HGAMN - Heterogeneous Graph Attention Matching Network for Multilingual POI Retrieval at Baidu Maps

Source: KDD '21

Speaker: Po-Jen, Wen

Advisor: Jia-Ling, Koh

Date: 2021/10/19

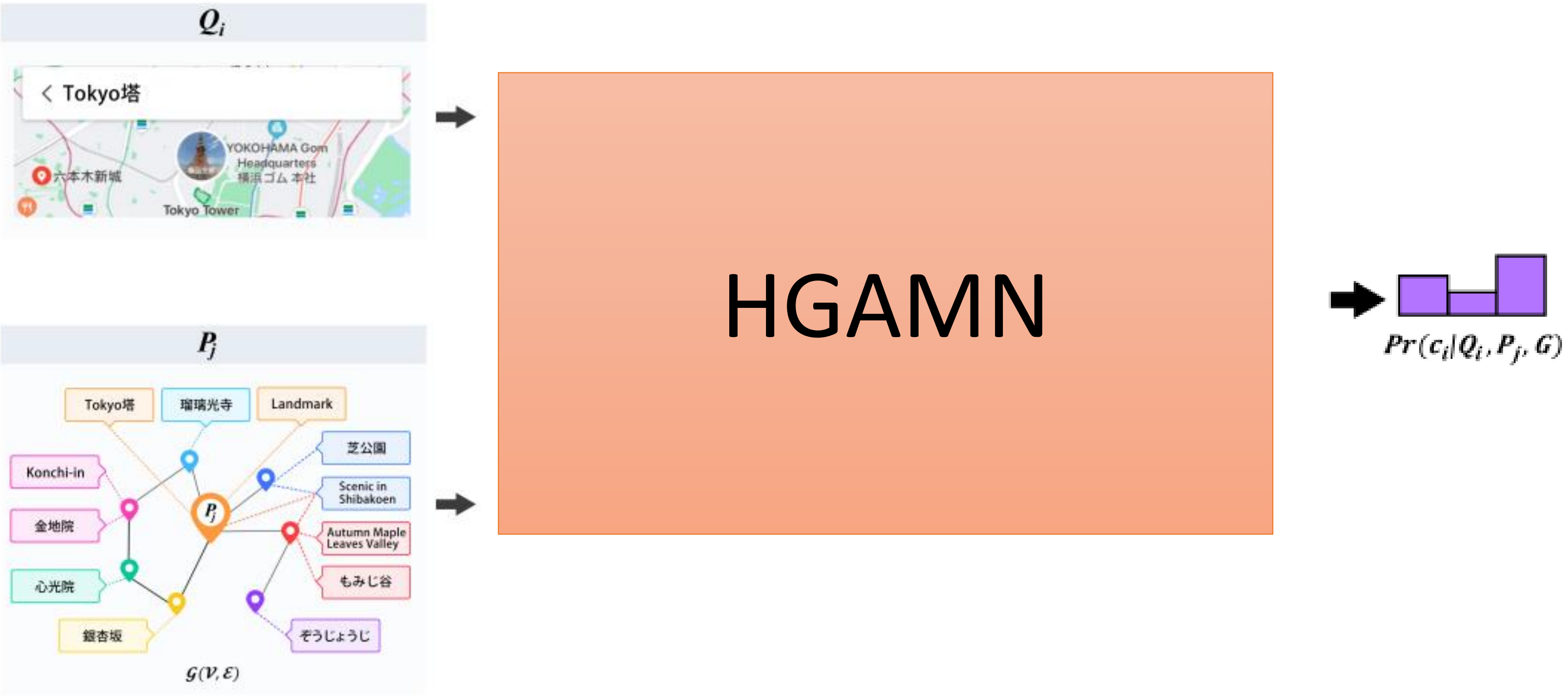
OUTLINE

- Introduction
- Method
- Experiment
- Conclusion

Introduction

The increasing interest in international travel has raised the demand of retrieving point of interests (POIs) in multiple languages.

Input & Output



Query
Candidate POIs
Historical queries

Correlation between p and q

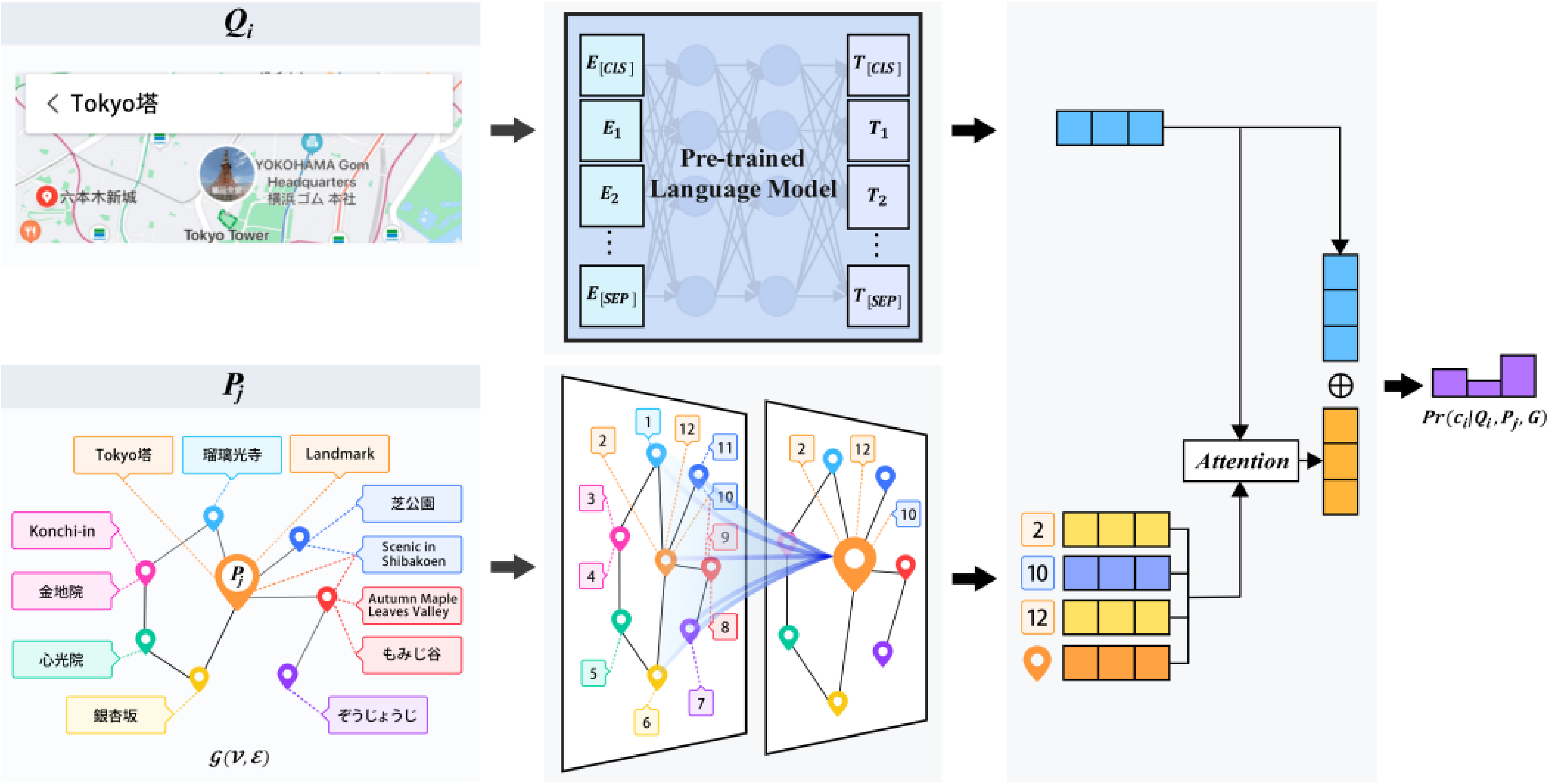
Challenges

1. Visiting Sparsity
    - Statistics show that only 6.4% of the POIs have been clicked by one or more users.
    - The effectiveness of a POI retrieval model would significantly decline when handling the majority of POIs that have sparse click logs.
2. Multilingual Query-POI Matching
    - Most of the users search the overseas POIs by their native languages, which are more likely to be inconsistent with the languages of the target POIs.
    - Queries are sometimes mixed keyboard inputs of multi-languages, which further necessitates multilingual POI retrieval.

OUTLINE

- Introduction
- Method
- Experiment
- Conclusion

Overall framework of HGAMN



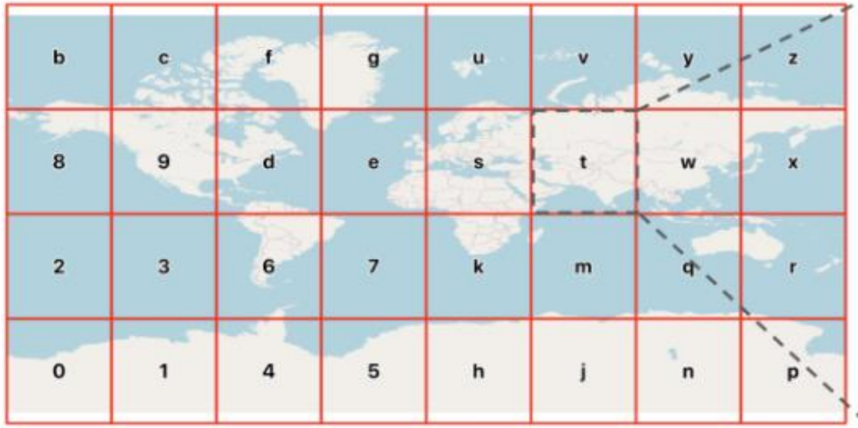
Heterogeneous Graph Attention Matching Network (HGAMN)

Multi-Source Information Learning

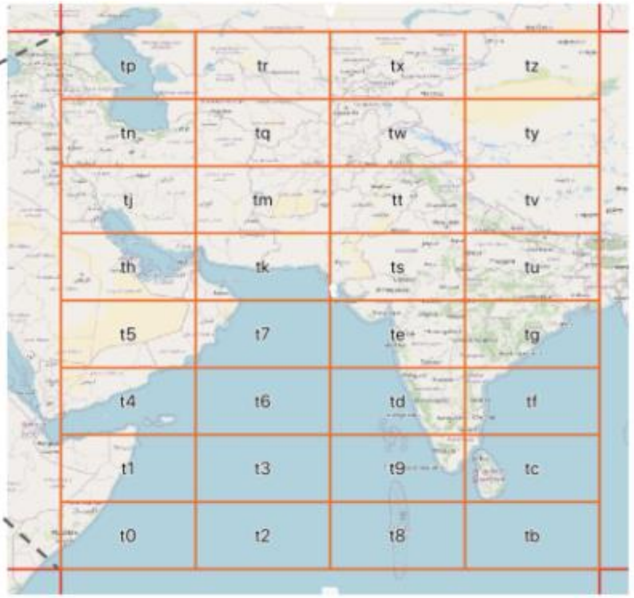
GPS Encoding

$s_{GPS} = \textbf{Geohash}((x_v, y_v))$ the length $|s_{GPS}| \in [1, 12]$

天安門39°54′26.4″N 116°23′27.9″E get s_{GPS} =“wx4g09np9p”



Geohash level-1



Geohash level-2

Multi-Source Information Learning

GPS Encoding

s_{GPS} = "wx4g09np9p"

X = ['[PAD]', '[PAD]', 'w', 'x', '4', 'g', '0', '9', 'n', 'p', '9', 'p']

Transform

character embeddings $\mathbf{X} \in \mathbb{R}^{12 \times d_c}$, where $d_c = 64$

$\widetilde{\mathbf{h}}_t = [\overrightarrow{\mathrm{GRU}}(\mathbf{X}_t); \overleftarrow{\mathrm{GRU}}(\mathbf{X}_t)]$. $\widetilde{\mathbf{h}}_{12}$ is used as the representation of the POI's GPS.

obtain an embedding matrix $\mathbf{G} \in \mathbb{R}^{|\mathcal{P}| \times d}$

Multi-Source Information Learning

Text Encoding

$\mathbf{q} = \text{ERNIE}([c_1, c_2, \ldots, c_L])$, where $[c_1, c_2, \ldots, c_L]$ is the character sequence of the query.

the final representation of a query is represented as: $\widetilde{\mathbf{q}} = \mathbf{q} + \mathbf{G}_u$.

$\mathbf{Q}_{P_i} = [\widetilde{\mathbf{q}}_1, \widetilde{\mathbf{q}}_2, \widetilde{\mathbf{q}}_3, \widetilde{\mathbf{q}}_4]$ the top-4 queries

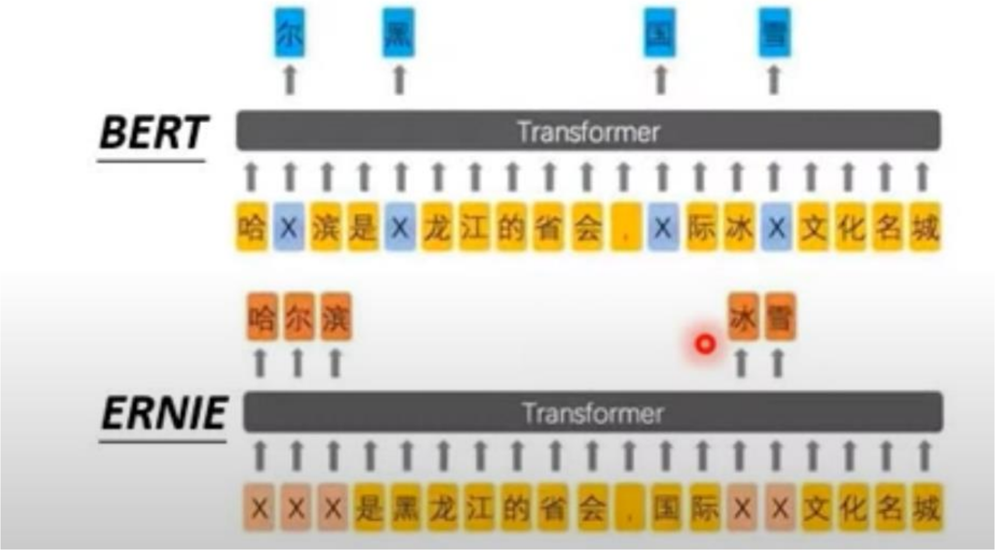
final representation $\mathbf{P}_i \in \mathbb{R}^{1 \times d}$

$\mathbf{P}_i = \text{ERNIE}([\boxed{x_1, x_2, \ldots, x_L,} \boxed{a_1, a_2, \ldots, a_L}]) + \mathbf{G}_i$

character sequence of name

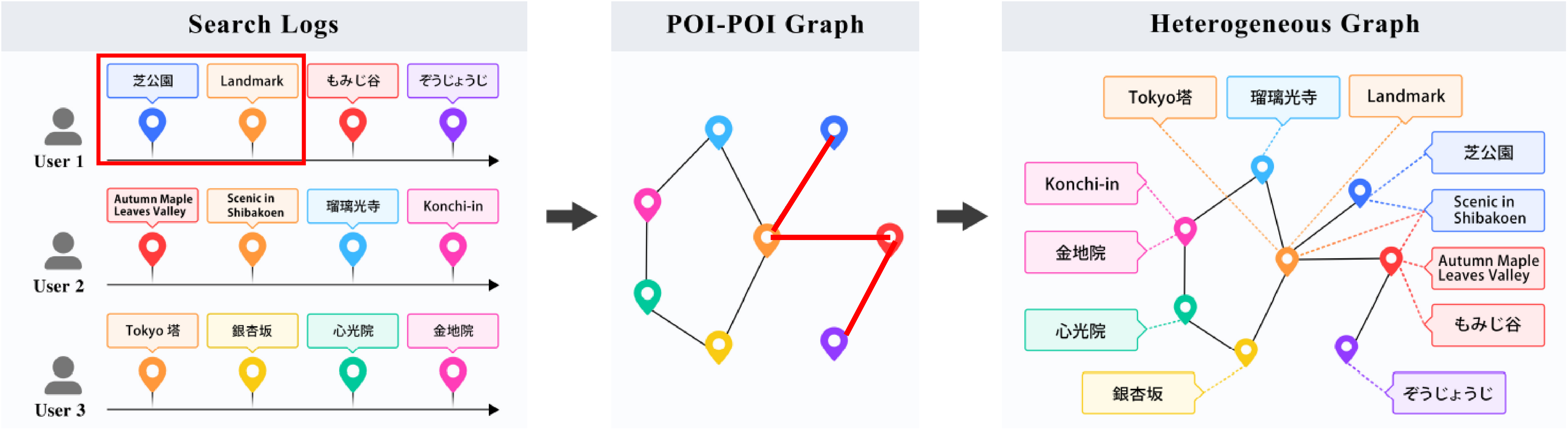
character sequence of address

- Designed for Chinese



Heterogeneous Graph Learning

POI-POI



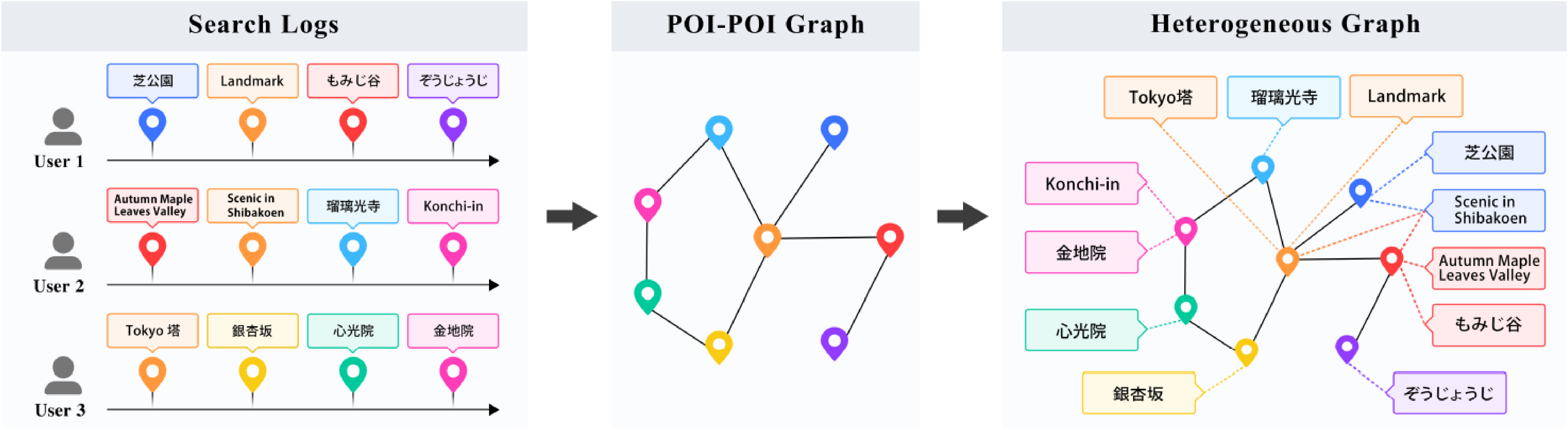
$$A_{ij}^{pp} = \mathrm{PMI}(P_i, P_j) = log \frac{Pr(P_i, P_j)}{Pr(P_i) \cdot Pr(P_j)}$$

$$Pr(P_i, P_j) = \frac{\#W(P_i, P_j)}{\#W},$$

$$Pr(P_i) = \frac{\#W(P_i)}{\#W},$$

Heterogeneous Graph Learning

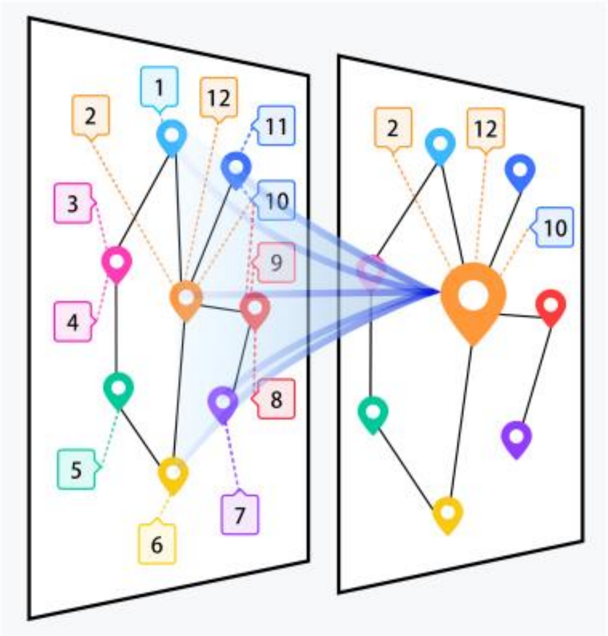
Graph Construction : POI-Query



$$A_{ij}^{pq} = \frac{c_{i,j}}{\sum_{k=1}^{|Q_{P_i}|} c_{i,k}}$$

$c_{i,j}$ is the frequency of query-POI pair (q_j, P_i), $q_j \in Q_{P_i}$

Heterogeneous Graph Learning



Heterogeneous Graph Learning

$$\mathbf{n}_i \in \mathbb{R}^{d_n}$$

$$\mathbf{e}_j^{(k)} = \sigma(\max(\{\mathbf{W}^{(k)}\mathbf{n}_t, n_t \in \mathcal{N}_j\}))$$

$$\mathbf{E}_i = (\mathbf{e}_{i,1}, \mathbf{e}_{i,2}, \ldots, \mathbf{e}_{i,m}) \quad \mathbf{E}_i \in \mathbb{R}^{m \times d_e}$$

$$\boldsymbol{\alpha}_i = \mathbf{softmax}(\mathbf{n}_i \tanh(\mathbf{W}_r \mathbf{E}_i^T) \odot \mathbf{A}_i^{(r)}) \quad \boldsymbol{\alpha}_i \in \mathbb{R}^m \quad \mathbf{W}_r \in \mathbb{R}^{d_n \times d_e} \quad \mathbf{A}_i^{(r)} \in \{\mathbf{A}_{i,1:m}^{pp}, \mathbf{A}_{i,1:m}^{pq}\}$$

$$\tilde{\mathbf{e}}_i = \boldsymbol{\alpha}_i^T \mathbf{E}_i \,, \; \tilde{\mathbf{e}}_i \in \mathbb{R}^{d_e}$$

$$\widetilde{\mathbf{n}}_i = \mathbf{n}_i + \mathbf{W}_1 \tilde{\mathbf{e}}_i + \mathbf{W}_2 \mathbf{n}_i' \qquad \mathbf{W}_1 \in \mathbb{R}^{d_n \times d_e}, \mathbf{W}_2 \in \mathbb{R}^{d_n \times d} \quad \mathbf{n}_i' \in \{\mathbf{Q}_{P_i}, \mathbf{P}_i\}$$

$$\text{POIs: } \widetilde{\mathbf{P}} \in \mathbb{R}^{|\mathcal{P}| \times d_n} \qquad \text{queries: } \widetilde{\mathbf{Q}} \in \mathbb{R}^{|Q| \times d_n}$$

POI Ranker

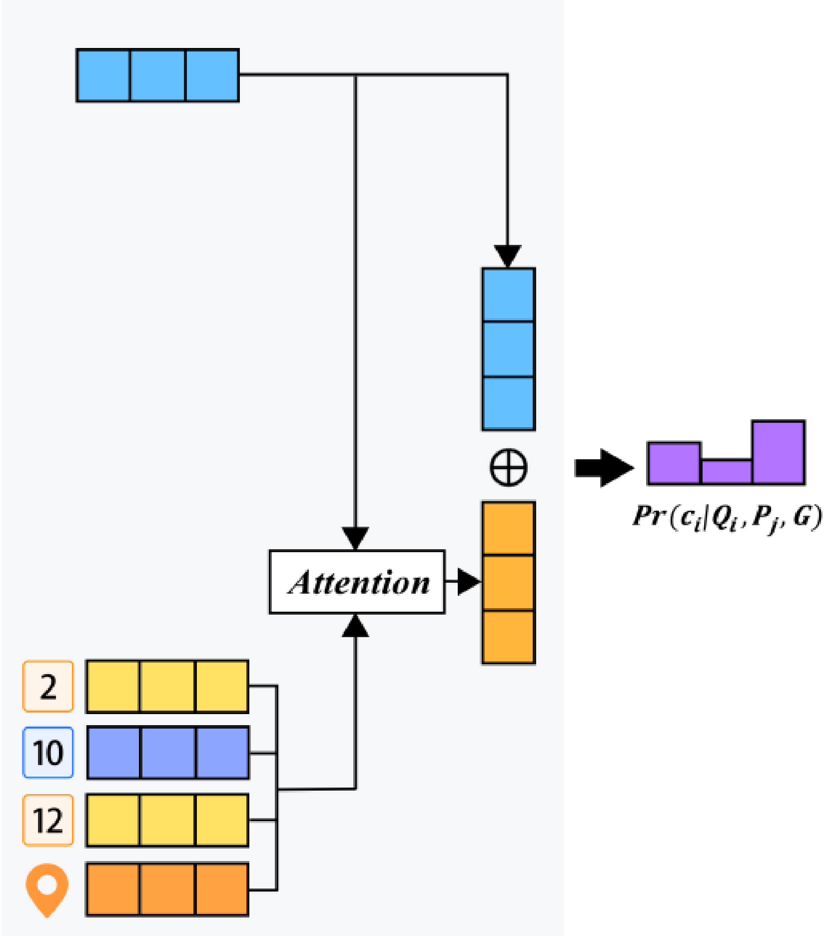
$$\mathbf{M} = [\widetilde{\mathbf{P}}_i, \widetilde{\mathbf{Q}}_{P_i}]$$

$$s_k = \mathbf{W}_4 \tanh([\widetilde{\mathbf{q}}; \mathbf{M}_k]\mathbf{W}_3 + b) \quad \mathbf{W}_3 \in \mathbb{R}^{2d_n \times d_n} \text{ and } \mathbf{W}_4 \in \mathbb{R}^{1 \times d_n}$$

$$\phi_k = \frac{exp(s_k)}{\sum_{j=1}^{|\mathbf{M}|} exp(s_j)},$$

$$\mathbf{m} = \sum_{k=1}^{|\mathbf{M}|} \phi_k \mathbf{M}_k \qquad \mathbf{m} \in \mathbb{R}^{d_n}$$

$$Pr(c_i|q, P_i, \mathcal{G}) = \mathbf{softmax}([\widetilde{\mathbf{q}}; \mathbf{m}]\mathbf{W}_v) \qquad \mathbf{W}_v \in \mathbb{R}^{2d_n \times 2}$$



Loss function

$$\mathcal{L} = -\sum_{i=1}^{|\mathcal{P}|} y_i \mathbf{log} Pr(c_i|q, P_i, \mathcal{G})$$

OUTLINE

- Introduction
- Method
- Experiment
- Conclusion

Datasets and baselines

- Datasets: Baidu Maps

Subset	#(Queries)	#(Candidate POIs)/#(Queries)
Train	11,935,730	2.7
Valid	73,255	2.8
Test	181,589	11.6
Total	12,190,574	2.8

Experiment results

Model	**Evaluation Metrics (Offline)**						
	MRR	nDCG@1	nDCG@3	nDCG@10	SR@1	SR@3	SR@10
DSSM [16]	0.6681	0.5235	0.6634	0.7258	0.5235	0.7705	0.9324
ARC-I [8]	0.6604	0.5127	0.6561	0.7206	0.5127	0.7665	0.9337
Conv-DSSM [26]	0.6485	0.4985	0.6420	0.7097	0.4985	0.7515	0.9281
DPSM [35]	0.6181	0.4900	0.6084	0.6655	0.4900	0.7010	0.8491
PALM [34]	0.6921	0.5488	0.6902	0.7500	0.5488	0.7993	0.9521
HGAMN	0.7663	0.6539	0.7653	0.8097	0.6539	0.8528	0.9636
w/o POI-POI Graph	0.7655	0.6527	0.7648	0.8091	0.6527	0.8526	0.9628
w/o POI-Query Graph	0.7573	0.6408	0.7557	0.8030	0.6408	0.8455	0.9640
w/o Heterogeneous Graph	0.6924	0.5451	0.6921	0.7507	0.5451	0.8052	0.9540
LTR	0.8253	0.7323	0.8294	0.8582	0.7323	0.9030	0.9721
LTR + HGAMN	**0.8307**	**0.7393**	**0.8347**	**0.8627**	**0.7393**	**0.9072**	**0.9743**

Experiment results

Model	Evaluation Metrics (Online)		
	SR@1	SR@3	SR@10
DSSM	0.4847	0.7130	0.8358
ARC-I	0.4668	0.7024	0.8349
PALM	0.4900	0.7010	0.8491
LTR	0.6647	0.8189	0.8802
LTR + HGAMN	**0.7173**	**0.8807**	**0.9437**

OUTLINE

- Introduction
- Method
- Experiment
- Conclusion

Conclusion

- Propose a heterogeneous graph attention matching network (HGAMN) to address the visiting sparsity and multilingual query-POI matching problems
- HGAMN is composed of three modules:
  1. A multi-source information learning module, which learns the text and location representations of the multilingual query, POI name, and POI address
  2. A heterogeneous graph learning module, which constructs the connections of different POIs and historical queries, and learns the node representations from the heterogeneous graph
  3. POI ranker module, which calculates the relevance between a query and candidate POIs.