Collaborative Filtering with Social Local Models

# Abstract

矩阵分解（MF）是推荐系统非常受欢迎的一种方法。它假定下面的评级矩阵是低秩的。但是，这种假设对于捕捉用户和项目之间复杂的关系和交互过于严格。最近，局部LOw-秩矩阵近似（LLORMA）已被证明在解决这个问题方面非常成功。它只是假定评分矩阵由许多类似用户和项目的子集构建的低阶子矩阵组成。虽然LLORMA优于MF，但如何构造这样的子矩阵仍然是一个大问题。受当今推荐系统中丰富的社交关系的推动，我们提出了一种新的框架，即Social LOcal低秩矩阵逼近（SLOMA）来解决这个问题。就我们所知，SLOMA是第一个将社交关系融入局部低秩框架的工作。此外，我们通过将社会正则化应用于子矩阵因子分解来增强SLOMA，表示为SLOMA ++。因此，所提出的模型可以受益于社会推荐和局部的低秩假设。来自两个真实世界的数据集Yelp和Douban的实验结果证明了所提出的模型优于LLORMA和MF。

# INTRODUCTION

推荐系统（RS）已经成为大数据时代不可或缺的工具。它通过帮助用户根据以前的行为获得有趣的项目来解决信息过载问题。协作过滤（CF）是一种最先进的RS技术，它试图根据类似的用户或项目预测用户对看不见项目的评分（或偏好）。在各种基于CF的方法中，矩阵分解（MF）因其良好的性能和可扩展性而最受欢迎[1]，[2]，[3]，[4]。MF基于以下假设：用户对项目的偏好由少量潜在因素控制。因此，大用户项目评分矩阵可以分解为两个较小的矩阵，分别代表用户特定的和项目特定的潜在因素。换句话说，评级矩阵是低秩的。

尽管RS在MF中取得了成功，但评级矩阵为低秩（称为全球低秩）的假设是有问题的，**因为现实世界中的评级矩阵非常大，并且由多种评级行为组成，使得低秩评级矩阵的假设不正确。**Lee等人[5]提出了一种称为局部LOw-Rank矩阵近似（LLORMA）的新框架来缓解全局低秩问题。基于观察到倾向于有一组用户对小组项目感兴趣，LLORMA认为评分矩阵由许多低秩子矩阵（称为局部低秩）组成，其被图示在图1的右侧。大量的实验已经证明了LLORMA在推荐系统中的有效性[5] [6] [7] [8]。除了RS，后续的工作已经在不同的领域进行了基于局部低秩假设，例如图像处理[9]，[10]，多标签分类[11]，文档分析[12]，证明了这个框架。

在局部低秩方法中，子矩阵的构造是一个根本性问题。Lee等人[5]建议首先从评级矩阵中选择一些随机锚点，即用户 - 项目对。然后，对于每个锚点，通过基于一些距离度量选择靠近锚点的剩余点来构造一个子矩阵。但是，这种方法会导致几个问题。首先，由于锚点是随机选择的，因此很难解释获得的子矩阵的含义，因此推荐结果也是无法解释的。其次，由于子矩阵是使用距离阈值在锚点周围构建的，因此它们对距离阈值非常敏感。因此，找到一个好的距离阈值是一项具有挑战性的任务。最后，**计算和空间成本都很高**，因为我们需要计算和存储所有子矩阵中用户和项目的成对相似度。

为了克服LLORMA的弱点，我们提出了一种新的框架，即Social Local低秩矩阵逼近（SLOMA），它将用户之间的社会联系结合到局部的低秩框架中。据我们所知，社交联系在基于MF的RS中被有效利用[13,14,16]，[17]，尽管SLOMA是利用社交联系的第一个工作，局部的低秩框架。就像在大多数社交网络研究中一样，SLOMA假设社交图形嵌入了许多社交群体，在这些群体中用户对彼此具有相似的偏好和影响。图1的左侧部分说明了SLOMA的基本思想。社会网络下有三个社会群体，根据这个群体我们可以在满足低级假设的评级矩阵中构建三个子矩阵。**社交联系帮助SLOMA解决重要的子矩阵构建问题。**具体而言，SLOMA可以选择有影响力的用户（称为连接符），而不是像LLORMA中那样随机选择锚点，而是选择社交图中某个距离内（即几跳内）的连接器的朋友，以构建有意义的子矩阵 因为他们可以被视为社交群体。见图1中的例子。因此，SLOMA解决了LLORMA所面临的所有三个问题，即它选择了有意义的锚点并构建了有意义的子矩阵，其距离度量是直观的，并且不需要保持每个子矩阵的用户和项目的成对相似度，这避免了高计算和空间成本。

# Motivation