Graph Fusion Across Languages using Large Language Models

Kaung Myat Kyaw*, Khush Agarwal*, Jonathan Chan
Innovative Cognitive Computing Research Center (IC2)
School of Information Technology, KMUTT
`{kaungmyat.kyaw, khush.agar}@mail.kmutt.ac.th, jonathan@sit.kmutt.ac.th`

1. ABSTRACT

Combining multiple knowledge graphs (KGs) across linguistic boundaries is a persistent challenge due to semantic heterogeneity and the complexity of graph environments. We propose a framework for cross-lingual graph fusion, leveraging the in-context reasoning and multilingual semantic priors of Large Language Models (LLMs).

The framework implements **structural linearization** by mapping triplets directly into natural language sequences (e.g., [head] [relation] [tail]), enabling the LLM to map relations and reconcile entities between an evolving fused graph $G_c^{(t-1)}$ and a new candidate graph G_t.

Evaluated on the **DBP15K** dataset, this exploratory study demonstrates that LLMs can serve as a universal semantic bridge to resolve cross-lingual discrepancies. Results show the successful sequential agglomeration of multiple heterogeneous graphs, offering a scalable, modular solution for continuous knowledge synthesis in multi-source, multilingual environments.

Keywords

Cross-lingual Graph Fusion · Large Language Models · Entity Alignment · Knowledge Graphs

2. INTRODUCTION

- Knowledge is fragmented across linguistic and cultural boundaries. **Cross-lingual entity alignment (EA)** and graph fusion are essential to build a universal knowledge base.
- Traditional embedding- and GNN-based methods rely on seed alignments, are sensitive to low-resource settings, and lack interpretability.
- LLMs encode multilingual semantic priors and enable zero-shot transfer via in-context reasoning.
- We introduce a **modular sequential agglomeration framework** that linearizes triplets, leverages LLM reasoning, and fuses N-graphs incrementally.
- Results show high precision **(88.0% @ τ=0.90)** without costly seed alignments, though recall is limited by context-window constraints.

3. PRELIMINARIES & PROBLEM FORMULATION

Knowledge Graph (KG)

$$G = (E, R, T) \text{ where } T \subseteq E \times R \times E \text{ are triplets } (h, r, t).$$

Cross-lingual Entity Alignment

$$\text{Find } A_{ij} = \{(e_i, e_j) \in E_i \times E_j \mid e_i \equiv e_j\} \text{ across } L_i, L_j.$$

N-Graph Fusion

$$\text{Sequentially fuse } \{G_1, \dots, G_N\} \text{ into } G_c^{(t)} \text{ via agglomeration.}$$

Objective at iteration t

$$\max_{(u_i,u_j) \in G_i \times G_k} P(u_i \equiv u_j \mid L(G_i), L(G_k))$$

Accept if $\sigma(u_i, u_j) \geq \tau$ (e.g., $\tau = 0.90$).

4. CONCEPTUAL FRAMEWORK

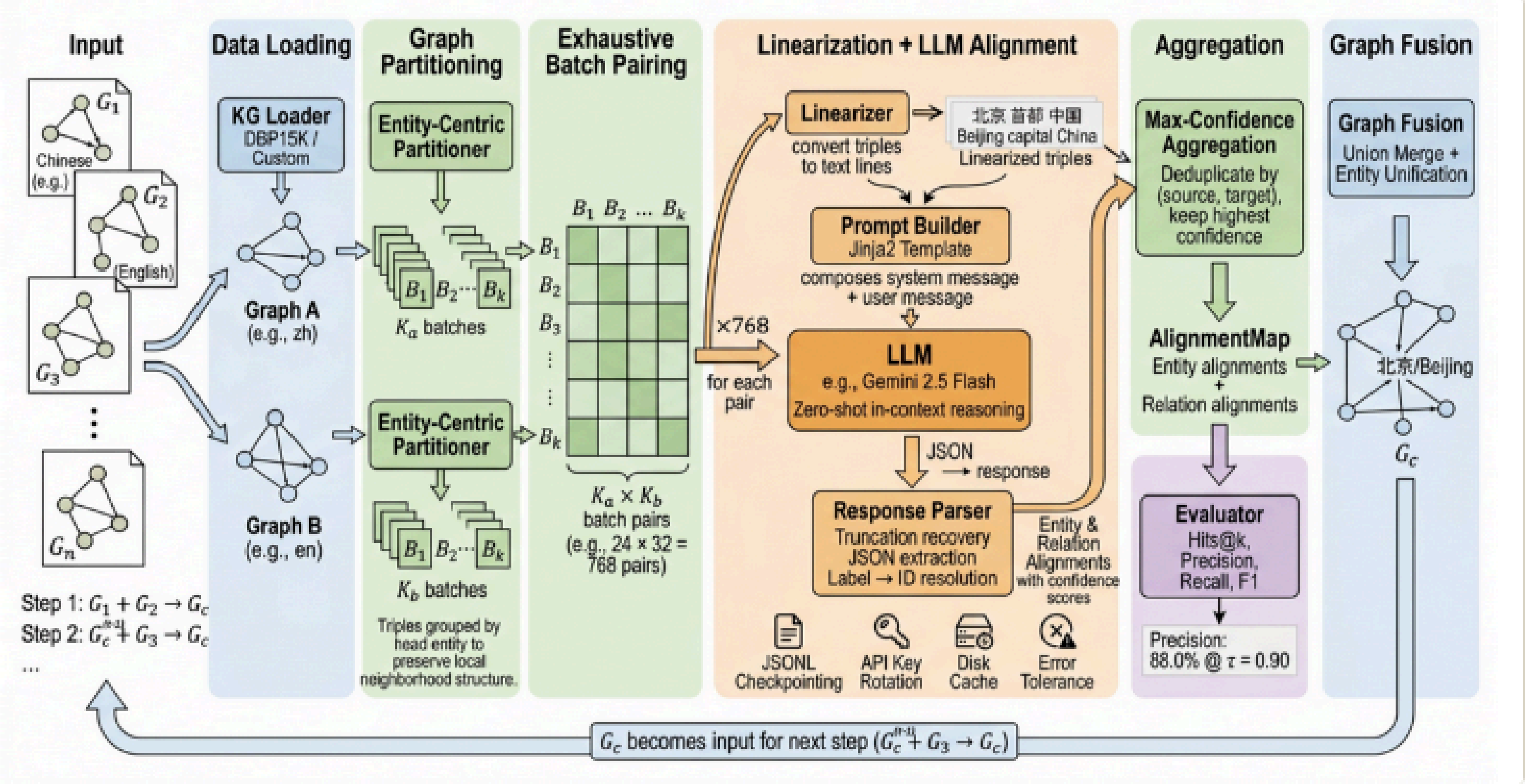

Figure 1: End-to-end pipeline of the modular framework for N-graph fusion. Input KGs are partitioned by entity, paired exhaustively, linearized for the LLM (Gemini 2.5 Flash, zero-shot), parsed and deduplicated by max-confidence, then merged into the evolving fused graph G_C. The fused graph becomes input to the next step ($G_C^{(t-1)} + G_t \rightarrow G_C$).

5. METHODOLOGY DETAILS

1. **Entity-Centric Partitioning:** Preserve topological context by batching triplets sharing the same head entity (K_a and K_b batches).
2. **Exhaustive Batch Pairing:** Compare every batch pair (B_i, B_j) — Cartesian product yields $K_a \times K_b$ pairs (e.g., 24 × 32 = 768).
3. **Linearization:** ⟨h, r, t⟩ → "[head] [relation] [tail]" with surrounding context for the LLM.
4. **LLM Reasoning:** Zero-shot in-context reasoning via Jinja2 prompt template; output constrained to JSON with confidence scores.
5. **Response Parsing:** Truncation recovery, JSON extraction, label↔ID resolution; JSONL checkpointing & API key rotation for resilience.
6. **Max-Confidence Aggregation:** Deduplicate by (source, target), keep highest-confidence prediction → AlignmentMap.
7. **Graph Fusion:** Union merge + entity unification produces fused graph G_c; G_c becomes input to the next step.

6. RELATED WORK

Traditional EA Methods

Embedding/translation models (TransE, MTransE, JAPE) and GNN-based aligners (GCN-Align, RDGCN) rely on seed alignments and suffer in low-resource settings.

LLM-based EA (Recent)

- **ZeroEA** (Huo et al., 2024): Zero-training EA with PLMs.
- **Seg-Align** (Yang et al., 2024a): Sample segmentation + zero-shot prompts.
- **ProLEA** (Munne et al., 2025): Entity profile generation with LLMs.

Multi-Graph Alignment / Our Work

MultiEA (Yang et al., 2024b) aligns multiple KGs in one pass but faces transitive error propagation. We instead tackle **sequential N-graph fusion** via a modular pipeline that combines LLM reasoning, entity-centric partitioning, exhaustive pairing, and max-confidence aggregation — without any training data.

7. EXPERIMENTS AND RESULTS

7.1 Experimental Setup

Dataset:	DBP15K (zh_{en}) — 15,000 ground-truth pairs
Zero-Shot:	No training data used
Model:	Gemini 2.5 Flash ($T = 0.0$)
Runtime:	5–6 hours end-to-end
Infra:	Google Cloud (Vertex AI)

7.2 Primary Alignment Metrics

Table 1: Primary alignment metrics (DBP15K zh_{en})

Metric	Value
Total Batch Pairs	768
Unique Predictions	5,416
True Positives (TP)	3,543
False Positives (FP)	1,873
Precision	65.4%
Recall	23.6%
F1 Score	34.7%

7.3 Hits@k

Table 2: Hits@k

k	Hits	Acc.
1	3,516	88.3%
5	3,543	–
10	3,543	89.0%

7.4 τ Sweep

Table 3: Confidence threshold sensitivity

τ	Prec.	Rec.	F1
0.00	65.4%	23.6%	34.7%
0.80	82.2%	23.6%	36.6%
0.90	**88.0%**	**23.5%**	**37.1%**
0.95	92.4%	21.4%	34.8%

★ Key Takeaways

- At τ = 0.90, framework achieves **88.0% precision** with negligible recall loss.
- When the model predicts, it is correct at **88.3% (Hits@1)**.
- TP mean confidence **0.980**; FP mean confidence **0.738**.
- Recall remains the main challenge due to context-window limits.

8. CONCLUSION

We presented a modular, zero-shot framework for multilingual N-graph fusion. On DBP15K zh_{en}, it achieves **88.0% precision at τ = 0.90** with negligible recall loss — indicating LLMs can act as a universal semantic bridge for synthesizing global knowledge bases without costly seed alignments.

9. FUTURE WORK

- Replace exhaustive pairing with heuristic blocking + semantic indexing.
- Hierarchical prompting for global structural context.
- Hybrid: LLM alignments as silver seeds for GNN aligners.
- Evaluate across more datasets, languages, and evolving LLMs.

* Equal Contribution

Repository: `https://anonymous.4open.science/r/KG-Fusion-1A7D/README.md`