**基于位置的社会网络中利益点推荐研究综述**

**摘要**----随着移动设备，全球定位系统（GPS）和Web 2.0技术的快速发展，基于位置的社交网络（LBSN）吸引了数百万用户共享丰富的信息，例如经验和技巧。兴趣点（POI）推荐系统在LBSN中起着重要作用，因为它可以帮助用户探索有吸引力的位置，并帮助社交网络服务提供商为兴趣点设计位置感知广告。在本文中，我们对LBSN中的兴趣点推荐任务进行了简要的调查，并讨论了兴趣点推荐的一些研究方向。我们首先描述兴趣点推荐的独特特征，它将兴趣点推荐方法与传统推荐方法区分开来。然后，根据POI推荐算法将登记数据与哪种类型的附加信息集成，我们将POI推荐算法分为四类：基于纯登记数据的POI推荐方法，地理影响增强POI推荐方法，社会影响增强 POI推荐方法和时间影响增强了POI推荐方法。 最后，我们讨论了兴趣点推荐的未来研究方向。

**Introduction**

随着移动设备，全球定位系统（GPS）和Web 2.0技术的快速发展，基于位置的社交网络（LBSN）已经变得非常流行并且吸引了来自工业界和学术界的大量关注。典型的基于位置的社交网络包括Foursquare，Gowalla，Facebook Place和GeoLife等。在LBSN中，用户可以通过登记感兴趣的点（例如，餐馆，旅游景点和商店等）与他们的朋友建立联系，上传照片并共享他们的位置。 除了为用户提供社交互动平台外，更希望LBSN利用丰富的信息（社交关系，登记历史等）来挖掘用户对地点的偏好，并推荐用户可能感兴趣的新地方。推荐新的有趣地点的任务被称为兴趣点（POI）推荐。POI推荐系统在LBSNs中发挥了重要作用，因为它们不仅可以满足用户访问新地点的个性化偏好，还可以通过为用户提供智能位置服务（如位置感知广告）来帮助LBSN增加收入。

虽然推荐系统已经被许多电子商务网站广泛研究和成功采用，例如亚马逊，Netflix，Last.fm和淘宝等，但POI推荐系统最近刚刚出现。与传统推荐系统不同，POI推荐系统具有以下独特特征。

* **Geographical Influence.** 正如Tobler的“地理第一定律”所述，“一切都与其他事物有关，但近处事物与遥远的事物相关”（Tobler 1970）。 对于LBSNs，Tobler的第一地理定律意味着用户更喜欢访问附近的位置而不是远距离的位置，并且用户可能对围绕用户喜欢的POI的POI感兴趣。地理影响是区分POI推荐系统与传统推荐系统的最重要特征，严重影响用户的访问行为。
* **Frequency Data and Sparsity.** 在传统的推荐系统中，用户通常通过明确地提供项目（例如，书籍，电影，音乐等）的评级来表达他们的偏好，这些项目被转换为用户项目评级矩阵。额定值通常是数值，并且属于数值范围，例如[1,5]。评级越高越好。与传统的推荐系统不同，用户的偏好通过登记位置的频率来重新选择，这些频率通常被转换为用户位置登记频率矩阵。与额定值相比，频率数据具有较大的范围。例如，用户可能在某些位置签入数千次，而对于其他位置则可能签入几次。此外，用户位置检查在频率矩阵中的稀疏性显著高于用户项评等矩阵，这给POI推荐带来了更大的挑战。例如Netflix数据集的稀疏度约为99%，而Gowalla的稀疏度约为2.08×10−4。
* **Social Influence.** 基于朋友倾向于分享更多共同兴趣并且用户经常倾向于他们的朋友寻求建议的假设，传统推荐系统将社交关系与评级相结合以提高推荐质量。 一些研究（Ma et al.2008; Jamali and Ester 2010）表明社会关系被证明对推荐系统有益。然而，在POI推荐系统中，先前的研究（Ye，Yin和Lee，2010）显示，大约96％的用户分享的共享访问兴趣低于10％，表明大量朋友在POI方面没有任何共享。因此，社会影响对用户的检查行为产生有限的影响。

在本文中，我们提出了对现有POI推荐算法的回顾，并讨论了POI推荐的一些研究方向。根据POI推荐算法与登记数据相结合的附加信息的类型，我们将POI推荐算法分为四类：基于纯登记数据的POI推荐方法，地理影响增强POI推荐方法，社会影响增强POI推荐方法，时间影响增强了POI推荐方法。基于纯登记数据的POI推荐方法将登记频率作为评级，并假设两个用户在检查了大量常见POI时是相似的。然后，采用传统的协同过滤方法，通过平均大多数类似用户对候选POI的偏好来制定POI推荐。在地理影响增强的POI方法中，在POI推荐的过程中考虑用户和位置之间的距离或用户访问的POI与作为用户的新地点的POI之间的距离。 地理增强的POI推荐方法通常假设用户倾向于访问附近的POI，并且随着距离的增加，访问新地点的概率也会减少。社交影响力增强POI推荐方法利用朋友之间的社交关系来增强POI推荐并假设LBSN的朋友比非朋友分享更多共同的兴趣。 时间影响增强POI推荐方法假设用户的兴趣随时间变化，并且用户的访问行为经常受到时间的影响，因为用户在一天中的不同时间访问不同的地方。

本文的其余部分安排如下。我们在第2节中正式确定了LBSN中POI推荐的问题。在第3节中调查了基于纯粹的登记数据的POI推荐方法，地理影响增强的POI推荐方法，社会影响增强的POI推荐方法和时间影响增强的POI推荐方法。第4节，我们讨论未来的研究方向，并总结这项调查。

**Formalization of POI Recommendation**

在典型的LBSN中，POI推荐系统包括一组N个用户U = {u1，u2，...，uN}，以及一组M个位置L = {ll，l2，...，lM}，也是 称为POI。用户u访问的POI集合由Lu表示。每个位置都由<经度，纬度>进行地理编码。用户的登记信息被转换为用户位置登记频率矩阵C。C的每个条目cui表示用户u登记位置i的频率。签到的频率反映了用户对不同位置的偏好。通常，用户仅访问LBSN中存在的一小部分位置，因此矩阵C非常稀疏。此外，每个用户保持一个信任朋友列表，用户的社交关系被转换为社交关系矩阵S，其中suv表示社会信任的价值u。在大多数情况下，社交关系是二元的，suv = 1 意味着用户u和v之间存在社会关系，零意味着他们之间没有社会关系。

POI推荐系统的目标是根据用户“历史登记历史”来学习用户“隐式偏好”，并向用户提供用户可能感兴趣的新位置。

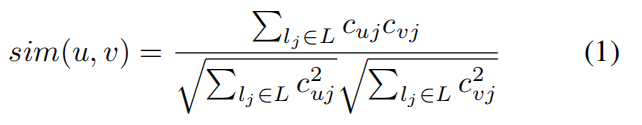
**The Taxonomy of POI Recommendation（POI推荐的分类学）**

在本节中，我们首先回顾基于POI推荐方法的纯登记数据。然后，我们将POI推荐方法划分为地理影响增强，社会影响增强和时间影响增强，根据哪种类型的附加信息与登记信息相结合，以提高POI推荐的质量。

**Pure Check-in Data Based POI Recommendation**

传统的推荐系统通过利用LBSN中没有的项目的显式评级来提出建议。 但是，LBSN记录的登记频率隐含地反映了用户对POI的偏好。 因此，为了产生POI推荐，一些研究（Berjani和Strufe 2011; Ye等人2011）采用传统的推荐算法，通过挖掘用户的登记模式来推断用户对POI的个性化品味。

通过可用的登记信息，现有的推荐方法可用于LBSN中的POI推荐，将POI视为项目，例如基于用户的协同过滤（Breese，Heckerman和Kadie 1998）和基于项目的协同过滤（Sarwar et al.2001; Linden，Smith和York 2003）。In（Ye et al.2011），Ye等人建议的基于用户和基于项目的POI推荐算法。基于用户的POI推荐方法假设类似用户对位置具有相似的品味，并基于大多数相似邻居的意见提出POI推荐。另一方面，基于项目的POI推荐方法假设用户对类似的POI感兴趣。基于用户的POI推荐算法的核心组成部分是如何计算用户u和v之间的相似权重sim(u, v)。在（Ye et al.2011）中，作者采用余弦相似性度量来估计sim(u, v)，定义如下。

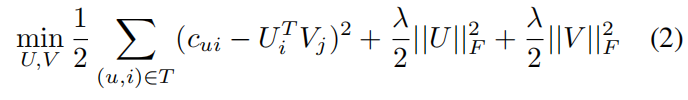


注意，如果用户u访问过位置j，则将C的每个条目cuj设置为1，否则为0。用这种方式表示的用户位置签入矩阵忽略了签入的频率。另外，要克服数据稀疏性问题。POI推荐，叶等人融合了基于用户和基于项目的方法来进行POI推荐。 他们的实验结果表明，基于用户的POI方法比基于项目的方法表现更好，因为少数用户访问了LBSN的许多POI，这导致与用户相似性相比不准确的项目相似性。

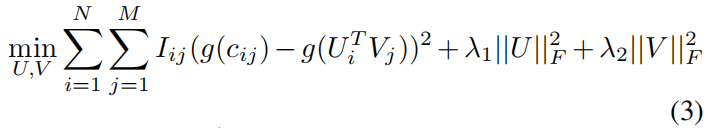
此外，上述基于内存的协作过滤也适用于POI推荐，LBSNs中的POI推荐也采用了基于模型的协同过滤方法。

自Netflix奖竞赛取得巨大成功以来，矩阵分解（Koren，Bell和Volinsky 2009; Mnih和Salakhutdinov 2007; Lee和Seung 2001）基于推荐算法因其在处理非常大的用户方面的有效性和效率而获得了很大的普及-项目评级矩阵。直观的是，矩阵因子分解技术可用于LBSN中的POI推荐。 Berjani等人。（Berjani和Strufe 2011）仅通过利用用户位置登记数据，在LBSN中提出了基于正则矩阵分解的POI推荐算法。他们认为，缺乏明确的评级是LBSN中POI推荐的主要问题。因此，他们使用二元和分箱偏好定义从登记数据中导出伪评级，并采用正则化矩阵分解（RMF）（Koren，Bell和Volinsky 2009）进行POI推荐。具体地，首先根据二进制或分箱优选定义来变换矩阵C的每个条目。然后，将导出的用户定位等级矩阵C“分解为两个低秩的潜在特征矩阵

它们提出的POI推荐算法通过最小化目标函数2来学习潜在特征矩阵U和V。



其中T表示已知派生评级的(u，i)对的集合，λ表示正则化参数。程等人（Cheng等人，2012）提出了基于POI推荐算法的基于概率矩阵分解（PMF）（Mnih和Salakhutdinov 2007）和概率因子模型（PFM）（Chen等人2009; Ma等人2011）。基于PMF的推荐方法假设高斯分布在观察到的登记数据上，并将高斯先验放置在潜在特征矩阵U和V上。基于PMF的POI推荐算法的相应目标函数定义如下



其中g（x）是逻辑函数. Iij是指标函数。如果用户i已检入位置j，则Iij取值1，否则为0。另一方面，基于PFM的POI推荐方法直接对登记频率进行建模，并将Beta分布作为潜在特征矩阵U和V的先验。在Gowalla数据集上的实验结果表明，基于PFM的POI推荐算法比基于PMF的POI推荐算法取得了更好的结果。

请注意，有一些研究纯粹利用用户的登记信息进行POI推荐。 随着LBSN中提供了丰富的附加信息，大多数研究工作将登记数据与其他信息结合起来产生POI推荐，因为据报道其他信息有助于提高传统推荐系统的性能，例如社会关系增强推荐方法（ Jamali和Ester 2010; Ma et al.2008; Yang，Steck和Liu 2012），内容增强推荐方法（Agarwal和Chen 2010; Wang和Blei 2011; Bao和Zhang 2014），标记增强推荐方法（Zhen，Li和Yeung 2009; Wu等人，2012）和项目属性增强推荐方法（Nguyen和Zhu 2013; Yu，Wang和Gao 2014）。

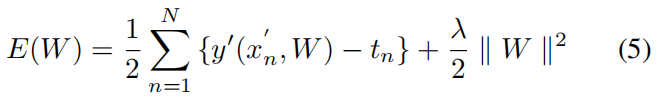
**Geographical Influence Enhanced POI Recommendation**

在LBSN中，用户和POI之间存在物理交互，这是区分POI推荐和传统项目推荐的独特属性。此外，Tobler的第一地理定律报告说“一切都与其他事物有关，但近处事物与遥远的事物有关”（Tobler 1970）。Tobler的第一个地理定律也表现为用户登记活动中的地理聚类现象。两种直觉导致了这种现象：（1）用户更喜欢访问附近的POI而不是遥远的POI;（2）用户可能对围绕用户喜欢的POI的POI感兴趣。一些研究（Ye et al.2011; Yuan et al.2013; Zhang，Chow，and Li 2014; Gao，Tang，and Liu 2012; Liu et al.2013）认为用户登记活动中的地理聚类现象，被称为地理影响，可用于改进POI推荐系统。

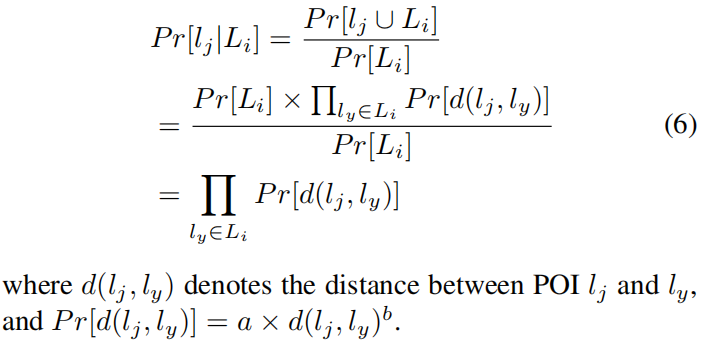
In（Ye et al.2011），Ye等人幂律分布（PD）对用户登记行为进行建模，并提出了一种基于地理影响的协作POI推荐算法，通过朴素贝叶斯算法。由同一用户访问的两个POI的登记概率y定义如下。



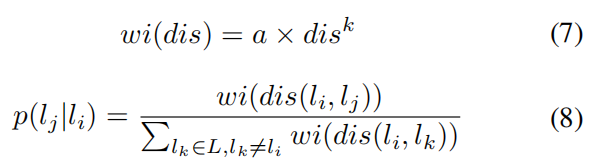
其中x表示两个POI之间的距离。a和b是幂律分布的参数，它们是从观察到的登记数据中学习的。然后，他们应用线性曲线拟合方法，通过最小化损失函数E（W）来学习从a和b导出的线性系数W。



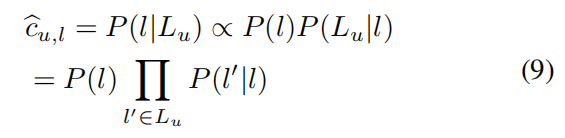
通过采用朴素贝叶斯方法，ui检查lj的似然概率如下。



另一方面，Yuan等人（Yuan et al.2013）通过做出不同的假设将地理影响整合到POI推荐中。 在（Ye et al.2011）中，提出的POI推荐算法假设用户检查新POI的概率是通过访问所有成对POI的概率的乘积估计的，每对由新POI组成，每个 之前访问的POI。虽然Yuan等人（Yuan et al.2013）假设用户从POI转移到另一个POI的意愿是他们距离的函数。具体而言，用户愿意访问远离POI的diskm以及用户将登记的概率lj，给定用户当前处于POI li，分别由等式7和8定义

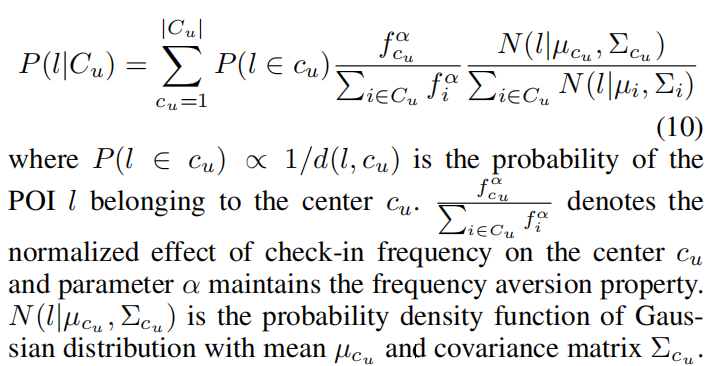


其中a和k是幂律分布的参数。最后，袁等人应用贝叶斯规则来计算给定用户u的历史POI Lu的每个新POI的排名分数P（l | Lu）。从形式上看

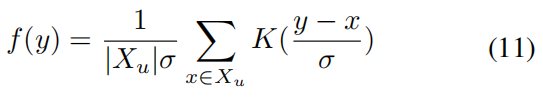


其中P(l)是数据集中所有用户检入POI 1的先验概率。他们的实验结果表明，他们提出的POI推荐在精确度和召回率方面比（Ye et al.2011）表现更好。

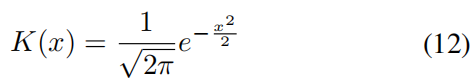
而不是制定幂律分布（PD）假设（Ye et al.2011; Yuan et al.2013），Cheng等人（Cheng etal.2012）假设用户倾向于在几个中心附近办理登机手续并通过建模捕捉地理影响 用户在某个位置登记的概率为多中心高斯模型（MGM）。 给定多中心集合Cu，用户u登记POI的概率定义为：

此外，MGM对登记数据采用贪婪的聚类算法来寻找中心。 他们的实验结果表明，MGM优于基于PMF和PFM的POI推荐算法。

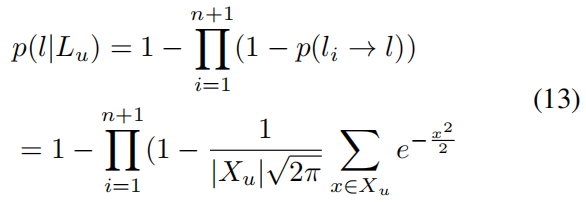
此外，Zhang et al。（Zhang，Chow和Li 2014）认为，当LBSN向用户推荐新的POI时，应该对个人用户的登记行为的地理影响进行个性化，并且不应将其建模为共同分布，例如PD（Ye et al.2011; Yuan et al.2013）和MGM（Cheng et al.2012）。到最后，张等人 使用核密度估计（KDE）（Silverman 1986）将地理影响模型化为每个用户的个性化距离分布。具体地，距离y的概率密度函数由等式11定义。



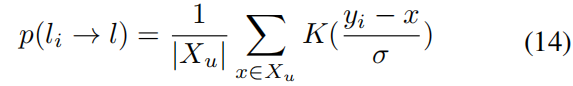
其中Xu是用户u住宅和Lu的联合集中每对POI之间距离的样本。K（.）是正常的内核：



给定联合集hu∪Lu的用户u访问新POI l的概率定义为：



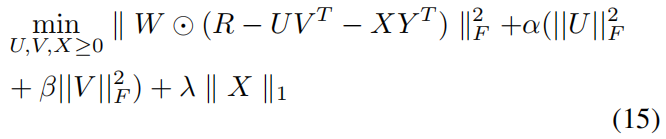
其中p(li→l)是由POIli的地理影响触发的用户u访问POIl的概率



他们的实验结果表明，与基于PD的POI推荐方法（Ye等人2011; Yuan等人2013）和基于MGM的POI推荐方法（Cheng等人2012）相比，他们提出的POI推荐方法提供了显着优越的性能。

另外，刘等人（Liu et al.2013）通过将地理影响与贝叶斯非负矩阵分解（BNMF）相结合，提出了POI推荐的地理概率因子分析框架。具体而言，他们使用高斯分布来表示采样区域上的POI，反映了地理第一定律。此外，贝叶斯非负矩阵分解模型用于从签入数据中捕获用户偏好。实验结果表明，所提出的POI推荐方法优于非负矩阵分解（NMF），RMF，PMF和BNMF模型。同时，他们发现基于NMF和BNMF的方法比基于PMF和RMF的方法表现更好。

最近，Lian等人（Lian et al.2014）提出了一种基于加权矩阵分解的POI推荐方法，命名为GeoMF。特别是，GeoMF分别利用用户的活动区域向量增加用户和POI的潜在特征向量，并影响POI的区域向量。此外，GeoMF对用户潜在特征向量和POI潜在特征向量施加稀疏和非负约束。在这个增强模型中，GeoMF不仅可以从二维核密度估计（KDE）的角度捕获地理聚类现象，而且可以解释为什么将地理影响整合到矩阵分解中有利于POI推荐系统的原因。GeoMF的目标函数定义如下

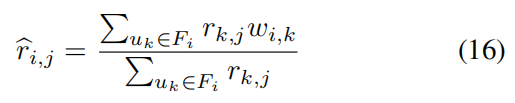


其中W是加权矩阵，其条目wui表示用户u对POI i的置信度。 X和Y分别表示用户的活动区域矩阵和POI的影响区域矩阵。 注意，POI的影响区域矩阵Y是通过使用二维核密度估计来计算的。实验结果表明，基于0/1评级矩阵的基于矩阵分解的POI推荐方法比基于频率矩阵的相同分解模型表现更好。此外，加权矩阵分解优于其他类型的矩阵分解模型，即矩阵分解（Berjani和Strufe 2011; Gao等人2013）和贝叶斯非负矩阵分解（Liu et al.2013; Cheng et al.2012））。

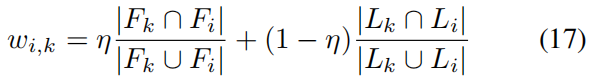
**Social Influence Enhanced POI Recommendation**

社会影响力增强推荐方法已在传统推荐系统中得到广泛探索，包括基于记忆的方法（Jamali和Ester 2009; Massa和Avesani 2007; Golbeck 2006）和基于模型的方法（Jamali和Ester 2010; Ma等人2008）。受到LBSN的朋友比非朋友更多共同兴趣的假设的启发，一些POI推荐方法通过考虑社会影响来提高推荐质量（Ye Yin和Lee，2010; Cheng等人2012）。

在（Ye，Yin和Lee，2010）中，Ye等人基于朋友的共同访问签到，提出了基于朋友的协同过滤（FCF）方法用于POI推荐。在进行POI推荐时，FCF仅考虑朋友的偏好，而不是LBSN的每个用户。在FCF中，根据等式16计算用户ui对lj的预测评级。

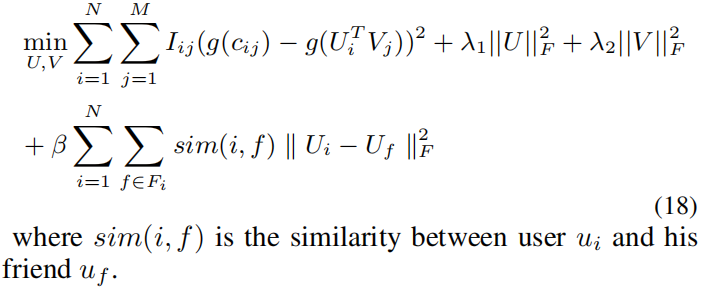


其中Fi是具有前n个相似性的朋友集合，wi，k是对英国的定向社会影响权重ui。请注意，FCF侧重于效率而非POI推荐系统的有效性。因此，在精确度方面，FCF对基于用户的POI推荐带来了微小的改进。在他们后来的工作中（Ye et al.2011），Ye等人根据他们的社交关系和他们的相似性来确定两个朋友之间的社会影响权重。办理登机手续的活动。通常情况下，



其中η是一个调优参数，Fk表示用户uk的朋友集。

Cheng等人（Cheng等人，2012）提出了具有社会正规化（PMFSR）的概率矩阵分解，其将社会影响整合到PMF中。 PMFSR通过最小化以下目标函数来学习用户的潜在偏好或POI的潜在特征。

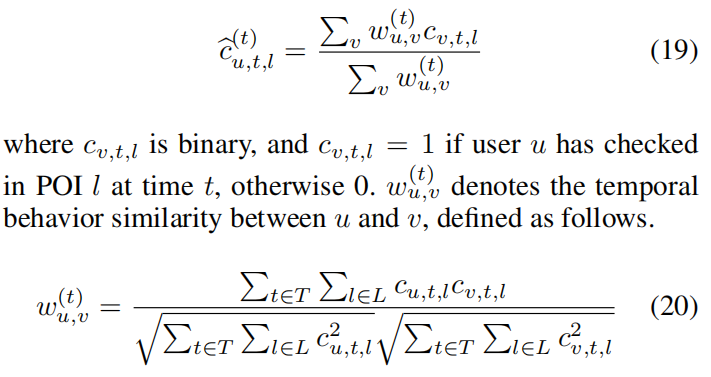


虽然社会影响对传统推荐系统的绩效有重要影响，但上述社会影响力增强的POI推荐方法的实验结果表明，社会影响力比地理影响和登记活动更重要。 我们认为用户的登记活动需要用户和位置之间的物理交互，这限制了社会影响力的贡献。在传统推荐系统中观看电影，听音乐和购买产品时，不受物理交互的限制，因为用户可以进行 这些活动通过网站进行。

**Temporal Influence Enhanced POI Recommendation**

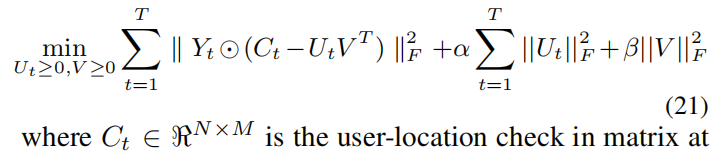
存在考虑传统推荐系统中的时间影响的研究，例如基于矩阵分解的方法（Koren 2010），基于随机游走的方法（Xiang等人2010）。然而，在传统推荐系统中，时间影响被用作衰减评级权重的因素。相反，POI推荐系统通常使用时间影响来针对特定时间状态进行POI推荐。

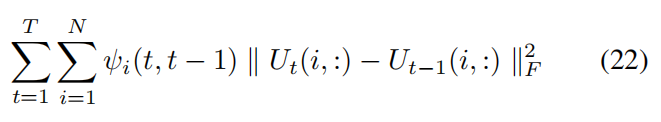
Yuan等（Yuan et al.2013）假设用户在不同时间访问不同位置并提出了时间感知POI推荐算法。具体地，他们提出POI推荐算法通过在计算两个用户之间的相似性时利用时间因素来扩展基于用户的POI推荐算法，以及在时间t考虑历史签到，而不是在所有时间进行POI推荐。 用户在时间t检查新POI 1的推荐分数被定义为



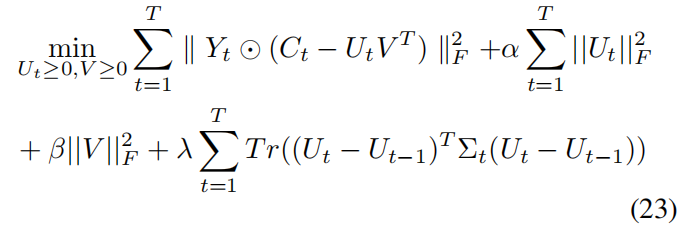
在Foursquare和Gowalla的实验中，他们发现时间信息增强的POI推荐方法总是优于基于用户的POI推荐方法，忽略时间信息。

Gao等人（Gao et al.2013）提出了基于矩阵分解的POI推荐算法，该算法具有基于两个时间属性的时间影响：（1）non-uniforms：用户在一天的不同时间表现出明显的登记偏好; （2）连续性：用户在连续时间内往往具有更多类似的登记偏好。为了模拟非一致属性，他们定义了Ut∈<N×K和V∈<M×K来描述时间状态t下的时间依赖的用户登记偏好并分别表示POI的特征。通过求解方程21的优化问题，他们获得了时间相关的用户潜在特征矩阵Ut和时间无关的POI潜在特征矩阵V

是在时间状态t的矩阵中的用户位置检查，并且Yt是对应的指标矩阵。 此外，他们通过将时间正则化项引入等式21来建模时间连续性。形式上



其中ψi（t，t -1）是测量ui在时间状态t和t-1之间的登记偏好的相似性的时间系数，由余弦相似性度量计算。最后，他们提出POI推荐方法通过求解以下目标函数来学习潜在用户特征矩阵Ut和潜在POI特征矩阵V。



他们的实验结果表明，他们提出的POI推荐算法比基于用户和NMF（Lee和Seung 2001）的POI推荐算法表现更好。

**Conclusion and Future Research Directions**

随着基于位置的社交网络的普及，个性化POI推荐技术引起了业界和学术界的广泛关注，因为它们不仅帮助用户探索新的地方，而且还增加了LBSN提供商的收入。 在本文中，我们简要介绍了LBSN中POI推荐的任务。我们首先描述POI推荐中存在的独特属性，它将POI推荐系统与传统推荐系统区分开来。此外，我们将POI推荐算法分为四类：基于纯登记数据的POI推荐方法，地理影响增强POI推荐方法，社会影响增强POI推荐方法和基于与其集成的附加信息类型的时间影响增强POI推荐方法 通过POI推荐算法签入数据。

从现有研究中，我们总结了以下观察结果：（1）尽管 附加信息对提高POI推荐系统的推荐质量有用，但登记数据，地理影响和时间影响对比社会影响力POI推荐表现出更显着的影响。特别是，地理影响在POI推荐中起着最重要的作用。（2）对于用户登记行为的建模，每个用户的个性化距离分布优于通用分布，例如PD和MGM。（3）基于模型的POI推荐方法比基于记忆的POI推荐方法更有效和有效，这与它们在传统推荐系统中的性能一致。 此外，基于用户的POI推荐方法更适合于LBSN中的POI推荐而不是基于项目的方法。（4）在基于矩阵分解的POI推荐方法中，NMF和BNMF模型的性能优于RMF和PMF模型。此外，基于加权MF的分解模型优于NMF和BNMF模型。

尽管一些研究工作已经研究了LBSN中POI推荐的问题，但最近出现了POI推荐，并且值得探讨几个有趣的研究方向。 首先，在LBSN中，POI的登记频率变化很大，用户的登记频率直观地反映了用户对POI的偏好程度。然而，经过审查的工作报告称，基于0/1评级矩阵的POI推荐优于登记频率矩阵。因此，期望POI推荐器系统采用合适的方法来建模登记频率数据。基于秩的协同过滤方法（Yi等人2013; Shi，Larson和Hanjalic 2010）可以适用于POI推荐，因为基于秩的协同过滤ap推断用户对成对比较的偏好而不是数字评级。其次，据报道，与最先进的POI推荐方法相比，社交影响力增强的POI推荐方法没有取得重要的改进。除了POI选择的决策过程受到POI的地理属性的影响，可能的原因是社会影响力增强POI推荐方法将所有社会关系视为同质的社会关系并忽略不同类型的社会关系。 LBSN中的社会关系是异质的，由不同类型的社会关系（如友谊和会员等）组成，用户倾向于不同的社交圈根据不同的目标提出建议。因此，POI推荐系统考虑不同类型的社会关系将是有用的。 第三，由于用户经常涉及多个社交网络，因此从其他社交网络获得的信息将有益于LBSN中的POI推荐。 在这种情况下，基于转移学习的POI推荐（Pan and Yang 2010）是一个潜在的研究方向。 第四，LBSN提供了丰富的附加信息，用于增强POI推荐系统的性能，即签到，地理信息，社交关系和时间信息。POI推荐系统需要统一的推荐框架，通过联合各种附加信息来提高POI推荐系统的性能。最后，随着LBSN中可用的用户和POI数量的快速增长，POI推荐系统严重受到可扩展性问题的困扰。 因此，并行计算方法，例如MapReduce（Dean和Ghemawat 2008）和Spark（Zaharia等人，2012），值得利用来加速POI推荐的计算过程。