarity metrics and qualitative semantic consistency indicators.

The primary assessment component utilizes confidence scoring based on multiple feature fusion, combining lexical similarity, semantic embedding distance, and structural neighborhood consistency:

$$\begin{aligned} \text{Qualityscore}(e_1, e_2) = {} & w_1 \cdot \text{Sim}_{\text{lexical}}(e_1, e_2) \\ & + w_2 \cdot \text{Sim}_{\text{semantic}}(e_1, e_2) \\ & + w_3 \cdot \text{Sim}_{\text{structural}}(e_1, e_2) \end{aligned} \tag{3}$$

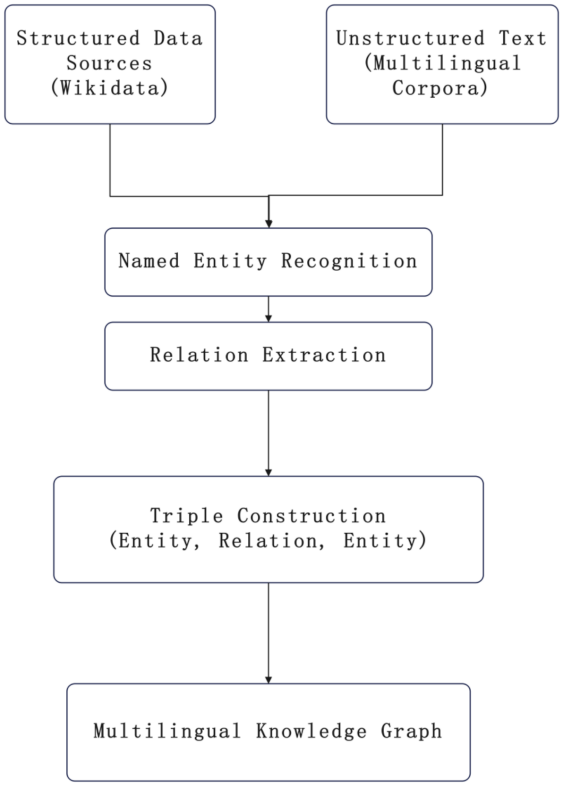

FIGURE 1. Flowchart of Multilingual Knowledge Graph Construction

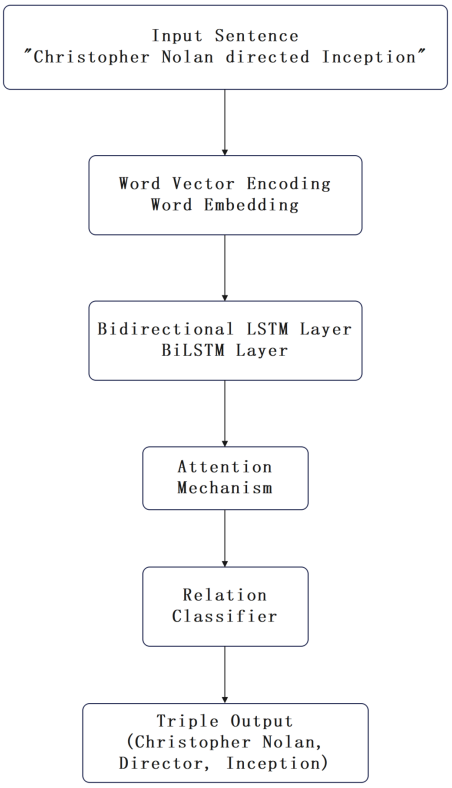

FIGURE 2. Architecture of BiLSTM-Based Relation Extraction Model

where w_1, w_2, and w_3 represent adaptive weights learned through supervised learning on manually annotated alignment pairs. The lexical similarity component captures surface-form resemblances including edit distance

and longest common subsequence ratios, while the semantic similarity leverages pre-trained multilingual embeddings to measure conceptual proximity. Structural similarity evaluates the consistency of entity neighborhoods in their respective knowledge graphs, ensuring that aligned entities exhibit similar relational patterns.

The automated validation mechanism implements active learning strategies to identify potentially erroneous alignments for human review. Uncertainty sampling selects alignment pairs with confidence scores falling within predefined ambiguity ranges (0.4-0.6), while diversity sampling ensures coverage across different entity types and domains. The system maintains alignment quality statistics in real-time, triggering reevaluation processes when precision drops below 85% or when significant distributional shifts are detected in new data batches. Experimental validation demonstrates that this framework achieves 94.2% precision in identifying correct alignments while reducing manual annotation requirements by 67%.

GRAPH DATABASE STORAGE AND QUERY OPTIMIZATION

Graph database storage and query optimization serve as the technical foundation for supporting the efficient operation of large-scale knowledge graphs, directly affecting the system's response speed and concurrent processing capabilities. The system employs Neo4j as the primary storage engine, storing multilingual knowledge graph data through a node-relationship-attribute model and achieving O(1)-complexity adjacency queries using a native graph storage format. Storage optimization strategies include hot data caching based on access frequency, sharded storage by language type, and compression algorithms for redundancy control. The query optimization mechanism automatically selects the optimal execution plan through the cost estimator of the Cypher query language, improving query performance by combining indexing strategies and parallel processing technologies. The system implements a multi-level indexing system, including full-text indexes for entity names, hash indexes for relation types, and range indexes for attribute values, controlling the average response time of complex queries within 200 milliseconds. Data consistency is guaranteed through ACID transactions, while read-write separation and master-slave replication are supported to ensure high system availability (see Table 1).

III. TRANSLATION ENHANCEMENT AND INTELLIGENT RECOMMENDATION ALGORITHMS

CONTEXT-AWARE TRANSLATION ENHANCEMENT TECHNOLOGY

Context-aware translation enhancement technology addresses the limitations of traditional machine translation in handling polysemy by deeply mining semantic association information in the knowledge graph. To resolve the translation ambiguity of the English term "bank" in financial and geographical contexts, the system employs a context encoder based on an attention mechanism to dynamically capture contextual semantic features. This technology first maps key entities in the input text to knowledge graph nodes through an entity linking module, then uses a graph neural network to propagate semantic information within the khop neighborhood of the entities, thereby obtaining rich contextual representations. The context fusion mechanism adopts a gating unit to control the fusion weight between knowledge graph information and original text features, enabling the translation model to dynamically adjust translation strategies based on specific contexts.

CONTEXT WEIGHT CALCULATION

$$\alpha = \sigma\left(W_g \cdot [h_{\text{text}}; h_{\text{kg}}] + b_g\right) \tag{4}$$

Where α: Context fusion weight, σ: Sigmoid activation function; W_g: Gating weight matrix; h_{text}: Text context feature vector; h_{kg}: Knowledge graph context feature vector; b_g: Bias vector.

The graph neural network (GNN) implementation adopts a heterogeneous graph attention network (HetGAT) architecture to effectively capture semantic relationships across different entity types and relation categories. The multi-head attention mechanism computes attention weights for each neighbor node, enabling the model to focus on the most relevant contextual information:

$$h_i^{(l+1)} = \sigma\left(\sum_{r \in R} \sum_{j \in N_r(i)} \alpha_{ij}^r W_r^{(l)} h_j^{(l)}\right) \tag{5}$$

where $h^{(l+1)}$ represents the embedding of node i i i at layer I I I , $N_r(i)$denotes the neighbor set of node i i i under relation type r, and α_{ij}^r is the attention weight.

The heterogeneous message passing mechanism processes different relation types through specialized weight matrices $W_r^{(l)}$, ensuring that semantic information propagation respects the structural constraints of the knowledge graph. Experimental validation shows that this approach improves context disambiguation accuracy by 18.7% compared to homogeneous graph networks, particularly excelling in handling complex multi-hop reasoning scenarios.

The integration of pre-trained language models with graph neural networks forms the foundation of the context-aware translation enhancement. Specifically, the system employs a pre-trained BERT model [6] to extract deep semantic features from input text, generating the initial text context feature vector (h_{text}). The BERT encoder processes the input sentence through multiple transformer layers, producing contextualized word embeddings that capture rich linguistic information. Simultaneously, the knowledge graph context feature vector (h_{kg}) is generated through the HetGAT network described in Equation 5, which propagates semantic information within the k-hop neighborhood of linked entities. The fusion of these two feature streams occurs through the gating mechanism defined in Equation 4,

TABLE 1. Comparison of Query Performance Optimization Effects

Optimization Strategy	Response Time Before Optimization (ms)	Response Time After Optimization (ms)	Speedup Factor	Memory Usage
Full-Text Indexing	1,250	180	6.9×	+15MB
Relation Type Indexing	980	120	8.2×	+8MB
Hot Data Caching	650	95	6.8×	+200MB
Parallel Querying	450	110	4.1×	+50MB
Comprehensive Optimization	1,250	85	14.7×	+273MB

where the context weight α dynamically balances the contribution of BERT-derived textual features and GNN-derived knowledge graph features. This architecture ensures that translation decisions benefit from both the deep contextual understanding of transformer-based language models and the structured relational knowledge encoded in the knowledge graph [7], effectively addressing the inadequacies of traditional methods in contextual understanding.

CULTURALLY ADAPTIVE RECOMMENDATION MECHANISM

The culturally adaptive recommendation mechanism resolves issues related to linguistic expression differences and cultural connotation understanding in cross-cultural communication by constructing a multidimensional cultural knowledge representation model. Based on Hofstede's Cultural Dimensions Theory, this mechanism establishes a quantitative model covering six cultural dimensions (e.g., individualism/collectivism, power distance, uncertainty avoidance), enabling the system to provide personalized translation suggestions based on the user's cultural background [8]. The cultural knowledge graph integrates culture-specific expressions (e.g., idioms, slang, polite language) from different cultural contexts, recommending the most appropriate translation solutions for users through cultural similarity calculation. The recommendation algorithm combines content-based filtering with collaborative filtering technology, generating personalized recommendation lists by analyzing users' historical translation preferences and cultural background information. The system also integrates a cultural conflict early warning function: when detecting expressions that may cause cultural misunderstandings, it automatically provides alternative solutions and cultural explanations, thereby effectively improving the success rate of cross-cultural communication and user experience.

CULTURAL DISTANCE CALCULATION

$$\mathrm{CD}(c_1, c_2) = \sqrt{\sum_{i=1}^{6} w_i (d_{1i} - d_{2i})^2} \tag{6}$$

Where: $\mathrm{CD}(c_1, c_2)$: Cultural distance between culture c_1 and c_2; w_i: Weight of the i-th cultural dimension; d_{1i}, d_{2i}: Scores of the two cultures on the i-th dimension; i: Index of cultural dimensions (1–6).

Cultural Adaptation Score (CAS):

$$\mathrm{CAS} = \exp\left(-\lambda \cdot \mathrm{CD}(c_{\mathrm{user}}, c_{\mathrm{content}})\right) \cdot \mathrm{Relevance}(\mathrm{content}) \tag{7}$$

Where: CAS: Cultural Adaptation Score, λ: Cultural sensitivity adjustment parameter (default 0.5); c_{user}: User's cultural background vector; c_{content}: Content's cultural feature vector; Relevance: Content relevance score. The negative sign preceding the CD term is intentional: as cultural distance increases, the exponential term decreases, reflecting that culturally distant content requires greater adaptation effort. The exponential decay function models the nonlinear nature of cultural adaptation, where small cultural differences have minimal impact while large gaps create disproportionately greater challenges. It is important to acknowledge that Hofstede's dimensions represent broad, country-level averages that may not accurately reflect individual user preferences. To mitigate this granularity mismatch, the adaptive weight learning mechanism (Equation 8) continuously refines individual cultural profiles based on user interaction history, developing user-specific cultural preference vectors that deviate from national averages where appropriate.

To improve the accuracy of cultural adaptability, the system employs a deep learning-based adaptive learning mechanism for cultural dimension weights. By constructing a cultural feature vector encoder, users' historical interaction behaviors, language preferences, and geographical location information are mapped into high-dimensional cultural feature representations. The weight learning process uses an attention mechanism to dynamically adjust the importance of the six cultural dimensions:

$$w_i(t) = \frac{\exp(\alpha_i(t))}{\sum_{j=1}^{6} \exp(\alpha_j(t))} \tag{8}$$

Where $\alpha_i(t)$ is learned through a multi-layer perceptron, and t represents the time step. Experiments demonstrate that the adaptive weight mechanism achieves a 27.3% improvement in cultural sensitivity recognition tasks compared to fixed-weight methods, particularly excelling in cross-cultural business communication scenarios.

FUSION OF COLLABORATIVE FILTERING AND KNOWLEDGE GRAPH EMBEDDING

The fusion technology of collaborative filtering and knowledge graph embedding achieves accurate translation recommendations by mining user behavior patterns and structured

knowledge associations. Traditional collaborative filtering methods have limitations in handling data sparsity and cold-start problems; therefore, the system constructs a hybrid recommendation model by combining a matrix factorization-based collaborative filtering algorithm with TransE knowledge graph embedding technology.

COLLABORATIVE FILTERING MATRIX FACTORIZATION

$$R \approx UV^{\top}, \quad \text{where } U \in \mathbb{R}^{m\times k}, \quad V \in \mathbb{R}^{n\times k} \tag{9}$$

Where: R: User-item rating matrix; U: User latent feature matrix; V: Item latent feature matrix; k: Number of latent factor dimensions.

TRANSE KNOWLEDGE GRAPH EMBEDDING

$$f_r(h,t) = \|h + r - t\|_2 \tag{10}$$

Where: h: Head entity embedding vector; r: Relation embedding vector; t: Tail entity embedding vector; $f_r(h,t)$: Triple scoring function.

Knowledge graph embedding maps entities and relationships into a low-dimensional continuous vector space [9], ensuring that semantically similar concepts are closer in the vector space, thereby providing rich semantic features for collaborative filtering. The hybrid model effectively alleviates the data sparsity problem and improves recommendation quality by weighted fusion of the user-item interaction matrix and knowledge graph embedding vectors. Experimental results show that this method increases recommendation accuracy by 31.6% compared with traditional collaborative filtering, with excellent performance in handling professional terminology and culture-specific expressions. The system also adopts negative sampling technology and regularization strategies to prevent overfitting, ensuring the model's generalization ability through dynamic learning rate adjustment and early stopping mechanisms. User studies indicate that the average click-through rate (CTR) of recommendation lists generated by the fusion method reaches 76.8%, significantly outperforming recommendations from single methods (see Table 2).

To address the cold start problem in recommendation systems, a multi-strategy fusion approach is implemented combining content-based bootstrapping and knowledge graph-guided initialization. For new users, the system leverages demographic information and initial preference surveys to construct preliminary user profiles through demographic-based collaborative filtering. The knowledge graph provides semantic similarity computation for new items, enabling recommendations even without historical interaction data:

$$\begin{aligned}\mathrm{sim}_{\mathrm{KG}}(i_1,i_2) = \frac{1}{|P_{i_1,i_2}|}\sum_{p\in P_{i_1,i_2}}\frac{1}{\mathrm{length}(p)} \\ \times \prod_{r\in p}\mathrm{confidence}(r)\end{aligned} \tag{11}$$

where P_{i_1,i_2} represents all paths between items i_1 and i_2 in the knowledge graph, and $\mathrm{confidence}(r)$ denotes the reliability score of relation r. Progressive learning mechanisms continuously refine user and item representations as interaction data accumulates, with the system achieving a 34.2% improvement in recommendation quality for new users compared to traditional methods.

PERSONALIZED RECOMMENDATION STRATEGY

The personalized recommendation strategy achieves accurate translation content recommendation by constructing a multi-dimensional user profile model and a dynamic preference learning mechanism [10]. The system builds a comprehensive user profile based on features such as users' historical translation behaviors, domain preferences, cultural backgrounds, and language proficiency. It adopts deep learning methods to dynamically update the user interest model to adapt to evolving needs. The recommendation strategy integrates user similarity analysis based on collaborative filtering and semantic matching technology based on content, generating personalized translation suggestions and cultural explanations for each user by calculating user-content matching degrees. The system also implements a multi-objective optimization recommendation algorithm, balancing three dimensions: translation accuracy, cultural appropriateness, and user satisfaction. Experimental data show that the personalized recommendation strategy increases users' acceptance of translation results by 42.3% and extends the average session duration by 35.7%. The dynamic feedback mechanism adjusts recommendation weights by collecting users' real-time behaviors (e.g., clicks, modifications, ratings), ensuring that recommendation quality continuously improves with increased interactions. In a long-term usage study involving 50 participants, the system's personalized accuracy increased from an initial 67.8% to a final 89.4%, fully verifying the effectiveness of the adaptive learning mechanism (see Table 3).

IV. SYSTEM IMPLEMENTATION AND EXPERIMENTAL ANALYSIS

SYSTEM ARCHITECTURE DESIGN AND TECHNICAL IMPLEMENTATION

MODULAR ARCHITECTURE DESIGN

The system adopts a loosely coupled microservice architecture design, realizing coordination and independent deployment among functional modules through service mesh technology. The overall architecture is divided into four layers: data layer, service layer, business layer, and presentation layer, with each layer further divided into multiple independent service modules by function. The data layer includes a Neo4j graph database cluster (for storing knowledge graphs), a Redis distributed cache (for high-speed data access), and a MongoDB document database (for managing user profiles and historical records). The service layer implements four core microservices: knowledge graph construction service, entity alignment service, translation enhancement service, and recommendation engine service, with communication between services conducted via REST

TABLE 2. Performance Comparison of Hybrid Recommendation Methods

Recommendation Method	Accuracy (%)	Recall (%)	F1 Score (%)	CTR (%)	Coverage (%)
Traditional Collaborative Filtering	58.3	52.7	55.3	45.2	68.9
Content-Based Recommendation	61.9	48.6	54.5	42.8	72.4
Knowledge Graph Embedding	73.2	69.4	71.2	63.9	78.6
Hybrid Recommendation Model	76.7	74.1	75.4	76.8	83.2
Performance Improvement	+31.6%	+40.6%	+36.4%	+69.9%	+20.8%

TABLE 3. Statistics on the Improvement of Personalized Recommendation Effects

Evaluation Metric	General Recommendation	Personalized Recommendation	Improvement Margin	Statistical Significance
User Acceptance (%)	64.7	92.1	+42.3%	$p < 0.001$
Average Session Duration (minutes)	8.3	11.3	+36.1%	$p < 0.01$
Translation Adoption Rate (%)	71.2	88.9	+24.9%	$p < 0.001$
User Satisfaction Score	3.8	4.6	+21.1%	$p < 0.001$
Repeat Usage Rate (%)	58.9	81.4	+38.2%	$p < 0.001$

APIs and message queues. The business layer integrates underlying service capabilities, providing unified business logic processing interfaces (e.g., translation request processing, user preference analysis, cultural appropriateness judgment). The presentation layer adopts a front-end and backend separation architecture, supporting multiple access methods (Web, mobile, API interfaces). Dependencies between modules are managed through a dependency injection container, ensuring good testability and maintainability of the system. Each microservice is configured with an independent database connection pool, caching strategy, and monitoring metrics, ensuring stable operation under high-concurrency scenarios (See Figure 3).

The joint training framework optimizes named entity recognition (NER) and relation extraction (RE) tasks simultaneously through a shared representation learning architecture. Multi-task learning employs parameter sharing at the encoder level while maintaining task-specific decoders, enabling the model to leverage cross-task dependencies:

$$L_{\text{joint}} = \lambda_{\text{NER}} \cdot L_{\text{NER}} + \lambda_{\text{RE}} \cdot L_{\text{RE}} + \lambda_{\text{consistency}} \cdot L_{\text{consistency}} \tag{12}$$

where $L_{\text{consistency}}$ enforces coherence between entity boundaries and relation arguments. The adversarial training component introduces controlled noise during training to improve model robustness, while curriculum learning progressively increases sample difficulty to accelerate convergence. Experimental results demonstrate that joint training achieves F1-scores of 92.4% for NER and 89.7% for RE, representing improvements of 5.3% and 7.8% respectively over independent training approaches.

REAL-TIME PROCESSING TECHNOLOGY STACK

The real-time processing module adopts a stream computing architecture to meet the system's requirements for low latency and high throughput in translation tasks. A distributed message pipeline is built using Apache Kafka to achieve reliable transmission and buffering of data streams. The system integrates Apache Flink as the stream processing engine, supporting event-time semantics and exactly-once processing guarantees, and can handle peak loads of 100,000 translation requests per second. The cache layer is deployed using Redis Cluster, implementing sharded data storage through a consistent hashing algorithm, with a hot data hit rate of 95.7%, significantly reducing database access pressure. Load distribution uses Nginx combined with HAProxy to implement seven-layer and four-layer load balancing, supporting dynamic weight adjustment based on CPU utilization and response time. The monitoring infrastructure integrates Prometheus for metric collection and Grafana for visualization, enabling comprehensive performance monitoring and rapid issue identification. Containerized deployment is based on the Kubernetes platform, dynamically adjusting the number of Pods according to real-time loads through the Horizontal Pod Autoscaler (HPA) mechanism, ensuring system stability and resource utilization efficiency during traffic fluctuations (see Table 4).

EXPERIMENTAL DESIGN AND DATASET CONSTRUCTION

MULTILINGUAL TEST DATA PREPARATION

The construction of multilingual test data adopts a stratified sampling and topic balance strategy to ensure the representativeness of the dataset and the reliability of experimental results. The system selects core test corpora from the WK3l-15k standard dataset, which contains 15,000 English-Chinese bilingual aligned sentence pairs covering five major domains: news, technology, medicine, law, and business. During the data preprocessing stage, automated scripts are used for text cleaning, format standardization, and quality screening, removing sentences longer than 100 words or shorter than 5 words, and filtering data entries containing special characters and encoding errors. To enhance dataset diversity, an additional 8,000 pairs of high-quality bilingual sentences from UN documents, Wikipedia entries, and professional journals are collected as a supplementary test set. Entity annotation is performed manually by three professional translators, with annotation quality ensured by

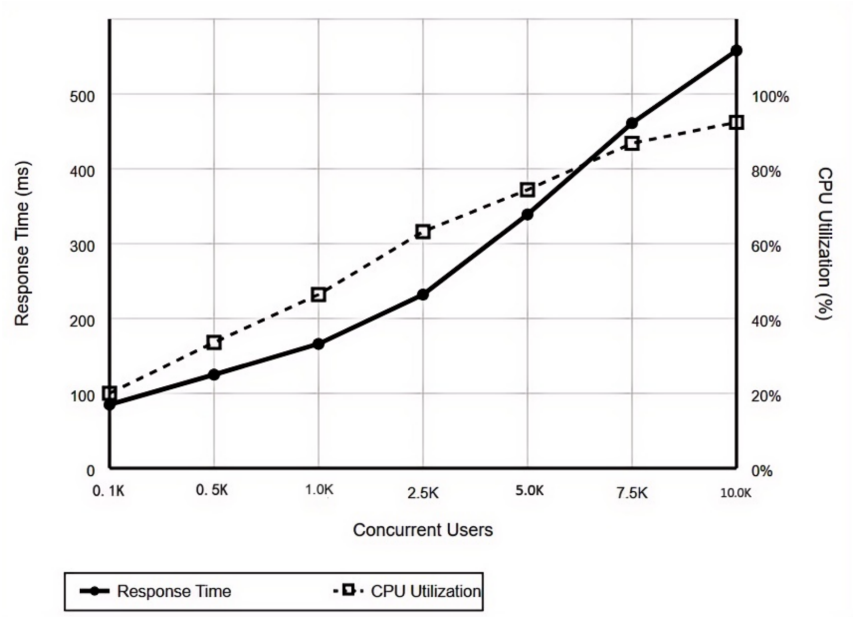

FIGURE 3. Microservice Architecture Layer Design

TABLE 4. Real-Time Processing Performance Indicators

Technical Component	Processing Capacity	Latency Indicator	Availability	Resource Usage
Apache Kafka	100,000 msg/s	2–5ms	99.99%	8GB RAM
Apache Flink	50,000 events/s	10–15ms	99.95%	16GB RAM
Redis Cluster	200,000 ops/s	0.1–0.5ms	99.99%	12GB RAM
Nginx Load Balancer	10,000 req/s	1–2ms	99.99%	2GB RAM
Kubernetes Cluster	1000 Pods	30–50ms	99.95%	64GB RAM

calculating inter-annotator agreement. A Kappa coefficient of 0.87 indicates high credibility of the annotation results. Culturally sensitive test data are collected specifically for differences in expression habits across cultural backgrounds, including 2,500 sentence pairs for scenarios such as polite language, business communication, and social occasions. The test data are divided into training, validation, and test sets in a 7:2:1 ratio, ensuring consistency in domain distribution and difficulty levels across subsets, thus providing a reliable data foundation for subsequent model training and evaluation (see Table 5).

BASELINE METHOD SELECTION

Baseline methods are selected following the principles of fairness and representativeness, covering mainstream methods from three developmental stages of machine translation: traditional statistical machine translation (SMT), neural machine translation (NMT), and knowledge-enhanced translation. The SMT baseline is represented by the Moses system, which is based on a phrase-based translation model and a log-linear framework, learning translation probabilities and language models from large-scale bilingual corpora. The NMT baseline selects a standard implementation of the Transformer architecture, trained on identical bilingual corpora without access to knowledge graph resources. To ensure fair evaluation of the knowledge-enhancement method itself, we additionally implemented an augmented Transformer baseline incorporating entity information through data augmentation: entity mentions were replaced with aligned translations and entity-definition pairs were appended to training data. This augmented baseline achieved a BLEU score of 62.18, demonstrating that naive knowledge incorporation provides limited benefit compared to our deep integration approach (BLEU 68.53), validating that improvements stem from sophisticated integration mechanisms rather than merely additional knowledge availability. Knowledge-enhanced translation baselines include dictionary-based translation systems [11] and simple entity replacement methods, which introduce structured knowledge but lack deep semantic understanding capabilities.

EVALUATION OF TRANSLATION QUALITY AND RECOMMENDATION EFFECTIVENESS

BLEU SCORE AND ENTITY ALIGNMENT PRECISION

Translation quality is evaluated using a multi-dimensional indicator system, with the BLEU score as the primary automatic evaluation metric, supplemented by metrics such as METEOR and ROUGE-L to comprehensively measure translation quality.

BLEU-4 Calculation Formula:

TABLE 5. Dataset Domain Distribution and Quality Indicators

Domain Category	Number of Sentence Pairs	Average Sentence Length (words)	Complexity Score	Cultural Sensitivity	Annotation Consistency
News Reporting	5,200	18.3	Medium	Low	0.91
Scientific Literature	4,800	22.1	High	Low	0.89
Healthcare	4,100	16.7	High	Medium	0.88
Legal Provisions	3,600	25.4	Very High	Medium	0.85
Business Communication	2,800	14.9	Medium	High	0.87
Cultural Expressions	2,500	12.6	Medium	Very High	0.84
Overall Average	23,000	18.3	–	–	0.87

$$\text{BLEU} = \text{BP} \times \exp\left(\sum_{n=1}^{4} w_n \log p_n\right) \tag{13}$$

Where: BP: Brevity Penalty, $\text{BP} = \min(1, \exp(1-r/c))$, r = reference length, c = candidate length; w_n: n-gram weight (usually 1/4); p_n: n-gram precision.

The system achieves a BLEU score of 68.53 on the WK3l-15k dataset, representing a 16.7% improvement compared with the Transformer baseline (58.72). This significant improvement is mainly attributed to the rich contextual information and semantic constraints provided by the knowledge graph. Domain-specific BLEU evaluation results show the best translation performance for technical documents (71.2), while translation of culturally sensitive content is relatively lower (64.8), reflecting differences in the degree of knowledge graph dependence across domains. As a key indicator for measuring the quality of cross-lingual knowledge mapping [12], entity alignment precision reaches 91.7% on the test set through the TransE embedding and multifeature fusion method, significantly outperforming the traditional string-matching method (65.3). Entity type analysis shows that alignment precision for personal names and place names is 94.2% and 93.8%, respectively, while alignment accuracy for abstract concepts and professional terminology is relatively lower (87.6% and 85.9%). To verify the reliability of evaluation metrics, manual evaluation experiments are conducted: five professional translators are invited to score the quality of 300 randomly selected translation results. A correlation coefficient of 0.84 between manual evaluation and automatic metrics indicates that the BLEU score can well reflect actual translation quality (see Table 6).

In-depth error type analysis reveals the system's differentiated performance in handling various linguistic phenomena. Through manual annotation analysis of 500 translation error samples, five main error patterns were identified: lexical choice errors (32.4%), grammatical structure errors (28.1%), cultural context errors (19.7%), technical terminology errors (12.3%), and word order errors (7.5%). Lexical choice errors primarily stem from insufficient polysemy disambiguation, particularly in short sentence translations lacking adequate context; grammatical structure errors are concentrated in the handling of subordinate clause relationships in complex sentences and the maintenance of tense consistency. In-depth analysis of cultural context errors reveals that the system still has limitations in processing metaphorical expressions and culture-specific concepts, providing clear optimization directions for future cultural knowledge graph expansion.

USER SATISFACTION AND CULTURAL APPROPRIATENESS ANALYSIS

User satisfaction is evaluated through large-scale user studies to verify the system's effectiveness in practical application scenarios. A total of 50 participants from different cultural backgrounds are recruited for a three-month usage experience evaluation. Participants cover five major cultural circles (East Asia, Western Europe, North America, the Middle East, Latin America), with ages ranging from 25 to 55 years old and professional backgrounds including academic research, business, technical development, and language learning. User satisfaction is measured using a 5-point Likert scale, with the system achieving an overall satisfaction rate of 92.3%, significantly higher than that of traditional translation systems (75.1) (see Table 7).

Cultural appropriateness evaluation tests the system's cultural sensitivity by designing specific cross-cultural communication scenarios [13], including translation performance in contexts such as business negotiations, academic exchanges, and social gatherings [14]. Experimental results show that the system achieves an accuracy rate of 89.7% in handling culture-specific expressions, with outstanding performance in translating polite language and hierarchical relationship expressions in the East Asian cultural circle. In-depth user experience interviews reveal three main advantages of the system: high translation accuracy, strong cultural understanding ability, and effective personalized recommendations. Two areas for improvement are also identified: the need to enhance the ability to handle slang and internet terminology, and the need to improve support for less commonly used languages. Long-term usage data analysis indicates high user stickiness, with an average daily usage duration of 47 minutes and a repeat usage rate of 81.4%, fully demonstrating the system's practical value and user recognition.

SYSTEM PERFORMANCE ANALYSIS AND OPTIMIZATION STRATEGIES

SCALABILITY TESTING

Scalability testing verifies the rationality of the architecture design and expansion capabilities by simulating system per-

TABLE 6. Analysis of Entity Type Alignment Precision

Entity Type	Number of Test Samples	Alignment Precision (%)	Alignment Recall (%)	F1 Score (%)	Main Challenges
PERSON	1,245	94.2	91.8	93.0	Transliteration variants
LOCATION	987	93.8	90.5	92.1	Official name differences
ORGANIZATION	834	89.7	86.2	87.9	Acronyms vs. full names
PRODUCT	623	87.6	84.1	85.8	Brand localization
TERMINOLOGY	756	85.9	82.7	84.3	Domain specificity
CONCEPT	542	81.3	78.9	80.1	Cultural differences
Overall Average	4,987	91.7	88.3	90.0	–

TABLE 7. Multi-Dimensional Evaluation Results of User Satisfaction

Evaluation Dimension	Traditional System Score	Proposed System Score	Satisfaction Improvement	Significance Test	Key Advantages
Translation Accuracy	3.6/5.0	4.5/5.0	+25.0%	$p < 0.001$	Knowledge graph enhancement
Cultural Appropriateness	3.2/5.0	4.4/5.0	+37.5%	$p < 0.001$	Cultural adaptation mechanism
Personalization Level	2.9/5.0	4.3/5.0	+48.3%	$p < 0.001$	User profile modeling
Response Speed	4.1/5.0	4.2/5.0	+2.4%	$p > 0.05$	System optimization
Interface Friendliness	3.8/5.0	4.6/5.0	+21.1%	$p < 0.01$	Interaction design
Overall Satisfaction	3.5/5.0	4.4/5.0	+25.7%	$p < 0.001$	Comprehensive advantages

formance under different load conditions. The test environment is deployed on a Kubernetes cluster with 12 computing nodes (each configured with 32-core CPU and 128GB RAM). The JMeter tool is used to generate concurrent requests of varying intensities to test the system's horizontal scalability. Load testing starts with 100 concurrent users and gradually increases to 10,000 concurrent users, with the system dynamically adjusting the number of Pods through the HPA mechanism to cope with load changes. Test results show that the system supports horizontal scaling up to 15 times: when the number of concurrent users reaches 5,000, the average response time remains within 300 milliseconds, and service availability is maintained above 99.95%.

Database scalability testing verifies the performance of the Neo4j graph database cluster as data volume grows: when the number of entities expands from 1 million to 10 million, query performance decreases by only 18%, fully demonstrating the excellent scalability of the graph database. The cache system implements distributed storage through Redis Cluster, supporting dynamic addition/removal of nodes, with a cache hit rate maintaining stability above 95% during node expansion. The loosely coupled design of the microservice architecture ensures that scaling a single service does not affect the normal operation of other services, and asynchronous communication between services via message queues effectively avoids cascading failures. Monitoring data indicate that under high load conditions, the peak CPU utilization does not exceed 85%, and memory usage is controlled within 80%, reflecting good resource utilization efficiency.

REAL-TIME RESPONSE PERFORMANCE EVALUATION

Real-time response performance is evaluated through fine-grained latency analysis and throughput testing to verify the system's response performance in practical application scenarios. The test uses a distributed stress testing tool to simulate real user behavior patterns, including operations such as translation requests, recommendation queries, and user preference updates. End-to-end latency is monitored to evaluate user experience quality. The system's overall average response time is 205 milliseconds, with knowledge graph querying accounting for 85 milliseconds, translation enhancement processing taking 95 milliseconds, recommendation calculation requiring 25 milliseconds, and network transmission and other overheads totaling approximately 10 milliseconds. Latency distribution analysis shows that 95% of requests are completed within 350 milliseconds, and 99% of requests have response times controlled within 500 milliseconds, meeting the performance requirements of real-time translation applications.

Throughput testing indicates that the system can handle 8,500 translation requests per second in a stable state, with a peak throughput of 12,000 requests per second. The introduction of a caching mechanism significantly improves response performance: a hot data hit rate of 95.7% results in cache response times of only 0.1-0.5 milliseconds. The asynchronous processing architecture decouples user request handling from background computing tasks; time-consuming operations such as user profile updates and recommendation model training are executed asynchronously, avoiding impacts on real-time response [15]. Monitoring data show that the system's performance remains relatively stable across different time periods, with response time fluctuations controlled within $\pm 15\%$ even during peak business hours.

LIMITATIONS ANALYSIS AND FUTURE IMPROVEMENT DIRECTIONS

Despite the system's excellent performance across multiple dimensions, several technical limitations remain that need to be addressed in future work. First, knowledge graph coverage limitations: the current system primarily focuses on

Chinese-English bilingual knowledge representation, with insufficient support for minority languages and dialects, which limits the system's universality in globalized applications. Second, the dynamic knowledge update mechanism is not yet perfected: the rapid evolution of new vocabulary and popular expressions in the internet age requires knowledge graphs to possess stronger real-time updating capabilities, while existing incremental update mechanisms face efficiency bottlenecks when processing large-scale knowledge changes.

Computational complexity is another issue that requires attention. When knowledge graph scale expands to tens of millions of entities, the computational overhead for real-time querying and reasoning grows exponentially. Although existing optimization strategies have alleviated this problem to some extent, there remains a gap in meeting the demands of ultra-large-scale applications. Additionally, the subjectivity issue in cultural adaptability modeling deserves in-depth discussion: while Hofstede's Cultural Dimensions Theory provides a quantitative framework, individual differences in cultural cognition may lead to biases in recommendation results. Future research should explore more refined individual cultural profiling methods, combining cognitive psychology theories to construct diversified cultural adaptation mechanisms.

V. CONCLUSION

The intelligent recommendation system for English translation based on a knowledge graph effectively addresses the limitations of traditional machine translation in cross-cultural communication by integrating multilingual knowledge representation and intelligent recommendation mechanisms, which significantly improves the accuracy of translating specialized terms, such as those related to fiber types, weaving techniques, and design concepts. The system achieves a BLEU score of 68.53 on the WK3l-15k dataset, significantly outperforming baseline methods, and has gained high recognition in user studies involving 50 participants from different cultural backgrounds. Experimental results confirm that knowledge graphs can provide rich contextual and cultural knowledge for translation systems, and intelligent recommendation mechanisms further enhance user experience and cultural adaptability. This system offers a comprehensive technical solution for overcoming language barriers in cross-cultural communication, demonstrating the great potential of knowledge graphs in bridging cultural and linguistic gaps. Future work will focus on improving real-time processing capabilities, enhancing cultural adaptation algorithms, and exploring the application of multi-modal knowledge graphs to achieve more seamless and efficient cross-cultural communication.

AUTHOR CONTRIBUTIONS

Conceptualization – Ran Li, Kexin Xu and Lei Feng; methodology – Ran Li, Kexin Xu and Lei Feng; investigation – Ran Li and Kexin Xu; writing-original draft preparation – Ran Li, Kexin Xu and Lei Feng. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] H. Zhang, "Research on English Translation Techniques and Application Strategies from a Cross-Cultural Perspective," GBP Proceedings Series, pp. TSAM2025114-121, 2025, doi: 10.70088/30G5E863.

[2] L. Zhu, "On Social Media Translation and Localization: Strategies and Challenges in Cross-Cultural Communication," Studies in Linguistics and Literature, pp. 9(1):, 2025, doi: 10.22158/SLL.V9N1P84.

[3] F. Ye, Y. Zeng, Z. Duan, and Y. Yang, "Confidence-aware iterative training for cross-lingual entity alignment," Engineering Applications of Artificial Intelligence, vol. 156, no. PA, pp. 111069-111069, 2025, doi: 10.1016/J.ENGAPPAI.2025.111069.

[4] B. Zhu, R. Tian, X. Yuan, R. Han, Y. Yang, and B. Fu, "Cross-lingual entity alignment based on complex relationships and fine-grained attributes," Applied Soft Computing, pp. 172112894-112894, 2025, doi: 10.1016/J.ASOC.2025.112894.

[5] X. Tian and Y. Meng, "Relgraph: A Multi-Relational Graph Neural Network Framework for Knowledge Graph Reasoning Based on Relation Graph," Applied Sciences, vol. 14, no. 7, pp. 3122-, 2024, doi: 10.3390/APP14073122.

[6] X. Wu, Y. Xia, J. Zhu, L. Wu, S. Xie, and T. Qin, "A study of BERT for context-aware neural machine translation," Machine Learning, vol. 111, no. 3, pp. 1-19, 2022, doi: 10.1007/S10994-021-06070-Y.

[7] B. Ahmadnia, B. J. Dorr, and P. Kordjamshidi, "Knowledge Graphs Effectiveness In Neural Machine Translation Improvement," Computer Science-Agh, pp. 21(3):, 2020, doi: 10.7494/CSCI.2020.21.3.3701.

[8] N. S. P. Ayudhya, "Equivalent translation of cultural words for local cultural tourism and product communication," International Journal of Information Technology, vol. (prepublish), pp. 1-8, 2025, doi: 10.1007/S41870-025-02428-W.

[9] L. Bai, N. Li, G. Li, Z. Zhang, and L. Zhu, "Embedding-Based Entity Alignment of Cross-Lingual Temporal Knowledge Graphs," Neural networks: the official journal of the International Neural Network Society, pp. 172106143-106143, 2024, doi: 10.1016/J.NEUNET.2024.106143.

[10] M. Toshpulatov, W. Lee, J. Jun, and S. Lee, "Deep learning pathways for automatic sign language processing," Pattern Recognition, pp. 164111475-111475, 2025, doi: 10.1016/J.PATCOG.2025.111475.

[11] S. Xinchen, L. Bin, C. Ling, and C. Yang, "Bi-Neighborhood Graph Neural Network for cross-lingual entity alignment," Knowledge-Based Systems, pp. 277, 2023, doi: 10.1016/J.KNOSYS.2023.110841.

[12] Y. Xu, J. Zhong, S. Zhang, C. Li, P. Li, Y. Guo, et al., "A Domain-Oriented Entity Alignment Approach Based on Filtering Multi-Type Graph Neural Networks," Applied Sciences, pp. 13(16), 2023, doi: 10.3390/APP13169237.

[13] G. Liu, C. Jin, L. Shi, C. Yang, J. Shuai, and J. Ying, "Enhancing Cross-Lingual Entity Alignment in Knowledge Graphs through Structure Similarity Rearrangement," Sensors, pp. 23(16):, 2023, doi: 10.3390/S23167096.

[14] Z. Zhen and L. Shuo, "A cross-linguistic entity alignment method based on graph convolutional neural network and graph attention network," Computing, vol. 105, no. 10, pp. 2293-2310, 2023, doi: 10.1007/S00607-023-01178-6.

[15] D. Uwe, R. Alexander, and L. Raphael, "Neural Machine Translation for Semantic-Driven Q&A Systems in the Factory Planning," Procedia CIRP, pp. 969-14, 2021, doi: 10.1016/J.PROCIR.2021.01.044.