**Point-of-Interest Recommendations: Learning Potential Check-ins from Friends**

# ABSTRACT

基于位置的社交网络（LBSN）服务的出现为构建个性化兴趣点（POI）推荐系统提供了绝佳的机会。虽然个性化的POI推荐系统可以显着方便用户的户外活动，但是面临很多挑战性的问题，比如用户POI决策过程模型的难度以及解决数据稀疏性和用户/位置冷启动问题的难度。为了应对这些挑战，我们在LBSN中定义了三种类型的朋友（即社交朋友，位置朋友和邻居朋友），并开发了一个两步框架，利用朋友的信息来提高POI推荐的准确性，启动问题。具体来说，我们首先要建议学习一下每个朋友之前签到的一个潜在的地点，这个人最感兴趣。然后，我们使用两种不同的损失函数（基于平方误差的损失和基于排名误差的方法）将三种类型的签入（即，观察签到，潜在签到和其他未观察签到）失利）。为了评估所提出的模型，我们用两个真实世界的数据集上的许多最先进的基线方法和评估指标进行了广泛的实验。实验结果证明了我们的方法的有效性。

# 1. Introduction

近年来，智能移动设备的普及和接入无线网络的便利性使得人们更容易获取其实时位置信息。这种发展刺激了Foursquare，Jiepang和Facebook Places等基于位置的社交网络（LBSN）服务的出现。这些LBSN允许用户彼此建立连接，并分享与兴趣点（POI）相关的经验和签到信息。与LBSN的各种这样的用户交互数据提供了开发个性化POI推荐系统的好机会。事实上，准确和个性化的POI建议是LBSN服务的一个关键要求。它不仅可以帮助用户探索新的地点，还可以方便用户找到相关的兴趣点，而不用花太多时间搜索，特别是在新的地区。

尽管开发个性化的POI推荐系统是一项至关重要的任务，并且可以使用户的户外活动受益，但是由于三个原因，它仍然是一个非常具有挑战性的问题。首先，用户的签到决策过程非常复杂，可能受到许多不同因素的影响。例如，很难模拟社交朋友对用户签到行为的影响。我们不知道哪个朋友会影响用户的POI决定，也不知道如何影响用户的选择。此外，地理距离可能会影响用户的POI决策。用户通常更喜欢附近的POI到另一个远的地方。其次，POI推荐系统通常面临极度稀疏的入住数据所带来的严峻挑战。在实际系统中，有数以百万计的POI。但是，单个用户通常会签到数量有限的POI，这显着增加了推荐难度。第三，当新的POI或新用户进入系统并且没有其访客信息或其历史签到信息时，向用户推荐新的POI或向新用户推荐POI是非常困难的。

在文献中，已经提出了一些将社交网络结合到POI建议中的相关工作。例如，[16]放置一个社会正则化术语来约束用户特征向量的估计，假设朋友会分享相似的兴趣。同时也有研究者提出将地理影响考虑在内以协助POI推荐。例如，[28]利用线性模型来结合用户兴趣，社交网络和POI预测的地理距离。另一方面，[15]模拟了地域邻里在实例和地区层面的影响。在实例级别中，一个用户对位置的偏好被预测为其在该位置上的特殊偏好与该位置的最近邻居的组合。在地区层面上，它放置了一组套索惩罚来学习位置特异性的潜在载体。然而，这些模型很少将地理上亲密的用户对彼此的签到活动的影响纳入矩阵分解。地理上密切的用户可能会有类似的兴趣，并应该有检查行为的潜在影响[23]。

而且，这些模型中很少能够解决位置推荐中的用户冷启动问题。受此推动，我们首先正式为每个用户定义三种类型的朋友：社交朋友，位置朋友和邻居朋友。用户的社交好友是指在LBSN中与该用户进行社交连接的用户组。用户的位置朋友表示与该用户签入相同位置的用户组。用户的邻居朋友是那些在地理上接近这个用户的用户。然后，**我们将他们的历史签到纳入具有不同损失函数的矩阵分解模型**。

通过对两个真实世界的数据集的分析，我们发现用户与他们的三类朋友有着相似的兴趣。因此，我们提出了一个两步框架来详细说明朋友的签到。在第一步中，我们设计了两种方法（即基于线性聚合和基于随机游走）来学习一组朋友的位置，每个用户最有可能喜欢的位置，她从来没有去过。因此，用户的签到分为观察签到，潜在签到和其他未观察签到。在第二步中，我们开发了两种损失函数来模拟这三种检查：基于平方误差的损失函数和基于排序误差的损失函数。具体而言，基于方差的损失将用户的签到视为具有不同置信度的积极，潜在和负面偏好的指示。基于排名错误的损失假定用户喜欢在任何可能的位置上观察到的位置，并且也倾向于在任何未观察到的位置上的潜在位置。我们用两个真实世界的数据集中的许多最先进的基准模型和不同的验证度量来广泛评估我们的模型。实验结果不仅证明了我们的模型对POI建议的改进，而且也显示了冷启动问题的有效性。

# 2. PRELIMINARIES

## 2.3 The Recommendation Framework

本文中的推荐任务被定义为：给定用户的历史签到位置，我们的目标是**向每个用户推荐她可能感兴趣但是之前没有访问过的top-K地点**。在本文中，我们提出了一个两步推荐框架。具体来说，第一步，我们从三类朋友那里学习一组潜在的地点，这将在第三节介绍。在第二步中，我们将每个人的学习潜在位置纳入具有不同误差损失函数的矩阵分解模型中，这将在第4节中介绍。最后，针对标准推荐，定位冷启动推荐和用户冷启动推荐，提出了不同的推荐策略。

# 3. LEARNING POTENTIAL LOCATIONS

社交网络在推荐中起着重要的作用[16,28,5,22]。然而，仅利用社交好友的历史位置，不能成功地模拟用户对于位置的偏好，这是因为**难以适当地模拟没有社交朋友的用户的偏好**，没有提及**处理用户冷启动问题（即，用户从未在任何地点登记过）**。为了解决这些问题，我们将利用三类朋友的特点：社交朋友，位置朋友和邻居朋友。前面的部分已经在LBSNs中表现出它们的重要性，即朋友们会分享类似的POI偏好。换句话说，用户可能对那些已经被他们的朋友签到的位置感兴趣，并且很有可能在下一次登记。

然而，这些位置数量非常大，随着位置的增加，计算效率低下，随着噪声的增加，预测的不准确性也会随之增加。因此，本节中的问题是为目标用户找到最可能的位置，定义为

为了获得每个用户i的潜在位置，我们提出了两种方法，即线性聚合和随机游走，以估计该用户在其朋友已经签到的每个位置j上的概率。然后我们用估计的概率对它们进行排序，并选择概率最高的S位置2。学习到的潜在地点将有助于在第4节做出准确的建议。

## 3.1 Linear Aggregation



## 3.2 Random Walk

随机行走重启已成功测量了图中两个节点之间的相关性[7,24]。在本节中，我们提出了一个随机游走的方法，表示为RW，以学习她的朋友已经访问过的位置j上的用户i的概率ppot ij。

我们构造了一个带有两种节点的有向图：用户（即目标用户和她的朋友）以及由她和她的朋友签入的位置。令y为列向量，其中yi表示随机游走在节点i的概率。A也是列标准化的转换矩阵，其中aij表示节点i跳转到节点j的概率。这里我们考虑三种类型的转移概率：（1）用方程（1）中的余弦相似度测量的用户之间的概率; （2）每个用户到每个位置的概率，如果用户签到相应的位置，则其为1，否则为零; （3）由归一化幂律函数测量的一对位置（j和k）之间的相似性，其定义如下



因此，用于更新每个节点的稳态概率的迭代公式如下给出



X为i的列向量，对j有签到为1

# 4. RECOMMENDATION MODELS

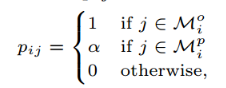
ε是一个调整参数，用于指示用户有一个小概率来选择另一个类别的一个位置。最近的许多工作都假设只对观察到的评分进行建模，该评分适应于明确的反馈数据集。然而，**签入数据集是隐式反馈数据集，其中我们没有明确的反馈用户偏好的位置**。换句话说，我们缺乏关于用户不喜欢哪个位置的实质证据。为了解决用户的冷启动问题和数据稀疏问题，我们建议**同时对用户的观测偏好，潜在偏好和不可观测偏好进行建模**。令j，k，h分别表示观测位置，潜在位置和未观测位置。

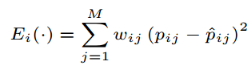
## 4.1 The Square Error based Model

在本节中，我们介绍基于增广平方误差的矩阵分解（ASMF）模型约束下的平方误差损失函数及其优化方法。

### 4.1.1 The ASMF Model

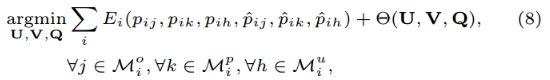
由于朋友之间的相似兴趣，**一个用户可能有机会去访问她的朋友之前访问过的那些可能的位置**，但是她从不签到。我们将每个用户的登记信息作为与不同信心相关的正面，潜在和负面偏好的指示。一个用户对于其签入的POI具有高度的自信。然而，她对这些潜在地点的潜在偏好以及对其他未访问地点的负面偏好具有低信心。相应地，我们将二元偏好变量pij增大到三元值如下



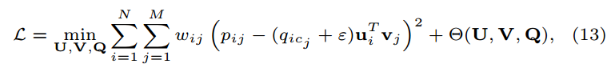
目标函数 

### 4.1.2 The Parameter Estimation









## 4.3 Incorporating Geographical Influence