

# GPR

이은경

# 목차

1. Abstract
2. Introduction
3. Related work
4. Classification to Semantic Segmentation
5. Experiments
6. Conclusion

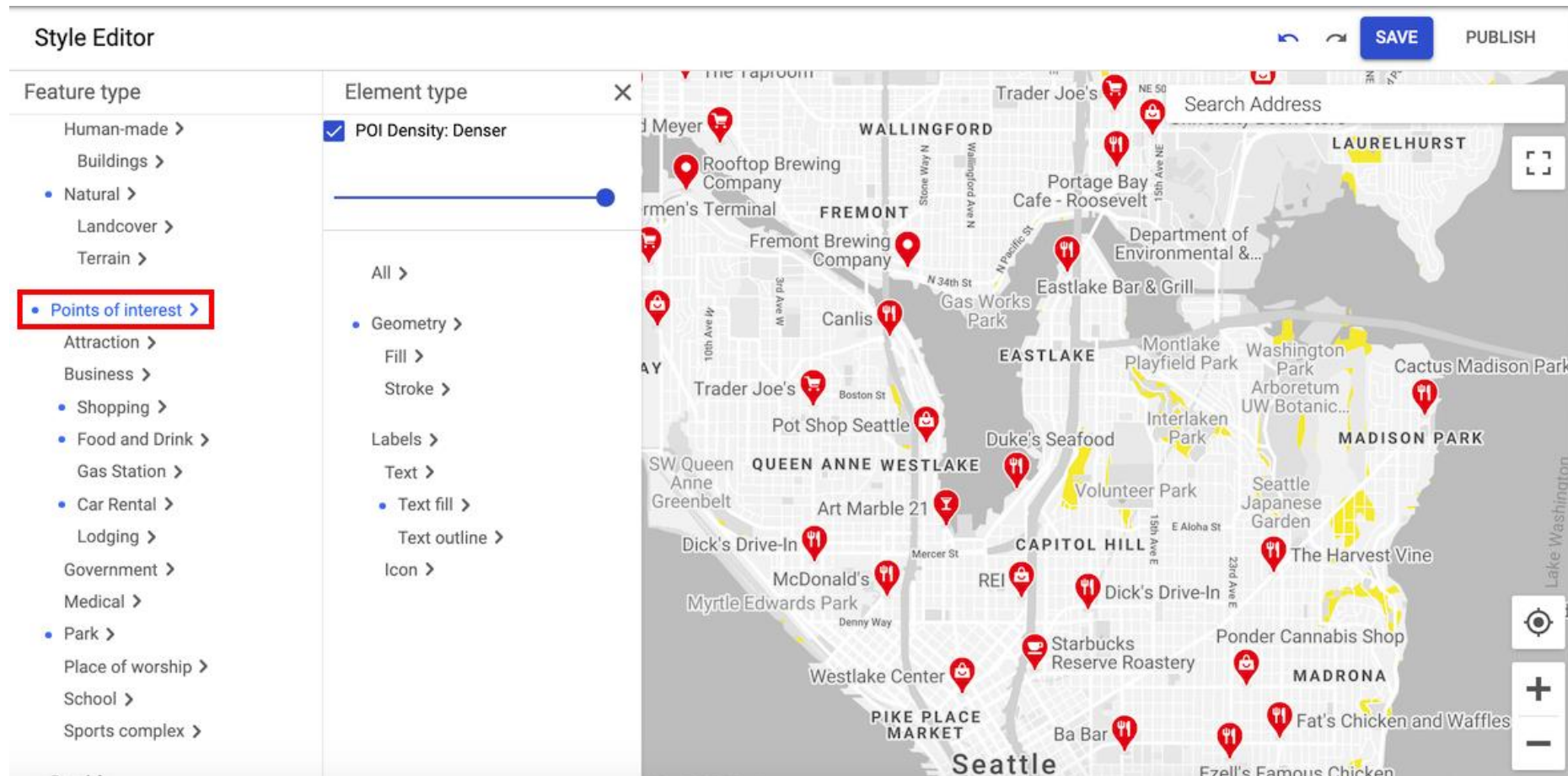
# 1. Abstract

# 1) GGLR

## 1. Abstract

- points-of-interest (POIs) recommendation

1. graph-based geographical latent representation model (**GGLR**) 은 복잡한 user-POI networks에서 non-linear geographical influences를 capture 가능



## 2) GPR

### 1. Abstract

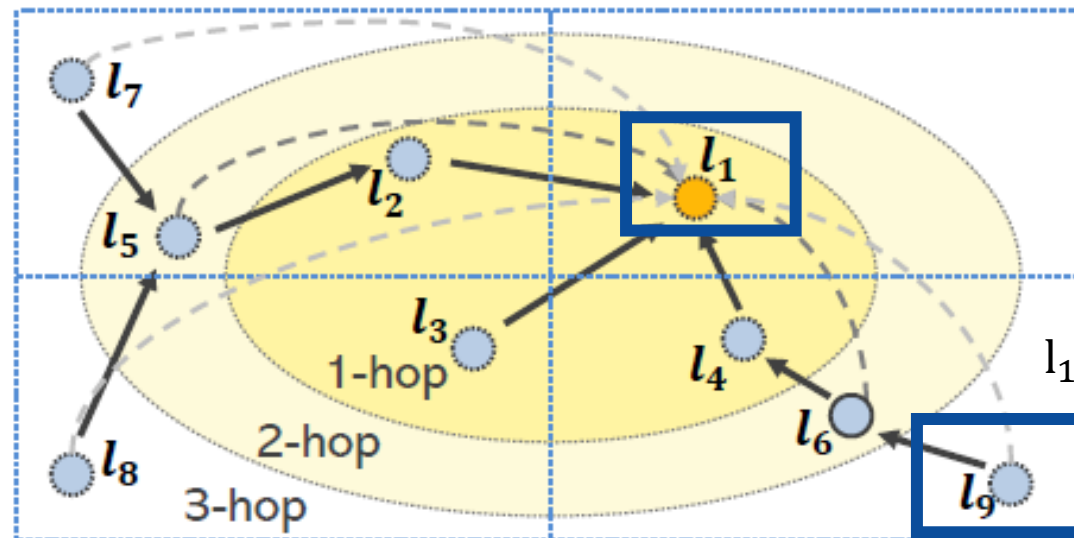
- ingoing influences와 outgoing influences를 모두 고려.
  1. graph auto-encoder 기반으로 체크인 이력에 자주 사용되는 두 개의 연속 POI 사이의 지리적 영향을 증가시키도록 ingoing 와 outgoing influences 의 지리적 잠재 표현을 훈련
  2. user preferences를 추정하기 위해 ingoing과 outgoing influences의 geographical latent representations을 사용하는 그래프 신경망 기반 POI 권장 모델(GPR)을 제안

## 2. Introduction

## 1) 사전연구

## 2. Introduction

- POI 간 지리적 영향을 표현하는 것이 핵심 : explicit / implicit models 두 종류
- Explicit models
  - USG[26], Rank-GeoFM [8], GeoSoCa [29], and STGCN [30] : power-law and normal kernel functions 등을 활용하여 geographical influences 를 캡처
  - GeoMF[9]와 IRenMF[13]는 지리적 위치에 따라 POI를 region으로 나누고 잠재된 표현을 훈련
- implicit models
  - SG-CWARP [11] 및 CAPE [2]는 체크인 이력이 동시에 발생하는 POI가 서로 멀리 떨어져 있다고 가정 > co-occurring POIs 를 고려하여 neighbors 의 latent representations에 기초하여 지리적 latent representations 를 훈련
- 문제 : shallow methods (e.g., inner-product and dividing regions) 에 의존

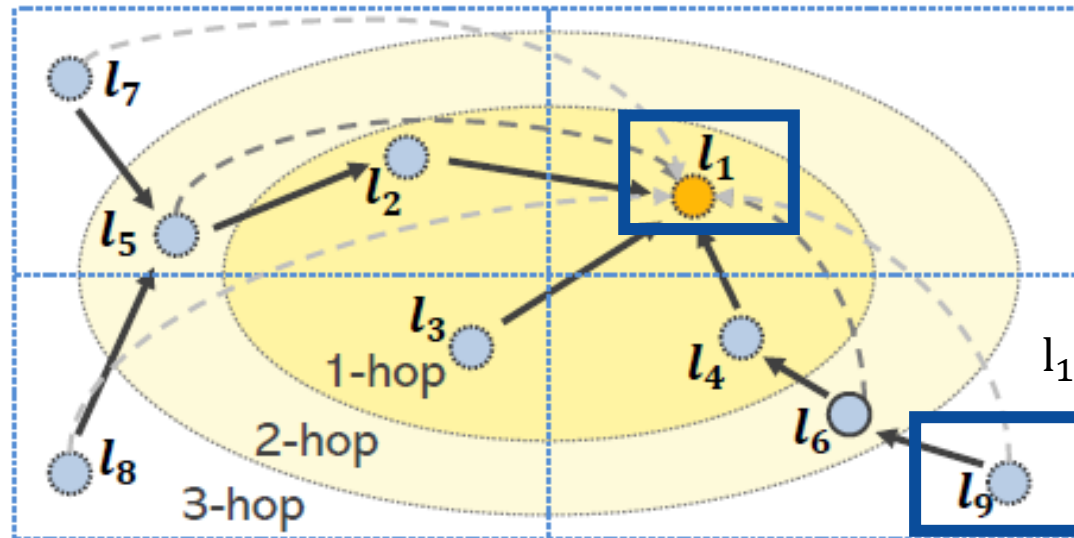


$l_1$ 과 직접 연결 없는  $l_9$  도 영향을 줌

## 1) 사전연구

## 2. Introduction

- inner-product는 이렇게 간접연결된 POI들간 지리적 영향을 포착 불가
  - Region-based models [9, 13] : 동일한 지역의 POI가 동일한 지리적 영향을 미친다고 가정
  - 이전의 연구[19]는 컴플렉스 네트워크의 구조가 매우 비선형적이라는 것을 발견
  - shallow methods은 complex user-POI networks [14, 20]에서 non-linear geographical influences를 capture하지 못함



$l_1$ 과 직접 연결 없는  $l_9$  도 영향을 줌



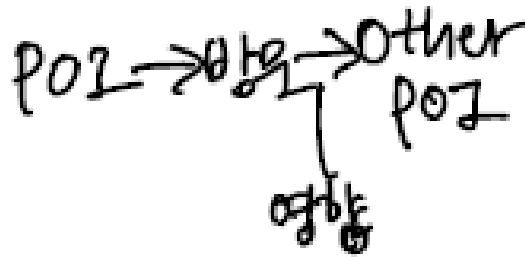
## 2) GGLR

## 2. Introduction

- geographical influences에는 ingoing 과 outgoing influences가 많이 포함됨

1) Ingoing : POI가 다른 POI의 방문자에게 미치는 영향

2) Outgoing : POI가 다른 POI를 방문하는 방문자에게 미치는 영향



- POI의 ingoing / outgoing 영향을 반영하는 directed POI-POI 그래프를 사용. Graph auto-encoder[1]를 기반으로, GGLR은 방문 빈도가 높은 두 연속 POI 사이의 지리적 영향을 증가시키기 위해 geographical latent representations을 훈련

## 2) GGLR

## 2. Introduction

- Graph auto-encoder : GNN(Graphneural Network)[7]을 통해 neighboring POI의 latent representations 으로부터 latent representations 의 표현을 캡처
- GNN은 complex POI-POI networks로부터 고도의 non-linear geographical influences를 포착하는데 적합
- geographical latent representations 학습을 위해 두 POI간 physical distance를 사용

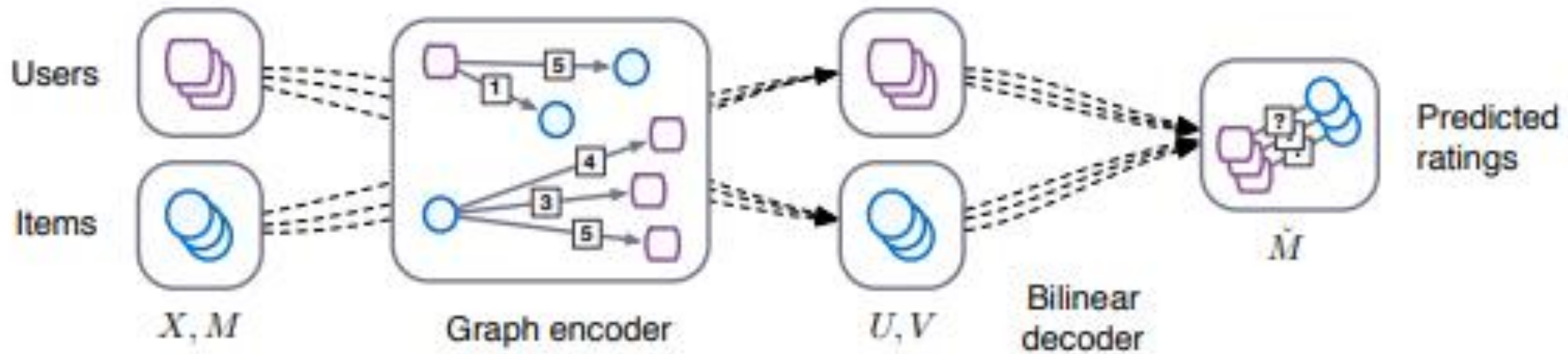


Figure 2: Schematic of a forward-pass through the GC-MC model, which is comprised of a graph convolutional encoder  $[U, V] = f(X, M_1, \dots, M_R)$  that passes and transforms messages from user to item nodes, and vice versa, followed by a bilinear decoder model that predicts entries of the (reconstructed) rating matrix  $\hat{M} = g(U, V)$ , based on pairs of user and item embeddings.

- 지수함수를 사용하여 physical distance feature 계산
  - graph auto-encoder의 encoder-decoder에 physical distance feature를 적용 (지리적 영향을 분명하게 해줌)
  - GGLR에 의해 훈련된 geographical latent representations을 사용하는 GNN-based POI Recommendation (GPR) 모델을 제안
1. 사용자 체크인 이력으로 사용자-POI 그래프를 구성
  2. 고차 GNN[17]을 기반으로 사용자가 방문한 POI의 outgoing 영향의 훈련된 geographical latent representations 에서 user latent representations 을 추출.
  3. 추출된 user latent representations과 POI의 ingoing 영향에 대한 geographical latent representations 을 기반으로 방문하지 않은 POI에 대한 사용자 선호도를 추정.
  4. POI에 대한 사용자 선호도는 사용자가 방문한 POI의 outgoing 영향에 대한 geographical latent representations과 사용자가 아직 방문하지 않은 POI의 ingoing 영향의 geographical latent representations 을 곱하여 추정

## 4. OUR APPROACH

# 1) Preliminary

1. Preliminary
2. Geographical Latent Representation Model
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

< Definition 1 (User-POI graph) >

- $\mathcal{U} = \{u_1, u_2, \dots, u_M\}$ ,  $\mathcal{L} = \{l_1, l_2, \dots, l_N\}$  사용자 sets과 POI sets
- bipartite user-POI graph  $\mathcal{G}_r = (\mathcal{U}, \mathcal{L}, \mathcal{E}_r)$ . 여기서  $\mathcal{E}_r$  : graph  $\mathcal{G}_r$ 의 edge sets .

User  $u_i$ , visited POI  $l_j$  일 때 user의 체크인 히스토리  $e_{r,i,j}$ 를 포함

< Definition 2 (POI-POI graph) >

- user 의 check-in history 는 사용자가 방문한 POIs 의 sequence. user's check-in history에서,  $\mathcal{E}_p$  가 directed POI-POI graph  $G_p = (\mathcal{L}, \mathcal{E}_p)$  에서의 edge set  $\mathcal{E}_p$  .
- edge  $e_{p,i,j}$  는  $\mathcal{E}_p$  users consecutively visiting POI  $l_j$  를 방문하고 연속해서 POI  $l_i$  를 방문하는 빈도
- 두 번의 연속 체크인 사이에 긴 시간 간격이 있는 경우 POI간 상관관계는 낮으므로  $h$  시간 안에 방문하는 건만 제약 :  $(h \in \{3, 6, 12, 24, \infty\})$ .

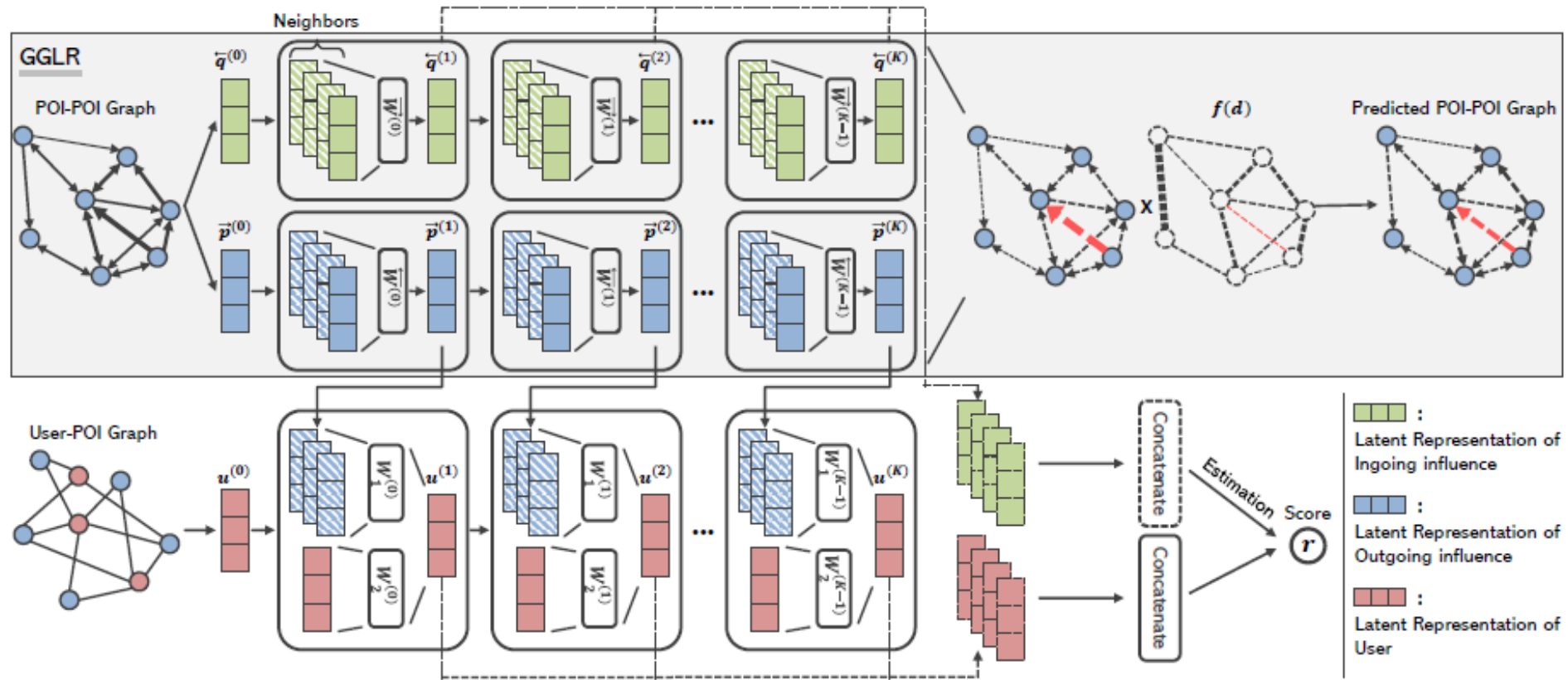
# 1) Preliminary

1. Preliminary
2. Geographical Latent Representation Model
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

< Definition 3 (POI recommendation) >

- POI recommendation에는 user  $u$  가 다음에 방문할 가능성이 높은 순위 POI 목록을 생성하는 작업이 포함.
- implicit feedback으로 정의.
- check-in histories을 기반으로 POI recommendation model은 방문하지 않은 POI에 대한 사용자 선호도를 추정
- 예상 사용자 선호도를 기준으로 POI 순위 ranking.



## 2) Graph-based Geographical Latent Representation Model

1. Preliminary
2. Geographical Latent Representation Model
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

각 POI  $i$  in 2 oam randomly  $\vec{p}_i^{(0)}$  or  $\vec{q}_i^{(0)} \in \mathbb{R}^d$   $\Leftarrow$   $d$ -dim  
real-valued vectors

$\rightarrow$  update ~~\*~~ training phase.

outgoing  $\vec{p}_i^{(K)}$  or ingoing  $\vec{q}_i^{(K)}$   $\Leftarrow$   $K$ -th layers oam  $\vec{p}_i^{(K)}$  or  $\vec{q}_i^{(K)}$  latent representation

연속 POI  $i, j$  is  $\vec{p}_i^{(K)}$  or  $\vec{q}_j^{(K)}$  :  $\hat{e}_{p,i,j} = (\vec{p}_i^{(K)} W) \cdot \vec{q}_j^{(K)}$

oam  $W \in \mathbb{R}^{d \times d}$  weight matrix

$\hat{e}_{p,i,j}$  is freq oam.

graph  $\rightarrow$  encoder  $\rightarrow$  bilinear decoder

$$\hat{e}_{p,i,j} = (\vec{p}_i^{(K)} W) \cdot \vec{q}_j^{(K)}, \quad (2)$$

### 3) Physical Distance Feature.

1. Preliminary
2. **Geographical Latent Representation Model**
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

- POI 사이의 물리적 거리는 지리적 영향에 영향을 미치는 중요한 요소
- physical distance information 을 이용하여 POIs의 순위를 매기고, 가까운 방문자 POIs의 순위를 높게 매김
- 성능저하 발생 가능성
- Exponential function를 사용해 physical distance feature
- exponential function : 두 POI가 서로 가까울 때 physical distance feature을 높은 값으로 계산하는 반면, 두 POI가 멀리 떨어져 있을 때 physical distance feature을 낮은 값으로 계산
- POI와 POI 사이의 물리적인 거리  $\theta_{i,j}$ 는 킬로미터로 표현되며  $[0.01m, 100k]$ 의 범위로 고정
- 100km 이상의 물리적 거리는 100km로 처리

$$f(x_{i,j}) = a * x_{i,j}^b * e^{cx_{i,j}}, \quad (3)$$

- 여기서 a,b,c 는 trainable variables

$$\hat{e}_{p,i,j} = f(x_{i,j}) \times (\overrightarrow{p}_i^{(K)} W) \cdot \overleftarrow{q}_j^{(K)}. \quad (4)$$



### 3) Physical Distance Feature.

1. Preliminary
2. **Geographical Latent Representation Model**
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

- GGLR의 모든 훈련 가능한 파라미터는 평균 제곱 오차로 정의된 다음의 손실 함수를 minimizing함으로써 최적화.

$$\mathcal{L}_p(\theta_p) = \frac{1}{|\mathcal{E}_p|} \sum_{e_p} (e_p - \hat{e}_p)^2, \quad (5)$$

- $\theta$ : GGLR의 훈련 가능한 매개변수 집합

# 1) User Latent Representation

1. Preliminary
2. Geographical Latent Representation Model
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

- POI recommendation은 체크인 기록에 따라 사용자가 방문할 가능성이 높은 POI를 추천.
  - 그림 3에서 빨간색 노드와 빨간색 상자는 각각 사용자  $u$ 와 사용자의 체크인 기록을 표시
  - POI간 geographical influences, 개별 POI의 geographical influences 고려 필요
- GGLR로 고차 GNN[17]을 확장하는 GPR을 제안. 고차 GNN의 각 레이어는 다음과 같이 계산.

$$f^{(k)}(v) = \sigma(f^{(k-1)}(v)W_1^{(k)} + \sum_{w \in \mathcal{N}(v)} f^{(k-1)}(w)W_2^{(k)} + b^{(k)}),$$

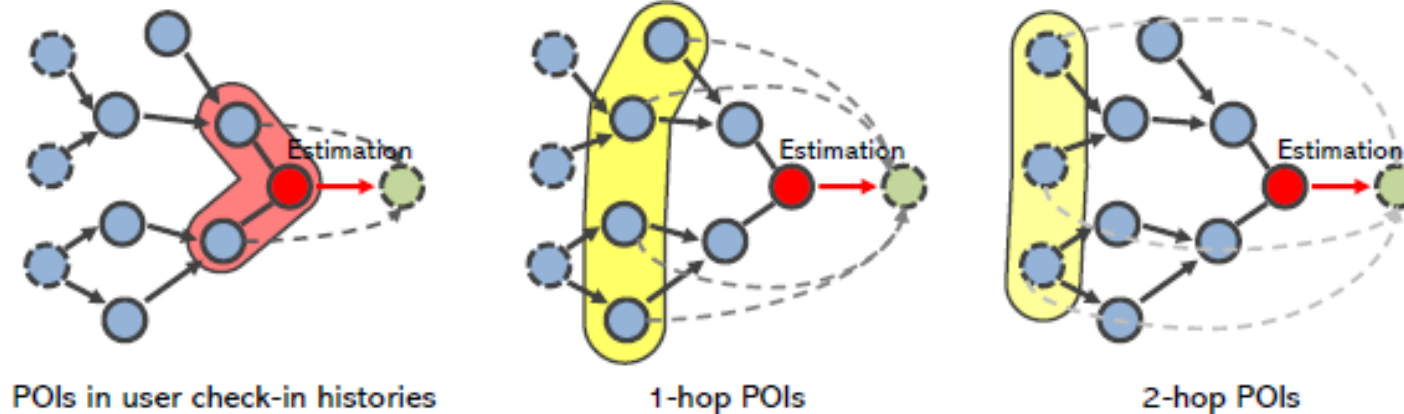


Figure 3: Motivation for the structure of GPR.

# 1) User Latent Representation

1. Preliminary
2. Geographical Latent Representation Model
3. Point-of-Interest Recommendation Model

## 4. OUR APPROACH

- 고차 GNN은 노드와 그 이웃 노드에 대해 서로 다른 가중치 매트릭스( $w_1$  및  $w_2$ )를 사용 > 이는 고차 GNN의 고유한 feature를 나타냄
- bipartite user-POI graph  $G_r$ 에서 이 feature를 사용하여 사용자의 latent representations를 추출 가능
- GGLR로 훈련한 POI의 outgoing influence  $\vec{p}_j$  를 사용

$$u_i^{(k)} = \sigma(u_i^{(k-1)} W_1^{(k)} + \sum_{l_j \in N(i)} \vec{p}_j^{(k)} W_2^{(k)} + b^{(k)}), \quad (7)$$

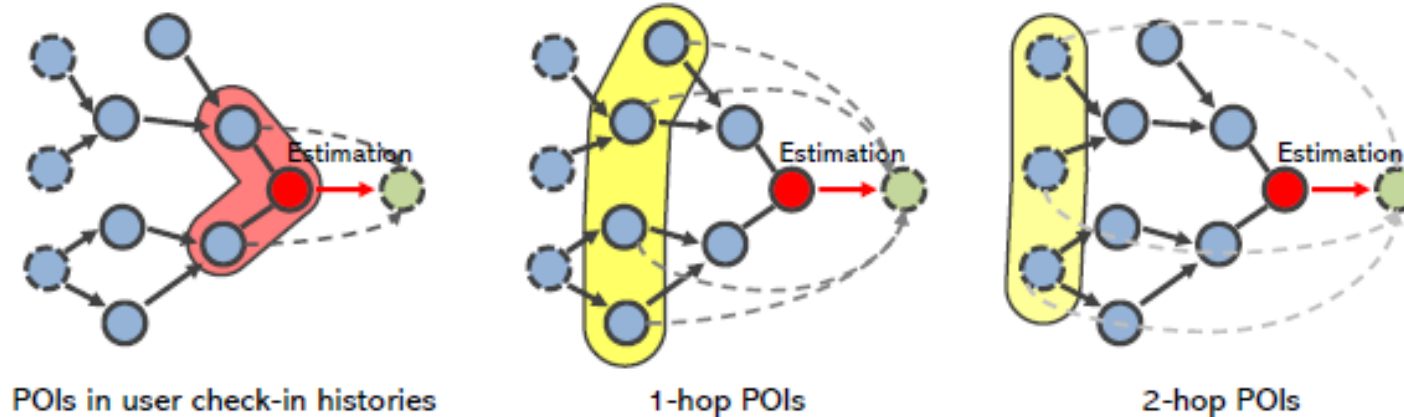


Figure 3: Motivation for the structure of GPR.

## 2) User Preference Estimation

1. Preliminary
2. Geographical Latent Representation Model
3. **Point-of-Interest Recommendation Model**

## 4. OUR APPROACH

- POI에 대한 사용자 선호도를 추정하기 위해 user latent representations  $u$ 와 ingoing influences  $\bar{q}$ 의 geographical latent representations를 사용한다.
- $K'$  GNN layers를 통해 각 사용자에게 대한  $K$  user latent representations을 얻음
- 각 GNN 레이어의 user latent representations  $u$ 는 해당 계층의 geographical latent representations outgoing influence  $\bar{p}_j$ 에서 캡처되기 때문에 different levels of multi-hop neighbor POIs를 다룸.
- 사용자 선호도를 추정하기 위해 모든 사용자 latent representations을 연결하고 최종 사용자 latent representations로 사용
- $u, l$ 는 POI  $i$ 에 대한 사용자  $u$ 의 예상 선호도 /  $K$ : layer 수
- POI에 대한 사용자의 선호도를 미리 가정하여 사용자가 방문한 POI의 outgoing influence으로 나타내고 사용자가 아직 방문하지 않은 POI의 ingoing influence으로 곱

$$u = [u^1; u^2; \dots; u^{K'}], \quad \bar{q} = [\bar{q}^1; \bar{q}^2; \dots; \bar{q}^{K'}], \quad (8)$$

$$r_{u,l} = u^T \bar{q}_l, \quad (9)$$

### 3) Optimization

1. Preliminary
2. Geographical Latent Representation Model
3. **Point-of-Interest Recommendation Model**

## 4. OUR APPROACH

- 사용자 선호도 추정에 사용되는 매개 변수  $\theta$  최적화
- pairwise ranking-based loss function
- 방문한 POI와 방문하지 않은 POI 사이의 상대적 선호도 고려

$$\mathcal{L}_r(\theta_r) = \sum_{(u,l,l') \in O} -\log(\hat{\sigma}(r_{u,l} - r_{u,l'})), \quad (10)$$

- GPR 최종 Loss 함수

$$\mathcal{L}(\theta) = \mathcal{L}_r(\theta_r) + \lambda_1 \mathcal{L}_p(\theta_p) + \lambda_2 (\|\theta_r\|_2^2 + \|\theta_p\|_2^2), \quad (11)$$

## 5. Experiments

# 1) Performance Comparison

## 5. Experiments

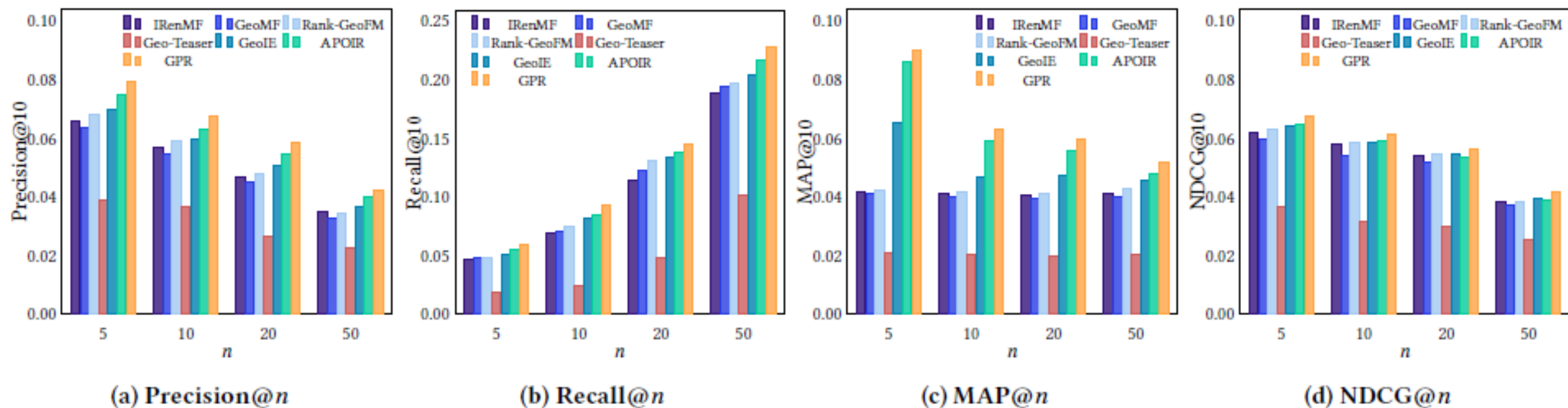


Figure 4: Performance comparison of GPR and the baselines on the Gowalla dataset as  $n$  increases.

# 1) Performance Comparison

## 5. Experiments

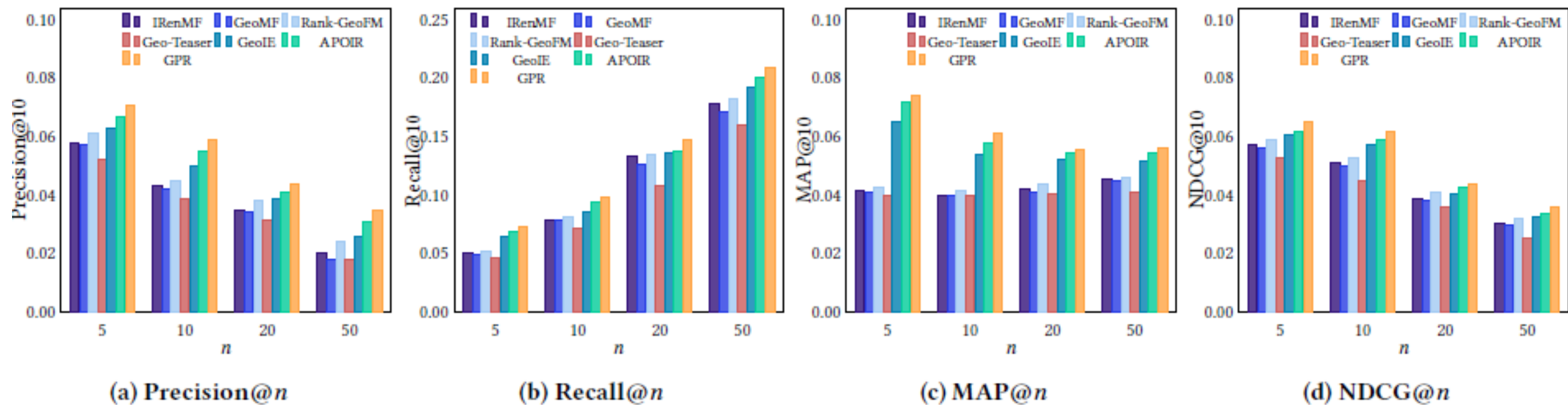


Figure 5: Performance comparison of GPR and the baselines on the Foursquare dataset as  $n$  increases.



# 1) Performance Comparison

## 5. Experiments

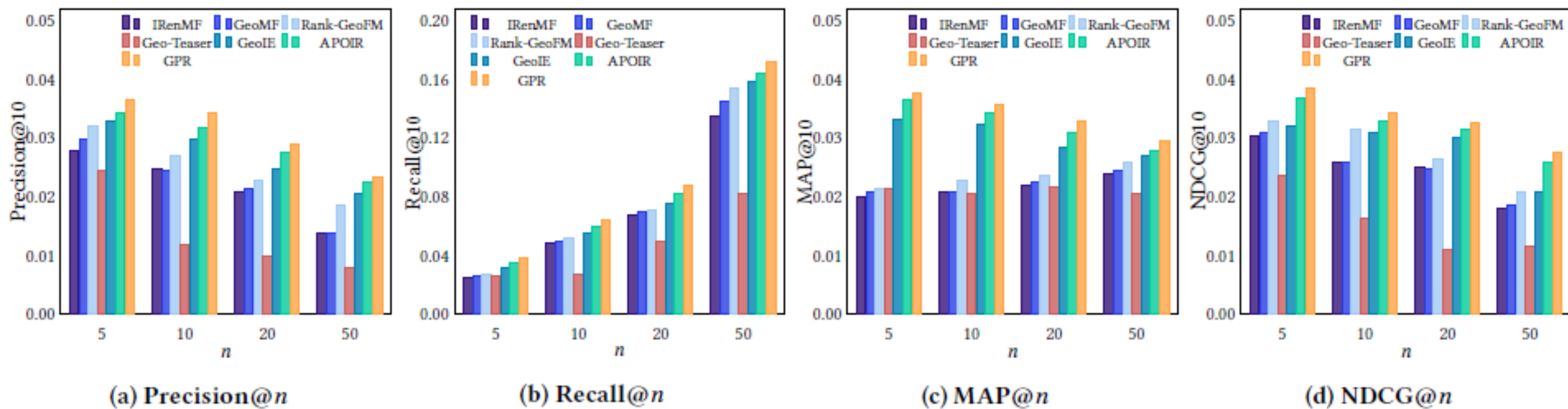
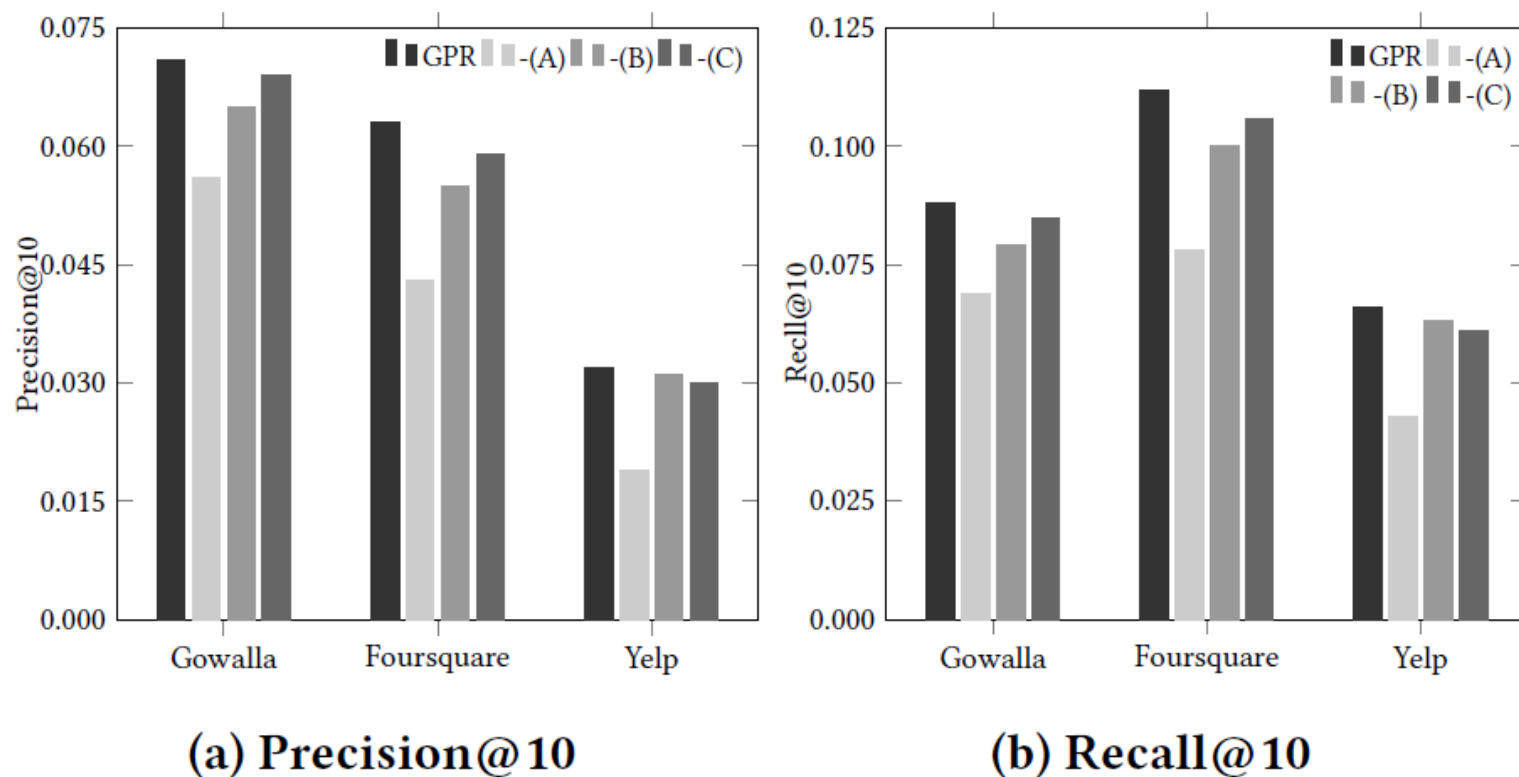


Figure 6: Performance comparison with the baselines on the Yelp dataset as  $n$  increases.

## 2) Effectiveness of Each Component.

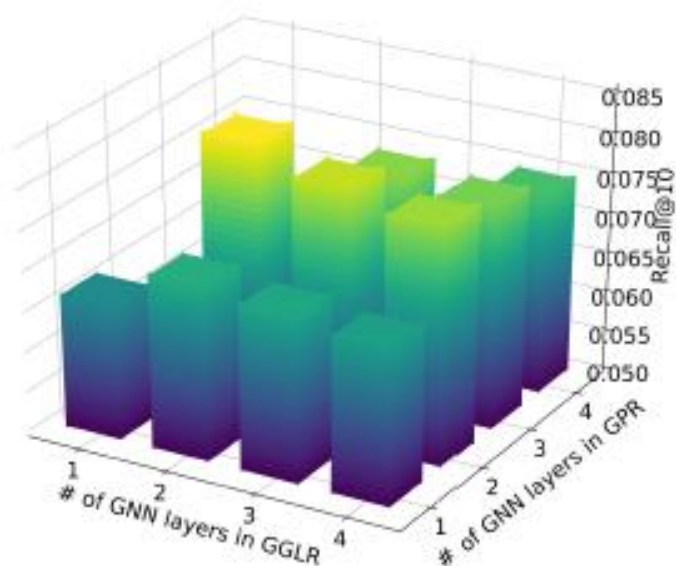
## 5. Experiments



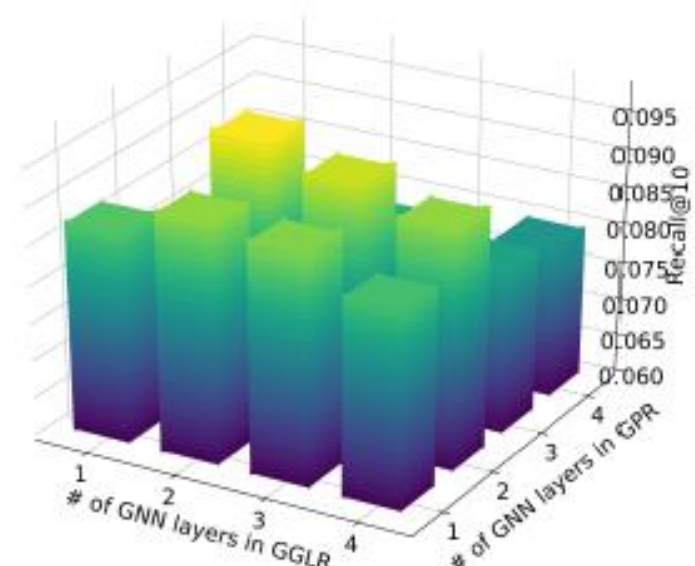
**Figure 8: Ablation test results on the three datasets. -(A), -(B), and -(C) denote the removed components GGLR, asymmetric features, and physical distance features, respectively.**

### 3) Hyperparameter Optimization.

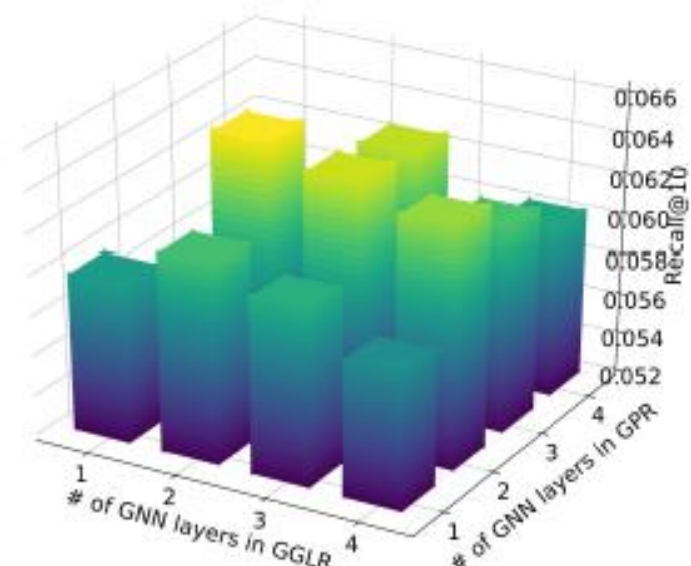
## 5. Experiments



(a) Gowalla



(b) Foursquare



(c) Yelp

Figure 7: Effect of the number of GNN layers  $K$  and  $K'$  in GGLR and GPR, respectively.

## 4) Non-linearity of Geographical Influences.

## 5. Experiments

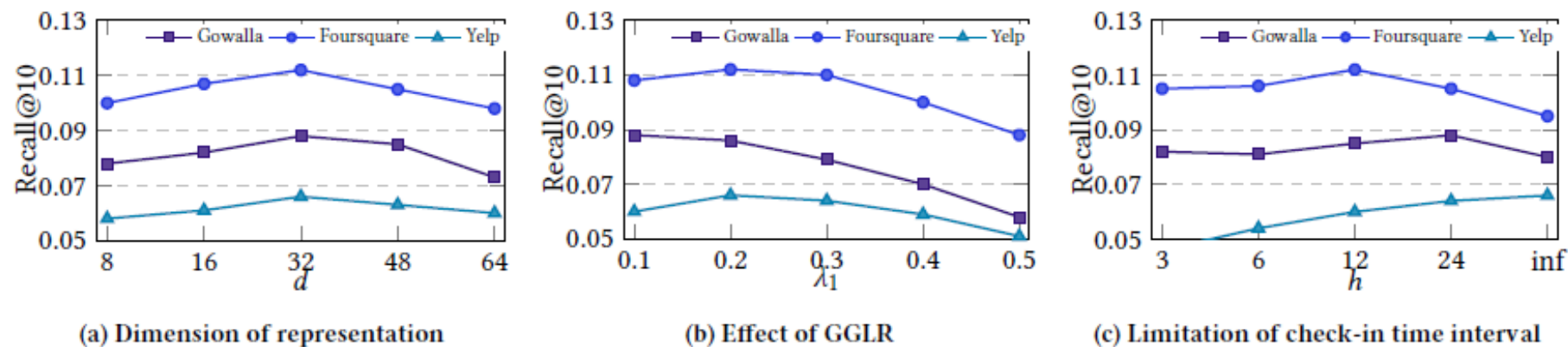


Figure 9: Effects of hyperparameters.

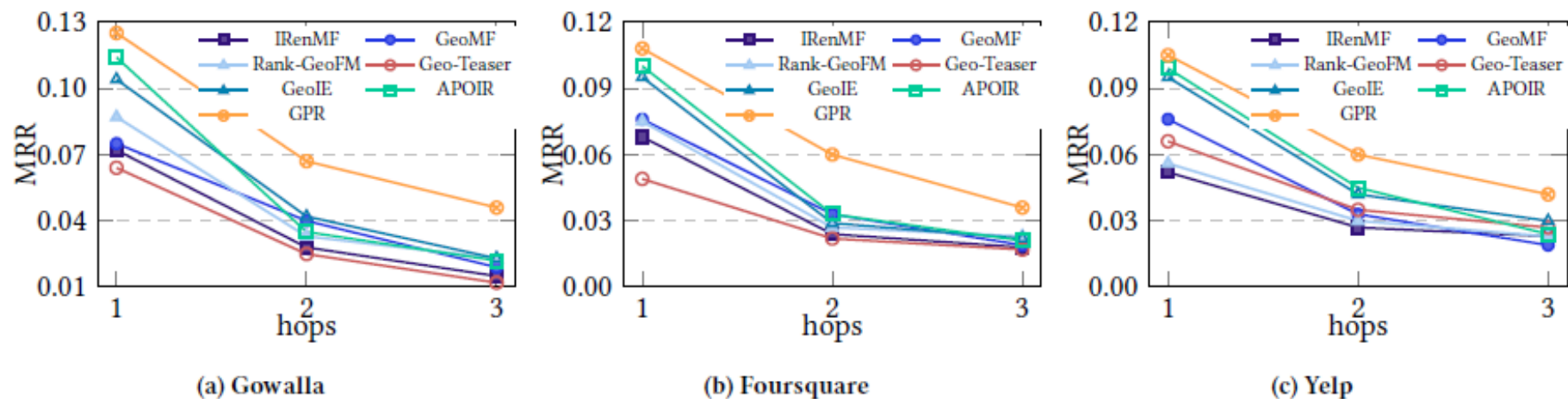


Figure 10: Performance evaluation results depending on the level of hops.

Q & A  
감사합니다