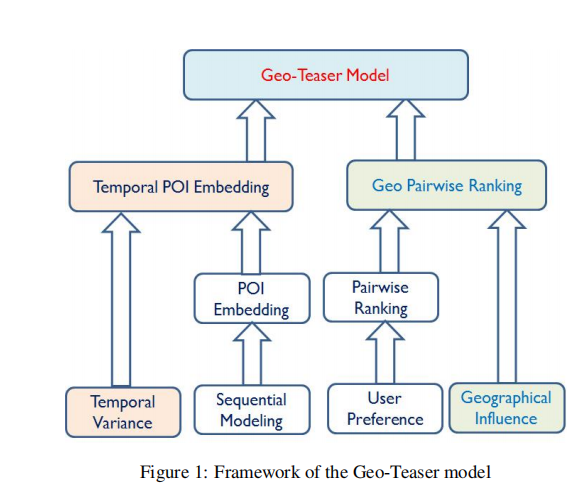
**Geo-Teaser：针对兴趣点推荐的地理时序顺序嵌入等级**

**摘要--**兴趣点（POI）推荐是基于位置的社交网络（LBSN）的重要应用，其从登记序列学习用户偏好和移动模式以推荐POI。以前的研究表明，对用户签入顺序模式进行建模对于POI推荐是必要的。在以往的工作中，使用马尔可夫链模型、递归神经网络和Word2vec框架对签入序列进行建模。然而，以往的序列模型都忽略了这样一个事实，即不同日的签入序列自然具有不同的时间特征，例如工作日的“工作”和周末的“娱乐”。本文针对这一挑战，提出了一种用于POI推荐的时间序列嵌入秩(Geo-teaser)模型。在Word2vec框架成功建模顺序上下文的启发下，我们提出了一个时态POI嵌入模型来学习特定时间状态下的POI表示。时态POI嵌入模型捕获序列中的上下文签入信息以及不同时间点的不同时间特征。此外，通过根据地理信息对未访问的POI进行判别，提出了一种将地理影响纳入成对偏好排序方法的新方法。然后我们开发一个地理上分层的成对偏好排名模型。最后，我们提出了一个统一的框架来推荐将这两种模型结合在一起的POI。为了验证该方法的有效性，我们在两个真实数据集上进行了实验。实验结果表明，Geo-teaser模型的性能优于现有的模型.与最好的基线竞争对手相比，Geo-teaser模型在两个数据集上都提高了至少20%。l指标。



**介绍**

总之，我们提出了一个用于POI推荐的Geo-时序嵌入秩(Geo-teaser)模型，如图1所示。一方面，我们提出了一种**时态POI嵌入模型**来捕获上下文签入信息以及不同的时态特征。特别是，我们将一个用户在一天内的签入顺序作为一个“句子”来处理。然后，我们考虑特定时间状态下的每个序列，并定义时态POI，指的是以特定时间状态作为上下文的POI。在此基础上，提出了**时态POI嵌入模型**来学习POI表示和时态表示。另一方面，我们将地理影响纳入成对偏好排序模型中，并建立了地理层次成对偏好排序模型。另一方面，我们将地理影响纳入成对偏好排序模型中，并建立了**地理层次成对偏好排序模型**。传统上，我们假设用户更喜欢访问的POI而不是未访问的POI，并建立成对排名模型以了解用户对POI的偏好[13,39]。以前的研究[2，31]表明，用户更喜欢地理上与其访问的POI相邻的POI。这种地理特征激励我们通过分层成对偏好关系来推动传统的成对排名模型，该关系根据POI的地理信息区分未访问的POI。最后，我们提出Geo-Teaser模型作为统一框架，推荐结合时间POI嵌入模型和地理分层成对排序模型的POI。我们将这些贡献归纳如下：

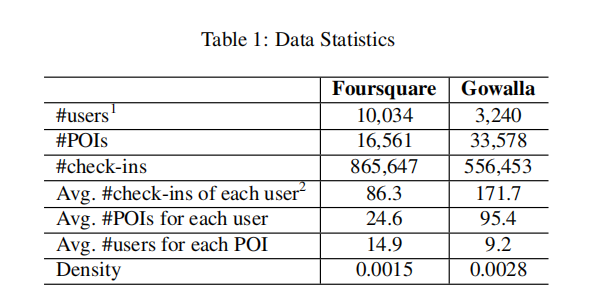
1. 我们提出了一种时态POI嵌入模型，该模型能捕捉到签入的顺序上下文和不同时间点的不同时间特征。特别是引入Word2vec框架，在嵌入空间中将每个POI映射为一个对象来学习POI之间的顺序关系。此外，在特定的时间状态下，我们从签入序列中学习时态POI表示。
2. 通过根据地理信息对未访问的POI进行判别，提出了一种将地理影响纳入成对偏好排序方法的新方法。特别地，我们为每个用户登记定义分层成对偏好关系：用户更喜欢所访问的POI而不是未访问的邻近POI，并且用户更喜欢未访问的邻近POI而不是未访问的非邻近POI。然后我们学习分层成对偏好来捕捉地理影响和用户偏好。
3. 结合时态POI嵌入模型和地理分层的双偏好排序模型，提出了Geo-teaser模型作为一个统一的框架。两个真实数据集的实验结果表明，Geo-Teaser模型优于最先进的模型。与最佳基线竞争对手相比，Geo-Teaser模型在所有指标的两个数据集上都至少提高了20％。

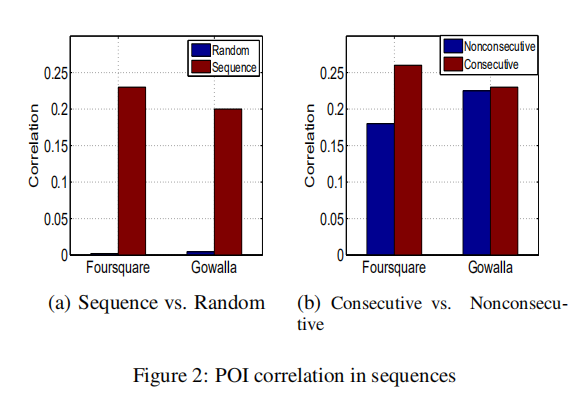
**3 数据集的描述与分析**

在这一部分中，我们首先介绍了两个真实的LBSN数据集，然后进行了实证分析，探讨了签入序列在一天内的性质。

**3.1 数据说明**

我们使用从现实世界LBSN中爬行的两个登记数据集进行数据分析。一个是从[8]中提供的Foursquare数据，另一个是[37]中提供的Gowalla数据。我们通过过滤POI签入不到5个用户和签入不到10次的用户对数据进行预处理。然后我们保留其余用户从2011年1月1日到2011年7月31日的记录。在预处理之后，数据集包含统计特性，如表1所示。





**3.2 经验分析**

我们进行数据分析，回答以下两个问题：1)一天序列中的POI是如何相互关联的？2)检入序列在不同的日子中表现如何？

我们研究了POI在一天的序列中的相关性，如图2所示。为了计算两个POI之间的相关性，我们根据登记记录构建用户POI矩阵。然后，我们测量了在这两个POI上签入的用户的Jaccard相似度与POI对的相关性。在图2(a)中，我们计算了所有用户序列中POI对的平均相关值，并将其与5，000个随机POI对的平均相关值进行了比较。我们观察到序列中POI的相关性比随机对高出100倍左右，对Gowalla来说是50倍左右，这激发了序贯建模的动机。在图2(B)中，我们比较了序列中连续对和非连续对的相关性。以(l1，l 2，l 3)序列为例，(l 1，l 2)和(l 2，l 3)是连续对，(l 1，l 3)是非连续对。我们还计算了所有序列的平均值，供所有用户进行比较。我们观察到，非连续对包含与连续对的可比相关性。因此，不仅连续的POI高度相关[3，39]，一个序列中的所有POI都与上下文属性高度相关。因此，仅用马尔可夫链模型来模拟连续检验员的传递概率或用张量因式分解对连续检验员的关联进行建模是不令人满意的。这一观察促使我们通过word2vec框架对整个序列进行建模。

我们探讨了不同日期的不同时间特征如何影响用户的签入行为。以前的工作[38，39]显示，用户签入在不同的日子表现出不同的模式，特别是在工作日和周末。图3显示了所有用户在一周的不同时间(周一到周日)不同时间的累积签入次数。根据图3中累积签到的统计数据，我们观察了不同时段的星期几登记模式：用户在下午晚些时候和周末晚上的工作日从下午16:00到上午3:00开始办理更多的办理登机手续。因此，周六和周日采取类似的模式，而从周一到周五的日子则采取与周末不同的类似模式。我们可以推断，工作日和周末对用户的签入行为有两种类型的影响。因此，对序列模式的建模应该包含这个时态特征。

**4 方法**

在这一部分中，我们首先提出了时态POI嵌入模型来捕捉时序建模的各种时态特征。接下来，我们描述了地理层次的两两偏好排序模型。然后，结合时态POI嵌入模型和地理分层成对偏好排序模型，提出GeoTeaser模型作为推荐POI的统一框架。最后，给出了Geo-teaser模型的学习过程。为了帮助理解这篇论文，我们在表2中列出了一些重要的符号。

**4.1 时态POI嵌入模型**

我们提出了一种时间POI嵌入方法来学习顺序模式，该模式从用户登记序列中捕获POI的上下文信息以及各种时间特征。与将用户的所有签入都视为“句子”的工作[18]不同，我们将用户一天的签入视为“句子”。 因为在不同日期的连续签到可能会持续很长时间并且不是高度相关的。 此外，我们假设不同日期的登记序列表现出各种时间特征。然后，我们学习POI嵌入在具有特定时间状态的序列中。为了更好地描述模型，我们提出了以下一些基本概念。

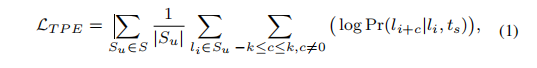
**定义1（登记入住）。**登记是一个三元组<u，l，t>，其描绘了在时间t访问POI 1的用户u。

**定义2(签入顺序)**。签到序列是一天内用户u的一组签到，表示为

Su = {<l1,t1>,…,<ln,tn>}，其中t1到tn属于同一天。 为简单起见，我们表示Su = {l1 ,…,ln}。

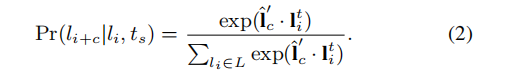
**定义3(目标POI和上下文POI)。**在序列Su中，选择的li是目标POI，而Su中的其他POI是上下文POIs。

我们在Skip-Gram模型[22]的基础上提出了时态POI嵌入模型。如图4所示，给定目标POI li和序列时态ts，我们学习从li- k到li + k的上下文POI的表示。这里k是控制上下文窗口大小的参数。此外，时间状态ts由两个选项组成，工作日和周末。因为我们想区分工作日和周末，这描绘了日级的各种时间特征，如图3所示。形式上，给定序列Su及其时间状态ts，我们的模型试图通过最大化以下函数来学习时间POI嵌入，

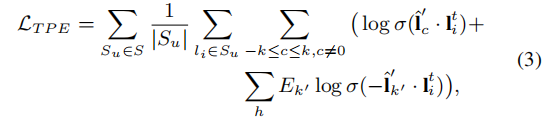


其中S是包含所有用户的所有序列Su的集合。LT P E的目的是最大化上下文POI对所有序列的条件出现可能性。

此外，我们用一个Softmax函数来表示概率Pr(li+c| li, ts)。为了更好地描述，我们引入了两个符号，定义如下，其中⊕是级联算子，、和ts分别是输出层上下文POI、目标POI和时态的潜在向量。因此，我们得到，因此，概率Pr(li+c | li，ts)可以表示为，



由于方程（2）中集合L的大小很大，我们利用负抽样技术[22]来有效地学习模型。这样，目标函数就可以用一种更容易优化的新形式来表示，



其中是采样的负POI，h是负样本的数量，σ（·）是sigmoid函数，E（·）表示计算所有生成的负样本的期望值。这里我们在[22]中采用相同的策略，即使用单字母分布来绘制负样本。

**4.2 地理层次成对排序**

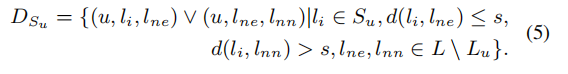
我们提出了地理层次的两两偏好排序模型，该模型将地理影响纳入成对排序模型。签入活动是一种类似于网页点击[14，20]的内隐反馈。为了学习这种隐式反馈，我们利用贝叶斯个性化排名(BPR)准则[26]来了解用户对POI的偏好。BPR是一种两两排序模型，它是在假设用户比未访问用户更喜欢访问的POI的基础上学习成对的用户偏好的。在我们的地理分层成对排名模型中，我们使用POI的地理信息区分未访问的POI。以前的研究[2,33,37]观察到用户更喜欢被访问的POI而不是远离POI，我们可以区分未访问的POI并定义相邻的POI和非相邻的POI如下。

**定义4(相邻POI和非邻接POI)。**对于每次登记<u，li>，相邻POI是与li的距离小于或等于阈值s的POI，而非相邻POI是距离大于s的POI。

此外，对于每次登记<u,li>，我们定义分层成对偏好关系：用户更喜欢被访问的POI li而不是未访问的邻近POI lne，并且优先于未访问的邻近POI lne而不是未访问的非邻近POI lnn 。将**d（li，lj）表示为两个POI li和lj的距离**，我们表示登记<u，li>的分层成对偏好关系，如下所示：



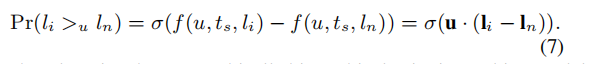
假设L是POI的集合，并且Lu是用户u的访问POI，满足等式（4）的序列Su的分层成对偏好关系集定义如下：



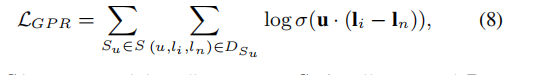
目前，学习地理层次成对排序模型与DSu中的偏好关系模型是等价的。我们采用MF模型来构造偏好评分函数。我们使用lti=li⊕ts来表示时态POI潜在向量，这与时态POI嵌入模型是一致的。此外，我们还定义了uˆ=u⊕u，然后将分数函数表述为：



其次，我们用Sigmoid函数表示成对的偏好概率。设pr(li>uln)表示用户u比ln更喜欢POI li的概率，σ(·)是Sigmoid函数。然后，偏好集中的每一对可以表示为，



因此，学习地理层次成对的排序模型相当于最大化以下功能，



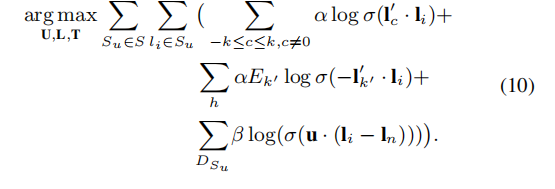
其中，S是包含所有用户的所有序列SU的集合，DSu是序列Su上的分层配对偏好关系。

**4.3 Geo-teaser模型**

我们建议将Geo-Teaser模型作为统一框架来推荐结合时间嵌入模型和成对排序模型的POI。学习Geo-teaser模型等价于最大LT P E和LGP R，

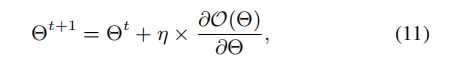


其中α和β是交换顺序建模和偏好学习模块的超参数。我们期望通过在Geo-Teaser模型中学习时间POI嵌入和地理上成对偏好关系来获得用户，POI和时间状态表示。分别用Eq(3)和Eq(8)代替LT P E和LGP R，我们可以通过以下目标函数学习Geo-Teaser模型：



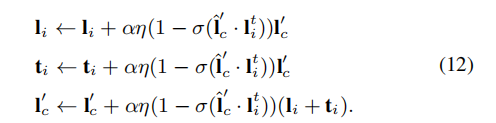
**4.4 Learning**

我们采用交替迭代更新程序，并使用随机梯度下降(SGD)来学习目标函数。为了学习模型，对于每个采样训练实例，我们分别计算LT P E和LGP R的导数，然后沿着上升梯度方向更新相应的参数，

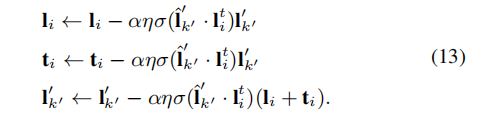


其中Θ是训练参数，η是学习率。

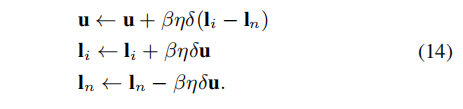
具体来说，对于登记<u，li>，我们计算LT P E的随机梯度下降。首先，我们得到上下文POI lc的更新规则，

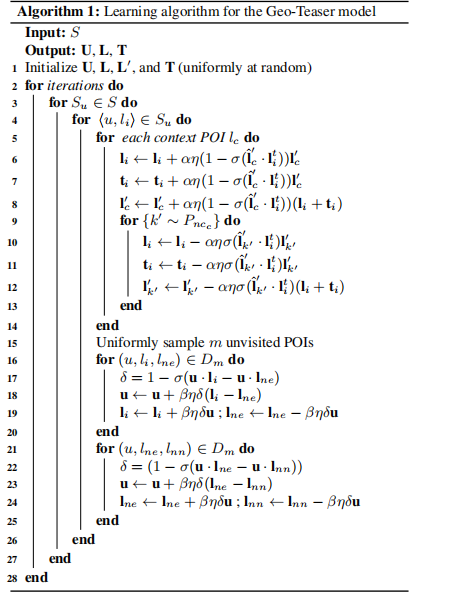


然后，我们更新负样本如下，



为了更新LGP R，我们计算DSu中每个偏好对(u, li, ln)的随机梯度。定义δ= 1-σ(**u**·li -**u**·ln)，我们更新如下参数，





算法1给出了学习Geo-teaser模型的细节。S是所有序列的集合，Su是用户u的序列。**U**、**L**和**T**是用户、POI和时态的特征矩阵。L’是Skip-Gram模型中的输出层POI矩阵，是辅助学习参数。我们使用标准方法[22]来学习序列中的POI表示，如算法1中从第5行到第14行所示。接下来，我们利用Bootstrap采样生成m个未访问的POI，然后根据它们与访问的POI的距离将未访问的POI分类为相邻POI和非相邻POI。在此基础上，我们建立了每一次签入的Dm配对偏好集<u, li>。这里。然后，我们学习DM中每个实例的参数，从算法1的第15行到第25行。

在学习Geo-teaser模型之后，我们得到了用户、POI和时态的潜在特征表示。然后，根据偏好评分函数，我们可以估计用户u在时态ts下对候选POI l的签入可能性。此外，我们使用方程（6）进行分数估计。最后，我们对候选POI进行排名，并为每个用户选择具有最高估计可能性值的前N个POI。

**可量测性。**对于一次登记，学习时间嵌入模型花费O(k·h·d)，其中k，h和d分别表示上下文窗口大小，负样本数量和潜在向量维度。对于算法1中从第15行到第25行的成对偏好学习，我们对未经访问的POI进行采样，这可以生成最大O(m2)成对偏好元组。每次入住，学习程序成本为O(m2·d)。因此，我们的模型的复杂度是，其中C是所有签入的集合。对于k，h，m和d是固定的超参数，所提出的模型可以在O(|C|) 中被视为线性的。此外，为了使我们的模型更有效，我们转向异步随机梯度下降（ASGD）[25]并以解锁方式并行运行算法。由于LBSNs中POI的登记频率分布遵循幂律[31]，这导致不常见的POI的长尾，这保证了使用ASGD来并行参数更新。

**5 实验评价**

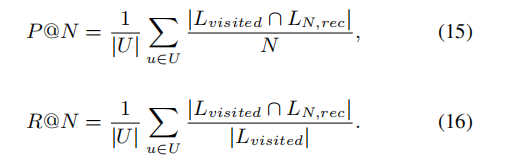
我们进行实验，寻求以下问题的答案：1)与最先进的推荐方法相比，Geo-teaser模型的性能如何？2)各分量(即不同的时间特征和地理影响)对模型性能的影响如何？3)参数对模型性能的影响如何？

**5.1 实验设置**

实验中使用了两个真实世界的数据集：一个来自[8]中提供的Foursquare数据集，另一个来自[37]中的Gowalla数据集。表1显示了数据集的统计信息。 为了使我们的模型满足推荐未来签到的情况，我们选择每个用户签到的前80％作为训练数据，其余20％作为测试数据，遵循[3,35]。

**5.2性能指标**

在本文中，我们通过精确性和召回性对模型的性能进行了比较，这两种方法通常用于评价POI推荐系统[6，13]。为了评估前N个推荐系统，我们将精度和召回分别表示为P@N和R@N。在我们的POI推荐任务中，P@N测量恢复的POI与N个推荐的POI的比率，R@N表示恢复的POI与测试数据中的POI集合的比率。然后我们计算平均精度并召回所有用户进行评估。假设Lvisited表示测试数据中相应访问的POI集合，LN,rec表示推荐的POI集合，则P@N和R@N的定义如下：



**5.3 模型的比较**

先前的工作[14，20]观察到，将签入作为内隐反馈处理更好地建模用户偏好。因此，我们将我们的模型与WRMF[10，24]和BPRMF[26]进行了比较，这是为捕获隐式反馈而设计的最先进的协作过滤模型。为了证明该模型的有效性，我们将其与四种最新的POI推荐方法进行了比较：LRT[6]、LORE[35]、Rank-GeoFM[13]和SG-CWARP[18]。

* **BPRMF[26]：**贝叶斯个性化排序矩阵分解(BPRMF)是一种流行的两两排序方法，它对隐式反馈数据建模以推荐顶级N项。
* **WRMF[10，24]：**加权正则矩阵分解(WRMF)模型是针对隐式反馈排序问题而设计的。我们将POI lj处的用户ui的权重映射函数设置为wi，其中是签到次数，遵循[20]中的设置。
* **LRT[6]：**带有时态效应模型的位置推荐框架(LRT)是一种最新的POI推荐方法，它捕捉POI推荐中的时间效应。
* **LORE[35]：**LORE是最先进的模型，利用顺序影响的位置推荐。与其他工作[3,30]相比，LORE采用了整个序列的贡献，而不仅仅是连续的登记顺序影响。
* **Rank-GeoFM [13]：**Rank-GeoFM是一种基于排名的地理分解方法，它将潜在排名模型中的地理和时间影响结合起来。
* **SG-CWARP[18]：**SG-CWARP是最新的工作，它利用Word2vec框架对顺序上下文的签入进行建模。

**5.4实验结果**

在下文中，我们展示了关于精确度和召回率的实验结果，表示为P@N和R@N，用于前N个POI推荐任务。由于模型的比较结果与N的不同值(如1、5、10和20)是一致的，因此在[6，7]之后，我们在5和10处显示了有代表性的结果。在适当的参数设置下，所有模型都能达到最佳的性能。

***5.4.1性能比较***

图5展示了不同模型的实验结果。我们发现，所提出的Geo-teaser模型比所有的基线都具有更好的性能。与Rank-GeoFM（包含地理影响和时间影响的最先进模型）相比，Geo-Teaser在所有指标的两个数据集上实现了至少28％的改进。这验证了我们的顺序建模的有效性，以及融合各种时间特征和地理影响的手段的有效性。SG-CWARP是最好的基线竞争对手，它通过Skip-Gram模型与马尔可夫链模型(即LORE模型)验证了顺序模式建模的优越性。我们的Geo-Teaser模型在所有指标的两个数据集上都优于SG-CWARP至少20％，这验证了我们将各种时间特征和地理影响纳入改进POI推荐的策略。此外，我们观察到，在Gowalla上模型的精度优于Foursquare，但在召回方面则更差。原因在于Gowalla中每个用户的测试数据大小都大于Foursquare。如表1所示，Gowalla中每个用户的平均签到大约是Foursquare的两倍。 根据方程（15）和方程（16）中的指标，结果是合理的。

***5.4.2 模型讨论***

在这一部分中，我们将探讨每个组件，即不同的时间特征和地理影响，如何影响模型的性能。Geo-teaser模型从两个方面对SG-CWARP进行了改进，捕捉了不同的时间特征和地理影响。忽略各种时间特征和地理影响，我们建议将SG-BPRMF模型作为我们提出的Geo-Teaser模型的基本版本。SG-BPRMF使用Skip-Gram模型对序列进行建模，使用BPRMF来捕获用户偏好，这相当于SG-CWARP。此外，我们将各种时间特征结合到SG-BPRMF中并提出Teaser模型。在下面，我们比较了SG-BPRMF，Teaser和Geo-Teaser，以显示不同的时间特征和地理影响对模型的影响。

图6显示了模型性能。 我们观察到Teaser模型对所有度量的两个数据集上的SG-BPRMF至少提高了约10％，这表明合并各种时间特征可以提高模型性能。此外，Geo-Teaser模型在两个数据集上的Teaser模型至少提高了约15％。这意味着我们通过区分未访问的POI来结合地理影响的策略是有效的。

***5.4.3 参数影响***

在本节中，我们将展示三个重要的超参数α，β和s如何影响模型性能。α和β平衡了顺序影响和用户偏好。s显示了我们地理模型的敏感性。

我们调优α和β以了解如何权衡顺序建模和用户偏好学习，如图7所示。在参数更新过程中，α和β都与学习速率η一起出现。没有必要分别调整这三个参数。我们能够将学习速率η吸收到α和β中。换句话说，我们设置了α←α·η，β←β·η。我们避免调整学习速率η，而是通过调优α和β来控制更新步骤的大小。因此，α和β应该足够小，以确保收敛。假设α和β的值相同，我们调优α以改变学习速度。当α=0.0 5时，模型的性能最好。然后我们设置α=0.0 5，并更改β以查看模型性能随β/α的变化情况。如果β/α∈[0.25，0.5]，

Geo-teaser就能获得最好的性能。

在Geo-teaser模型中，我们将未访问的POI划分为相邻的POI和非邻近的POI，根据阈值距离s构成一个新的偏好集。在这里，我们选择不同的s值来查看这个参数是如何影响模型性能的，如图8所示。这里是以公里为单位计算的。我们观察到，Geo-teaser模型在s=10时达到了最佳性能。

**6 结论和进一步工作**

本文研究了POI推荐问题。我们提出了时间POI嵌入模型来捕获签到的连续上下文和不同日期的各种时间特征。此外，我们还提出了地理层次成对排序模型，通过引入地理影响来提高推荐性能。最后，我们提出了Geo-teaser模型作为一个统一的框架，将这两部分结合起来推荐POI。两个数据集Foursquare和Gowalla的实验结果表明，我们的模型优于最先进的模型。与SG-CWARP模型相比，所提出的Geo-Teaser模型在所有指标的两个数据集上都至少提高了20％。

我们今后的工作如下：1）由于我们在本文中仅考虑一天的序列，我们可能会讨论将来的其他场景，例如，序列由连续签到组成，其间隔在固定时间阈值以下，例如，四小时或八 小时。2)在本系统中可以包含更多的用户评论和社会关系等信息，以提高系统的性能。