**GeoMF++: Scalable Location Recommendation via Joint**

**Geographical Modeling and Matrix Factorization**

位置推荐是帮助人们发现有吸引力的位置的重要手段。 然而，极端稀疏的用户位置矩阵导致严峻的挑战，因此有必要考虑用户移动性数据的隐式反馈特性并利用位置的空间信息。为此，我们在已有的GeoMF框架的基础上，提出了一个可扩展和灵活的框架GeoMF，用于联合地理建模和基于隐式反馈的矩阵因式分解。然后，我们提出了一种有效的参数学习优化算法，它与数据大小和所有位置的邻居网格总数成线性关系。GeoMF可以从两个角度得到很好的解释。首先，它包含了二维核密度估计，从而捕捉用户移动数据中的空间聚类现象；其次，它与广泛使用的邻域加性模型和图拉普拉斯正则化模型有很强的联系。我们最终在两个大型LBSN数据集上评估GeoMF，包括暖启动和冷启动两种情况。实验结果表明，GeoMF在NDCG和召回方面的表现一直好于最新的技术和其他相竞争的基线。更重要的是，GeoMF随着数据大小和潜在空间维数的增加，具有更高的效率和可扩展性。

**1 INTRODUCTION**

随着智能移动设备的普及和多种定位技术的融合，人们获得有关其位置的实时信息变得更加容易。这一发展引发了基于位置的社交网络(LBSNs)的出现，如Foursquare、Jiepang、FacebookPlace等。这种出现不仅使基于位置的社交成为社交互动的新形式，而且还帮助人们加快了对周围环境的熟悉。为了实现后一个目标，位置推荐已成为一个重要手段。

由于大规模用户移动性数据的易访问性和社交网络信息的包含性，位置推荐近年来得到了广泛的研究。来自LBSNs的用户移动性数据(即签入)只包括用户曾经去过的位置的历史，因此他们可能更喜欢。访问频率反映了她积极偏好的信心。然而，用户从未访问过的位置要么真的没有吸引力，要么未被发现但可能具有吸引力。如果没有提供辅助信息，这两种情况通常很难区别开来。移动行为的这些特征使我们可以将它们视为隐式反馈并利用加权正则化矩阵分解来进行位置推荐，因为它更好地处理稀疏性问题并且在经验上优于其他方法[11,25]。

然而，由于用户位置矩阵的极度稀疏性，该算法仍然面临着严峻的挑战，推荐性能相对较低。幸运的是，这些挑战可以通过考虑到地点的地理信息而得到进一步缓解。随着位置的地理信息的存在，空间聚类现象[29]表明个体访问位置倾向于聚集在一起，已经在LBSN上的用户移动行为中被揭示[33]。这一现象已通过地理建模被用于位置推荐。以往的地理建模算法包括参数化方法[33]和非参数方法[35]。任何两个访问地点之间的距离分布。为了提高地理建模的效率，文[3，18]提出了针对不同访问地点的二维地理聚类算法。为了避免设置n簇数，二维核密度估计已经发展[15]。然而，这些地理建模算法不依赖于协同过滤，因此通常基于启发式appro将它们集成在一起进行位置推荐。疼痛就像线性组合。

为此，我们扩展了二维核密度估计，提出了一种基于优化的地理建模算法。它以用户活动区域和位置影响区域为基础，通过内部产品经营者来估计用户对地点的地理偏好。此外，一个一致的客观目标开发了加权正则矩阵分解。因此，地理建模可以无缝地结合到矩阵分解中。特别是，我们建议GeoMF分别用户活动区域和位置影响区域来增强用户和位置的潜在因素，如图2所示。以这种方式，用户对位置的偏好被建模为在增强空间中它们之间的内部产品，包括用户的基于兴趣的偏好和该位置的地理偏好。 如果用户对该位置的地理偏好不为零，则用户的活动区域与该位置的有影响区域相交，使得该位置可从用户的活动区域到达。然后，我们提出交替优化用于学习用户/位置潜在因子与加权最小二乘和学习用户活动区域与稀疏和非负加权最小二乘。然而，基于时间复杂度的分析，它尤其在学习非负用户活动区域时遇到计算问题，尽管我们揭示了关于提高学习算法效率的两个重要属性。

为了提高效率和灵活性，我们进一步提出了GeoMF，通过将位置影响区域映射到与加权正则化矩阵因子形成的相同的潜在空间中。如图3所示。因此，我们通过用户的潜在因素的点积与位置 - 影响区域的潜在因子的加权和来表达用户对位置的地理偏好。 与GeoMF不同，用户活动区域不再用于GeoMF ++，但可以使用非负稀疏编码从用户潜在因子中恢复，因为用户潜在因素编码来自位置影响区域的传播地理影响。然后，我们开发了一种用于学习用户/项目/区域潜在因子的交替优化算法，并且表明它与数据大小和所有位置的有影响区域的总数线性地成比例。 有趣的是，在关于位置的空间相似性矩阵与位置影响区域近似的理论结果的支持下，GeoMF ++与广泛使用的邻域加法模型和图拉普拉斯正则化模型紧密相关。

最后，我们对两个大规模LBSN数据集的暖启动和冷启动场景进行了评估。实验结果表明，所提出的算法从地理建模中受益匪浅，并且在NDCG和Recall两个方案中，两个数据集的数据集始终优于几个竞争基线。然后研究了它们的训练效率，发现随着数据大小和隐空间维数的增加，GeoMF比基线具有更高的效率和可扩展性。

本文是我们之前的论文[16]的扩展，其中我们提出了GeoMF用于联合地理建模和基于隐式反馈的矩阵分解，因此地理建模可以无缝地结合到矩阵分解中。在本文中，我们还为位置建议做出了以下贡献：

* 为了提高GeoMF的效率和灵活性，我们建议使用GeoMF ++将位置影响区域映射到与加权正则化矩阵分解形成的潜在空间相同的潜在空间。在一种新的交替优化算法的基础上，GeoMF的尺度与数据大小和影响区域的总数呈线性关系。
* 我们建议利用非负稀疏编码从GeoMF学习的用户潜在因素中恢复用户活动区域，因为它们不再是GeoMF的一部分。案例研究表明，与GeoMF相比，恢复的用户活动区域是有意义和合理的。
* 我们提供了近似位置的空间相似性矩阵与位置影响区域的理论结果，并建立了GeoMF ++与另外两种广泛使用的联合地理建模和矩阵分解算法的强大联系。
* 我们揭示了GeoMF的两个重要性质，以提高学习非负用户活动区域的效率，从而加快学习用户活动区域的初始过程。
* 我们针对热启动和冷启动方案广泛评估了LBSN数据集上提出的算法。 除了我们之前的工作中使用的自爬行街罩数据集，我们还使用另一个大型公共Gowalla数据集，其中包含来自107K用户的6.4M签到。实验结果表明，新提出的GeoMF ++在两个数据集上始终优于GeoMF和其他竞争基线，不仅在培训效率方面，而且在NDCG和召回方面也具有推荐性能。

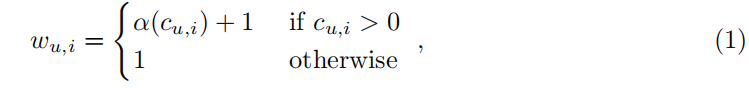
**2 RELATED WORK**

**3 PRELIMINARY**

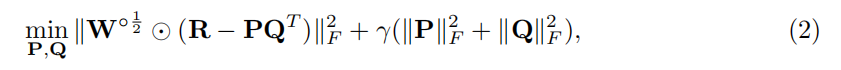
给定用户的移动性数据，位置推荐对用户位置偏好矩阵R = [𝑟𝑢，𝑖]∈{0,1}𝑀×N进行操作，其中有𝑀个用户，用𝒰= {𝑎1，···， 𝑎𝑀}和𝑁位置，用ℒ= {𝑙1，···，𝑙𝑁}表示。每个条目𝑟𝑢，𝑖表示用户𝑎𝑢是否已访问过某个位置𝑙𝑖。列向量r𝑢对应于矩阵R的第u行;列向量r𝑖对应于矩阵R的第i列。访问频率矩阵C∈N𝑀×𝑁与R的大小相同。ℒu= {𝑙𝑖∈ℒ|𝑟𝑢，𝑖> 0}表示用户all的所有访问位置。这里，大写粗体字母表示矩阵，小写粗体字母表示列向量，非粗体字母表示标量。

**3.1 Collaborative Filtering for Implicit Feedback**

给定偏好矩阵R，只有用户的正偏好被观察到，因为未访问的位置对用户来说要么是负面的，要么是未知的。访问频率表示积极偏好的置信度，因此较高的访问频率对应于更大的正偏好置信度。因此，给定偏好矩阵R，位置推荐是一类协同过滤(OCCF)问题[11，25]。在这种情况下，我们可以对每个用户随机采样一些负位置，并为它们分配比正位置更小的权重。为了更好地处理稀疏性挑战，我们甚至可以将所有未访问的位置视为负值，但在为它们分配较小权重的同时，为了提高效率，我们应该保证加权矩阵的特殊结构。例如，加权矩阵可以遵循稀疏和一阶结构，即加权矩阵W = [𝑤𝑢，𝑖]的每个条目𝑤𝑢，𝑖，



其中α（𝑐𝑢,i）> 0是相对于𝑐𝑢,i的单调递增函数。这样，它准确地编码了观察到的频率是用户积极偏好的信心。基于加权矩阵，隐式反馈协同过滤的目标函数称为加权正则矩阵分解(WRMF)，具体表现为：



其中⊙是Hadamard乘积算子，即矩阵和是W的Hadamard平方根。‖·‖𝐹是矩阵的Frobenius范数，只是矩阵中平方值之和的平方根。该目标函数涉及通过映射矩阵P∈R𝑀×𝐾和映射矩阵Q∈R𝑁×𝐾分别将用户和位置映射到具有维度𝐾«min（𝑀，𝑁）的联合潜在空间。在联合潜在空间中，用户对位置的偏好被建模为用户之间的内积。

值得注意的是，近似误差在用户位置偏好矩阵中的所有条目上求和，但可以通过交替最小二乘法有效地减少，并且每次迭代中的时间复杂度仍然与非零的总数成比例。 用户位置偏好矩阵中的条目。 我们将在后续章节中提供详细的分析。

**4 GEOMF**

加权正则矩阵因式分解在大多数隐式反馈数据集中显示出其优越性。然而，由于列入了地点的地理信息，仍然有空间对该算法的改进。尽管最近的一些研究已经利用空间聚类现象来改进位置建议[3，7，18，35]，但大多数研究几乎与协作过程无关。过滤，特别是矩阵因式分解。将地理建模与加权正则矩阵分解无缝地结合起来可能更有益处。首先，它更好地帮助应对稀疏的挑战和位置冷启动问题。其次，它有助于了解如何在提供地理信息时推荐位置，即如何在地理影响和个性化的基于兴趣的偏好之间自动平衡。为此，我们首先提出了GeoMF的联合地理建模和矩阵分解。

**4.1 Optimization-based Kernel Density Estimation**

矩阵因式分解是一个双线性模型，即给定一个映射矩阵固定，目标函数相对于另一个映射是线性的。在实际应用中，它采用交替最小二乘和随机梯度下降等优化方法来学习映射矩阵。为了它与地理的无缝结合在地理建模任务中，我们提出了二维核密度估计的加权线性回归方法。它涉及两个关键概念：用户活动区域和位置影响区域。 粗略地说，用户活动区域由用户将出现的空间区域组成，而位置影响区域是可以传播位置影响的那些空间区域。更具体和正式地，假设区域是通过将整个世界划分为𝐹空间网格的偶数大小来获得的，表示为𝒢= {𝑔1，𝑔2，...，𝑔𝐹}，我们有以下定义：

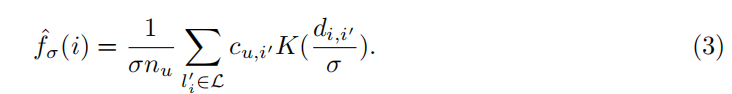
Definition 4.1 (User Activity Areas)：用户的活动区域包括一组空间网格，在这些网格中，用户可能会出现非负的可能性。

我们将用户au的活动区域表示为非负向量，每个条目𝑥𝑢，𝑗表示该用户出现在网格中的可能性𝑔𝑗∈𝒢。

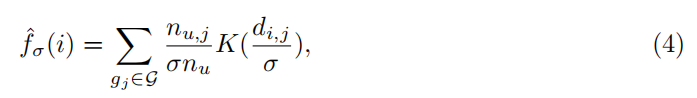
Definition 4.2 (Location Influential Areas)。位置的影响区域由一组空间网格组成，可以将该位置的影响传播到这些网格中。

我们还通过非负向量y𝑖∈R𝐹≥0表示位置的影响区域li，其中每个条目𝑦𝑖，𝑗表示从位置how传播到空间网格𝑔𝑗∈𝒢的影响程度。通常，有影响的区域因地点而异。 为简单起见，我们假设位置的影响区域是预先固定的，并且具有以该位置为中心的正态分布。特别是，从位置传播到网格的影响𝑙𝑖定义为，其中𝐾（·）是标准正态分布，σ是标准差，𝑑𝑖，𝑗代表地理距离 位置𝑙𝑖和空间网格的中心𝑔𝑗。图1显示了这样一个设置的示例。

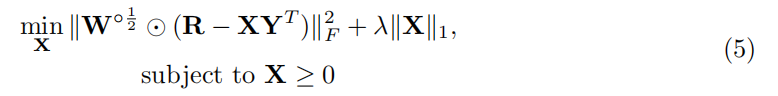
以这种方式设置影响区域的优点是x𝑢和y𝑖之间的内积表示用户访问位置上的二维核密度估计。特别是，根据核密度估计，在一位置处的𝑙𝑖用户𝑎𝑢的密度由估计



有如果ℒ中的位置被映射到空间网格𝒢，则此估计变为



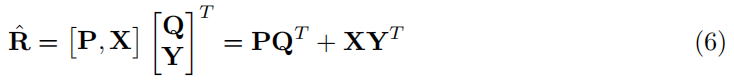
用户visit是空间网格的访问频率𝑔𝑗，而δ（𝑙𝑖'∈𝑔𝑗）表示位置𝑙𝑖落入空间网格𝑔𝑗。此时，通过将x𝑢与访问频率成比例地设置为相应的网格使得，用户au 在位置li的估计密度等于内积。为了考虑隐式反馈1的特性，我们将其视为变量并通过优化以下目标函数（称为GeoWLS [16]）来学习它，而不是将x𝑢经验频率赋值，



其中，我们逐行叠加每个用户的活动区域向量，以获得用户活动区域矩阵并逐行叠加各位置的影响面积向量，得到位置影响面积矩阵。‖X‖1是矩阵X的ℓ1范数，鼓励用户活动区域的稀疏解[23]。在用户活动领域实施稀疏性正规化的根本原因有两个：首先，用户通常被限制在几个长期停留的位置，例如家庭或工作场所; 第二，它也可以提高推荐的有效性和效率，如之前的会议论文[16]所示。

**4.2 Joint Model**

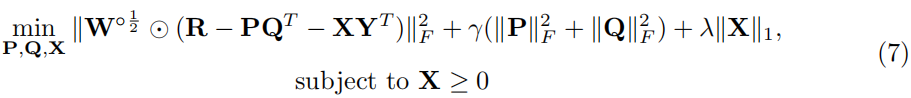
回想一下，在矩阵分解模型中，用户𝑎𝑢对位置𝑙𝑖的偏好表示为，因此地理建模与矩阵分解一致，使得无缝组合成为可能。特别是，我们利用X和Y分别在分解模型中增加用户潜在因子P和位置潜在因子Q，如图2所示。然后，提出GeoMF模型的估计偏好矩阵如下：



使用地理信息明确增加的一个原因是，仍然没有证据表明潜在空间已经包括这些信息。以这种方式，用户对位置的偏好被建模为增强空间中的内积，因此包括来自潜在空间的用户的基于兴趣的偏好和该位置的地理偏好。如果一个用户对一个位置的地理偏好不为零，则她的活动区域与该位置的有影响区域相交，以便该位置可从她的活动区域到达。

**4.3 Optimization**

在增强后，除了P和Q之外，我们还需要通过最小化下列目标函数来学习X：



通过交替优化方案实现该目标函数的最小化。它包括一个程序，用于在修复X时为用户和位置轮流学习潜在因素，另一个在修复所有潜在因素时涉及关于X的稀疏和非负加权最小二乘。在每个过程中，目标函数由于最小化而不增加，因此这种交替优化的迭代可以保证目标函数的不增加。因此，这种交替优化算法可以在几轮迭代之后收敛。

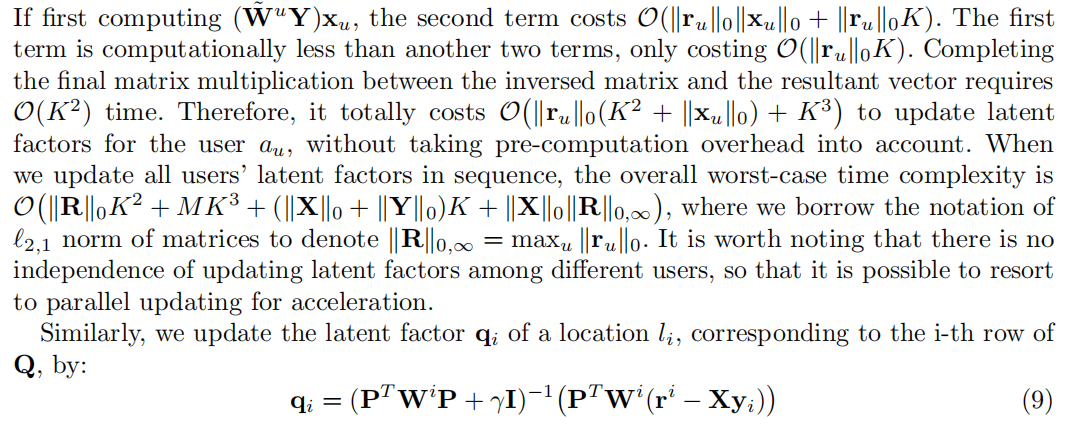
***4.3.1 Learning Latent Factors.*** 当固定用户活动区域矩阵X时，关于用户/项目潜在因子的方程（7）的优化类似于先前讨论的加权正则化矩阵因子分解中的交替最小二乘。更具体地，基于对应于第u行P的用户au的潜在因子Pu被更新



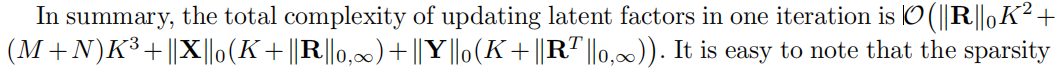
其中W𝑢是一个𝑁×𝑁对角线矩阵，条件是。在这里，由于我们设置了相同的权重，即对未访问的地点设置了1，因此有一个技巧可以通过使用来加速计算[11]，从而使是非零on，如果𝑟𝑢，𝑖0。特别是，。在这种情况下，第二部分独立于用户，因此可以预先计算，成本为𝒪（𝑁𝐾2），而第一部分仅需要𝒪（‖r𝑢‖0𝐾2），与用户的访问位置数量成比例𝑎𝑢。这里，矩阵(向量)的ℓ0范数是该矩阵(向量)中非零项的个数。对于𝐾×𝐾矩阵的逆，我们假设它需要𝒪（𝐾3）时间，即使存在更有效的算法，特别是对于正半定矩阵，但可能与通常较小的values值不太相关。 应用类似的技巧来计算其他部分



可以为所有用户预先计算最正确的术语，花费。



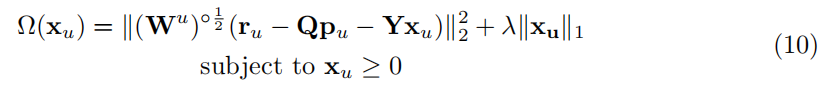
其中W𝑖是一个𝑀×𝑀对角线矩阵，条件是。应用类似的优化技巧，我们可以在中按顺序完成所有位置的潜在因子的更新。

总之，在一次迭代中更新潜在因素的总复杂性是

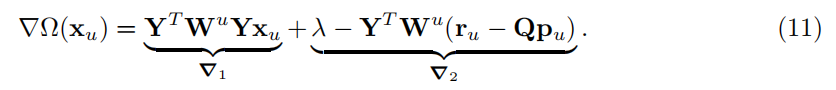
很容易注意到，X和Y的稀疏结构对更新这些潜在因素的效率也很重要。因此，在对用户活动区域矩阵X施加ℓ1范数的同时，我们还假设用于生成位置影响区域的二维正态分布被截断。换句话说，只有距离某一地点一定距离（即𝑑km）的区域才被视为其影响区域。这在一定程度上是合理的，因为正常分布通常随着距其中心的距离的增加而快速衰减。

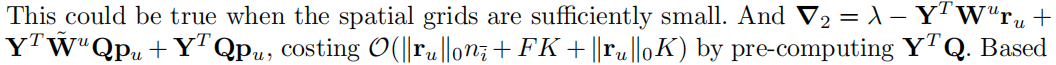
***4.3.2 Learning User Activity Areas.*** 现在让我们来学习用户活动区域矩阵X。当固定用户/项目潜在因素时，方程（7）中关于X的目标函数类似于稀疏和非负加权最小二乘问题，其可以进一步推广为有界变量最小二乘问题[12]。这类问题已经通过几种方法解决，包括有源集方法[12]，顺序坐标方法[6]和投影梯度下降方法[17]。 在这些方法中，投影梯度下降是高效的，并且已经在非负矩阵分解中进行了广泛研究，这也可以归结为与非负最小二乘相关的两个子问题[17]。投影梯度下降算法的一般思想是通过梯度下降更新参数，然后将更新的参数投影到由约束约束定义的可行区域中。然而，梯度下降中学习率的选择需要保证预测参数能够充分降低方程（7）中的目标函数。因此，我们利用[17]中提出的方法来更新用户活动区域矩阵。然而，由于加权矩阵的存在，Eq（7）相对于X的梯度是全矩阵。一次更新所有参数是不切实际的。因此，我们会独立更新每个用户的活动区域。

让我们重写关于用户𝑎𝑢的活动面积向量的目标函数，并放弃不相关的项，

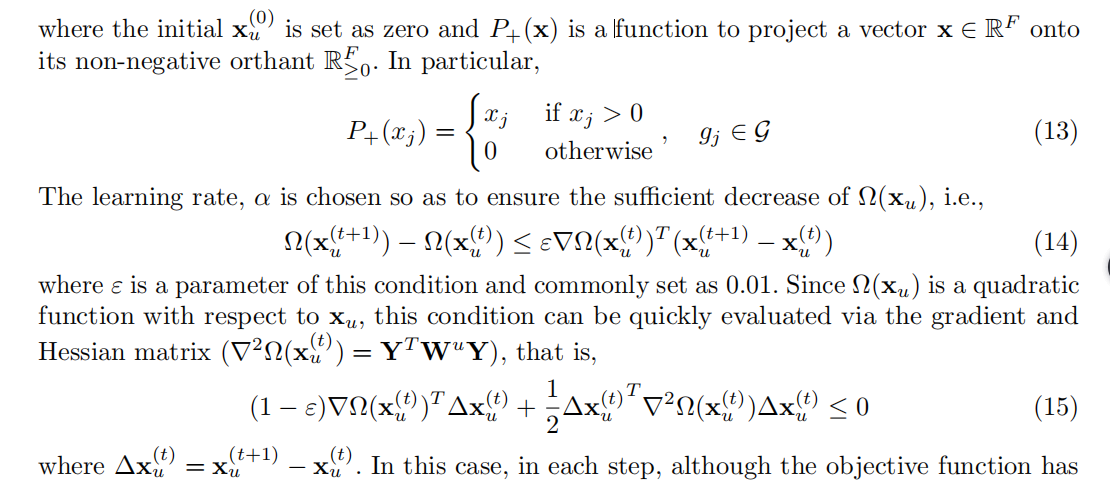


Ω(x𝑢)相对于x𝑢的梯度是



既然，其中)通过预计算Y𝑇Y第一次的花费，第二次花费。这里我们假设有相同数量的不同位置的有影响的区域，用𝑛𝑖来表示。当空间网格足够小时，这可能是正确的。，基于这个梯度，我们更新x𝑢如下：



在这种情况下，在每个步骤中，尽管目标函数已经充分减少，但是它需要基于一些启发式规则重复搜索有效的学习速率。 假设＃𝑡𝑟𝑖𝑎𝑙𝑠是搜索有效学习率的试验次数。 然后在搜索有效学习速率期间，以下两个观察结果可以使计算投影梯度变得更快。

