Points-of-Interest Relationship Inference with Spatial-enriched Graph Neural Networks

Proceedings of the VLDB Endowment
November 2021

Yile Chen[1], Xiucheng Li[1], Gao Cong[1], Cheng Long[1], Zhifeng Bao[2], Shang Liu[1],
Wanli Gu[3], Fuzheng Zhang[3]

[1]Nanyang Technological University [2]RMIT University [3]Meituan

Agenda

- Background & Related Work & Issue
- Problem & Methodology
- Experiments
- Q&A

1

Background & Related Work & Issue

Background – PoIs & PoI Relationship

PoIs

- A specific point location that someone may find useful or interesting.
- Have coordinates & attributes (e.g., category, human flow, rating, ……)
- Example:
  - Eiffel Tower
  - A hotel
  - A restaurant
  - A beverage shop
  - ……

Background – PoIs & PoI Relationship

PoI Relationship

- POIs are correlated under different relationships

- Example:
  - competitive relationship
  - complementary relationship

Background – PoI Relationship Inference

PoI Relationship Inference

- Bring significant benefits for different groups of people

- Example:
  - For business owners: Design targeted operation strategies
  - For customers: be recommended with places of their interests based on complementary POIs
  - For government: understand regional functionality, and hence make sustainable urban plannings

Intuitive Solutions & Related Work

Infer the relationships between entities

- Early Methods:

  - Solution 1: Focus on modeling text content

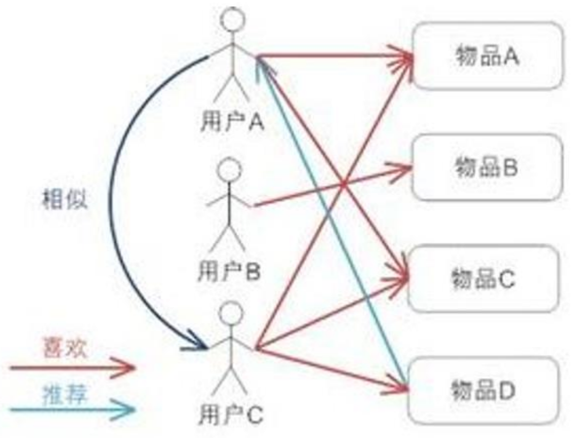
    Kaiquan Xu et al. “Mining comparative opinions from customer reviews for Competitive Intelligence” decision support systems (2011): n. pag.

  - Solution 2: Focus on using social networks

Intuitive Solutions & Related Work

Infer the relationships between entities

- Some Thought: Recall Recommendation System



Intuitive Solutions & Related Work

Infer the relationships between entities

- Recent Methods: Graph representation learning methods

  - Heuristic path constraints

    Zihan Wang et al. "A Path-constrained Framework for Discriminating Substitutable and Complementary Products in E-commerce" web search and data mining (2018): n. pag.

  - Node proximity

    Shijie Zhang et al. "Inferring substitutable products with deep network embedding" international joint conference on artificial intelligence (2019): n. pag.

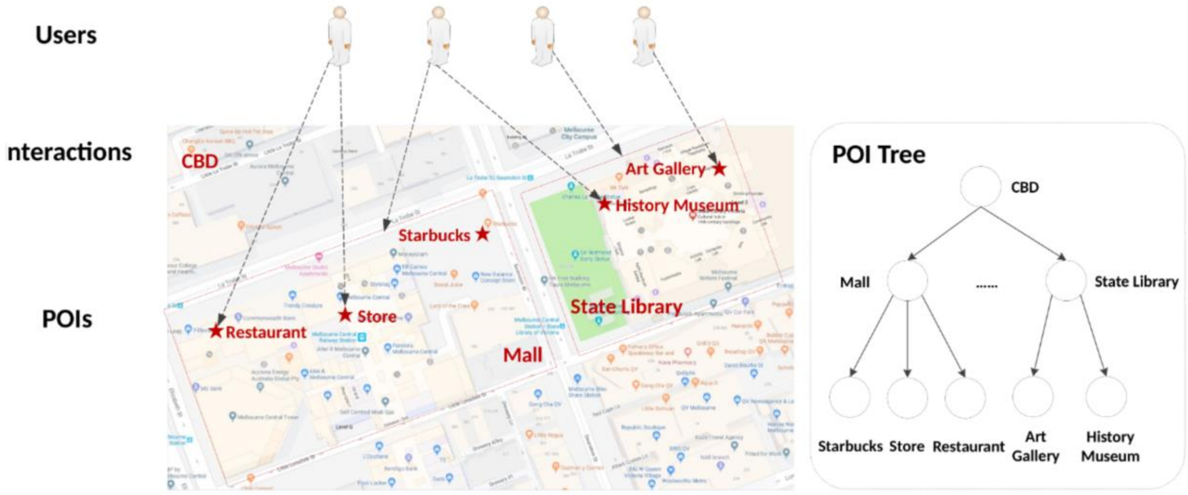
  - Graph Neural Networks

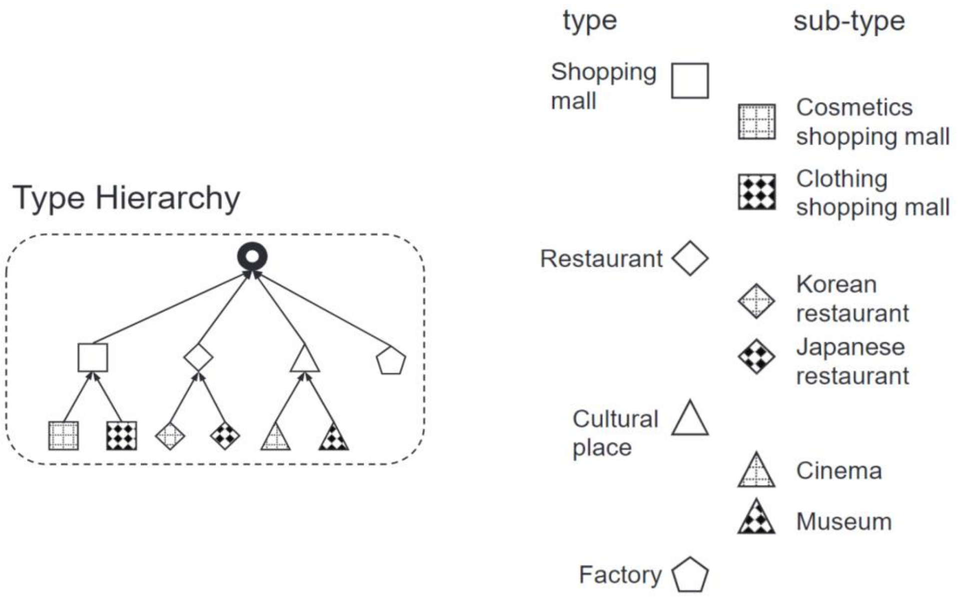
    Yiding Liu et al. "Decoupled Graph Convolution Network for Inferring Substitutable and Complementary Items" conference on information and knowledge management (2020): n. pag.

Intuitive Solutions & Related Work

Infer the relationships between entities

- Some Thought: Knowledge Graph?

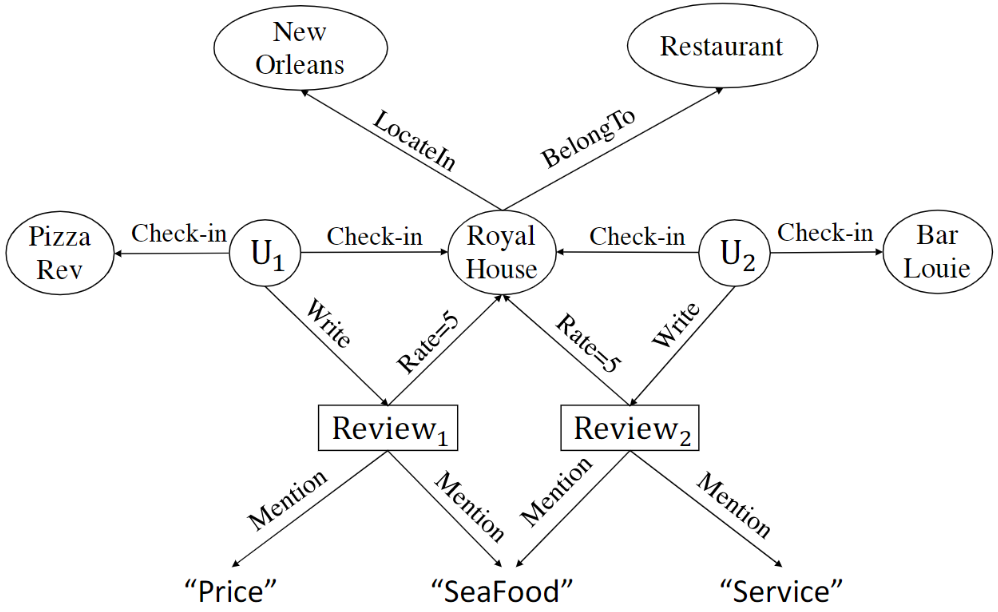




Intuitive Solutions & Related Work

Infer the relationships between entities

- Some Thought: Meta-path?



Intuitive Solutions & Related Work

Infer the relationships between entities

- Recent PoI-specific Methods
  - Extract POI pairwise features & send these features to neural networks to infer the relationship
    Jingbo Zhou et al. "Competitive Relationship Prediction for Points of Interest: A Neural Graphlet Based Approach"
    IEEE Transactions on Knowledge and Data Engineering (2021): n. pag.
  - Use GNN to aggregation different geographical sectors, brand and aspect knowledge
    Shuangli Li et al. "Competitive Analysis for Points of Interest"
    knowledge discovery and data mining (2020): n. pag.

Issue

Issue 1: Some methods only focus on a particular type of relationship

- Recent POI-specific methods consider the competitiveness relationship only.

- Intuitive Solutions:
  - Use multiple sub-graphs
  - Each sub-graph containing only one relation type
- Shortcoming of intuitive solutions:
  - Fail to model the inherent interactions of POIs under different types of relationships

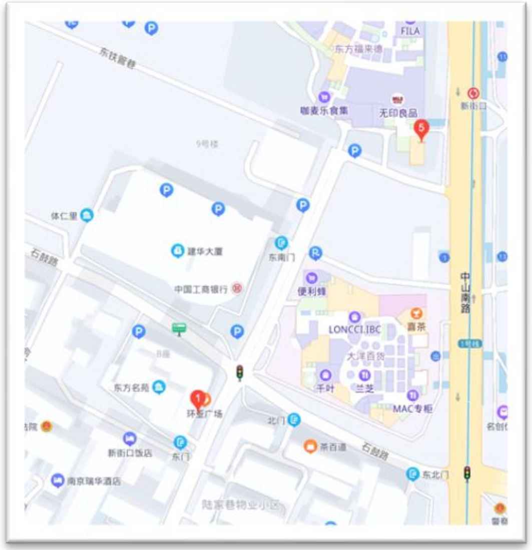
- Remarks:
  - We can use GNNs on Heterogeneous Graphs & multi-view joint learning

    Mingyang Zhang et al. "Multi-View Joint Graph Representation Learning for Urban Region Embedding" international joint conference on artificial intelligence (2020): n. pag.

Issue

Issue 2: Spatial features have important effects on POI relationships

- A strong competitive relationship if they are close
- But have no relationship if far away



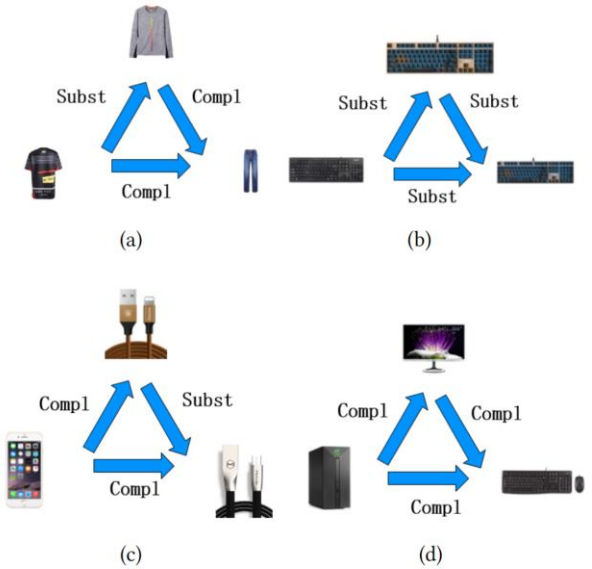
- Remarks: Maybe some previous methods have implicitly learned

Issue

Issue 3: Hard to use structural knowledge

- Enhance the model with structural knowledge
- Require either heavy processing or efforts on manually designed constraints
  - (e.g., fuzzy logics or graphlet patterns)

Zihan Wang et al. "A Path-constrained Framework for Discriminating Substitutable and Complementary Products in E-commerce" web search and data mining (2018): n. pag.



$$(Prod_A, Subst, Prod_B) \wedge (Prod_B, Subst, Prod_C) \Rightarrow (Prod_A, Subst, Prod_C),$$

$$(Prod_A, Compl, Prod_B) \wedge (Prod_B, Subst, Prod_C) \Rightarrow (Prod_A, Compl, Prod_C),$$

$$(Prod_A, Subst, Prod_B) \wedge (Prod_B, Compl, Prod_C) \Rightarrow (Prod_A, Compl, Prod_C),$$

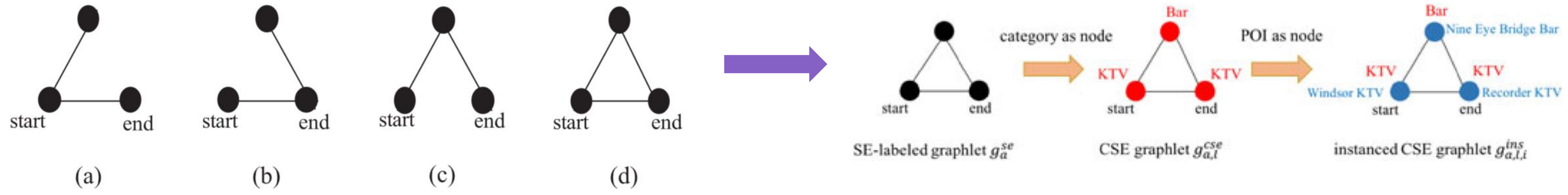
$$(Prod_A, Compl, Prod_B) \wedge (Prod_B, Compl, Prod_C) \Rightarrow (Prod_A, Compl, Prod_C).$$

Issue

Issue 3: Hard to use structural knowledge

- Enhance the model with structural knowledge
- Require either heavy processing or efforts on manually designed constraints
  - (e.g., fuzzy logics or graphlet patterns)

Jingbo Zhou et al. “Competitive Relationship Prediction for Points of Interest: A Neural Graphlet Based Approach”
IEEE Transactions on Knowledge and Data Engineering (2021): n. pag.



Issue

Issue 3: Hard to use structural knowledge

- Remarks:
  - Meta-path related method?
  - Optimizing with Meta-path Generation and Mining
  - Still need heavy efforts or lead to low efficiency

Issue

Issue 4: Lack considering the context information of target POIs

- Existing methods only focus on its connected POIs that have relationships with it.
- POIs that are spatially close to the target POI can also provide rich context information.

- A POI with shopping centers and entertainment spots nearby
- A residential area
- Different degree of competitive environment

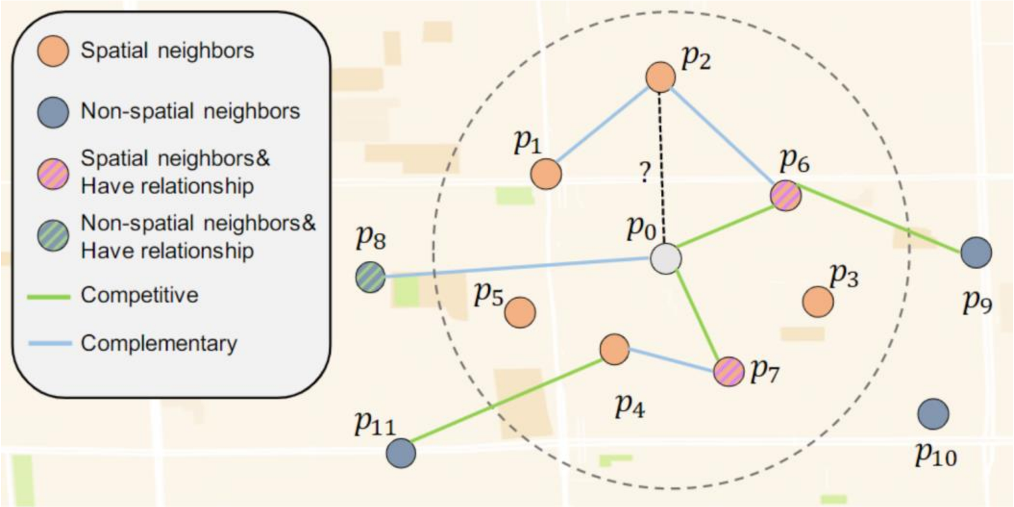
- Remarks:
  - Negative. Most work on urban computing takes this into account.
  - Maybe related to spatial eminence

2

Problem & Methodology

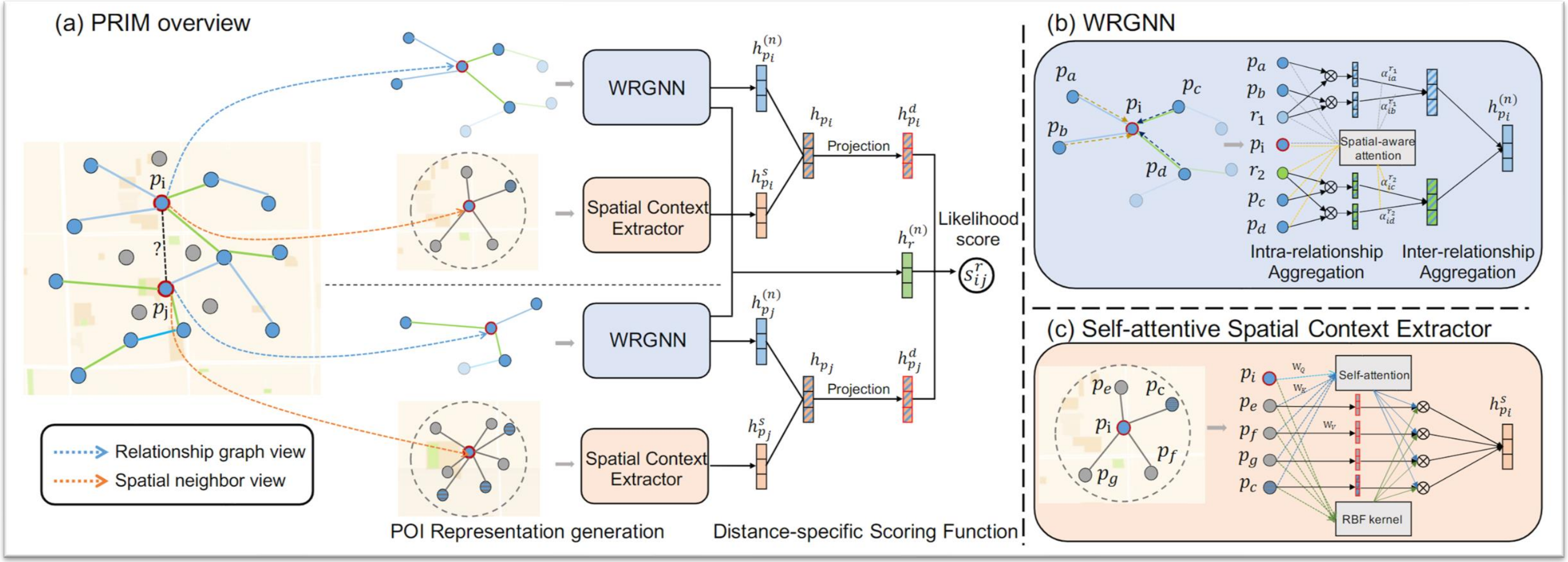
POI Relationship Inference Problem

Definition 3.4. **(POI Relationship Inference Problem).** Given a heterogeneous POI relationship graph G, a category taxonomy T, a distance threshold d, and candidate relation types $\mathcal{R}^* = \mathcal{R} \cup \{\phi\}$, we aim to learn a predictive function $f(\mathcal{P} \times \mathcal{P} | G, T, d) \rightarrow \mathcal{R}^*$ that maps a POI pair (p_i, p_j) to a certain relation type.



Methodology - Architecture

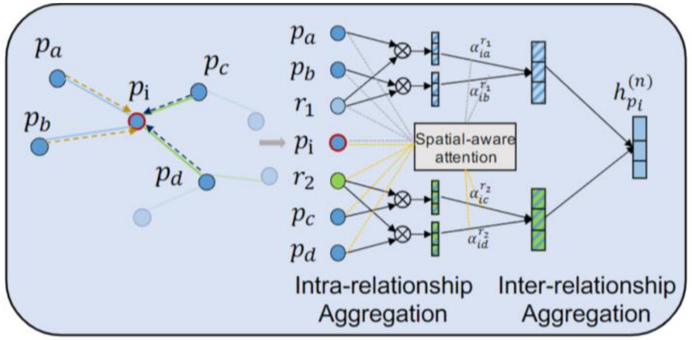
- PRIM consists of four components:
  - weighted relational graph neural network
  - taxonomy integration module
  - self attentive spatial context extractor
  - distance-specific scoring function.



Methodology - Weighted Relational GNN

- Propose a two-level aggregation process
  - Intra-relationship: neighbors connected by the same relation type are first aggregated
  - Inter-relationship: then results from different types of relationships are further aggregated

$$\mathbf{h}_{p_i}^{(l+1)} = \sigma\left(\sum_{r\in\mathcal{R}}\sum_{p_j\in\mathcal{N}_{p_i}^r} \alpha_{ij}^r \mathbf{W}^{(l)}\gamma(\mathbf{h}_{p_j}^{(l)}, \mathbf{h}_r^{(l)})\right) \qquad \mathbf{h}_r^{(l+1)} = \mathbf{W}_r^{(l)}\mathbf{h}_r^{(l)}$$



Here, $h^{(l)}_{pi}$, $h^{(l)}_{r}$ are the representations of pi and relationship of type r in the l-th layer

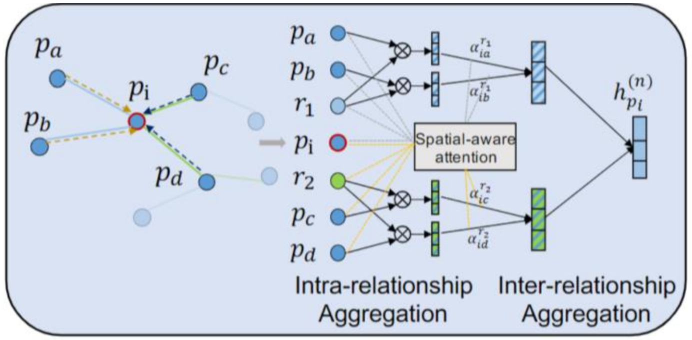
- Advantages:
  - These two types of representations are highly coupled
  - Can complement each other to learn rich semantic characteristics from the heterogeneous POI relationship graph.

Methodology - Weighted Relational GNN

- Spatial-aware attention mechanism
  - In traditional GNN, neighbors in each relationship are treated equally

$$\alpha_{ij}^{r} = \frac{1}{\left|N_{pi}^{r}\right|}$$

  - Use spatial-aware attention mechanism



$$e_{ij}^{r} = \sigma\left(\mathbf{a}_r^{\top} \cdot \left[\mathbf{W}_a \mathbf{h}_{p_i}^{(l)} \middle\| \mathbf{W}_a \mathbf{h}_{p_j}^{(l)} \middle\| \mathbf{W}_d \mathbf{d}_{ij}\right]\right),$$

$$\alpha_{ij}^{r} = \mathrm{softmax}\left(e_{ij}^{r}\right) = \frac{\exp\left(e_{ij}^{r}\right)}{\sum_{k \in N_{p_i}^{r}} \exp\left(e_{ik}^{r}\right)},$$

d_{ij} is spatial distance between pi and pj, ॥ denotes the vector concatenation operation

  - The above attention mechanism can be extended to multiple heads.

- Remarks: Over all, the Weighted Relational above is a very routine operation

Methodology - Taxonomy Integration Module

- Intuitions
  - POIs whose categories are close in category taxonomy tend to be more semantically similar and show higher degree of competitiveness

- Solutions
  - Derive the category representation by backtrack to retrieve all the nodes

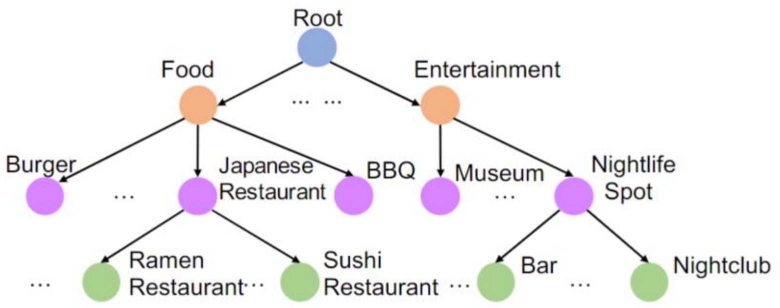
    E.g., the category path for bar is [root, entertainment, nightlife spot, bar]

$$\mathbf{q}_{p_i} = \sum_{t \in Q_{p_i}} \mathbf{e}_t$$

  - Concatenate the original representations and the category representations

$$\mathbf{h}_{p_i}^{*(l)} = [\mathbf{h}_{p_i}^{(l)} \| \mathbf{q}_{p_i}]$$

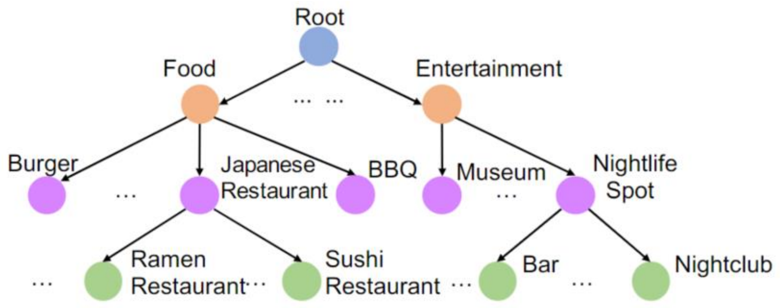
- Advantages
  - Do not apply more complicated techniques
  - more efficient



Methodology - Taxonomy Integration Module

- Authors:
  - Note that we do not apply more complicated techniques, such as Tree-LSTM and hyperbolic embedding, to model category taxonomy because they are less efficient and found to show no improvement over the proposed solution in our preliminary experiments.

- Remarks
  - How to get the representations of all categories in the tree?
  - Still need to build the tree
  - Indeed no need to manually define category relationships



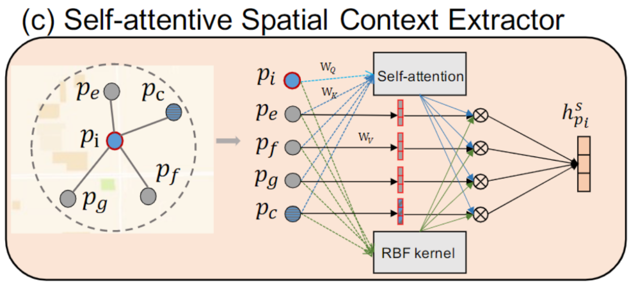
Methodology – Self-attentive Spatial Context Extractor

- Spatial neighbors can be leveraged to provide extra spatial context

- Solutions:
  - Calculate the spatial context as

$$\mathbf{h}^{s}_{p_i} = \sum_{p_j \in \mathcal{S}_{p_i}} \beta_{ij}(\mathbf{W}_V \mathbf{h}^{(L)}_{p_j})$$

$$\beta_{ij} = \mathrm{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{p_m \in \mathcal{S}_{p_i}} \exp(e_{im})}$$



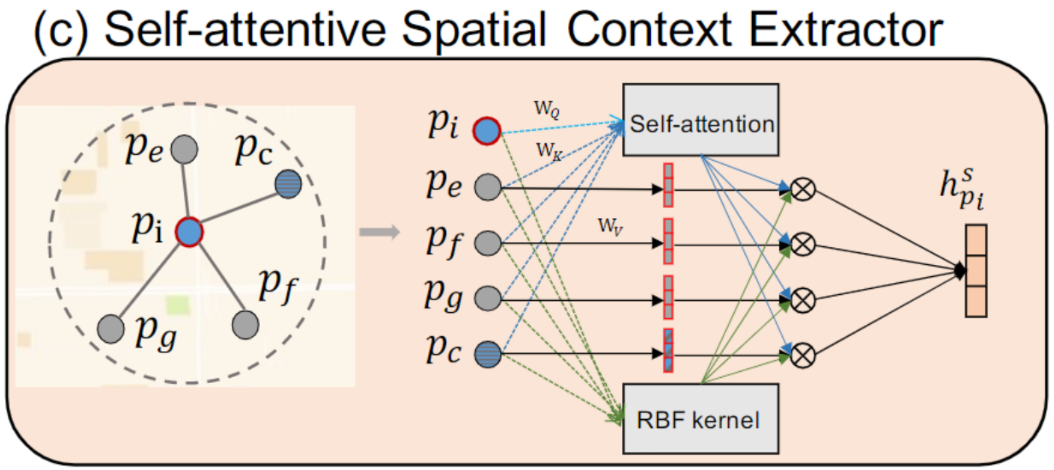
In which, considering semantic & spatial correlation

$$e'_{ij} = \frac{\left(\mathbf{W}_Q \mathbf{h}^{(L)}_{p_i}\right)^{\top} \cdot \left(\mathbf{W}_K \mathbf{h}^{(L)}_{p_j}\right)}{\sqrt{d_p}} \qquad D\left(l_{p_i}, l_{p_j}\right) = \exp\left(-\theta \left\| l_{p_i} - l_{p_j} \right\|^2\right) \qquad e_{ij} = e'_{ij} \cdot D\left(l_{p_i}, l_{p_j}\right)$$

  - For each PoI , fuse heterogeneous POI relationship graph view and the spatial context view

$$\mathbf{h}_{p_i} = \mathbf{h}^{(L)}_{p_i} + \mathbf{h}^{s}_{p_i}$$

Methodology – Self-attentive Spatial Context Extractor



- Remarks
  - In a heterogeneous graph, we can add edges representing geographical adjacency
  - I don't see how to solve the issue, This context is not explicitly used for relationship inference

Methodology – Distance-specific Scoring Function

- In the previous issue, spatial features have important effects on POI relationships

- Solutions:
    - Project the learned POI representations to a distance-specific hyperplane
        - Split the distance into non-overlapping bins (e.g., 0-1km, 1-2km, etc)
        - Project p_i into the hyperplane specified by distance

$$\mathbf{h}_{p_i}^{d} = \mathbf{h}_{p_i} - \mathbf{w}_{g(d_{ij})}^{\top} \mathbf{w}_{g(d_{ij})} \mathbf{h}_{p_i}$$

    - Compute the likelihood of different relationships based on the projected POI & relationship representations.

$$s_{ij}^{r} = \mathbf{h}_{p_i}^{d\top} \operatorname{diag}\left(\mathbf{h}_{r}^{(L)}\right) \mathbf{h}_{p_j}^{d}$$

- Remarks
    - Different category should have a different representations in different hyperplane. There should exist a category-related learnable parameter to be learned when doing projection.

Methodology – Training and Inference

- Training method: cross-entropy loss with negative sampling
  - For each observed positive triplet, sample ω negative pairs by replacing a POI with a randomly sampled one

$$\mathcal{L} = \sum_{(p_i, r, p_j) \in \mathcal{D}} y_{ij}^r \log \sigma(s_{ij}^r) + (1 - y_{ij}^r) \log \sigma(1 - s_{ij}^r)$$

- Inference
  - Calculate the likelihood score w.r.t. each relation type
  - Rank the scores and select the relationship with the highest score as the prediction result

$$\hat{r}_{ij} = \arg\max_{r \in \mathcal{R}^*} s_{ij}^r$$

- Remarks
  - What if the two POIs don't have any relationship?

Methodology – Overall Review

- Different issues and modules to solve them
  - Issue 1: Some methods only focus on a particular type of relationship
  - Module: Weighted Relational GNN

  - Issue 2: Spatial features have important effects on POI relationships
  - Module: Spatial-aware Weighted Relational GNN & Distance-specific Scoring Function

  - Issue 3: Hard to use structural knowledge
  - Module: Taxonomy Integration Module

  - Issue 4: Lack considering the context information of target POIs
  - Module: Self-attentive Spatial Context Extractor
- Authors: Although existing methods have made progress in inferring the relationships for POIs, they still suffer from at least one of the issues.

Experiments

3

Experiments – Datasets

- Datasets
  - Beijing and Shanghai, from Meituan
  - Construct relationships between POIs from user logs extracted from these two datasets
    - Note that the proposed methods do not use user logs for evaluation, and are applicable when the user logs are unavailable

Table 1: Statistics of the datasets

Dataset	Taxonomy		#POIs	#Relational Edges
	#Non-leaf nodes	#Categories		
Beijing	95	805	13,334	122,462
Shanghai			10,090	112,848

Experiments – Training Data Generation

- Training Data Generation
  - Generate the ground truth of competitive relationship for a POI pair if the two POIs in the pair are "viewed/clicked together" by users within a query session

  - Generate the ground truth of complementary relationships for a POI pair if the two POIs in the pair are "also viewed/clicked" by same users across different query sessions

Experiments – Baselines

- Five types of baseline methods
  - rule-based methods
    - CAT, CAT-D
  - random walk based graph embedding methods
    - Deepwalk, node2vec
  - vanilla GNN
    - GCN, GAT
  - heterogeneous GNN
    - HAN, HGT, R-GCN, CompGCN
  - SOTA relationship inference methods
    - DecGCN, DeepR

Experiments – Experiment Results

- Basic results

Table 2: Results on the two datasets in terms of Macro-F1 and Micro-F1 (with best in bold and second-best underlined

Dataset	Metric	Train%	CAT	CAT-D	Deepwalk	node2vec	GCN	GAT	HAN	HGT	R-GCN	CompGCN	DecGCN	DeepR	PRIM
BJ	Macro-F1	40%	0.464	0.519	0.638	0.640	0.707	0.724	0.782	0.779	0.789	0.794	0.757	0.783	**0.845**
		50%	0.464	0.519	0.691	0.692	0.737	0.748	0.811	0.814	0.814	0.832	0.801	0.815	**0.870**
		60%	0.464	0.519	0.731	0.734	0.755	0.776	0.839	0.842	0.820	0.860	0.811	0.842	**0.882**
		70%	0.464	0.519	0.757	0.761	0.770	0.795	0.857	0.857	0.828	0.870	0.823	0.861	**0.895**
	Micro-F1	40%	0.559	0.579	0.707	0.710	0.729	0.753	0.817	0.813	0.808	0.827	0.805	0.820	**0.879**
		50%	0.559	0.579	0.758	0.762	0.766	0.776	0.842	0.845	0.827	0.859	0.823	0.847	**0.895**
		60%	0.559	0.579	0.783	0.784	0.780	0.804	0.867	0.869	0.832	0.882	0.826	0.871	**0.907**
		70%	0.559	0.579	0.816	0.817	0.796	0.821	0.882	0.883	0.839	0.892	0.843	0.887	**0.913**
SH	Macro-F1	40%	0.443	0.509	0.652	0.655	0.673	0.692	0.788	0.790	0.777	0.776	0.765	0.786	**0.822**
		50%	0.443	0.509	0.682	0.684	0.704	0.711	0.807	0.811	0.802	0.808	0.797	0.808	**0.844**
		60%	0.443	0.509	0.698	0.702	0.713	0.729	0.834	0.838	0.828	0.837	0.817	0.839	**0.861**
		70%	0.443	0.509	0.707	0.711	0.731	0.735	0.853	0.858	0.839	0.857	0.832	0.852	**0.875**
	Micro-F1	40%	0.551	0.573	0.724	0.727	0.744	0.769	0.843	0.841	0.822	0.836	0.824	0.845	**0.886**
		50%	0.551	0.573	0.766	0.768	0.775	0.788	0.861	0.864	0.837	0.860	0.846	0.863	**0.896**
		60%	0.551	0.573	0.783	0.786	0.793	0.806	0.886	0.888	0.864	0.891	0.856	0.890	**0.909**
		70%	0.551	0.573	0.796	0.798	0.804	0.811	0.897	0.901	0.872	0.903	0.869	0.898	**0.920**

Experiments – Experiment Results

- Remarks: What kind of experiments we can do?
  - Use different fractions of datasets as training data
  - Evaluate Model Scalability
    - Randomly generate subdatasets of different sizes from a large dataset
  - Ablation Study

東南大學
SOUTHEAST UNIVERSITY

Q & A

Thank you