개요.

관심 장소(POI) 추천, 즉 사용자가 방문하지 않은 장소를 추천하는 것은 위치 기반 소셜 네트워크의 근본적인 문제입니다. POI 추천은 POI 간의 지리적 영향력을 통해 영화 추천과 같은 기존의 항목 추천과 차별화됩니다. 기존 방식은 두 지점 간의 지리적 영향력을 물리적 거리를 고려했을 때 동일한 사용자가 두 지점을 함께 방문할 확률 또는 성향으로 모델링합니다. 이러한 방법은 지리적 영향력이 물리적 거리에 의해 결정된다고 가정하기 때문에 지리적 영향력의 비대칭성과 지리적 영향력의 높은 변동성을 포착하지 못합니다. 이 페이지에서는 POI별 지리적 영향력을 활용하여 POI 추천을 개선합니다. 지리적 영향력, 지리적 민감도, 물리적 거리라는 세 가지 요소를 사용하여 두 POI 간의 지리적 영향력을 모델링합니다. 지리적 영향력은 다른 POI에 지리적 영향력을 행사할 수 있는 POI의 역량을 파악하고, 지리적 민감도는 다른 POI의 지리적 영향을 받는 POI의 성향을 반영합니다. 두 개의 실제 데이터 세트에 대한 실험 결과에 따르면 POI별 지리적 영향력이 POI 추천의 성능을 크게 향상시키는 것으로 나타났습니다.

소개

Foursquare, Gowalla와 같은 위치 기반 소셜 네트워크(LBSN)가 점점 인기를 끌면서 물리적 세계와 온라인 소셜 네트워킹 서비스 사이의 간극을 메우고 있습니다[Xiao et al., 2010; Sun et al., 2017]. LBSN에서 사용자는 자신의 위치와 위치 정보와 관련된 콘텐츠를 공유하여 사용자의 선호도와 행동에 대한 이해를 용이하게 합니다[Bao et al., 2012; Liu and Xiong, 2013; Gao et al., 2015; Wang et al., 2015a]. 관심 지점 추천, 즉 사용자의 체크인 기록에 따라 사용자가 방문하지 않은 지점(예: 레스토랑, 쇼핑몰, 극장)을 추천하는 것은 최근 몇 년 동안 큰 연구 관심을 받고 있습니다[Li et al., 2016; He et al., 2016; Zhang et al., 2016; Li et al., 2017].

POI 추천의 가장 두드러진 특징 중 하나는 POI와 타겟 사용자의 위치가 추천에 중요한 요소라는 점입니다. 예를 들어, 고왈라와 포스퀘어의 경우 사용자의 연속 체크인 중 90%가 50km 미만의 거리 내에 있습니다[Liu et al., 2017]. 따라서 연구자들은 기존 아이템 추천에서와 같이 사용자와 POI 간의 상호작용을 통해 사용자의 선호도를 모델링하는 것 외에도 POI 간의 지리적 근접성 또는 지리적 영향력을 활용하여 POI 추천의 성능을 개선하기 위해 노력하고 있습니다[Ye et al., 2011; Lian et al., 2014; Xie et al., 2016].

지리적 영향력을 POI 추천에 활용하는 기존 방식은 크게 두 가지 패러다임으로 나뉩니다. 첫 번째 방식은 지리적 근접성을 활용하여 사용자 선호도 학습을 개선하는 방식으로, 서로 가까운 곳에 있는 POI가 유사한 사용자 선호도를 공유한다고 가정합니다[Liu et al., 2014; Li et al., 2015; Xie et al., 2016; Feng et al., 2017]. 이러한 방법에서는 지리적 근접성이 사용자 선호도에 대한 일종의 공간 정규화로 사용됩니다. 두 번째 방법은 물리적 거리가 주어졌을 때 동일한 사용자가 두 POI를 함께 방문할 확률 또는 성향으로 POI 간의 지리적 영향력을 명시적으로 모델링합니다[Ye et al., 2011; Cheng et al., 2012; Zhang and Chow, 2013; Lian et al., 2014; Saleem et al., 2017]. 물리적 거리와 관련하여 POI의 공동 방문 확률 분포를 파악하기 위해 파워 법칙 함수, 가우스 분포 등 다양한 형태의 함수가 사용됩니다. 앞서 언급한 방법들은 지리적 영향력을 활용하는 데는 어느 정도 성공적이지만, POI 간의 지리적 영향력의 큰 변화를 포착할 수는 없습니다. 예를 들어, 그림 1에서 볼 수 있듯이 Foursquare 데이터 세트에서 무작위로 선택된 10개의 POI는 지리적 영향력이 상당히 다르게 나타나며, 이는 지리적 영향력이 물리적 거리만으로는 잘 포착될 수 없으므로 지리적 영향력을 POI별로 파악해야 한다는 것을 나타냅니다. 이 문서에서는 POI 추천을 개선하기 위해 POI별 지리적 영향력을 활용합니다. 지리적 영향력, 지리적 민감도, 물리적 거리라는 세 가지 요소를 사용하여 두 POI 간의 지리적 영향력을 모델링합니다. 지리적 영향력은 다른 POI로 방문자를 확산시키는 POI의 역량과 지리적 요소입니다.

그림 1: Foursquare 데이터 세트에서 무작위로 샘플링한 10개 POI의 거리에 따른 체크인 상관관계 히트 맵. 0.5km를 하나의 빈으로 간주하고 각 빈에 대해 선택한 각 POI와 해당 빈에 속하는 POI 간의 평균 상관관계를 계산합니다. 이 값은 가장 큰 값으로 정규화합니다.

민감도는 다른 POI로부터 방문자를 받는 POI의 성향을 반영합니다. 예를 들어, 지하철 역은 일반적으로 지리적 영향력이 높고 레스토랑은 일반적으로 지리적 민감도가 높습니다. 여기서 지리적 영향력과 지리적 민감도는 두 개의 저차원 벡터이며, 두 POI 간의 지리적 영향력은 한 POI의 지리적 영향력 벡터와 다른 POI의 지리적 민감도 벡터의 내적 곱으로 표현됩니다.

POI별 지리적 영향력 모델에는 두 가지 고유한 이점이 있습니다: (1) POI 간의 지리적 영향력은 비대칭적이기 때문에 POI 간 지리적 영향력의 높은 가변성을 포착할 수 있는 높은 유연성을 제공합니다. (2) POI 상호작용 행렬을 사용하여 POI별 지리적 영향력을 직접 모델링하는 대신, 우리 모델은 각 POI에 대해 두 개의 저차원 벡터로 지리적 영향력을 표현하여 자유 파라미터의 수를 크게 줄입니다[Wang et al., 2015b]. 따라서 데이터 희소성 문제가 심각한 POI 추천에 적합한 모델입니다.

마지막으로, 사용자의 선호도를 파악하는 표준 모델에 POI별 지리적 영향력을 통합하여 새로운 POI 추천 방식을 구축합니다. 사용자의 체크인 기록을 사용하여 모델을 학습시키고, 사용자가 가까운 미래에 방문할 가능성이 높은 POI를 '예측'하는 데 모델을 적용하여 추천 성능을 검증합니다. 포스퀘어와 고왈라의 두 가지 실제 데이터 세트에서 광범위한 실험을 수행하여 모델의 효율성을 입증했습니다. 실험 결과, POI별 지리적 영향력이 POI 추천의 성능을 크게 향상시켜 최신 POI 추천 방법을 능가하는 것으로 나타났습니다.

2.Related work

이 섹션에서는 POI 추천에 대해 간략하게 살펴봅니다. POI 추천은 사용자의 체크인 기록에 따라 사용자가 방문하지 않은 POI를 추천하는 기능입니다. 사용자의 체크인은 암묵적 피드백이라는 점을 고려할 때, 기존 방식은 체크인 횟수에서 변환된 점수를 피팅하여 체크인을 모델링하거나[Lian et al., 2014], 사용자의 선호도 쌍별 순위를 POI에 최적화하여 모델링합니다[Li et al., 2015; 2016; Zhao et al., 2017]. 사용자의 체크인 횟수가 드물기 때문에 체크인 횟수만 활용하는 방법은 성능이 떨어지는 경우가 많습니다. 이러한 상황을 완화하기 위해 보조 정보를 통합할 수 있습니다. 예를 들어, 지리적 영향은 가장 중요한 요소 중 하나이며 온라인에는 존재하지 않습니다.

지리적 영향력을 모델링하는 기존의 방법은 글로벌 방법[Ye et al., 2011; Cheng et al., 2012; Zhang and Chow, 2013; Lian et al., 2014]과 지역적 방법[Liu et al., 2014; Li et al., 2015; Xie et al., 2016; Feng et al., 2017]으로 분류할 수 있습니다. 글로벌 방법은 POI의 동시 발생과 지리적 좌표 사이의 관계를 모델링합니다. Ye 등[2011]과 Lian 등[2014]은 각각 전력법 분포와 가우스 분포를 사용하여 거리에 따른 지리적 영향력을 특성화합니다. [Cheng et al., 2012; Zhang and Chow, 2013]은 각 사용자의 체크인(예: 경도 및 위도)을 고정된 분포로 산점도를 캡처합니다. 지역적 방법은 같은 지역에 있는 POI가 사용자들에게 유사한 매력을 공유한다는 점을 고려합니다. 표현 기반 학습 방법[Xie et al., 2016; Feng et al., 2017; Zhao et al., 2017]은 같은 지역의 POI가 유사한 표현을 공유하도록 제한하는 방법을 사용합니다. [Liu 등, 2014; Li 등, 2015]는 지리적 이웃의 매력을 고려하여 대상 POI의 매력을 직접 계산합니다. 그러나 글로벌 방식과 지역 방식은 지리적 영향력에 대한 두 가지 거친 표현을 제공하며, 이는 POI 고유의 속성을 무시합니다. 이 문서에서는 이 문제를 다룹니다. 또한, 많은 연구에서 사회적 관계[Tang et al., 2013], 시간적 요인[Yuan et al., 2013], 카테고리[Zhang and Chow, 2015] 등과 같은 다른 정보를 통해 POI 추천 성능을 향상시킬 수 있는 방법을 모색했습니다.

3.preliminary

U와 I로 각각 사용자 집합과 POI 집합을 나타냅니다. 사용자 u와 POI i의 경우, 사용자 u가 POI i를 방문한 횟수를 cui로 표시하고, wui는 cui의 스케일링된 버전입니다. 사용자 u가 방문한 모든 POI는 체크인 기록을 구성하며, Hu로 표시됩니다. 각 POI i의 위치는 경도 loni 와 위도 lati 로 표시됩니다. 각 사용자 u의 선호도는 벡터 ~tu로 표시되며, 경위도 dij를 사용하여 POI i와 POI j 사이의 물리적 거리를 나타냅니다. 각 POI i에 대해 선호도 벡터를 ~zi로 표시하고 지리적 영향력 벡터와 지리적 민감도 벡터를 ~gi 및 ~hi로 표시합니다. 표 1에 표기법이 요약되어 있습니다.

POI 추천: 체크인 기록이 있는 사용자 집합 U와 위치 정보(경도, 위도)가 있는 POI 집합 I가 주어지면, POI 추천은 각 대상 사용자 u ∈ U에 대해 대상 사용자가 잠재적으로 관심이 있지만 추천까지 방문하지 않은 POI로 구성된 {i|i ∈ I}의 목록을 추천합니다.

4. POI -Specific Geographical Influence

목표 POI j의 경우, 사용자 u의 체크인 이력 Hu에서 각 POI i의 지리적 영향력을 고려합니다. Figure 1에서 볼 수 있듯이 사용자는 인접한 POI를 방문하는 것을 선호하며, 각 POI는 물리적 거리로는 잘 설명되지 않는 고유한 특성을 가지고 있습니다. 지리적 영향력의 큰 변화를 포착하기 위해, 지리적 영향력 yij를 POI i에서 POI j로 모델링하면 yij = ~gT i ~hj × f(dij )가 됩니다. (1) 여기서 벡터 ~gi는 POI i의 지리적 영향력, 즉 방문자를 다른 POI로 분산시키는 POI의 역량을 나타내고, 벡터 ~hj는 POI j의 지리적 민감도, 즉 다른 POI로부터 방문자를 받아들이는 POI의 성향을 반영하며, dij는 POI i와 POI j 사이의 물리적 거리를 나타냅니다.

첫째, f(dij )는 두 POI의 물리적 거리 dij 가 주어졌을 때 동일한 사용자가 두 POI를 방문할 확률을 반영합니다. 이 문서에서는 Table 2에 표시된 네 가지 유형의 함수를 고려합니다. 일반적으로 f(dij )는 dij 가 증가함에 따라 감소하며, 이는 사용자가 지리적으로 인접한 POI를 방문하는 것을 선호하는 현상을 포착합니다. 또한, f(dij )의 파라미터는 POI 추천 모델을 학습하기 전에 미리 학습하거나 POI 추천 모델 학습과 함께 학습할 수 있습니다.

방문한 POI i와 대상 POI j 사이의 상호작용을 ~gT i ~hj 로 모델링합니다. 이러한 방식으로 POI 간의 지리적 영향력은 비대칭적이므로, POI 간 지리적 영향력의 높은 변화를 유연하게 포착할 수 있습니다. 또한, POI 상호작용 행렬을 사용하여 POI별 지리적 영향력을 직접 모델링하는 대신 각 POI에 대해 두 개의 저차원 벡터로 지리적 영향력을 나타내므로 자유 변수 수가 크게 줄어듭니다. 따라서 데이터 희소성 문제가 심각한 POI 추천에 적합한 모델입니다.

POI별 지리적 영향력 yij는 기본적으로 두 POI의 물리적 거리와 고유한 특성 모두에서 발생하는 공동 효과를 포착합니다. 대상 POI j의 경우, 지리적으로 인접하고 영향을 미치는 POI는 높은 y를, 멀지만 영향을 미치는 POI(또는 인접하지만 영향을 덜 미치는 POI)는 상대적으로 작은 y를 갖게 됩니다. 영향력 점수 y는 대상 POI의 고유 특성 ~h에 따라 달라집니다. 다시 말해, 우리 모델의 영향력 점수는 관련된 POI에 따라 POI별로 달라집니다. 즉, 동일한 방문 POI i가 주어졌을 때 다른 타겟 POI에 대해 y가 달라집니다.

사용자 u가 방문한 POI의 집합 Hu와 식 (1)이 주어지면, 이러한 모든 방문 POI의 영향을 고려하고 대상 POI j에 대한 Hu의 전반적인 지리적 영향을 (2)와 같이 모델링합니다.

4.2 Preference Modeling and Recommendation

타겟 POI에 대한 사용자의 선호도를 추론하기 위해 사용자 선호도와 지리적 영향력의 영향을 모두 고려합니다. 구체적으로, 사용자 u의 POI j에 대한 선호도를 suj 로 나타내며, (그람 3 참고)

여기서 ~tu와 ~zj는 사용자 u의 선호도와 POI j의 선호도 간의 상호작용을 모델링하는 데 사용되는 행렬 인수 분해 방법에 따라 주어집니다. 사용자의 체크인 기록은 일종의 암묵적인 사용자 선호도인 POI의 방문 빈도를 기록합니다. 따라서, 기존 추천이 직접적으로 고객 선호도를 맞추는 것과 달리, 저희는 각 체크인을 모든 후보 POI 중에서 하나의 타겟 POI를 선택하는 과정으로 모델링합니다. 이러한 방식으로 방문 빈도를 직접 수치로 모델링할 때 발생하는 편향성을 피할 수 있습니다.

구체적으로, 사용자 u가 POI j를 선호할 확률 puj는 (그림4 참고)

여기서 I는 POI의 집합이고 분모는 주어진 사용자 u에 대한 모든 POI의 정규화입니다. 식 (4)에서 주어진 사용자 u의 경우 {puj}는 다항 분포의 매개 변수 집합으로 작용한다는 것을 확인할 수 있습니다. 따라서 사용자 u가 POI j를 방문하는 행동은 사용자가 모든 후보 중에서 하나의 POI j를 선택하는 의사결정 과정의 결과로 모델링됩니다.

체크인은 제27회 국제 인공 지능 공동 컨퍼런스(IJCAI-18) 3879 논문집에서 사용자의 선호도 분포 {puj}에서 추출한 샘플로 편리하게 해석할 수 있습니다. 사용자의 체크인을 관찰할 수 있는 로그 확률을 최대화합니다: (그림 5참고)

체크인 횟수 cuj는 종종 매우 왜곡된 분포를 보이기 때문에, 몇몇 인기 POI와의 상호작용이 로그 가능성을 지배할 수 있습니다. 따라서 이 문제를 완화하기 위해 cuj를 스케일링된 버전인 wuj로 대체합니다. 구체적으로, 우리는 로그 형식 함수[Lian et al., 2014]를 채택하는데, 여기서 wuj = 1+log(1+cuj ×10 )은 스케일링 매개변수입니다. 미방문 POI의 경우 wuj = 0으로 설정합니다. 마지막으로 사용자 u가 주어졌을 때, 사용자가 POI j를 방문할 확률 puj에 따라 미방문 POI를 추천합니다. 각 사용자의 추천 목록은 미방문 POI 중 확률 puj가 가장 높은 상위 n개의 POI로 구성됩니다.

4.3 Optimization

이제 ~tu, ~zj, ~gi, ~hj 및 지리적 함수의 매개변수를 포함한 네 가지 유형의 잠재 요인 최적화에 대해 설명합니다. 우리는 로그 가능성 L을 최대화하기 위해 [Mikolov et al., 2013]에서 제안한 네거티브 샘플링 방식을 채택합니다. 각 방문 POI에 대해 일부 잡음 분포에 따라 K개의 부정 체크인을 샘플링합니다. 구체적으로, 다음과 같은 목적 함수를 사용하여 L (그림 6 참고)

여기서 NEG(j)는 POI j에 대한 음의 POI 집합을 나타냅니다. δlj는 l = j이면 1, 그렇지 않으면 0이 되는 지표이고 σ(-)는 시그모이드 함수입니다. 새로운 목적 함수를 최적화하기 위해 확률적 경사 상승(SGA) 알고리즘을 채택합니다. 각 반복에서 최적화할 비율 ζ로 설정된 쌍의 미니 배치를 무작위로 샘플링합니다. 샘플링 확률은 스케일링된 체크인 횟수, 즉 wuj 에 비례합니다. 쌍(u, j)이 샘플링되면 잠재 벡터는 다음과 같이 업데이트됩니다: (그림10 참고) 여기서 i ∈ Hu, l ∈ j ∪ NEG(j이고 η는 학습 속도입니다.

5 Experiment

5.1 DataSets

평가를 위해 Foursquare [Cho et al., 2011]와 Gowalla [Yuan et al., 2013]의 두 가지 실제 데이터 세트를 사용합니다. Foursquare에서는 10개 미만의 POI를 방문한 사용자와 10명 미만의 사용자가 방문한 POI를 제거하여 체크인을 사전 처리합니다. 고왈라에서는 제거 임계값이 40으로 설정되어 있습니다. 전처리 후, Foursquare 데이터 세트에는 6,118명의 사용자가 88,193개의 POI에 대해 생성한 172,961건의 체크인과 1,624명의 사용자가 3,585개의 POI에 대해 생성한 115,890건의 체크인이 있습니다. 두 데이터 세트의 각 POI는 경도 및 위도와 연관되어 있습니다. 또한 Foursquare 데이터 세트에서 각 POI는 8개의 카테고리와 240개의 하위 카테고리로 표시되어 있습니다. 각 사용자 u에 대해 시간순으로 체크인을 정렬하고, 초기 70%의 체크인을 학습 데이터로, 다음 15%는 검증 데이터로, 마지막 15%는 테스트 데이터로 사용합니다.

5.2 Evaluation Metrics

우리는 평가를 위해 널리 사용되는 두 가지 메트릭을 채택합니다[Lian et al, 2017], 즉 precision@n과 recall@n을 사용하며, 여기서 n은 추천 목록에 있는 POI의 수입니다. 여기서 P n u 는 사용자 u의 추천 목록에서 상위 n 개의 POI 집합이고 Tu는 사용자 u의 기준 진실 POI 집합입니다. |x|는 집합 x의 카디널리티를 나타냅니다. 각 메트릭에 대해 실험에서는 n의 7개 값(즉, 1, 2, 3, 5, 10, 15, 20)을 고려합니다.

5.3 Methods in Comparison

편의상 제안된 POI 추천 방법의 명칭은 GeoIE로 사용합니다. 다음과 같은 대표적인 POI 추천 방법과 비교하여 GeoIE의 성능을 평가합니다: - UCF+G [Ye et al., 2011]: 전력법 함수를 사용하여 거리에 따른 체크인 확률을 파악한 후 사용자 기반 협업 필터링 방법을 결합한 방식입니다. - MGM+PFM [Cheng et al., 2012]: 다중 센터 가우스 모델을 사용하여 사용자의 체크인 분포를 파악한 다음 확률적 요인 모델을 결합합니다. - GeoMF [Lian et al., 2014]: 기존 사용자와 POI의 잠재 요인을 사용자의 활동 지역과 POI의 영향력 영역으로 보강하여 MF를 확장합니다. - RankGeoFM [Li et al., 2015]: 주변 POI의 매력을 고려하여 지리적 영향력을 포함하는 랭킹 기반 MF 모델입니다. 최근 POI 추천에 대한 리뷰[Liu et al., 2017]에 따르면 RankGeoFM은 가장 성능이 우수한 방법 중 하나이며, 최신 POI 추천 방법 중 하나로 사용되고 있습니다. - 지오 티저[Zhao et al., 2017]: 시간적 POI 임베딩 모델과 지리적으로 계층화된 쌍별 랭킹 방법을 결합한 방법입니다.

5.4 Experimental Setting

스케일링 매개변수를 10으로 설정합니다. 최적화를 수행할 때 각 잠재 벡터에 대해 L2 정규화 항을 배치하고 정규화 계수는 0.02로 설정합니다. 잠재 벡터의 차원 수는 32입니다. 각 반복에서 음수 샘플 K의 수는 10, 샘플링 비율 ζ는 0.2, 학습률은 0.001로 설정합니다. 우리 방법을 기준선과 비교하기 전에 먼저 표 2에 설명된 대로 네 가지 지리적 함수를 사용하여 우리 방법의 성능을 평가합니다. 해당 지오아이의 이름은 GeoIE-PL-PRE, GeoIE-PL, GeoIE-EXP 및 GeoIE-HB입니다. 그림 2와 그림 3은 각각 포스퀘어 데이터 세트와 고왈라 데이터 세트에서 네 가지 변종 GeoIE의 성능을 보여줍니다. 두 그림 모두 지수 함수가 가장 우수한 성능을 달성하는 것을 볼 수 있으며, 이는 지수 함수가 이 백서에 사용된 두 데이터 세트 모두에서 지리적 영향과 거리 간의 관계를 묘사하는 데 가장 적합한 선택임을 의미합니다. 지수 함수의 우수성은 지수 함수가 더 많은 매개 변수를 가지고 있어 지리적 영향의 높은 변화를 유연하게 포착할 수 있기 때문입니다. GeoIE-PL-PRE와 GeoIEPL을 비교하면 비슷한 성능을 달성한다는 것을 알 수 있습니다. 그러나 지리적 함수의 매개 변수를 모델의 다른 잠재 요인과 함께 최적화하면 수렴 속도가 더 빨라질 수 있다는 점을 지적할 필요가 있습니다.

5.5 Effectiveness of GeoIE

저희는 GeoIE-EXP를 대표적인 방법으로 선정하여 최신 방법과 비교했습니다. 포스퀘어 데이터 세트와 고왈라 데이터 세트의 성능 비교는 각각 그림 4와 그림 5에 나와 있습니다.

GeoIE-EXP는 경쟁 기준 방법보다 일관되게 우수한 성능을 보이는 것으로 관찰되었습니다.

그림 4는 포스퀘어에서는 MGM+PFM이 UCF+G보다 성능이 더 우수하다는 것을 보여줍니다. 그러나 그림 5에 표시된 것처럼 Gowalla에서는 k = 1인 경우를 제외하고는 UCF+G가 MGM+PFM보다 성능이 뛰어납니다. 이들은 서로 다른 지리적 함수를 가정하고 두 데이터 세트에서 각각 좋은 성능을 달성합니다. 두 개의 개별 모델을 선형 보간으로 융합하는 UCF+G 및 MGM+PFM과 달리 GeoMF는 사용자 선호도와 지리적 영향을 하나의 통합된 모델로 통합합니다. 매개변수 학습 과정에서 잠재 요인이 상호 영향을 미쳐 더 나은 성능을 이끌어낼 수 있습니다. 그러나 지오엠에프는 제로가 아닌 체크인 외에도 사용자의 제로 체크인을 요인화합니다. 이로 인해 성능이 어느 정도 약화될 수 있습니다. RankGeoMF는 인접한 POI의 지리적 영향을 고려하고 순위 기반 방법을 사용하여 음수 샘플을 탐색합니다. 두 데이터 세트 모두에서 GeoMF보다 성능이 뛰어납니다. 그러나 인접한 POI의 집객력을 목표 POI의 집객력으로 직접적으로 받아들일 수는 없습니다. 지오 티저는 표현 학습 기반 방식과 선형 보간을 통한 지리적으로 계층화된 쌍별 랭킹 방식을 통합합니다. RankGeoMF보다 약간 더 나은 것을 관찰할 수 있습니다.

요약하자면, 실험 결과는 POI별 지리적 영향력이 POI 추천을 향상시킨다는 것을 분명히 보여줍니다. 지리적 방법을 활용하는 다른 방법보다 POI별 지리적 영향력의 우월성은 두 가지 중요한 시사점을 제공합니다. (1) POI 추천을 위한 지리적 영향력은 노이즈가 너무 커서 파워 법칙 함수나 지수 함수 같은 단순한 함수로 정량화하기 어렵다는 점, (2) POI 추천이 데이터 희소성 문제로 어려움을 겪지만, 두 가지 벡터를 추가로 도입하여 POI별 지리적 영향력을 특성화하면 현격한 개선 효과를 얻을 수 있다는 점 등입니다.

5.6 POI’s Geo-Influence and Geo-Susceptibility

POI별 지리적 영향력을 파악하기 위해 두 가지 잠재 벡터, 즉 각 POI에 대한 지리적 영향력과 지리적 민감도를 도입합니다. 다음에서는 다양한 범주의 POI에 대한 속성 강도를 연구합니다. 지면이 제한되어 있기 때문에 사례 연구를 위해 각 카테고리에서 하나의 하위 카테고리만 선택합니다. 구체적으로, 각 하위 카테고리에 대해 두 가지 유형의 잠재 벡터의 평균 규범을 계산하고 이 규범을 가장 큰 규범으로 나누어 0에서 1 사이의 스케일로 나눕니다.

이러한 규범을 그림 6에 제시합니다. 버스 정류장이 가장 큰 지리적 영향력을 가지고 있음을 알 수 있으며, 이는 다른 POI로 사용자를 확산시키는 강력한 능력을 시사합니다. 그러나 이 영향력 있는 하위 범주의 지리적 민감도는 가장 크지 않습니다. 한편, 술집과 미술관의 두 속성은 모두 낮은 위치에 있어 사용자의 여행 선택에서 중요한 장소가 아님을 나타냅니다. 이러한 관찰은 각 POI가 특정 지리적 영향력과 지리적 민감도가 비대칭적이라는 것을 의미합니다. 이는 또한 POI별 지리적 영향력을 모델링할 필요가 있다는 가정을 검증합니다.

6.conclusion

이 문서에서는 POI별 지리적 영향력을 활용하여 POI 추천을 개선합니다. 지리적 영향력, 지리적 민감도, 물리적 거리라는 세 가지 요소를 사용하여 두 POI 간의 지리적 영향력을 모델링합니다. 지리적 영향력은 다른 POI에 지리적 영향력을 행사할 수 있는 POI의 역량을 파악하고, 지리적 민감도는 다른 POI의 지리적 영향을 받는 POI의 성향을 반영합니다. 이 모델은 POI 간의 비대칭적인 지리적 영향력을 자연스럽게 포착하여 POI 간 지리적 영향력의 높은 변동성을 포착할 수 있는 높은 유연성을 제공합니다. 향후 작업으로 POI 간의 상호작용으로부터 지리적 영향력에 대한 매니폴드(롤처럼 말린 형태로 데이터 분포)를 직접 학습하는 것이 유망합니다.