推荐系统中的二部图嵌入算法

**问题背景与定义**

推荐系统是环节信息过载问题的关键工具。其目的是通过分析观察到的用户-项目关系，从数百万的候选对象中预测用户的首选项目。为了缓解推荐系统遇到的稀疏性和冷启动问题，利用侧边信息(side information)和知识(knowledge)来丰富观测信息，揭示隐藏(间接)用户-项目关系。其性能很大程度上取决于面对高复杂性和大规模侧边信息和知识的推荐模型的可扩展性。图嵌入技术在有效利用复杂、大规模数据方面的研究是一个重大课题。

**二部图**：形式化的定义为,包含两种节点和一种边, 其中，边只存在于具有不同类型的节点之间，这是用户-项交互的常用表示。具体来说，二部图表示具有两种类型节点的用户和项目，在推荐系统中隐式用户-项目交互对应的用户-项目节点对之间添加边；如果交互是显式的，则可以进一步通过评级进行加权。同时，通过用矩阵的行和列来表示用户和项，其中其元素记录用户-项交互(例如：0/1或评分)，一个二部图也可以直接转换为一个矩阵，这使推荐模型能够建立在代数技术上，如矩阵分解。如图1所示，说明了如何分别用二部图和矩阵来表示显式的用户-项交互。

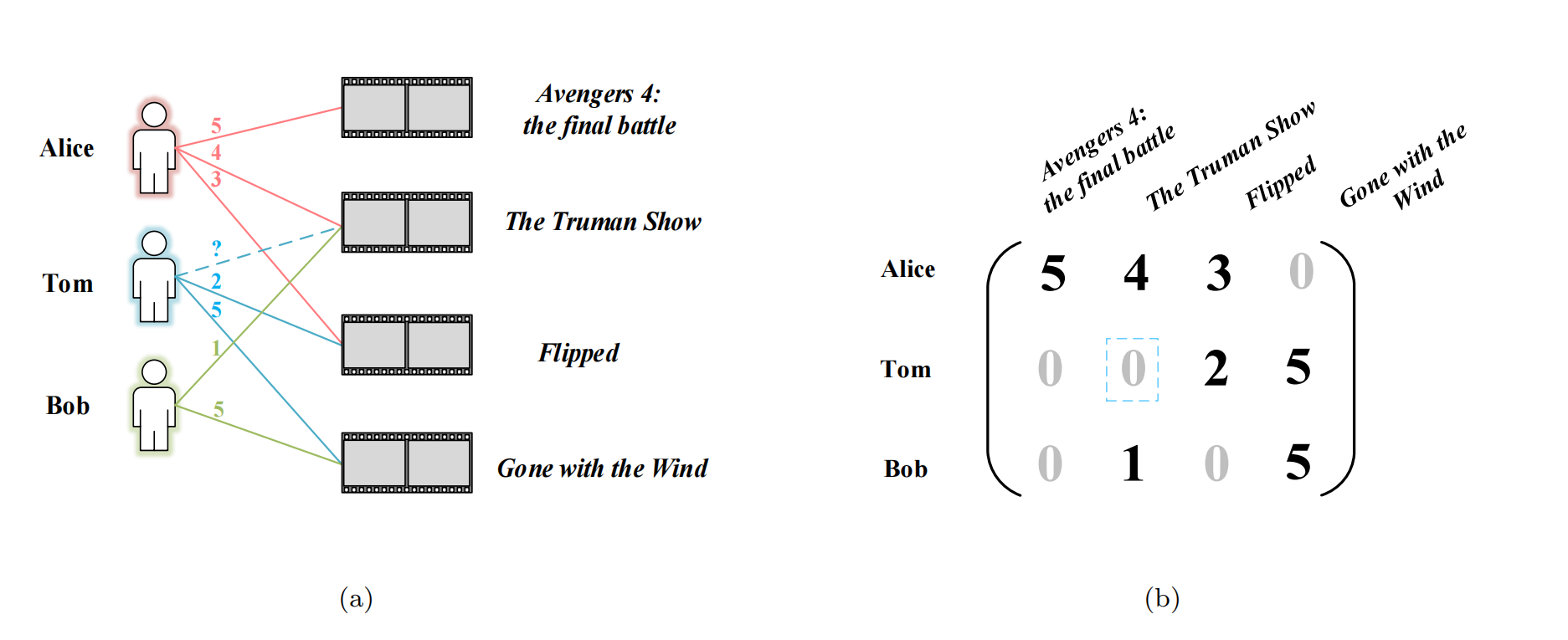


图1 分别用一个二部图和一个矩阵来表示用户-项之间的交互

推荐系统通过分析观察到的用户-项目交互来运行，其中的侧信息和知识被用来发现更多隐藏的信息。在这种情况下，基于二部图的推荐模型是研究中最优先考虑的模型，被认为是一般图和知识图的模型扩展的基础。

下面将进行分析基于静态用户-项目交互的推荐算法：

**推荐算法**

一般来说，基于二部图嵌入技术的静态用户-项交互的推荐模型可以分为基于矩阵分解、贝叶斯分析和深度学习方法三类。综上所述，二部图嵌入的先驱—矩阵分解方法具有很好的可扩展性。作为一个概率版本的矩阵分解方法，通过设置正则化项的模型通过先验知识等误差遵循高斯分布，贝叶斯分析方法在某种程度上减轻了非凸优化问题的矩阵分解方法引起的稀疏性问题。深度学习方法具有有效学习和保存数据中的非线性模式的能力，一直是近年来研究的主导地位。

**基于矩阵分解的方法**，奇异值分解(SVD)是基于矩阵分解的推荐模型背后的基本原理。简而言之，SVD将一个矩阵分解为，其中U、V是两个标准正交特征矩阵，是由a的奇异值组成的对角矩阵；对U、V和进行矩阵积可以重构A。根据这一思想，通过SVD，用这些分解元素表示用户项评级矩阵，在此基础上获得用户和项目的嵌入，通过一些操作重构用户项评级矩阵。

为此，潜在语义分析(LSA)被认为是SVD在文本信息检索中应用的先驱。详细地说，LSA基于至少两个文档中出现的一些有编号的文档和术语，构造了一个术语-文档矩阵A，其中其元素表示术语i在文档j中出现的频率。通过截断的SVD，A由进行分解，其中项i的嵌入用矩阵的第i行表示，文档j的嵌入是在公共k维向量空间中矩阵的第j行。对于对用户的查询q（一组单词）的信息检索，LSA通过计算q的嵌入。将用于通过对其嵌入进行操作来度量查询与所有文档的接近性。

可行LSA理论上，当推荐系统的用户和项目的数量通常是数亿，LSA变得不可行的分解这样一个非常巨大的用户项交互矩阵R由于SVD的高复杂性和U V的稀疏性，导致NP-hard问题。为了打破这一限制，SimonFunk在他的博客上提出，FunkSVD继承了LSA的思想，并基于优化算法实现了大规模矩阵分解的推荐。从反来看，FunkSVD不通过直接分解R，而是假设R可以用两个矩阵U、V表示，即用户和项的嵌入。在初始化U，V后，FunkSVD继续通过优化算法搜索其值，得到尽可能近似地满足的最优U，V。

FunkSVD最显著的优点之一是它具有很强的可扩展性，使其能够将用户偏差和项目偏差等辅助信息纳入模型中，以获得更好的推荐精度。从那时起，FunkSVD启发了随后的各种变体。例如，BiasSVD通过将评分偏差定义为，将用户偏差和项目偏差纳FunkSVD，这可以线性地附加到FunkSVD的公式中。基于BiasSVD，SVD++通过将每个用户的偏好建模为一个术语，进一步适应了用户的隐式交互，可以减少和之间的偏差。除此之外，全面的数据分析以挖掘更多与建议相关的辅助信息似乎是一个突破点。

FunkSVD及其变体的强可扩展性也探索了它们与基于k-最近邻居的推荐模型(KNN)的组合。例如，Slim在用户-项目矩阵上实现了因子分解来学习KNN的项目-项目接近矩阵

参考文献

G. W. Stewart, On the early history of the singular value decomposition, SIAM review 35 (1993) 551–566.

S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, R. Harshman, Indexing by latent semantic analysis,Journal of the American society for information science 41 (1990) 391–407.

P. C. Hansen, Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank, SIAM Journal on Scientific and Statistical Computing 11 (1990) 503–518.

G. Davis, S. Mallat, M. Avellaneda, Adaptive greedy approximations, Constructive approximation 13 (1997) 57–98.