RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems

Hongwei Wang , Fuzheng Zhang , Jialin Wang , Miao Zhao , Wenjie Li , Xing Xie , Minyi Guo

Shanghai Jiao Tong University,

Microsoft Research Asia

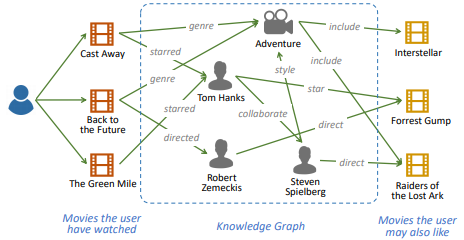
Meituan AI Lab,

The Hong Kong Polytechnic University

CIKM 2018 B

1. 主要内容

本文将知识图视为辅助信息的来源。为了解决现有嵌入式和基于路径的知识图谱知识推荐方法的局限性，提出了一个端到端框架RippleNet，它自然地将知识图谱合并到推荐系统中。类似于在水上传播的实际涟漪，RippleNet通过自动地并且沿着知识图谱中的链接迭代地扩展用户的潜在兴趣来刺激用户偏好在知识实体集上的传播。由用户的历史点击项激活的多个“涟漪”因此被叠加以形成用户关于候选项的偏好分布，其可用于预测最终点击概率。



知识图谱提供了丰富的事实和项目之间的联系，有助于提高推荐结果的精确性、多样性和可解释性。

1. 创新点

* 将基于嵌入和基于路径的方法结合到KG感知推荐中的第一项工作。
* 提出了RippleNet，一个利用KG协助推荐系统的端到端框架。 RippleNet通过在KG中迭代传播用户的偏好来自动发现用户的等级潜在兴趣。

1. 相关知识

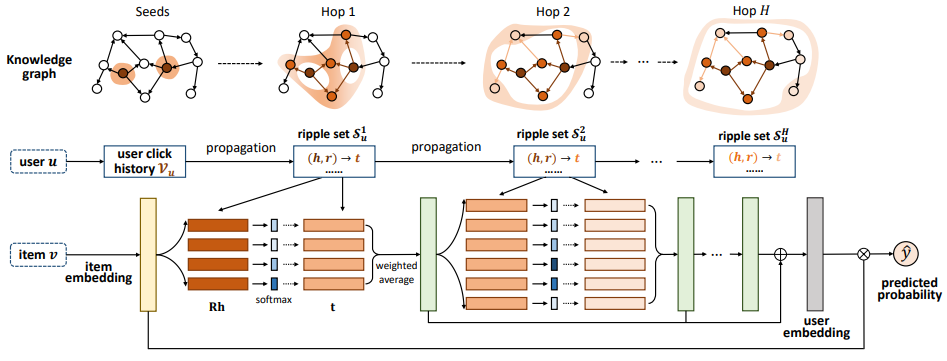
利用知识图谱（KG）的好处：（1）KG引入项目之间的语义相关性，有助于找到潜在的联系，提高推荐项目的准确性; （2）KG由各种类型的关系组成，有助于合理地扩展用户的喜好，增加推荐项目的多样性; （3）KG连接用户的历史记录和推荐的记录，从而为推荐系统带来可解释性。

一般而言，现有的KG感知推荐可分为两类：基于嵌入，基于路径。

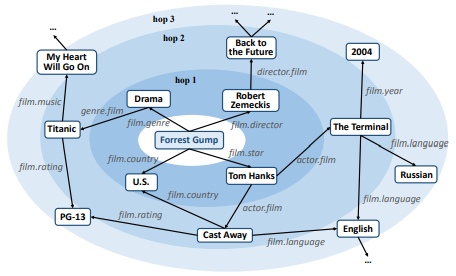
RippleNet设计用于点击率（CTR）预测，其将用户 - 项目对作为输入并输出用户参与（例如，点击，浏览）项目的概率。关键思想是偏好传播：对于每个用户，RippleNet将其历史兴趣视为KG中的种子集，然后沿着KG链接迭代地扩展用户的兴趣，以发现他对候选项目的等级潜在兴趣。将偏好传播与由在水上传播的雨滴产生的实际波纹进行类比，其中多个“波纹”叠加以形成用户在知识图上的最终偏好分布。 RippleNet优点：

（1）RippleNet通过偏好传播自然地将KGE方法纳入推荐; （2）RippleNet可以自动进行发现从用户历史记录中的项目到候选项目的可能路径，没有任何手工设计。

1. 整体架构



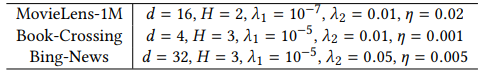
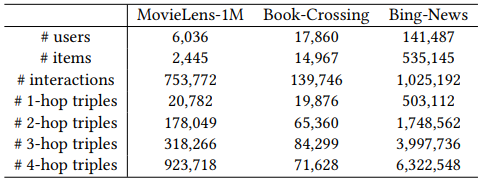
RippleNet将用户u和项目v作为输入，并输出用户u将点击项目v的预测概率。对于输入用户u，他的历史兴趣集Vu被视为 在KG中的种子，然后沿着链接延伸以形成多个波纹组S ku（k = 1,2，...，H）。 纹波集S k u是知识三元组的集合，其是远离种子集Vu的k跳。 这些波纹集用于迭代地与项目嵌入（黄色块）交互以获得用户u关于项目v（绿色块）的响应，然后将其组合以形成最终用户嵌入（灰色块）。 最后，我们使用用户u和项v的嵌入来计算预测概率yuv。

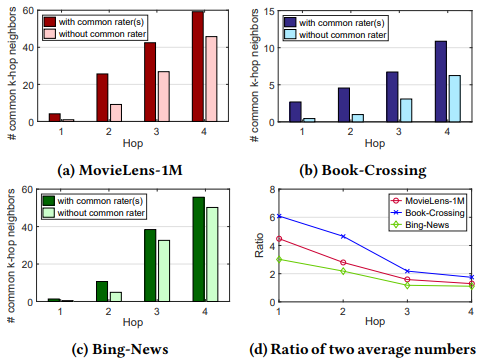


同心圆表示具有不同跳跃的波纹组。淡蓝色表示中心与周围实体之间的相关性降低。不同跳跃的波纹组在实践中不一定是不相交的。

1. 实验

数据集：

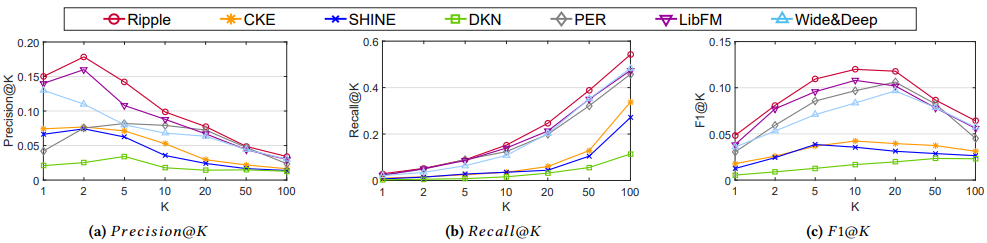




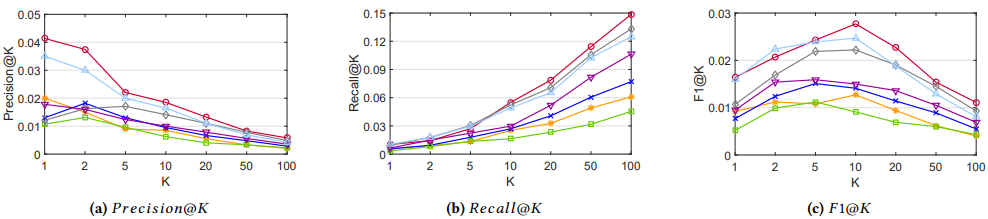
两个项目在KG w.r.t中共享的k-hop邻居的平均数量。 他们是否在（a）MovieLens-1M，（b）Book-Crossing和（c）BingNews数据集中有共同的评价者。 （d）具有不同跳数的两个平均数的比率。

实验结果：

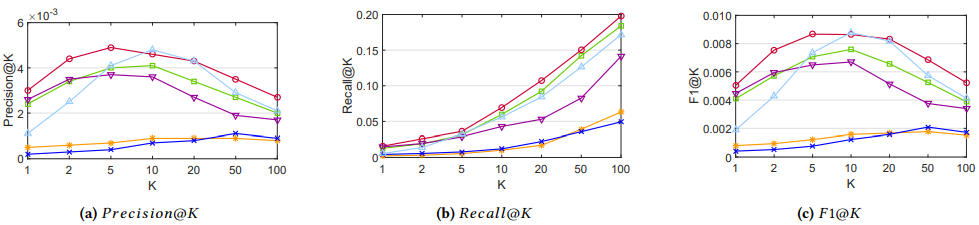
MovieLens-1M：



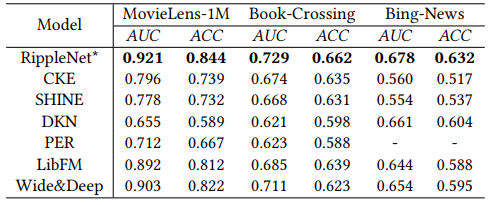
Book-Crossing：



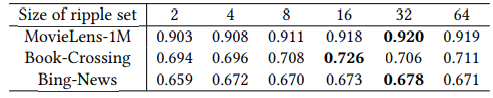
Bing-News：



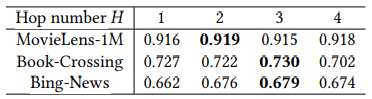
CTR预测中的AUC和准确度的结果：



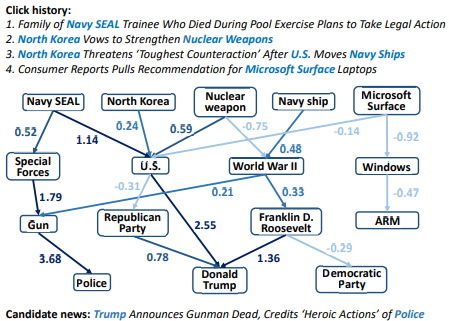
AUC的结果w.r.t. 用户纹波集的不同大小：



AUC的结果w.r.t. 不同的跳数：



随机抽样用户的相关概率的可视化w.r.t. 带有标签1的候选新闻。省略了低于-1.0的值的链接：



1. 总结

本文提出了RippleNet，一种端到端的框架，自然地将知识图合并到推荐系统中。 RippleNet通过引入偏好传播克服了现有基于嵌入和基于路径的KG感知推荐方法的限制，该传播自动传播用户的潜在偏好并探索他们在KG中的等级兴趣。 RippleNet将偏好传播与贝叶斯框架中的KGE正则化结合起来进行点击率预测。

对于未来的工作，计划（1）进一步研究表征实体 - 关系相互作用的方法; （2）在偏好传播过程中设计非均匀采样器，以更好地探索用户的潜在兴趣，提高性能