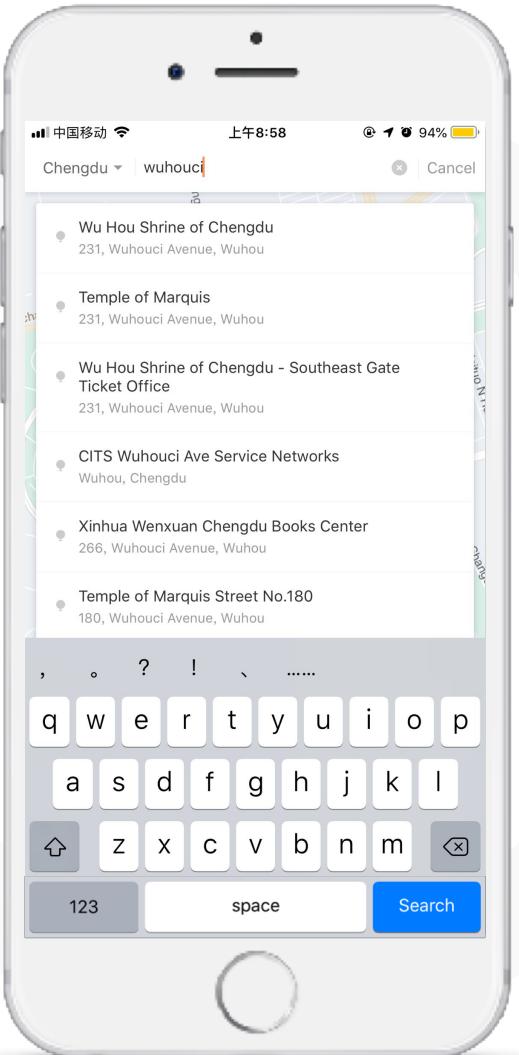


# Incorporating Semantic Similarity with Geographic Correlation for Query-POI Relevance Learning

Ji Zhao, Dan Peng, Chuhan Wu, Huan Chen, Meiyu Yu, Wanji Zheng,  
Li Ma, Hua Chai, Jieping Ye, and Xiaohu Qie

**DiDi Chuxing**

# POI Retrieval on ride-hailing App



MORE THAN A JOURNEY

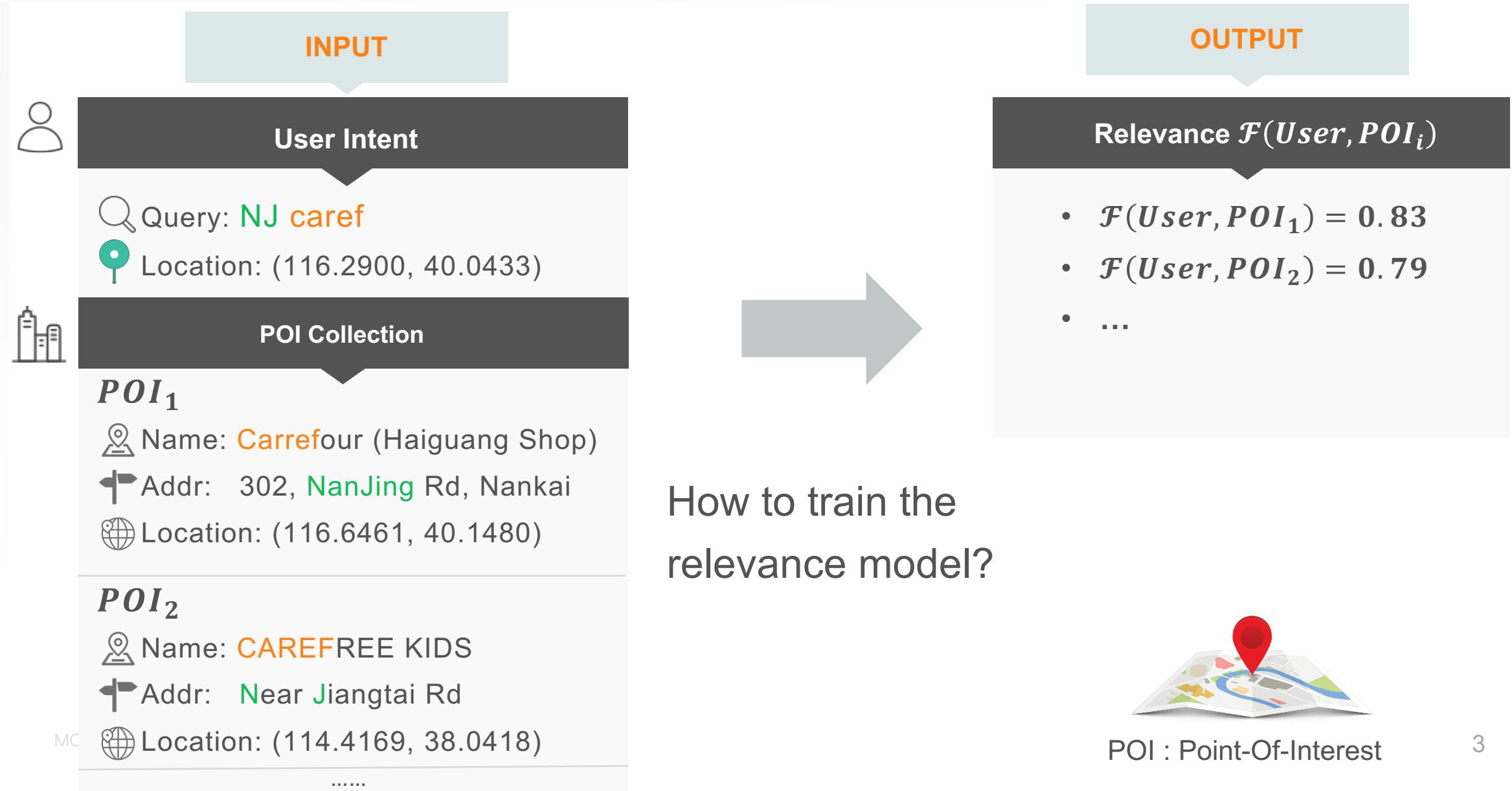
## Important: Critical Steps to Service Delivery

- Finding destination is the **first** and key step of rides
- Affecting **billions** of rides
- **Extendable** to other Location-Based Service
  - eg. Hotel&Travel, Food Delivery, Package Delivery

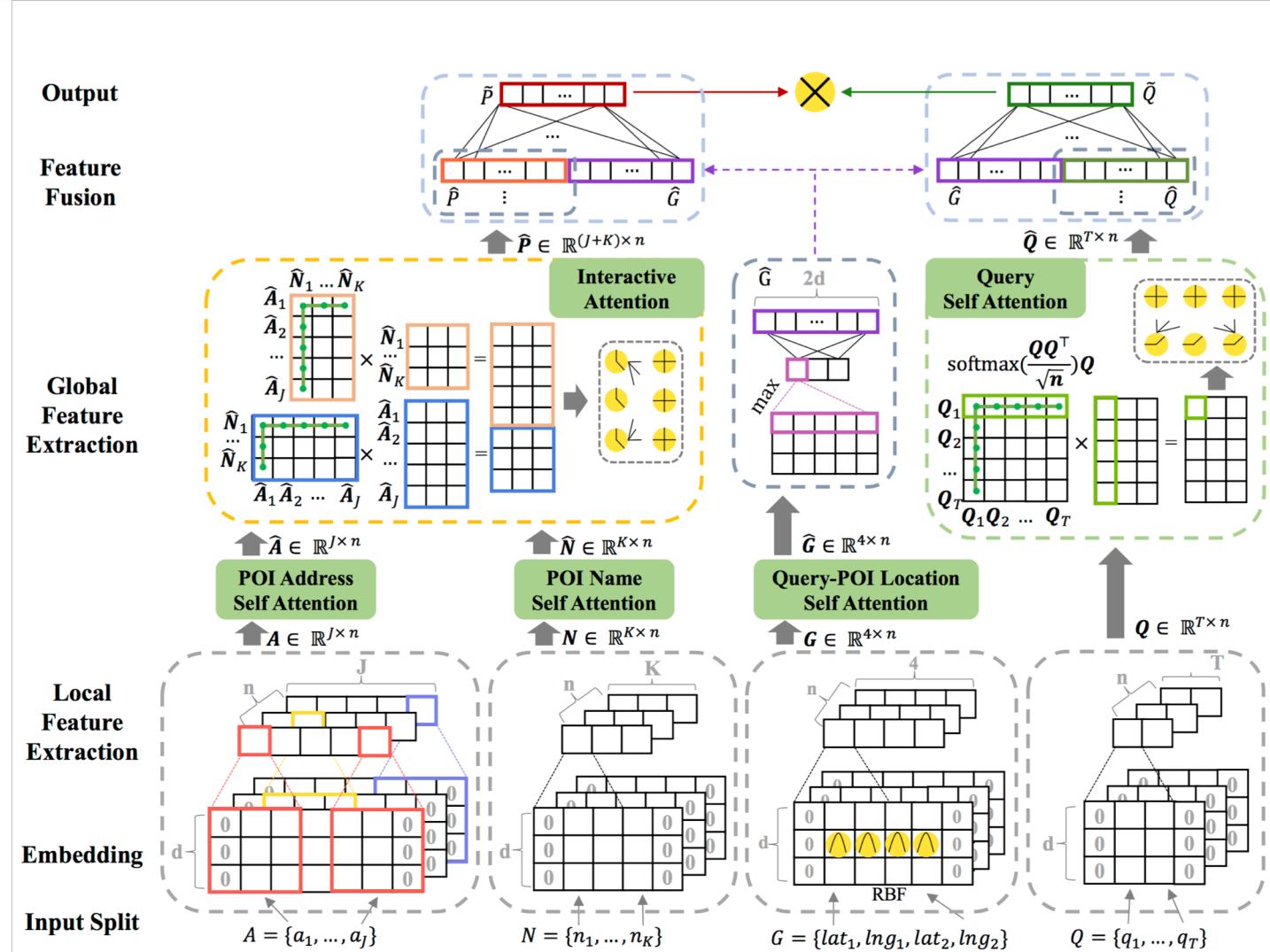
## Challenging: Gap between User Intent & POI Info

- **Short text learning:** incomplete Query and various order of semantic unit
- **Multi-text fusion and representing:** POI matching based on multi-field textual attributes
- **Geographic learning:** location correlation of User and POI

# Problem Definition



# POI Attention Location Model+(PALM+)



## Semantic Similarity

- Multi-field texts
- Local & Global features

## Geographic Correlation

- Location embedding
- Geographic features

# Semantic Representation - I

## Multi-field Textual Attributes

- POI Address + POI Name + Query

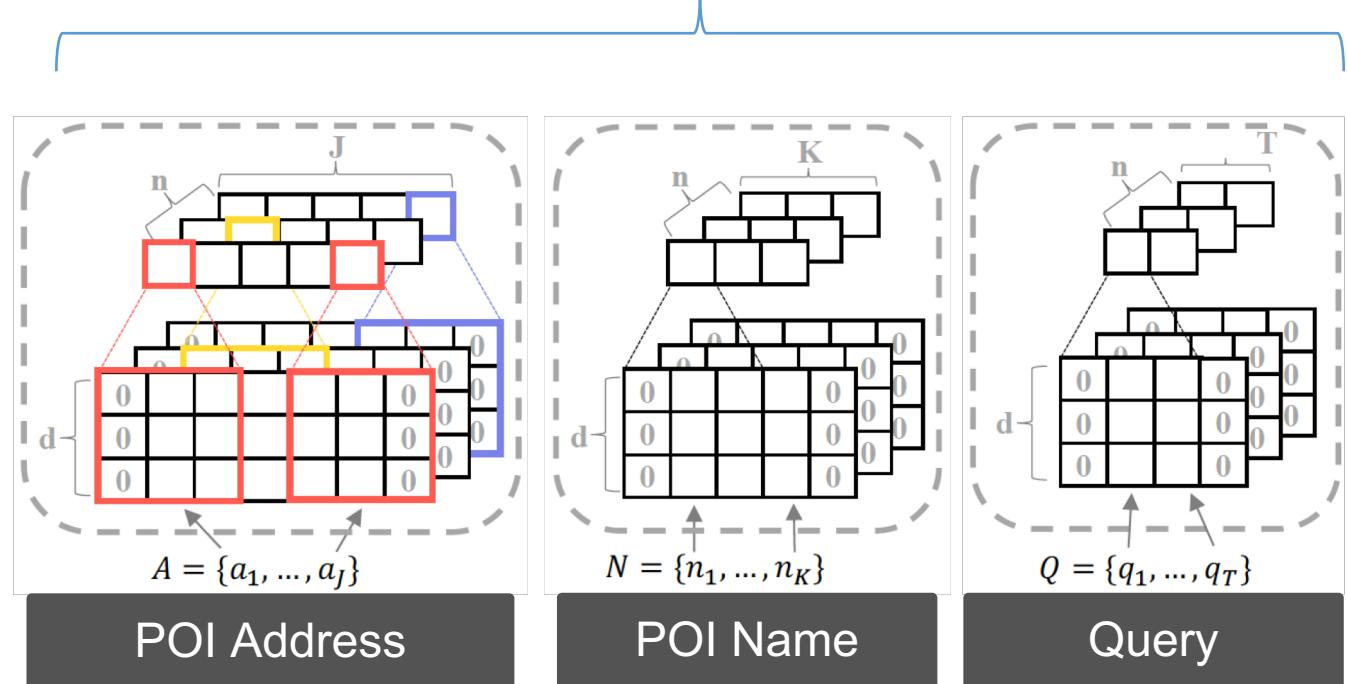
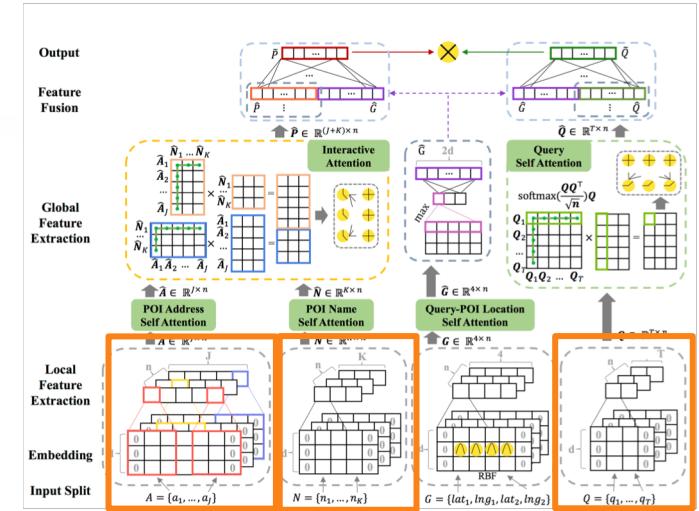
## Multi-granularity Embeddings

- Letter + Word

	Chinese	English
letter	中	a
word	中国	apple

## Convolution

- Capture the local semantic features



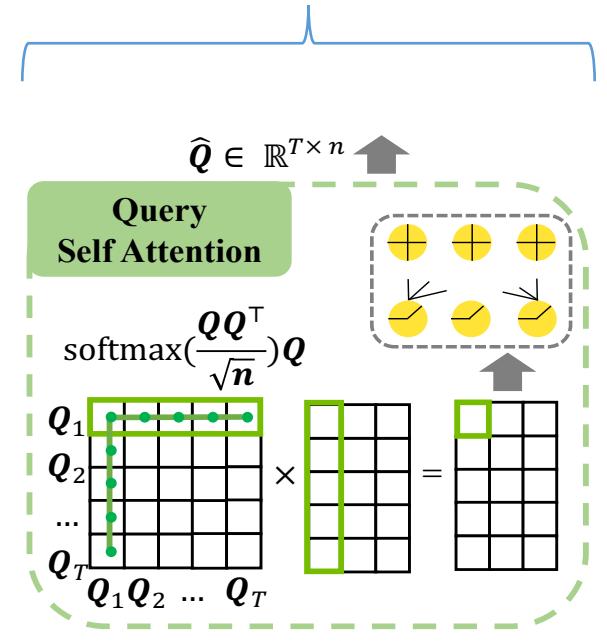
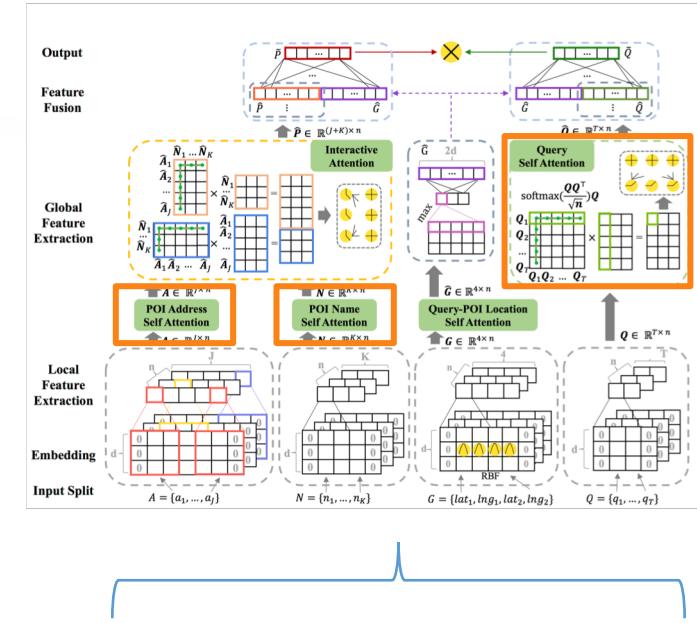
# Semantic Representation - II

## Self-Attention mechanism

- Identify the key information
- Intra-dependency within the texts

$$X_q = \text{softmax} \left( \frac{Q Q^T}{\sqrt{n}} \right) Q$$

$$\widehat{Q} = \max(0, X_q \times W_{q1} + B_{q1}) \times W_{q2} + B_{q2}$$



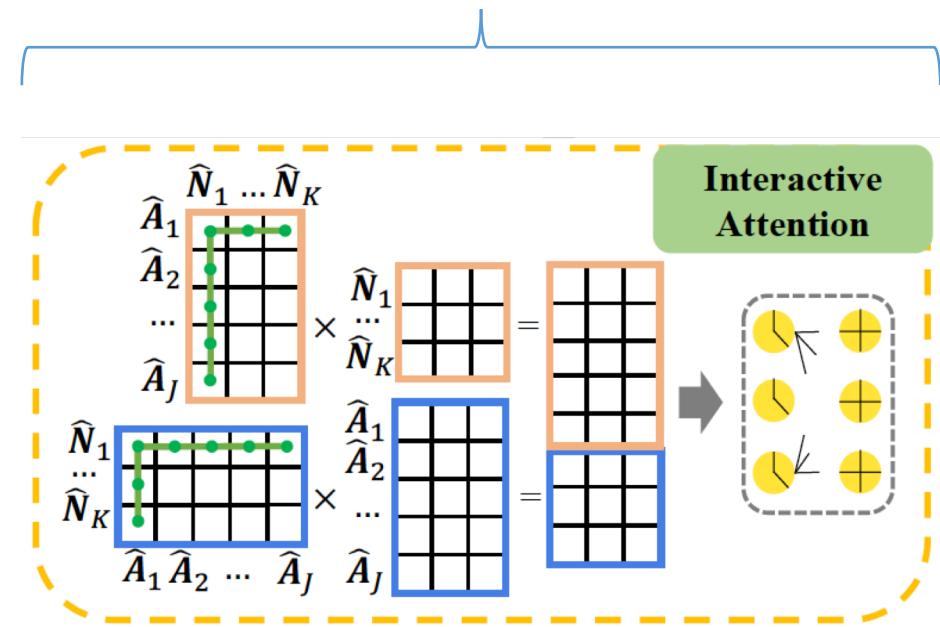
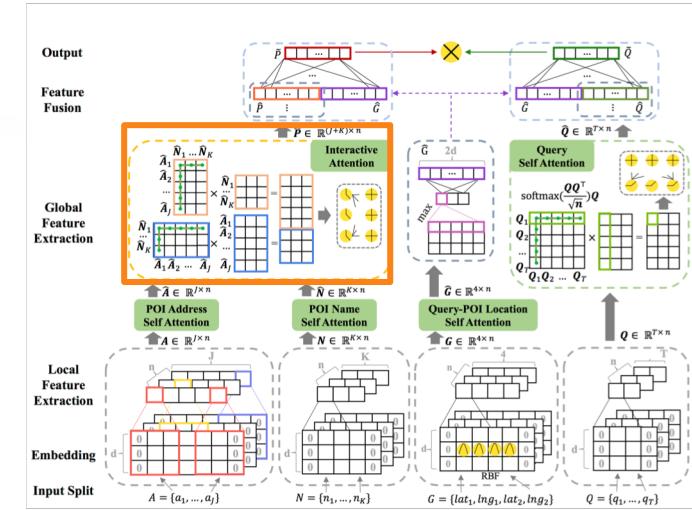
# Semantic Representation - III

## Interactive Attention mechanism

- Fusion information from multi-field attributes
- Highlights keywords based on correlation between Address and Name

$$\widehat{N2A} = \text{softmax} \left( \frac{\widehat{A}\widehat{N}^T}{\sqrt{n}} \right) \widehat{N}$$

$$\widehat{A2N} = \text{softmax} \left( \frac{\widehat{N}\widehat{A}^T}{\sqrt{n}} \right) \widehat{A}$$



# Example: Interactive Attention

Strengthen semantic expression by  
highlighting keywords according to another field

■ Example:

POI Name: **Hutaoli** Music Restaurant & Bar (Joy City)

POI Address: Dayue Rd No. 518, Joy City, 1F-**J01**

	Hutaoli	Music	Rest- aurant	Bar	Joy City
Dayue Rd	0.35	0.42	0.07	0.03	0.10
No. 518	0.66	0.06	0.01	0.00	0.25
Joy City	0.56	0.12	0.03	0.00	0.26
1F	0.73	0.00	0.00	0.00	0.25
J01	0.65	0.09	0.04	0.01	0.18

(a) Interactive attention: Name2Addr

	Dayue Rd	No. 518	Joy City	1F	J01
Hutaoli	0.28	0.02	0.11	0.03	0.55
Music	0.76	0.00	0.05	0.00	0.17
Restaurant	0.57	0.00	0.07	0.00	0.34
Bar	0.61	0.00	0.03	0.00	0.34
Joy City	0.27	0.02	0.16	0.03	0.49

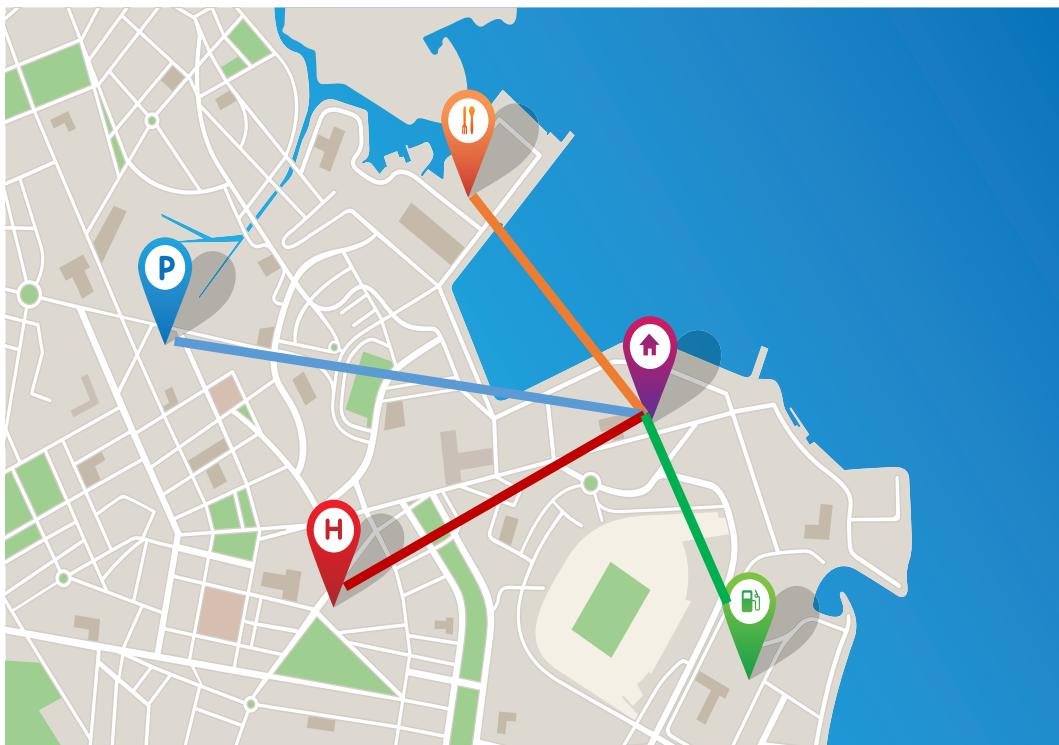
(b) Interactive attention: Addr2Name

# Geographic Correlation - I

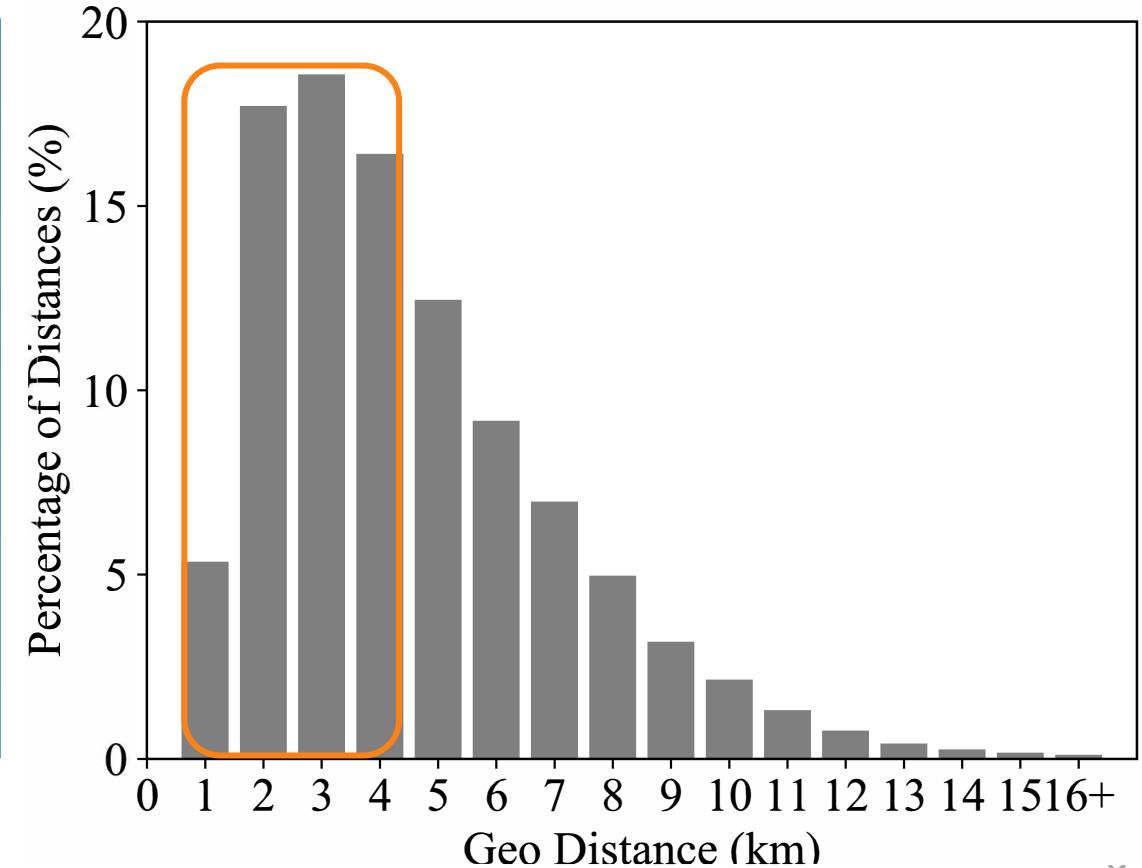
Users are highly sensitive to **Origin-Destination distances**

Hidden correlation of clicked Query-POI pairs

- Over 50% O-D distances are **less than 4km**



MORE THAN A JOURNEY



# Geographic Correlation - II

**Split Coordinates to reduce embedding size**

## Location Embeddings

- One-hot Vector
- Coordinate Embedding
- Kernel Embedding to avoid Boundary Effect

$$\text{Latitude: } \widehat{\Phi}_i = w_{i-1} \widehat{\Phi}_{i-1} + w_i \widehat{\Phi}_i + w_{i+1} \widehat{\Phi}_{i+1}$$

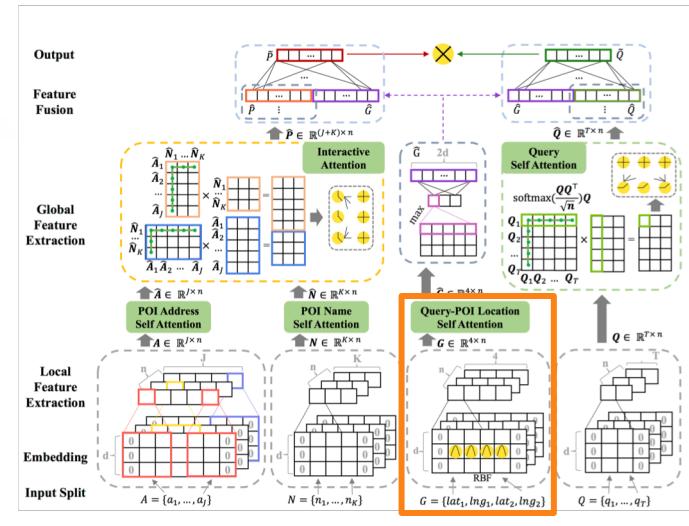
$$\text{Longitude: } \widehat{\Psi}_j = w_{j-1} \widehat{\Psi}_{j-1} + w_j \widehat{\Psi}_j + w_{j+1} \widehat{\Psi}_{j+1}$$

$$\text{RBF: } w = \frac{(dis - \mu)^2}{\sigma^2}, \mu = 0, \sigma^2 = 0.3$$

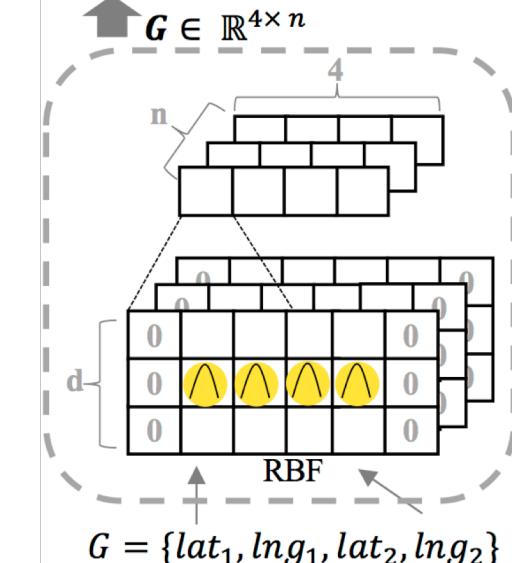
## Geographic Features

- Convolution + Self-Attention

MORE THAN A JOURNEY



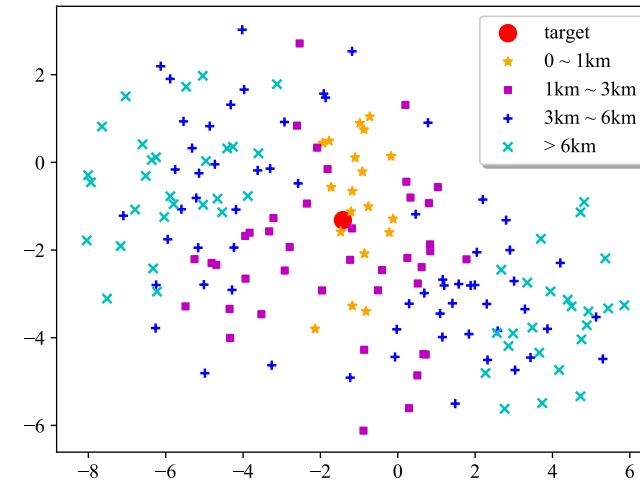
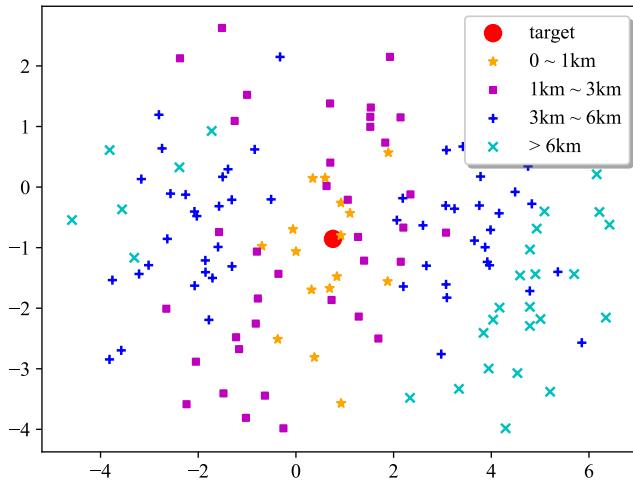
### Query-POI Location Self Attention



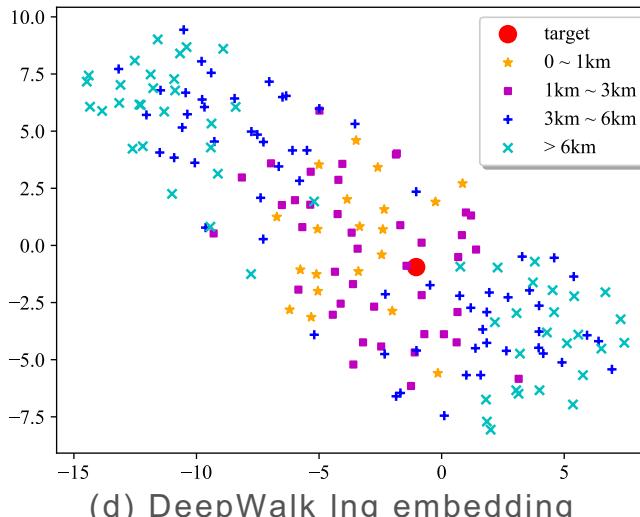
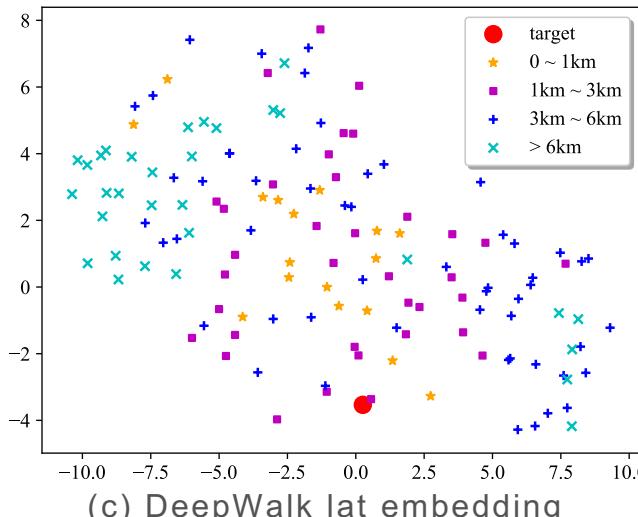
# Example: Location Embedding

Faithfully preserve the physical relation

- PALM



- Deep Walk



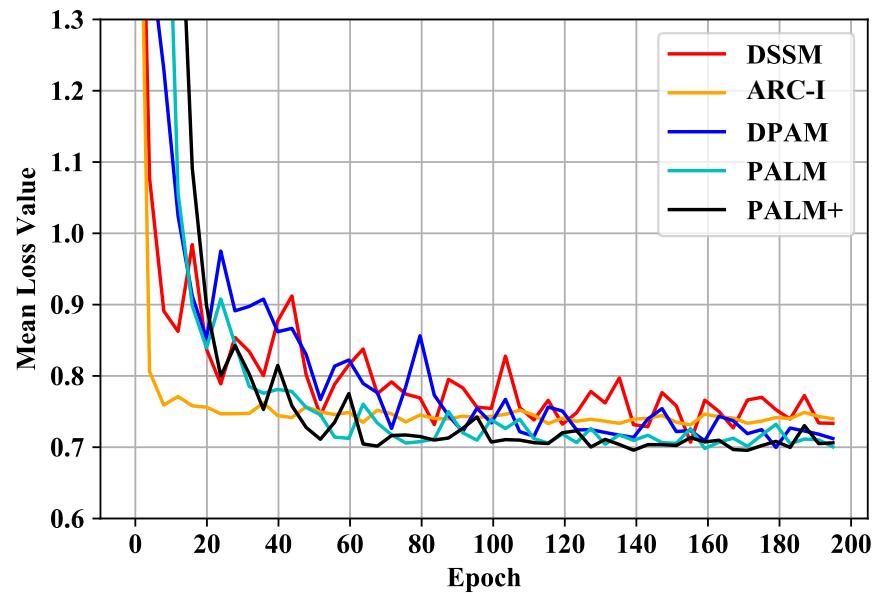
# Dataset

POI retrieval data and click behavior collected by DiDi APP for one month

As part of GAIA Initiative, Dataset A is **publicly available** to the academic community  
<https://outreach.didichuxing.com/appEn-vue/POI?id=11>

	Dataset A (Chengdu)		Dataset B (Nationwide)	
	Training	Testing	Training	Testing
Total Num of Query	115,724	15,261	1,476,645	112,747
Total Num of POIs	711,824	95,369	12,654,847	947,878
Avg Num of Recalls	6.15	6.24	8.57	8.41
Avg Len of Query	4.61	4.65	3.21	3.24
Avg Len of POI Addr	19.04	19.05	18.93	18.89
Avg Len of POI Name	9.38	9.42	7.87	7.96
Avg P-Q Distance (km)	4.05	4.04	7.28	8.53

# Experiments: Results



	Dataset A (Chengdu)		Dataset B (Nationwide)	
	NDCG@3	NDCG@10	NDCG@3	NDCG@10
DSSM	0.8246	0.8989	0.7617	0.8810
ARC-I	0.8298	0.9024	0.7558	0.8788
DPAM	0.8383	0.9058	0.7822	0.8907
PALM	0.8407	0.9050	0.8116	0.9022
<b>PALM+</b>	<b>0.8465</b>	<b>0.9110</b>	<b>0.8124</b>	<b>0.9027</b>

**DSSM:** DNN, semantic + word hashing

**ARC-I:** CNN, semantic + pre-trained Word2Vec

**DPAM:** CNN, semantic + attention mechanisms + text embedding

**PALM:** DPAM + geographic + coordinate embedding

**PALM+:** PALM + kernel embedding

# Summary

## A novel Query-POI relevance model for effective POI retrieval

- Enriched semantic similarity

- via attention mechanism

- Integrated geographic correlation

- with location kernel embedding

- Extensive experiments

- achieve 5pp NDCG@3 improvement on real-world large-scale click-through datasets

- Open dataset

- dataset is publicly available to the academic community

# THANKS

Q & A

