

基于知识图谱的推荐系统研究综述

秦川^{1,2}, 祝恒书^{2*}, 庄福振^{3,4}, 郭庆宇^{5,3,4}, 张琦^{1,2}, 张乐¹, 王超^{1,2},
陈恩红¹, 熊辉^{1*}

1. 中国科学技术大学计算机科学学院, 合肥 230022
2. 百度公司, 北京 100085
3. 中国科学院计算技术研究所中国科学院智能信息处理重点实验室, 北京 100190
4. 中国科学院大学, 北京 100049
5. 香港科技大学计算机学院, 香港 999077

* 通信作者. E-mail: zhuhengshu@gmail.com, xionghui@gmail.com

收稿日期: 2019-12-05; 接受日期: 2020-02-19; 网络出版日期: 2020-07-14

国家自然科学基金(批准号: 91746301, 71531001, 61836013, U1836206, 61773361)资助项目

摘要 推荐系统旨在为用户推荐个性化的在线商品或信息, 其广泛应用于众多 Web 场景之中, 来处理海量信息数据所导致的信息过载问题, 以此提升用户体验。鉴于推荐系统强大的实用性, 自 20 世纪 90 年代中期以来, 研究者针对其方法与应用两方面, 进行了大量广泛的研究。近年来, 很多工作发现知识图谱中所蕴含的丰富信息可以有效地解决推荐系统中存在的一系列关键问题, 例如数据稀疏、冷启动、推荐多样性等。因此, 本文针对基于知识图谱的推荐系统这一领域进行了全面的综述。具体地, 首先简单介绍推荐系统与知识图谱中的一些基本概念。随后, 详细介绍现有方法如何挖掘知识图谱不同种类的信息并应用于推荐系统。此外, 总结了相关的一系列推荐应用场景。最后, 提出了对基于知识图谱的推荐系统前景的看法, 并展望了该领域未来的研究方向。

关键词 知识图谱, 推荐系统, 协同过滤, 异质信息网络, 图嵌入

1 引言

伴随着互联网的快速发展, 目前我们正处于一个信息爆炸的时代。我们在享受着互联网带来获取知识、信息的便利性的时候, 同样苦恼于如何从浩瀚的知识海洋中快速有效地汲取到我们所需要的信息。近年来, 推荐系统作为一种有效地解决信息泛滥的方法受到了大量国内、国外研究者的广泛研究。其核心目标是通过分析用户行为、兴趣、需求等信息, 在海量的数据中挖掘用户感兴趣的信息, 如商品^[1]、新闻^[2]、POI (point of interest)^[3] 和试题^[4] 等。目前, 推荐系统在各种各样的信息系统中都

引用格式: 秦川, 祝恒书, 庄福振, 等. 基于知识图谱的推荐系统研究综述. 中国科学: 信息科学, 2020, 50: 937–956, doi: 10.1360/
SSI-2019-0274
Qin C, Zhu H S, Zhuang F Z, et al. A survey on knowledge graph-based recommender systems (in Chinese). Sci Sin Inform, 2020, 50: 937–956, doi: 10.1360/SSI-2019-0274

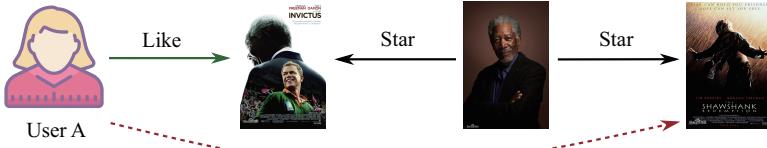


图 1 (网络版彩图) 一个基于可解释性图谱的推荐示例 (图片素材来源为百度百科)
Figure 1 (Color online) An illustration of explainable knowledge graph-based recommendation

扮演着至关重要的角色, 用来促进商务业务发展以及辅助决策过程, 并且广泛地应用于各大电子商务 (淘宝、Amazon、eBay 等) 和多媒体 (MovieLens、豆瓣) 等网站中.

个性化推荐算法是推荐系统的核心, 其主要可以被分为 3 类, 即基于内容的推荐算法、基于协同过滤的推荐算法和混合推荐算法^[5]. 其中, 协同过滤方法通过利用用户历史的行为偏好数据构建模型, 取得了巨大的成功^[6~8]. 这种方法的优势在于不需要类似基于内容的推荐算法那样对物品进行复杂的特征提取与建模. 尽管这类协同过滤方法通常是有效且普适的, 但其依旧存在着多种问题. 主要包括用户和物品之间的行为关系数据的稀疏问题和对新用户或者新物品进行推荐时存在的冷启动问题. 为此, 研究者尝试将协同过滤推荐算法和其他辅助信息相结合 (例如, 用户与物品的属性特征、用户社交网络信息等) 搭建混合推荐系统来解决以上问题, 从而提升推荐效果^[9~11].

知识图谱作为混合推荐系统中一类有效的辅助信息, 近年来受到了大量研究者的关注^[2, 10, 12~15]. 知识图谱是一种有向信息异构网络. 其包含节点 – “实体” 与有向边 – “关系”, 因而蕴含了推荐系统中对于物品的大量背景信息, 以及物品之间的关系^[10, 13]. 并且它可以与用户行为数据构成的用户 – 物品网络集成起来^[16, 17], 从而扩展了用户与商品之间存在的隐藏的关联关系, 补充了用户与物品交互数据, 因此可以更进一步地提升推荐效果. 另外, 基于知识图谱的推荐方法为构建可解释推荐系统提供了一些新的思路^[18~20]. 例如图 1 所示, 通过利用知识图谱中影视相关的辅助信息, 我们可以推测用户 A 可能因为喜欢由摩根·弗里曼主演的电影《成事在人》而同样喜欢其主演的电影《肖生克的救赎》. 目前, 学术界与工业界的研究者已经构建了一系列知识图谱, 例如 DBpedia KG¹⁾、AceKG²⁾、Microsoft Satori³⁾、百度知识图谱⁴⁾和 OwnThink KG⁵⁾等. 并且覆盖了不同的领域, 如学术、历史人物、影视、音乐等. 因此, 为我们在不同应用场景中基于知识图谱构建推荐系统提供了充足的数据基础.

为此, 本文将对基于知识图谱的推荐系统从方法、应用两方面进行系统性的综述. 具体地, 我们通过调研分析近年来的相关文章, 从两类技术方法: 基于嵌入的方法^[10~12] 和基于路径的方法^[16, 21] 展开详细的介绍. 随后, 针对推荐物品所属领域 (如, 新闻^[2, 22]、音乐^[23, 24] 等) 和推荐任务 (如, 序列推荐^[12]、可解释性推荐^[19, 20]) 两个方面对现有工作的应用场景进行讨论分析. 最后, 根据我们对基于知识图谱的推荐系统这一领域的理解给出一些未来发展方向的展望. 例如构建动态的推荐系统和联合迁移学习下的推荐系统等. 据我们所知, 这篇文章是第 1 篇对基于知识图谱构建推荐系统的综述, 它不仅总结了近年来的国内外的研究工作, 同时也提出了未来相关研究方向的展望.

本文第 2 节将介绍相关研究背景, 包括传统的推荐系统与知识图谱的基本知识; 第 3 节将系统地从方法、应用场景两方面介绍近年来基于知识图谱的推荐系统的相关研究工作; 第 4 节将从不同方面对基于知识图谱的推荐系统的未来研究工作进行展望; 最后第 5 节对全文进行总结.

1) <https://wiki.dbpedia.org/>.
 2) <https://www.acemap.info/app/AceKG/>.
 3) <https://searchengineland.com/library/bing/bing-satori>.
 4) <https://kgopen.baidu.com/index>.
 5) <https://github.com/ownthink/KnowledgeGraphData>.

2 研究背景

在深入介绍基于知识图谱的推荐系统之前,首先对推荐系统的基础概念和常用技术进行介绍。随后介绍知识图谱的一些基础概念,并且汇总近年来学术界和工业界提出的知识图谱。最后讨论应用知识图谱到推荐系统中的动机。

2.1 推荐系统

推荐系统这一思想在 1994 年被 Resnick 等^[25]引入之后,逐步成为了一个相对独立的研究方向,并备受学术界和工业界的研究者关注。推荐系统的主要目的是向用户 (user) 推荐其可能感兴趣的物品 (item)^[5],如商品^[1]、新闻^[2]、POI (point of interest)^[3]、试题^[4]、音乐^[24]、手机应用^[26]等。推荐系统的形式化定义^[5]如下:给定用户集合 U , 物品集合 V , 并用 $R_{i,j}$ 表示用户 U_i 对 V_j 的喜好程度, $R \in \mathbb{R}^{|U| \times |V|}$ 。例如在电影推荐网站 MovieLens 中, $R_{i,j}$ 就可以是用户 i 对电影 j 的打分。随后我们定义一个效用函数 $f : U \times V \rightarrow R$, 则推荐系统所研究的问题是给定任意一个用户 U_i , 希望找到其喜好程度最大的物品 V_k , 即

$$\forall U_i \in U, \quad V_k = \arg \max_{V_j \in V} f(U_i, V_j). \quad (1)$$

在真实的推荐场景中 $|U|$ 和 $|V|$ 的量级往往非常大,如淘宝网的日活跃用户上千万,每天在线商品上亿,但是我们可以观测的 R 却极其稀疏,因此给研究推荐系统带来了极大的挑战。

目前,我们通常将推荐系统中的推荐算法分为 3 类,分为基于内容的推荐算法、基于协同过滤的推荐算法和混合推荐算法^[5]。

- **基于内容的推荐算法**^[27]. 基于内容的推荐算法源于一个基本假设:“用户可能会喜欢与他曾经喜欢过的物品相似的物品”,其通过建模计算用户曾经有过的显示反馈(打分、点赞等)和隐式反馈(搜索、点击、购买等)的物品集合与所有物品的相似度,按照相似度的大小排序到推荐的列表。并且可以考虑时间因素、地点因素和推荐场景因素等来优化推荐的结果。此外,还可以使用基于统计和机器学习方法从用户历史反馈数据中建模学习用户的喜好,替代传统计算相似度的方法。基于内容的推荐算法的主要优点在于可以解决新物品冷启动的问题,不受 R 的稀疏性所影响;推荐结果有不错的可解释性。其缺点在于需要复杂的特征工程构造物品属性的特征,否则会严重影响推荐结果;缺乏多样性,推荐结果会与用户曾有过显示和隐式反馈的物品高度相似;新用户的冷启动问题。

- **基于协同过滤的推荐算法**. 基于协同过滤的推荐算法是目前应用最成功的推荐方法,其利用用户和物品历史的反馈数据,挖掘用户和物品本身的相关联性,并基于此进行推荐。具体地,这类方法可以被分为 3 类:基于用户的推荐、基于物品的推荐和基于模型的推荐^[28]。基于用户的协同过滤方法是基于假设“用户可能喜欢与他相似用户喜欢的物品”,通过用户历史反馈记录计算用户间的相似度,利用其相似的用户对物品的反馈,来预测对应用户的反馈情况,并进行推荐。这类方法的主要优点在于避开了对物品自身属性的特征挖掘,缺点在于在用户数量变化很大的情况下,算法效率较低,并且面临新用户的冷启动问题。基于物品的协同过滤方法与基于内容的推荐算法假设类似,不同在于其使用物品历史被反馈的数据来判断物品之间相似性。其优点在于计算简单,因为物品反馈结果变化比用户要低很多,相较于基于用户的协同过滤算法,更可以通过离线计算,定时更新来完成,其缺点则是无法在不离线更新物品相似性时推荐新的物品给用户。基于模型的协同过滤方法是为了解决基于用户、物品的协同过滤方法所面临的数据稀疏、难以在大数据量级上返回即时结果的问题。其通过历史数据利用机器学习方法训练得到一个预训练模型 f ,从而可以实时预测任意用户对某一物品的喜好。

- **混合推荐算法** [29]. 混合推荐算法是指将多种推荐技术混合, 从而互相弥补缺点. 其中混合方法包括简单的推荐结果加权融合、切换、混合, 不同数据源的特征组合, 复杂的多模型级联、特征递增和元层次混合. 其中, 最常见的混合推荐系统就是将协同过滤推荐方法与其他推荐方法结合, 从而解决冷启动和稀疏性等问题 [30]. 此外, 混合推荐系统的优点在于可以针对特定的推荐场景定制化建模, 从而合理有效地利用额外的数据信息. 例如, Konstas 等 [31] 利用音乐网站 Last.FM⁶⁾特有的用户之间的社交网络信息构建了一个有效的混合推荐系统. Wang 等 [2] 通过引入知识图谱的知识信息, 学习新闻内容之中的潜在知识关联, 从而构建了一个在新闻推荐场景下的高效混合推荐系统.

2.2 知识图谱

知识图谱是一种知识库, 其概念最早由 Google 公司提出, 目的在于提升搜索引擎的搜索质量, 增强用户的搜索体验 [32]. 知识图谱的本质是一个存储知识实体与实体之间关系的结构化网络, 可以帮助形式化描述、理解现实世界的事物及其相互关系. 随着互联网和物联网技术的飞速发展, 知识图谱的应用也从搜索引擎逐步扩展到各个领域, 例如推荐系统 [2, 10, 12~14]、智能问答 [33, 34] 和文本分析 [35] 等.

知识图谱通常以三元组的形式存储实体及其关系, 形式化表示为 $G = (E, R, S)$, 其中 $E = \{E_1, E_2, \dots\}$ 表示知识图谱中所有种类实体的集合; $R = \{R_1, R_2, \dots\}$ 表示知识图谱中所有关系的集合; $S = \{S_1, S_2, \dots\}, S \subseteq E \times R \times E$ 表示所有三元组的集合, 每一个三元组 S_i 都由一个头实体、关系和尾实体构成. 例如三元组 (Donald Trump, president_of, America) 表示 “Donald Trump” 是 “America”的总统, 其中 Donald Trump 是头实体, America 是尾实体, president_of 是两者的关系名称. 知识图谱中的三元组不仅可以帮助我们理解知识实体之间的关系, 也可以存储知识实体的属性, 例如 (Donald Trump, born_in, New York) 和 (Donald Trump, nationality, America) 存储实体 Donald Trump 的两个属性, 即出生地和国籍, 这时尾实体也被称作属性值. 此外, 在特定场景下知识图谱的内容会频繁地动态更改, 最近也有学者尝试构建动态知识图谱, 将原有的三元组变更为包含时间戳的四元组, 从而进一步建模知识的演化性 [36]. 表 1 [37~43] 汇总了部分常见的知识图谱, 其中包括诸如 YAGO KG^[37], DBpedia KG^[38], Freebase KG^[39] 和 OwnThink KG⁷⁾ 等通用知识图谱, 也有如 WordNet^[40], UMLS KG⁸⁾, Douban's movie KG⁹⁾ 和 MusicBrainz¹⁰⁾ 等特定领域的知识图谱. 这类特定领域下的知识图谱往往拥有比通用化知识图谱更完整的领域内实体和关系. 此外, 知识图谱在不同领域, 如搜索引擎、推荐系统等上的应用前景, 大大吸引了工业界在知识图谱构建上的投入, 并应用到不同商业产品中.

2.3 为什么应用知识图谱在推荐系统中

基于知识图谱的推荐系统的相关研究近几年来发展迅速, 在最近短短几年间大量研究者在如 SIGKDD [10, 17, 21, 44, 45], SIGIR [12, 46], WWW [2, 47~51], ICDM [52] 和 CIKM [53] 等顶级会议上发表了相关的研究工作. 在我们开始深入介绍这些相关技术之前, 很有必要理解其应用知识图谱在推荐系统的动机.

首先, 现有推荐系统技术, 尤其是基于协同过滤的推荐方法一直面临着一些严峻的问题, 包括用户—物品的交互数据稀疏问题和冷启动问题, 会导致推荐结果不准确 [28]. 在知识图谱中蕴含着大量实体

6) <http://www.last.fm>.

7) <https://www.ownthink.com/>.

8) <https://www.nlm.nih.gov/research/umls/index.html>.

9) <http://openkg.cn/dataset/douban-movie-kg>.

10) <https://musicbrainz.org/>.

表 1 常见知识图谱汇总
Table 1 A summary of common knowledge graphs

Name	Organization	Data source	Domain	Is open-source
YAGO KG [37]	Max Planck Institute	WordNet, Wikipedia	General	Yes
DBpedia KG [38]	DBpedia Association	Wikipedia, Expert knowledge	General	Yes
Freebase KG [39]	Google	Wikipedia	General	Yes
NELL KG [41]	Carnegie Mellon University	Web data	General	Yes
Wikidata [42]	Wikimedia Deutschland	Wikidata, Freebase	General	Yes
Google's Knowledge Graph	Google	Freebase, Web data	General	Yes
Microsoft Satori	Microsoft	Web data	General	No
Baidu's Knowledge Graph	Baidu	Web data	General	No
OwnThink KG	OwnThink	Web data	General	Yes
CN-DBpedia [43]	Fudan University	Chinese encyclopedia website	General	Yes
WordNet [40]	Princeton University	Expert knowledge	Linguistics	Yes
UMLS KG	National Library of Medicine	Medical literature	Medical	Yes
Douban's movie KG	Zhejiang University	Douban data	Movie	Yes
MusicBrainz	MetaBrainz Foundation	Web data	Music	Yes

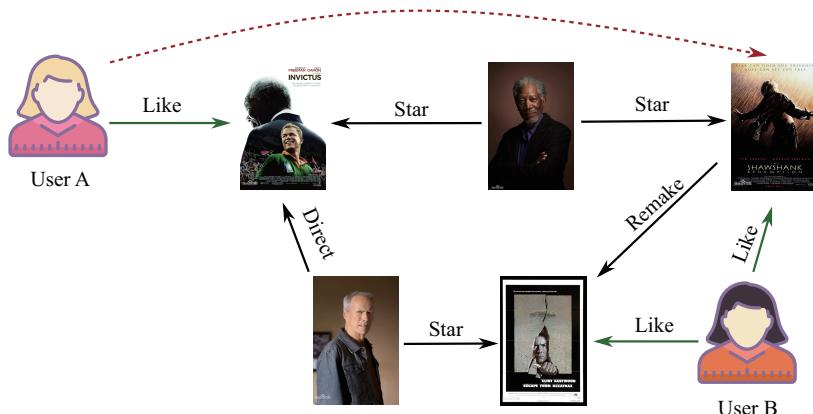


图 2 (网络版彩图) 一个基于知识图谱的异构信息网络图在推荐中的应用示例 (图片素材来源为百度百科)
Figure 2 (Color online) An illustration of leveraging knowledge graph based heterogeneous information network (HIN) in recommender system

以及实体之间关系的信息,这可以作为一种有效的辅助信息丰富用户和物品的描述、提升推荐算法的准确程度以及缓解、解决上面两个问题。例如图 1 所示,即使原有数据中我们并没有关于《肖申克的救赎》任何的交互信息,我们依旧可以通过图谱提供的主演知识进行关联,并且深层次地发现用户可能的兴趣点。具体地,一些研究通过对知识图谱中的实体和关系进行表征,进而将知识图谱中蕴含的语义信息带入到推荐系统的用户和物品上,从而提升推荐系统的效果 [2, 10, 12, 22, 44]。另外,一些研究者通过将知识图谱和原有的用户 - 物品的交互数据构成一个异质信息网络图,利用挖掘图上用户到物品之间不同的路径,提升推荐效果 [14, 19, 20]。图 2 展示了这样一个异质信息网络图应用到推荐场景中的示例,我们可以发现从用户 A 到电影《肖申克的救赎》的 3 条路径,分别是

- (用户 A, 喜欢, 《成事在人》) \wedge (《成事在人》, 主演是, 摩根 · 弗里曼) \wedge (摩根 · 弗里曼, 主演, 《肖

表 2 一个相关文献的查询表
Table 2 A lookup table for relevant publications

Category	Year	Ref.
Embedding-based methods	Before 2018	[10, 62, 63]
	2018	[2, 11~13, 22, 52, 64~67]
	2019	[44, 47~51, 61, 68~73]
Path-based methods	Before 2018	[16, 21, 53]
	2018	[18]
	2019	[14, 19, 20, 45, 74]

申克的救赎》) \Rightarrow (用户 A, 喜欢, 《肖申克的救赎》)

- (用户 A, 喜欢, 《成事在人》) \wedge (《成事在人》, 导演是, 克林特·伊斯特伍德) \wedge (克林特·伊斯特伍德, 主演, 《逃出亚卡拉》) \wedge (《逃出亚卡拉》, 被翻拍, 《肖申克的救赎》) \Rightarrow (用户 A, 喜欢, 《肖申克的救赎》)

- (用户 A, 喜欢, 《成事在人》) \wedge (《成事在人》, 导演是, 克林特·伊斯特伍德) \wedge (克林特·伊斯特伍德, 主演, 《逃出亚卡拉》) \wedge (《逃出亚卡拉》, 被喜欢, 用户 B) \wedge (用户 B, 喜欢, 《肖申克的救赎》) \Rightarrow (用户 A, 喜欢, 《肖申克的救赎》)

因此, 我们可以基于上面的推理关系去推荐《肖申克的救赎》给用户 A. 这样的异构信息网络图可以帮助我们挖掘大量源数据集中未出现的用户 - 物品交互的关系, 解决数据稀疏与冷启动的问题.

此外, 近年来构建可解释性推荐系统成为趋势 [54~58]. 可解释性推荐算法的目的是在给用户合适地推荐物品的同时, 给出推荐的原因使得用户更易理解系统为什么做出这样的推荐. 研究发现给出推荐的原因会使得用户的信任度和满意度提升, 并且更易于使用户选择被推荐的物品 [58, 59]. 知识图谱作为非常容易理解的外部知识库可以帮助我们将用户、物品之间构建合理的关联, 从而可以帮助我们构建可解释的推荐系统, 其吸引了研究者的广泛关注 [14, 19, 20, 60].

3 基于知识图谱的推荐系统: 方法与应用

本节首先介绍基于知识图谱的推荐算法的分类, 并且重点介绍其中一些前沿的研究方法. 随后, 我们对基于知识图谱的推荐系统的不同应用场景进行讨论, 并归纳了不同应用场景使用到的数据集.

3.1 基于知识图谱的推荐算法分类

整体来讲, 现有的基于知识图谱的推荐方法可以被分为两类^[61], 包括: 基于嵌入的方法 (embedding-based methods) 和基于路径的方法 (path-based methods). 为了便于读者对文献的查询, 在表 2 [2, 10~14, 16, 18~22, 44, 45, 47~53, 61~74] 中, 我们汇总了近年来的相关文章.

3.1.1 基于嵌入的方法

基于嵌入的应用知识图谱在推荐系统的方法主要是通过图嵌入的方法对实体和关系进行表征, 进而扩充原有物品和用户表征的语义信息. 其中包括基于 Trans 系列的图谱嵌入方法和基于异质信息网络的图嵌入方法.

- **基于 Trans 系列的图谱嵌入方法.** Trans 系列的嵌入方法是比较典型的一类对知识图谱中的实体、关系进行表征的方法. 这类方法目的是将实体和关系映射到连续的向量空间中, 获得低维稠密表

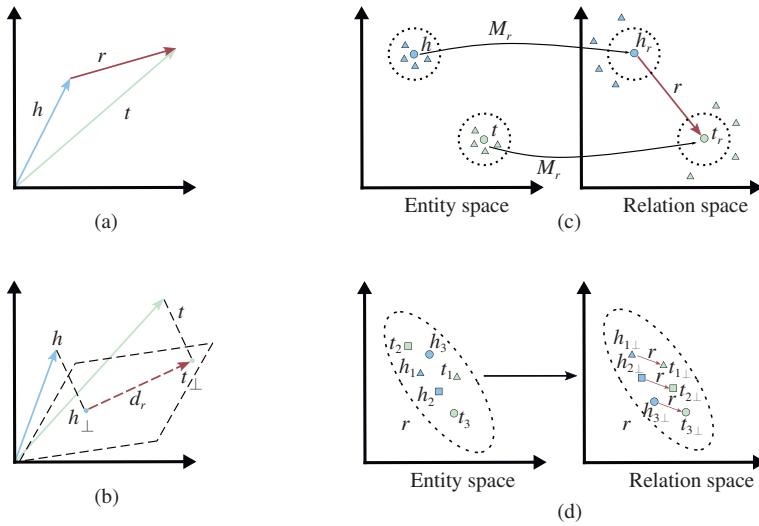


图3 (网络版彩图) 4种常用传统图谱嵌入方法原理图

Figure 3 (Color online) The motivations of four traditional knowledge graph embedding approaches. (a) TransE; (b) TransR; (c) TransH; (d) TransD

示. 主要包括 TransE^[75], TransH^[76], TransR^[77] 和 TransD^[78]. 如图 3 所示, 可以看出这些方法在假设上的区别. TransE 是这类方法的第 1 篇, 其基本想法是头实体的向量 h 与关系向量 r 之和与尾实体向量 t 越接近越好 ($h + r \approx t$), 这里使用 L1 或 L2 范数来衡量“接近”. 具体地, 其构造一个基于负采样方法的损失函数, 即

$$L(h, r, t) = \max(0, d_{\text{pos}} - d_{\text{neg}} + \text{margin}), \quad (2)$$

其中 $d = \|h + r - t\|$ 为 L1 或 L2 范数. 这里 d_{pos} 和 d_{neg} 分别表示正样本和负样本的三元组, 并且使用 margin 表示正负样本最大距离的一个常数. 然而 TransE 很难解决一对多的情况. 例如, TransE 通过优化两个正样本 (《肖申克的救赎》, 类别, 剧情片) 和 (《十二怒汉》, 类别, 剧情片), 会使得《肖申克的救赎》与《十二怒汉》的表征向量过于相近, 但这却不一定符合实际的认知.

为此, TransH 被提出处理这样的一对多或多对一的关系, 其将头实体 h 和尾实体 t 的向量投影到由 W_r 确定的超平面上得到向量 h_{\perp} 和 t_{\perp} , 则在这个超平面上存在一个表示向量 r 满足 $h_{\perp} + r \approx t_{\perp}$, 其中

$$h_{\perp} = h - w_r^T h w_r, \quad t_{\perp} = t - w_r^T t w_r. \quad (3)$$

TransR 则认为实体具有不同方面, 而不同的关系侧重于不同的方面, 因此其先通过一个关系矩阵 M_r 将实体映射到关系空间, 再优化目标 $h_r + r \approx t_r$,

$$h_r = h M_r, \quad t_r = t M_r. \quad (4)$$

TransD 则进一步认为 TransR 的映射关系应该由实体和关系共同决定, 即映射矩阵通过 $M_{rh} = r_p h_p^T + I$ 和 $M_{rt} = r_p t_p^T + I$ 计算.

在通过 Trans 系列的图谱嵌入方法学习得到实体、关系的表征之后, 就可以用来扩充原有推荐系统中物品、用户的表征. 例如, Zhang 等^[10] 使用 TransR 的方法在知识图谱中学习物品相关实体的结构化知识, 得到物品的结构化语义嵌入表征, 从而对推荐系统中原有基于协同过滤中每个物品 V_j 的

隐向量 e_j 进行扩充。此外作者同时引入物品文本、图像的嵌入表征，进一步扩充物品语义向量最终得到 e'_j 。最后构造 pair-wise 的形式：

$$p(j > j'; i | \theta) = \sigma(u_i^T e'_j - u_i^T e'_{j'}), \quad (5)$$

其表示对于用户 U_i 来说相比物品 $V_{j'}$ 更喜欢物品 V_j 。其中 u_i 表示用户 U_i 基于协同过滤的隐向量， θ 为模型中的参数。作者通过一系列的实验在 MovieLens 和 IntentBooks 两个数据集上有效地证明了其通过 TransE 引入知识图谱中的边信息，可以有效地提升推荐效果。在文献 [12] 中，作者使用 TransE 生成图谱中的实体和物品的表征，并更进一步使用键值对记忆网络 (key-value memory network)，基于用户历史交互数据中关联的实体得到用户的细粒度动态特征，从而有效地提升了推荐效果。Tang 等 [44] 则使用 TransR 得到的用户历史交互数据中关联的实体表征，并利用自注意力机制 [79] 学习它们的加权和，进而更有效地获得用户的特征。

此外，一些研究者将用户 – 物品的交互数据与图谱一起构成一个新的信息网络图，在一些文献中被称作协同知识图 (collaborative knowledge graph) [17]，图 2 展示一个简单的示例。其在原有的实体节点中再加入用户和物品节点，并把它们之间的交互数据，如浏览、购买、收藏和打分等行为构成图中新的边。Zhang 等 [66] 在这样的一个信息网络图上基于 TransE 的思想对图上的节点和关系进行表征，最后通过计算用户节点 e_i 和物品节点 u_j 在关系“购买” r_{buy} 上的距离，即 $d(e_i + r_{buy}, u_j)$ ，对所有物品进行排序，从而得到推荐列表。He 等 [69] 进一步对这个异质信息网络图进行扩充，增加了用户的一些属性作为节点，例如性别、年龄和职业等，从而构建用户与属性之间的边关系，然后利用 TransD 的方法进行图嵌入的学习，从而解决了用户和物品两个方面的冷启动问题。

在实际推荐系统中，很多物品并不在知识图谱中存在。针对这一问题，文献 [65] 中的作者通过构建一个概念层，对推荐系统的物品与知识图谱的实体之间进行关联，在基于 TransR 对实体进行表征后，通过概念层对可能与推荐系统中的物品相关的实体进行整合，从而得到物品的表征。Wang 等 [2]。针对新闻推荐中物品的特殊性，首先利用 TransD 对实体进行表征之后，再通过对新闻文本先进行实体抽取，结合实体嵌入表征和原有的词级别嵌入表征对每条新闻进行表征，从而提升了新闻推荐系统中物品语义的关联性。

- **基于异质信息网络的图嵌入方法。** 知识图谱因其节点和边具有不同的类别，又可以被称作一种异质信息网络图，因此可以使用一些异质信息网络图的嵌入的方法对图上的实体和关系进行表征。Metapath2Vec [80] 是其中的一种，其通过元路径下的随机游走方式构造异质信息图上节点的邻居节点集合，并基于 skip-gram 模型对节点进行表征。Yang 等 [52, 71] 利用 Metapath2Vec 对知识图谱上的节点进行表征，进而扩充推荐系统中物品的表征。类似的，Palumbo 等 [63] 基于 node2vec [81] 的思想，设计出一种针对异质信息网络图的 entity2vec 方法对知识图谱中的实体进行表征，并应用到推荐系统中。

基于 Trans 系列的图谱嵌入方法以及上述几种异质信息网络图的嵌入方法在应用到推荐系统时存在着一定的问题，即他们更适合知识图谱内的相关应用，因为他们主要的优化目标是知识图谱的补全任务或是边预测任务，而并不是推荐任务。Wang 等 [13, 61] 从异质信息网络图中抽取用户节点相连的 N-hop 实体节点，设计了一种 RippleNet 的方法利用这些实体节点的嵌入表征更新用户的表征，从而利用用户和物品表征的点积去预测推荐结果。此外，随着图卷积网络的发展 [82~86]，一些学者尝试利用它构建对知识图谱的建模，并使其优化目标与推荐系统一致，从而提升了推荐效果。Wang 等 [48] 受图卷积网络的启发，设计了一种 knowledge graph convolutional networks (KGCN)，其每层网络通过对异质图中每个实体采样得到其部分邻居节点，利用这些邻居节点上一层的状态更新该实体所在层的表征，通过多层 KGCG 的迭代进而对应物品的表征 e_u ，最后利用 $f(e_u, e_v), f : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ 预测用

户 u 与物品 v 的交互行为。文献 [48] 对图谱的嵌入建模的优化目标为实际的推荐任务,因此其推荐效果比直接使用传统知识图谱嵌入方法要好。在文献 [17] 中,作者在用户-物品交互数据与知识图谱构造的协同知识图上,首先利用 TransR 的方法对图上的节点和边的表征进行预训练,并基于注意力机制下的图网络^[83]的方法在协同知识图上进行正向反向的信息传播,从而迭代得到用户与物品的表征 e_u 和 e_v ,最后通过计算 $e_u^T e_v$ 估计预测的推荐结果。此外,大部分现有应用图卷积网络在推荐系统的方法存在两个问题,其一是在迭代计算网络中节点的表征时需要固定邻居的数量或者对邻居节点随机选取,从而限制了对图信息的利用;其二是随着图特征的增长,导致的模型难以收敛的问题。Tai 等^[87]针对上述两个问题,提出了 GraphSW 训练方法,提升了原有基于图卷积网络模型的推荐效果。最近,也有研究者通过使用多任务训练的方式解决嵌入方法优化目标与推荐任务不同的问题,例如文献 [47, 50] 通过联合图嵌入与推荐任务的损失函数,联合学习两个任务。

3.1.2 基于路径的方法

基于路径的应用知识图谱在推荐系统的方法主要是挖掘基于图谱用户、物品之间多种连接关系。例如我们在第 2.3 小节引入的例子,从图 1 中看到通过利用知识图谱中有关物品的知识,可以得到 3 条从用户 A 到达物品《肖申克的救赎》的路径,进而基于挖掘路径的信息构造推荐算法。由于知识图谱可以和推荐系统中的用户-物品交互数据构成一个异质信息网络 (heterogeneous information networks),因此可以在推荐系统中引入传统的对异质信息网络进行挖掘的元路径 (Meta-path) 的方法。具体地,首先给出异质信息网络的形式化定义。

定义1 (信息网络) 一个异质信息网络定义一个有向图 $G = (V, E)$,包含节点类型的映射函数 $\phi : V \rightarrow A$ 和链接类型映射函数 $\varphi : E \rightarrow B$,其中对于任意的节点 $V_i \in V$ 存在一个特定的节点类型 $\phi(V_i) \in A$,并且每个链接 $E_j \in E$ 属于一个特定的关系类型数 $\varphi(E_j) \in R$ 。

这里如果有节点类型数量 $|A| > 1$ 或者链接类型数量 $|B| > 1$,我们则称这个信息网络为异质信息网络。这里为了区分表示节点的类型和表示链接的边的类型,我们用 $A_i R_k A_j$ 表示类型 $A_i \in A$ 到类型 $A_j \in A$ 中存在联系 $R_k \in R$,反过来则用 $A_j R_k^{-1} A_i$ 表示。元路径是定义在不同类型节点之间的一系列关系序列组成的一条路径。其形式化定义如下。

定义2 (元路径) 一个元路径 $P = A_0 \xrightarrow{R_1} A_1 \xrightarrow{R_2} \dots \xrightarrow{R_k} A_k$ 是网络 G 中定义的在节点类型 A_0 和 A_k 中的一种复合关系 $R_1 R_2 \dots R_k$ 。

例如在图 1 中,存在元路径 $P_1 = \text{user} \xrightarrow{\text{like}} \text{movie} \xrightarrow{\text{StarredIn}^{-1}} \text{actor} \xrightarrow{\text{StarredIn}} \text{movie}$, $P_2 = \text{user} \xrightarrow{\text{like}} \text{movie} \xrightarrow{\text{DirectedIn}^{-1}} \text{actor} \xrightarrow{\text{StarredIn}} \text{movie} \xrightarrow{\text{RemakedInto}} \text{movie}$ 和 $P_3 = \text{user} \xrightarrow{\text{like}} \text{movie} \xrightarrow{\text{DirectedIn}^{-1}} \text{actor} \xrightarrow{\text{StarredIn}} \text{movie} \xrightarrow{\text{like}^{-1}} \text{user} \xrightarrow{\text{like}} \text{movie}$ 。

在文献 [16] 中,作者关注基于知识图谱构造推荐系统中形如 $\text{user} - \text{item} - \dots - \text{item}$ 的元路径,然后基于文献 [88] 提出的 PathSim 方法给出用户 u_i 到物品 e_j 之间的基于用户偏好的特征值:

$$s(u_i, e_j | P) = \sum_{e \in I} \frac{2 \times R_{u_i, e} \times |p_{e \sim e_j} : p_{e \sim e_j} \in P'|}{|p_{e \sim e} : p_{e \sim e} \in P'| + |p_{e_j \sim e_j} : p_{e_j \sim e_j} \in P'|}, \quad (6)$$

其中 $p_{e \sim e_j}$, $p_{e \sim e}$ 和 $p_{e_j \sim e_j}$ 分别表示 e 与 e_j , e 与 e 和 e_j 与 e_j 之间的路径, P 和 P' 分别是元路径表示复合关系 $R_1 R_2 \dots R_k$ 和 $R_2 \dots R_k$ 。作者通过计算所有用户与物品之间这一特征值得到用户的偏好特征矩阵,最后利用矩阵分解的方法对这一特征矩阵进行分解得到每种元路径的特征向量,从而类似文献 [89] 的方法通过计算不同元路径下所有用户与物品之间的特征向量点积的加权累加和得到推荐结果。在文献 [45] 中,作者使用卷积神经网络对每种不同元路径采样得到的从用户到物品路径进行

嵌入表征, 进而构造基于元路径的用户偏好特征, 并结合 NeuMF [8] 的算法构建推荐系统。在文献 [21] 中, 作者使用 Meta-Graph 的方式替代 Meta-path 对异质信息网络进行特征提取。Meta-Graph 相比 Meta-path 可以描绘异质信息网络中更复杂的特征信息, 例如在图 1 中, 我们可以合并之前两条元路径 P_2 和 P_3 中相同的节点、关系, 得到一个 Meta-Graph。类似的, 作者定义了用户的偏好特征矩阵的计算方式, 并设计了一种基于因子分解机的方法 Factorization Machine with Group lasso (FMG) 求解这个推荐问题。

传统基于元路径的方法虽然可以有很好的推荐效果以及可解释性, 但是存在着一定的问题。首先, 这类方法在构建推荐算法前需要先从数据中抽取、构造大量的 Meta-path 或 Meta-Graph, 因此其并不是一个端到端的方式, 并且当推荐场景或是图谱发生改变时, 需要重新构造。此外, 针对一些特定的推荐场景, 例如新闻推荐问题, 其每条新闻与图谱中多个实体关联, 从而很难直接应用元路径相关的方法。为此, 一些方法不采用构造元路径的方式, 而是直接对异质信息网络中的用户与物品之间存在的语义路径进行挖掘。在文献 [18] 中, 作者对每条在用户 u 与物品 i 中存在的语义路径使用循环神经网络进行建模, 从而得到每条语义路径的表征, 并且使用基于池化层的方法得到 u 到 i 的整体语义路径的特征, 并使用全连接网络预测推荐结果。此外, 类似文献 [89], 作者也对语义路径的长度设置了阈值, 其认为长度低于一定阈值的语义路径已经可以很好建模推荐系统, 并且过长的语义路径因为引入了过多的实体信息会包含一定的噪声。文献 [14, 20] 中也使用了类似的循环神经网络进行建模, 不同的是他们分别基于传统推荐任务和序列推荐任务设计了两种不同的注意力机制, 用于量化不同语义路径对推荐结果的权重, 从而提升了推荐模型的可解释性。

3.2 基于知识图谱的推荐应用分类

目前已有的利用知识图谱的推荐系统可以按照应用场景分为几类, 包括传统推荐系统、序列化推荐系统与可解释推荐系统, 下面对每个推荐场景进行介绍。

3.2.1 传统推荐系统

传统推荐系统是指在给定用户 - 物品的交互数据以及可能包含用户、物品的属性数据下, 对用户推荐其可能感兴趣的商品。在应用知识图谱的时候, 绝大多数情况下可以将物品或物品的相关属性直接映射到知识图谱上, 从而利用 3.1 小节的方法进行推荐。针对传统的推荐任务, 目前绝大多数的研究关注在电影、图书、新闻、电商商品、POI、音乐和药物这 7 大类推荐任务上。下面我们将介绍这几类推荐任务中常用的数据集和部分任务的特殊性, 表 3^[2, 10, 12~14, 16, 18~21, 44~48, 50~53, 61~64, 66~68, 71, 72, 74, 90, 91] 汇总了这些数据集并且给出了在相应数据上对利用知识图谱的推荐系统进行实验验证的文献。

- 电影。** 电影推荐是传统推荐任务之一, 目前相关的研究主要在 MovieLens 和 DoubanMovie 两个数据集上进行。MovieLens 数据集¹¹⁾是从 MovieLens 网站¹²⁾收集得到的, 包括电影、用户、用户打分数据和标签数据。其根据打分数据的数量, 被分为 3 个数据集, 即 MovieLens-100K, MovieLens-1M 和 MovieLens-20M。DoubanMovie 数据是从豆瓣电影¹³⁾上收集的, 同样包括用户评分和标签数据。在使用知识图谱推荐电影时, 因为电影名容易在知识图谱上匹配到, 因此研究者可以直接将推荐系统中的物品与知识图谱进行映射。通常在 MovieLens 数据上研究者使用公开的 DBpedia 知识图谱^[38] 或是公司私有的图谱 (例如 Microsoft Satori) 对电影进行匹配, 在 DoubanMovie 数据上使用 CN-DBpedia 图

11) <https://grouplens.org/datasets/movielens/>.

12) <http://movielens.org>.

13) <https://movie.douban.com/>.

表3 一个相关数据集的查询表
Table 3 A lookup table for relevant datasets

Category	Data	Ref.	Category	Data	Ref.
Movie	MovieLens-100K	[16, 45, 90]	Product	Amazon Electronics	[21]
	MovieLens-1M	[10, 12, 13, 18, 19, 44, 46, 47, 50, 63, 74, 91]		Amazon e-commerce	[66, 67]
	MovieLens-20M	[12, 20, 20, 48, 61, 72, 91]	POI	Yelp challenge	[14, 16, 18, 21, 45, 46, 53, 90]
	DoubanMovie	[52, 53, 71]		Dianping-Food	[72]
Book	BDbook2014	[19, 47]		CEM	[51]
	Book-Crossing	[13, 44, 50, 61, 72, 74, 91]	Music	Last.FM	[12, 14, 19, 45, 46, 48, 50, 72]
	Amazon-Book	[12, 14, 91]		NetEase Cloud Music	[64]
News	Intent Book	[10]	Medicine	TCM	[68]
	Bing-News	[2, 13, 50, 61]		MIMIC-III	[62]

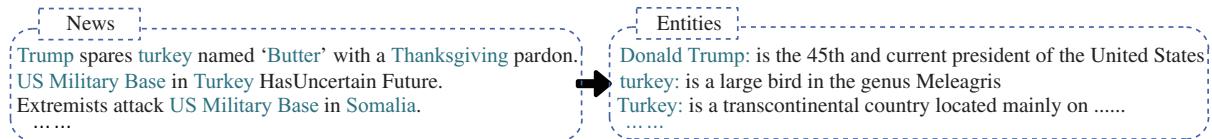


图4 (网络版彩图) 一个从新闻文本中进行实体链接的示例
Figure 4 (Color online) An illustration of entity linking for news data

谱^[43]进行匹配。在电影数据中,知识图谱上关于电影的相关补充特征(例如:导演、演员和电影类型等)可以作为辅助信息有效地帮助我们提升推荐系统的效果。

- **图书.** 图书推荐任务也是常见的推荐任务,目前常用的数据集有4个,包括BDbook2014, Book-Crossing^{[92]14)}, Amazon-Book¹⁵⁾和Intent Book^[10]。其中BDbook2014, Book-Crossing和Amazon-Book均为用户对阅读过的图书的打分数据,可以使用知识图谱DBpedia或Freebase^[39]找到对应的物品。Intent Books数据是从微软Bing搜索引擎日志中构造出来的,将用户点击或搜索书名的行为作为推荐系统中的交互数据,并与知识图谱Microsoft Satori中的数据进行关联。

- **新闻.** 新闻推荐属于一类特殊的推荐场景,因为新闻更新十分频繁造成其具有冷启动、交互数据稀疏的问题,并且新闻中的文本是语义高度浓缩的,充满实体和常识,如图4所示。为此,在应用知识图谱到新闻的推荐场景中时,首先会对它做实体链接^[93, 94],从而将新闻文本中的实体提取出来,并和图谱中的实体关联,从而进行后续的推荐算法。这类任务主要应用的是Bing-News^[2]数据集,可以使用一些通用类知识图谱如Wikidata^[42]和Microsoft Satori进行实体关联。

- **电商商品.** 电商网站对商品的主要展示形式高度依赖于商品的推荐结果,因此不少研究者对这个实际的推荐场景进行研究。其主要的研究数据是从电商网站Amazon上收集的,包括Amazon Electronics¹⁶⁾数据集以及Amazon e-commerce^[95]数据集。其利用商品的一些相关属性特征,如类别和品牌等,以及这些属性在图谱中的语义信息构成辅助信息提升推荐结果。

- **POI.** POI是指“兴趣点”(point of interest),也就是地图上任何非地理意义的有意义的点,例如商店、酒吧和餐厅等,其属于一种特殊的推荐场景,因为在POI推荐场景中通常会考虑用户和POI当

14) <http://www2.informatik.uni-freiburg.de/~cziegler/BX/>.

15) <http://jmcauley.ucsd.edu/data/amazon/>.

16) <http://jmcauley.ucsd.edu/data/amazon/>.

前所处的地理位置。其主要应用的数据集为 Yelp challenge¹⁷⁾, 其包含用户对去过的 POI 的评价数据, 包括文本和分数, 通常我们使用这些 POI 的属性信息数据, 如种类、地理位置、能否外带和停车位信息等构造包含 POI 的知识图谱。此外也有研究者在 Dianping-Food^[72] 和 CEM^[51] 等数据集上进行研究。

- **音乐.** 音乐推荐中, 主要应用的数据集是 Last.FM¹⁸⁾, 其包含用户在 Last.FM 在线音乐系统上对音乐家的欣赏历史数据, 并可以使用知识图谱 Freebase 或 Microsoft Satori 对音乐家进行匹配。

- **药物.** 随着智能医疗的发展, 近年来有研究者开始探索药品推荐这一应用, 通过对医疗图谱中病症、药品成分等相关辅助信息的利用, 可以有效地提升推荐结果与解释性。目前主要的数据集包括 Medical Information Mart for Intensive Care III (MIMIC-III)^[96] 和 Traditional Chinese Medicine (TCM)^[68] 数据集。

3.2.2 序列化推荐系统

区别于传统的推荐场景, 序列化推荐系统的目的在于捕捉用户兴趣的变化模式, 从而基于用户历史的交互信息预测当前的兴趣点^[97]。目前经典的建模方法包括基于马尔科夫 (Markov) 的模型^[97, 98] 和循环神经网络 recurrent neural network (RNN) 的模型^[99~101], 其中基于循环神经网络的方法因其在数据量大的情况下有较好的拟合效果, 因此近几年成为主流建模序列化推荐系统的方法。

知识图谱作为一种通用的推荐系统辅助信息, 近年来也开始被应用到序列化推荐任务中^[12, 20, 64]。其中文献 [64] 中, 作者使用一种比较直观的应用方法引入知识图谱的语义信息到基于 RNN 的序列化推荐算法中, 作者使用 3.1 小节中介绍的嵌入方法引入知识图谱的语义信息到物品的表征中, 然后作为双向 RNN 的输入, 最后基于网络的输出进行推荐预测。文献 [12] 针对 RNN 网络无法记忆过长时间序列的问题, 设计一种基于记忆网络^[102] 的结构, 在 RNN 网络迭代每个时间节点的信息时, 能够更好地记忆、更新使用 TransE 得到的实体语义表征信息, 从而有效地提升了对知识图谱语义信息的利用效果, 从而优化了推荐结果。文献 [20] 使用 3.1.2 小节中介绍的基于路径的方法引入知识图谱的信息, 区别于传统的推荐任务, 在每个时间节点构建路径时考虑了用户历史交互数据的时间, 避免了信息泄露的问题, 最后使用自注意力网络建模用户交互数据序列进行预测。

3.2.3 可解释推荐系统

可解释推荐系统的目的是在推荐物品给用户的同时给出推荐的原因, 从而提升推荐系统的可靠性^[58, 59]。知识图谱是一种非常高效的外部知识库, 帮助我们建立用户、物品之间的联系, 因此近年来基于知识图谱的可解释系统备受研究者关注^[14, 19, 20, 60]。目前研究者设计的知识图谱下的可解释推荐系统主要是基于路径推荐的方法, 如图 5(a) 所示, 通过知识图谱中实体关联的数据将物品进行关联, 并通过后续不同建模的方式进行可解释性的分析。例如, 文献 [14, 20] 设计不同的注意力机制来计算通过不同路径得到预测的推荐结果的权重, 通过权重的差异性, 可以给出其推荐结果主要的几条推荐原因。例如在图 5(a) 中, 对用户 A 来讲, 系统推荐给她电影《肖申克的救赎》的原因可能是其观看过剧情类的电影《阿甘正传》, 或者是其观看过同时由摩根·弗里曼主演的电影《成事在人》。此外, Ma 等^[60] 提出一种基于规则归纳与规则指导的神经网络推荐系统, 其基于物品相关的知识图谱中归纳常见的多跳关系模式, 推断不同的物品间的关联, 并且可以提出易于理解的可解释推荐系统。图 5(b) 给出了这样一个推荐系统的示例, 通过物品相关的知识图谱以及用户的交互行为, 我们可以学习到相关

17) <https://www.yelp.com/dataset/challenge>.

18) <https://grouplens.org/datasets/hetrec-2011/>, <http://ocelma.net/MusicRecommendationDataset/>.

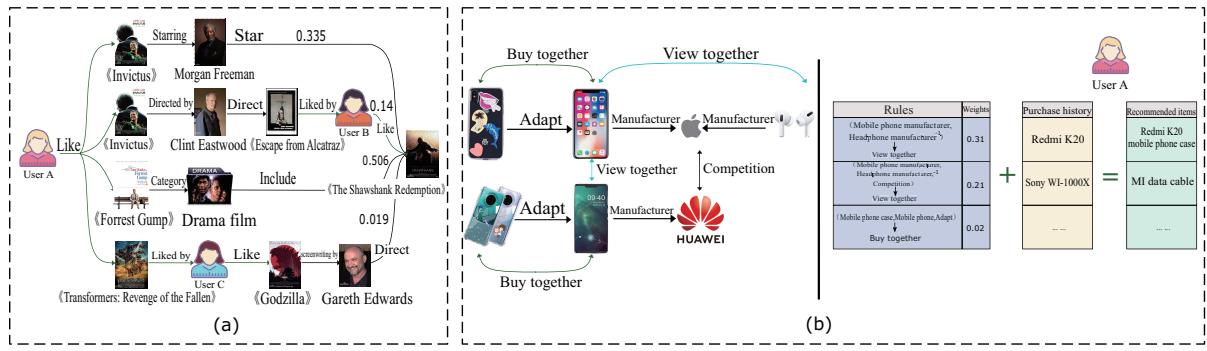


图 5 (网络版彩图) 一些基于知识图谱的可解释性推荐系统示例 (图片素材来源为百度百科)

Figure 5 (Color online) An illustration of two explainable knowledge graph-based recommender system

的规则,例如物品的手机制造商是竞争关系的情况下,用户可能同时浏览相关的物品;手机壳适配相关对应手机时,用户可能同时购买相关的物品。随后基于规则以及模型学习到的权重,我们就可以结合用户的历史数据进行推荐了。

4 未来的研究

基于以上的讨论,我们可以看到使用知识图谱在推荐系统时所带来的优势,可以帮助解决推荐系统中的交互数据稀疏与冷启动问题,从而提升推荐的准确性、推荐结果的多样性,并且很容易应用到不同的推荐场景中,从传统的商品推荐系统到序列化推荐任务,并且可以增强推荐系统的可解释性。本节将讨论一些我们认为会成为未来发展趋势的开放问题。

4.1 动态知识图谱

目前我们研究的知识图谱仅仅是静态,是某一时间片段上的数据,然而实际上,知识图谱之间的关系是具有时效性的。例如,(Bill Clinton, president of, USA)这个三元组仅在1993~2001年之间成立。近几年,一些学者已经开始展开对动态知识图谱的图嵌入方法^[103, 104]。具体地,例如文献[103]将原有的 (h, r, t) 三元组转化为 $(h, r, t, [\tau_s, \tau_e])$,其中 τ_s 和 τ_e 分别表示这组关系有效的时间戳。随后通过对时间戳的离散化,可以将 $\forall \tau \in [\tau_s, \tau_e]$ 对应的 (h, r, t) 重新定义为正样本集合 D_τ^+ ,相应可以基于时间维度、实体与关系中构造负样本集合 D_τ^- 。文献[103]类似TransH的思想,希望 $h'_\tau + r'_\tau \approx t'_\tau$,其中

$$h'_\tau = h_\tau - w_\tau^\top h_\tau w_\tau, \quad r'_\tau = r_\tau - w_\tau^\top r_\tau w_\tau, \quad t'_\tau = t_\tau - w_\tau^\top t_\tau w_\tau, \quad (7)$$

最后通过优化

$$L = \sum_{\tau \in [T]} \sum_{x \in D_\tau^+} \sum_{y \in D_\tau^-} \max(0, d_\tau(x) - d_\tau(y) + \text{margin}), \quad (8)$$

求解我们模型中的参数,其中 $d_\tau = \|h'_\tau + r'_\tau - t'_\tau\|$ 为相应的L1或L2范数。

进而一些研究者^[105, 106]将动态知识图谱的图嵌入方法运用到知识补全、知识推理^[36]、机器阅读^[107]等领域,取得了很好的效果。但目前还极少有研究者结合动态知识图谱的方法。事实上,在例如新闻推荐中这类对时效性要求较高的推荐系统中,我们可以运用动态知识图谱的方法,学习实体的动态演化规律,从而更有效地利用结构化知识中的边信息,进而提升原有对物品的表征效果。另一方面,在基于序列化推荐系统的问题中,其构造的协同知识图中,用户实体在不同时间具有不同的历史购买

行为, 因此可以将这个协同知识图更进一步考虑为动态的, 从而结合动态知识图嵌入的思想, 优化推荐效果.

4.2 动态推荐系统

在一些对知识数据的实时性要求较高的推荐场景中, 如音乐推荐、电影推荐等, 其物品更新频率较快, 因而我们的知识图谱也应会有较好的实时性. 但现有的两种方法在图谱发生频繁更新的情况下, 均存在训练时间成本过高的问题. 在嵌入方法中, 如果原始交互数据与知识图谱关联得到的异质信息网络图过大, 那么每次训练需要的时间成本是很大的, 面对图谱发生数据或数据类型新增的情况下, 我们需要每次进行重新训练. 而在基于路径的方法中, 针对图谱数据重新归纳元路径、抽取语义路径也需要较大的时间成本, 因此目前方法无法应对一些对时效性要求较高的推荐系统. 而在线学习算法中的增量学习为了解决这个问题可以提供有力的支持, 增量学习具有以下特点, 包括: (1) 可以从新的数据中学习知识, 更新模型; (2) 以往处理过的数据不必全部重新处理; (3) 学习新的数据知识时可以保留原有大部分的知识. 我们可以融合近几年来被提出的增量图嵌入方法^[108,109] 对我们原有推荐场景进行扩展, 以适应实时性要求较高的推荐问题, 降低对知识图谱与用户交互数据训练的时间成本.

此外, 当前的推荐场景是静态的, 即我们在训练推荐系统时往往假设用户的数据已经获取好, 并且在运行期内长期保持稳定, 但这与实际情况往往不符. 而在实际场景中, 例如电子商务中, 我们可以实时获得一系列的反馈数据. 如何基于实时的反馈数据结合知识图谱动态优化我们的推荐算法, 也是未来的研究趋势之一.

4.3 基于迁移学习

我们已经了解到知识图谱可以作为一种有效的辅助数据帮助我们解决冷启动的问题, 但这也仅适用于已经稳定积累了一段时间数据的推荐系统中出现一些全新的用户或者物品时的情况. 而面对没有过多数据积累的推荐系统, 目前往往使用一些基于内容的推荐算法, 但这样的结果往往容易产生推荐结果单一、准确率低的问题. 目前有很多研究者尝试使用迁移学习解决这类问题^[110,111]. 其主要是利用从源推荐系统学习到通用的模型参数, 通过共享这部分模型, 联合目标系统的私有模型参数共同训练, 从而提升模型的实际效果, 并且解决目标推荐系统训练数据不足的问题. 事实上, 知识图谱很多情况下在不同的推荐系统中拥有相似的规律, 因此利用迁移学习将源推荐系统对知识图谱中学习到的规律迁移到刚启动的目标推荐系统中, 可以进一步解决推荐系统的冷启动问题, 提升推荐结果, 并且可以提升应用迁移学习到推荐系统中时的可解释性.

5 结束语

知识图谱作为一种辅助数据可以有效地提升推荐系统的效果, 并且解决传统推荐算法面临的多种问题. 本文针对近年来应用知识图谱的推荐系统的相关论文进行了调研, 在简单回顾传统的推荐算法和知识图谱知识之后, 对基于知识图谱的推荐系统的研究方法与应用场景进行了系统性的总结和介绍. 最后给出了未来这个研究方向存在的发展趋势, 希望可以促进该领域的持续发展和进步.

参考文献

- 1 Schafer J B, Konstan J, Riedl J. Recommender systems in e-commerce. In: Proceedings of the 1st ACM Conference on Electronic Commerce, 1999. 158–166
- 2 Wang H, Zhang F, Xie X, et al. DKN: deep knowledge-aware network for news recommendation. In: Proceedings of the 2018 World Wide Web Conference, 2018. 1835–1844
- 3 Liu B, Fu Y, Yao Z, et al. Learning geographical preferences for point-of-interest recommendation. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2013. 1043–1051
- 4 Huang Z, Liu Q, Zhai C, et al. Exploring multi-objective exercise recommendations in online education systems. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019. 1261–1270
- 5 Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans Knowl Data Eng*, 2005, 17: 734–749
- 6 Sarwar B M, Karypis G, Konstan J A, et al. Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International Conference on World Wide Web, 2001. 285–295
- 7 Koren Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008. 426–434
- 8 He X, Liao L, Zhang H, et al. Neural collaborative filtering. In: Proceedings of the 26th International Conference on World Wide Web, 2017. 173–182
- 9 Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: Proceedings of the 4th ACM Conference on Recommender Systems, 2010. 135–142
- 10 Zhang F, Yuan N J, Lian D, et al. Collaborative knowledge base embedding for recommender systems. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. 353–362
- 11 Wang H, Zhang F, Hou M, et al. SHINE: signed heterogeneous information network embedding for sentiment link prediction. In: Proceedings of the 11th ACM International Conference on Web Search and Data Mining, 2018. 592–600
- 12 Huang J, Zhao W X, Dou H, et al. Improving sequential recommendation with knowledge-enhanced memory networks. In: Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018. 505–514
- 13 Wang H, Zhang F, Wang J, et al. RippleNet: propagating user preferences on the knowledge graph for recommender systems. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018. 417–426
- 14 Wang X, Wang D, Xu C, et al. Explainable reasoning over knowledge graphs for recommendation. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2019. 5329–5336
- 15 Qin C, Zhu H, Zhu C, et al. DuerQuiz: a personalized question recommender system for intelligent job interview. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019. 2165–2173
- 16 Yu X, Ren X, Sun Y, et al. Personalized entity recommendation: a heterogeneous information network approach. In: Proceedings of the 7th ACM International Conference on Web Search and Data Mining, 2014. 283–292
- 17 Wang X, He X, Cao Y, et al. KGAT: knowledge graph attention network for recommendation. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019. 950–958
- 18 Sun Z, Yang J, Zhang J, et al. Recurrent knowledge graph embedding for effective recommendation. In: Proceedings of the 12th ACM Conference on Recommender Systems, 2018. 297–305
- 19 Song W, Duan Z, Yang Z, et al. Explainable knowledge graph-based recommendation via deep reinforcement learning. 2019. ArXiv: 190609506
- 20 Huang X, Fang Q, Qian S, et al. Explainable interaction-driven user modeling over knowledge graph for sequential recommendation. In: Proceedings of the 27th ACM International Conference on Multimedia, 2019. 548–556
- 21 Zhao H, Yao Q, Li J, et al. Meta-Graph based recommendation fusion over heterogeneous information networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,

2017. 635–644
- 22 Gao J, Xin X, Liu J, et al. Fine-grained deep knowledge-aware network for news recommendation with self-attention. In: Proceedings of the IEEE International Conference on Web Intelligence, 2018. 81–88
- 23 Passant A. dbrec-music recommendations using DBpedia. In: Proceedings of the International Semantic Web Conference. Berlin: Springer, 2010. 209–224
- 24 Oramas S, Ostuni V C, Noia T D, et al. Sound and music recommendation with knowledge graphs. ACM Trans Intell Syst Technol, 2017, 8: 1–21
- 25 Resnick P, Iacovou N, Suchak M, et al. GroupLens: an open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, 1994. 175–186
- 26 Zhu H, Xiong H, Ge Y, et al. Mobile app recommendations with security and privacy awareness. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014. 951–960
- 27 Pazzani M J, Billsus D. Content-based recommendation systems. In: Proceedings of the Adaptive Web. Berlin: Springer, 2007. 325–341
- 28 Su X, Khoshgoftaar T M. A survey of collaborative filtering techniques. Adv Artif Intell, 2009, 2009: 1–19
- 29 Burke R. Hybrid recommender systems: survey and experiments. User Modeling User-Adapted Interact, 2002, 12: 331–370
- 30 Bellogín A, Cantador I, Díez F, et al. An empirical comparison of social, collaborative filtering, and hybrid recommenders. ACM Trans Intell Syst Technol, 2013, 4: 1–29
- 31 Konstas I, Stathopoulos V, Jose J M. On social networks and collaborative recommendation. In: Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2009. 195–202
- 32 Singhal A. Introducing the knowledge graph: things, not strings. Official Google Blog, 2012. <https://www.blog.google/products/search/introducing-knowledge-graph-things-not/>
- 33 Huang X, Zhang J, Li D, et al. Knowledge graph embedding based question answering. In: Proceedings of the 12th ACM International Conference on Web Search and Data Mining, 2019. 105–113
- 34 Hao Y, Zhang Y, Liu K, et al. An end-to-end model for question answering over knowledge base with cross-attention combining global knowledge. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017. 221–231
- 35 Kumar A, Kawahara D, Kurohashi S. Knowledge-enriched two-layered attention network for sentiment analysis. 2018. ArXiv: 180507819
- 36 Trivedi R, Dai H, Wang Y, et al. Know-evolve: deep temporal reasoning for dynamic knowledge graphs. In: Proceedings of the 34th International Conference on Machine Learning, 2017. 3462–3471
- 37 Suchanek F M, Kasneci G, Weikum G. Yago: a core of semantic knowledge. In: Proceedings of the 16th International Conference on the World Wide Web, 2007. 697–706
- 38 Lehmann J, Isele R, Jakob M, et al. DBpedia — a large-scale, multilingual knowledge base extracted from Wikipedia. Semantic Web, 2015, 6: 167–195
- 39 Bollacker K, Evans C, Paritosh P, et al. Freebase: a collaboratively created graph database for structuring human knowledge. In: Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, 2008. 1247–1250
- 40 Miller G A. Wordnet: a lexical database for english. Commun ACM, 1995, 38: 39–41
- 41 Carlson A, Betteridge J, Wang R C, et al. Coupled semi-supervised learning for information extraction. In: Proceedings of the 3rd ACM International Conference on Web Search and Data Mining, 2010. 101–110
- 42 Vrandečić D, Krötzsch M. Wikidata: a free collaborative knowledge base. Commun ACM, 2014, 57: 78–85
- 43 Xu B, Xu Y, Liang J, et al. Cn-dbpedia: a never-ending chinese knowledge extraction system. In: Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Berlin: Springer, 2017. 428–438
- 44 Tang X, Wang T, Yang H, et al. Akupm: attention-enhanced knowledge-aware user preference model for recommendation. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019. 1891–1899

- 45 Hu B, Shi C, Zhao W X, et al. Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018. 1531–1540
- 46 Sha X, Sun Z, Zhang J. Attentive knowledge graph embedding for personalized recommendation. 2019. ArXiv: 191008288
- 47 Cao Y, Wang X, He X, et al. Unifying knowledge graph learning and recommendation: towards a better understanding of user preferences. In: Proceedings of the International Conference on World Wide Web, 2019. 151–161
- 48 Wang H, Zhao M, Xie X, et al. Knowledge graph convolutional networks for recommender systems. In: Proceedings of the International Conference on World Wide Web, 2019. 3307–3313
- 49 Chowdhury G, Srilakshmi M, Chain M, et al. Neural factorization for offer recommendation using knowledge graph embeddings. In: Proceedings of the SIGIR 2019 eCom Workshop, 2019
- 50 Wang H, Zhang F, Zhao M, et al. Multi-task feature learning for knowledge graph enhanced recommendation. In: Proceedings of the International Conference on World Wide Web, 2019. 2000–2010
- 51 Dadoun A, Troncy R, Ratier O, et al. Location embeddings for next trip recommendation. In: Proceedings of the International Conference on World Wide Web, 2019. 896–903
- 52 Yang D, Guo Z, Wang Z, et al. A knowledge-enhanced deep recommendation framework incorporating gan-based models. In: Proceedings of the 2018 IEEE International Conference on Data Mining, 2018. 1368–1373
- 53 Shi C, Zhang Z, Luo P, et al. Semantic path based personalized recommendation on weighted heterogeneous information networks. In: Proceedings of the 24th ACM International Conference on Information and Knowledge Management, 2015. 453–462
- 54 McAuley J, Leskovec J. Hidden factors and hidden topics: understanding rating dimensions with review text. In: Proceedings of the 7th ACM Conference on Recommender Systems, 2013. 165–172
- 55 Zhang Y, Lai G, Zhang M, et al. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In: Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, 2014. 83–92
- 56 He X, Chen T, Kan M Y, et al. Trirank: review-aware explainable recommendation by modeling aspects. In: Proceedings of the 24th ACM International Conference on Information and Knowledge Management, 2015. 1661–1670
- 57 Gao J, Wang X, Wang Y, et al. Explainable recommendation through attentive multi-view learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2019. 3622–3629
- 58 Chen Z, Wang X, Xie X, et al. Co-attentive multi-task learning for explainable recommendation. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence, AAAI Press, 2019. 2137–2143
- 59 Rago A, Cocarascu O, Toni F. Argumentation-based recommendations: fantastic explanations and how to find them. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018. 1949–1955
- 60 Ma W, Zhang M, Cao Y, et al. Jointly learning explainable rules for recommendation with knowledge graph. In: Proceedings of the International Conference on World Wide Web, 2019. 1210–1221
- 61 Wang H, Zhang F, Wang J, et al. Exploring high-order user preference on the knowledge graph for recommender systems. ACM Trans Inf Syst, 2019, 37: 1–26
- 62 Wang M, Liu M, Liu J, et al. Safe medicine recommendation via medical knowledge graph embedding. 2017. ArXiv: 171005980
- 63 Palumbo E, Rizzo G, Troncy R. Entity2rec: learning user-item relatedness from knowledge graphs for top-n item recommendation. In: Proceedings of the 11th ACM Conference on Recommender Systems, 2017. 32–36
- 64 Lin Q, Niu Y, Zhu Y, et al. Heterogeneous knowledge-based attentive neural networks for short-term music recommendations. IEEE Access, 2018, 6: 58990–59000
- 65 Zhou Z, Liu S, Xu G, et al. Knowledge-based recommendation with hierarchical collaborative embedding. In: Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining. Berlin: Springer, 2018. 222–234
- 66 Zhang Y, Ai Q, Chen X, et al. Learning over knowledge-base embeddings for recommendation. 2018. ArXiv:

- 180306540
- 67 Ai Q, Azizi V, Chen X, et al. Learning heterogeneous knowledge base embeddings for explainable recommendation. *Algorithms*, 2018, 11: 137
- 68 Wang X, Zhang Y, Wang X, et al. A knowledge graph enhanced topic modeling approach for herb recommendation. In: Proceedings of the International Conference on Database Systems for Advanced Applications. Berlin: Springer, 2019. 709–724
- 69 He M, Wang B, Du X. HI2Rec: exploring knowledge in heterogeneous information for movie recommendation. *IEEE Access*, 2019, 7: 30276–30284
- 70 Huang H H. An MPD player with expert knowledge-based single user music recommendation. In: Proceedings of the International Conference on Web Intelligence, 2019. 318–321
- 71 Yang D, Wang Z, Jiang J, et al. Knowledge embedding towards the recommendation with sparse user-item interactions. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2019. 325–332
- 72 Wang H, Zhang F, Zhang M, et al. Knowledge-aware graph neural networks with label smoothness regularization for recommender systems. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019. 968–977
- 73 Zhu G, Bin C, Gu T, et al. A neural user preference modeling framework for recommendation based on knowledge graph. In: Proceedings of the Pacific Rim International Conference on Artificial Intelligence. Berlin: Springer, 2019. 176–189
- 74 Zhong Y, Song X, Yang B, et al. An interpretable recommendations approach based on user preferences and knowledge graph. In: Proceedings of the International Conference on Swarm Intelligence, Springer, 2019. 326–337
- 75 Bordes A, Usunier N, Garcia-Duran A, et al. Translating embeddings for modeling multi-relational data. In: Proceedings of the 2013 Advances in Neural Information Processing Systems, 2013. 2787–2795
- 76 Wang Z, Zhang J, Feng J, et al. Knowledge graph embedding by translating on hyperplanes. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence, 2014. 1112–1119
- 77 Lin Y, Liu Z, Sun M, et al. Learning entity and relation embeddings for knowledge graph completion. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence, 2015. 2181–2187
- 78 Ji G, He S, Xu L, et al. Knowledge graph embedding via dynamic mapping matrix. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language, 2015. 687–696
- 79 Lin Z, Feng M, dos Santos C N, et al. A structured self-attentive sentence embedding. 2017. ArXiv: 170303130
- 80 Dong Y, Chawla N V, Swami A. Metapath2Vec: scalable representation learning for heterogeneous networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017. 135–144
- 81 Grover A, Leskovec J. node2vec: scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. 855–864
- 82 Ying R, He R, Chen K, et al. Graph convolutional neural networks for web-scale recommender systems. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018. 974–983
- 83 Veličković P, Cucurull G, Casanova A, et al. Graph attention networks. 2017. ArXiv: 171010903
- 84 Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering. In: Proceedings of the Advances in Neural Information Processing Systems, 2016. 3844–3852
- 85 Ye Y, Zhu H, Xu T, et al. Identifying high potential talent: a neural network based dynamic social profiling approach. In: Proceedings of 2019 IEEE International Conference on Data Mining, 2019. 718–727
- 86 Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks. 2016. ArXiv: 160902907
- 87 Tai C Y, Wu M R, Chu Y W, et al. GraphSW: a training protocol based on stage-wise training for GNN-based recommender model. 2019. ArXiv: 190805611
- 88 Sun Y, Han J, Yan X, et al. Pathsim: meta path-based top-k similarity search in heterogeneous information networks. In: Proceedings of the VLDB Endowment, 2011. 992–1003

- 89 Yu X, Ren X, Sun Y, et al. Recommendation in heterogeneous information networks with implicit user feedback. In: Proceedings of the 7th ACM Conference on Recommender Systems, 2013. 347–350
- 90 Catherine R, Cohen W. Personalized recommendations using knowledge graphs: a probabilistic logic programming approach. In: Proceedings of the 10th ACM Conference on Recommender Systems, 2016. 325–332
- 91 Qu Y, Bai T, Zhang W, et al. An end-to-end neighborhood-based interaction model for knowledge-enhanced recommendation. 2019. ArXiv: 190804032
- 92 Ziegler C N, McNee S M, Konstan J A, et al. Improving recommendation lists through topic diversification. In: Proceedings of the 14th International Conference on World Wide Web, 2005. 22–32
- 93 Milne D, Witten I H. Learning to link with wikipedia. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management, 2008. 509–518
- 94 Sil A, Yates A. Re-ranking for joint named-entity recognition and linking. In: Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, 2013. 2369–2374
- 95 He R, McAuley J. Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering. In: Proceedings of the 25th International Conference on World Wide Web, 2016. 507–517
- 96 Johnson A E W, Pollard T J, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*, 2016, 3: 160035
- 97 Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized markov chains for next-basket recommendation. In: Proceedings of the 19th International Conference on World Wide Web, 2010. 811–820
- 98 Cheng C, Yang H, Lyu M R, et al. Where you like to go next: successive point-of-interest recommendation. In: Proceedings of the 23rd International Joint Conference on Artificial Intelligence, 2013. 2605–2611
- 99 Hidasi B, Quadrana M, Karatzoglou A, et al. Parallel recurrent neural network architectures for feature-rich session-based recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems, 2016. 241–248
- 100 Quadrana M, Karatzoglou A, Hidasi B, et al. Personalizing session-based recommendations with hierarchical recurrent neural networks. In: Proceedings of the Eleventh ACM Conference on Recommender Systems, 2017. 130–137
- 101 Yu F, Liu Q, Wu S, et al. A dynamic recurrent model for next basket recommendation. In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2016. 729–732
- 102 Miller A, Fisch A, Dodge J, et al. Key-value memory networks for directly reading documents. 2016. ArXiv: 160603126
- 103 Dasgupta S S, Ray S N, Talukdar P. Hyte: hyperplane-based temporally aware knowledge graph embedding. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018. 2001–2011
- 104 Tang X, Yuan R, Li Q, et al. Timespan-aware dynamic knowledge graph embedding by incorporating temporal evolution. *IEEE Access*, 2020, 8: 6849–6860
- 105 García-Durán A, Dumančić S, Niepert M. Learning sequence encoders for temporal knowledge graph completion. 2018. ArXiv: 180903202
- 106 Jiang T, Liu T, Ge T, et al. Towards time-aware knowledge graph completion. In: Proceedings of the 26th International Conference on Computational Linguistics, 2016. 1715–1724
- 107 Das R, Munkhdalai T, Yuan X, et al. Building dynamic knowledge graphs from text using machine reading comprehension. 2018. ArXiv: 181005682
- 108 Goyal P, Kamra N, He X, et al. Dyngem: deep embedding method for dynamic graphs. 2018. ArXiv: 180511273
- 109 Goyal P, Chhetri S R, Canedo A. dyngraph2vec: capturing network dynamics using dynamic graph representation learning. *Knowl Based Syst*, 2019, 187: 104816
- 110 Zhao L, Pan S J, Xiang E W, et al. Active transfer learning for cross-system recommendation. In: Proceedings of the 27th AAAI Conference on Artificial Intelligence, 2013. 1205–1211
- 111 Pan W, Xiang E W, Liu N N, et al. Transfer learning in collaborative filtering for sparsity reduction. In: Proceedings of 24th AAAI Conference on Artificial Intelligence, 2010. 230–235

A survey on knowledge graph-based recommender systems

Chuan QIN^{1,2}, Hengshu ZHU^{2*}, Fuzhen ZHUANG^{3,4}, Qingyu GUO^{5,3,4},
Qi ZHANG^{1,2}, Le ZHANG¹, Chao WANG^{1,2}, Enhong CHEN¹ & Hui XIONG^{1*}

1. School of Computer Science, University of Science and Technology of China, Hefei 230022, China;
2. Baidu Inc., Beijing 100085, China;
3. Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China;
4. University of Chinese Academy of Sciences, Beijing 100049, China;
5. School of Computer Science, The Hong Kong University of Science and Technology, Hong Kong 999077, China

* Corresponding author. E-mail: zhuhengshu@gmail.com, xionghui@gmail.com

Abstract Recommender system (RS) targets at providing accurate item recommendations to users with respect to their preferences; it has been widely employed in various online applications for addressing the problem of information explosion and improving user experience. In the past decades, while tremendous efforts have been made in enhancing the performance of RSs, some long-standing challenges, such as data sparsity, cold start, and result diversity, are unaddressed. Along this line, an emerging research trend is to exploit the rich semantic information contained in the knowledge graph (KG); it has been proven to be an effective way to enhance the capability of RSs. To this end, we provide a focused survey on KG-based RS via a holistic perspective of both technologies and applications. Specifically, firstly, we briefly review the core concepts and classical algorithms of the RSs and KGs. Secondly, we comprehensively introduce the representative and state-of-the-art works in this field based on different strategies of exploiting KGs for RSs. Meanwhile, we also summarize some typical application scenarios of KG-based RSs, for facilitating the hands-on practices of corresponding algorithms. Finally, we present our opinions on the prospects of KG-based RS and suggest some future research directions in this area.

Keywords knowledge graph, recommender system, collaborative filtering, heterogeneous information network, graph embedding



Chuan QIN was born in 1993. He received his B.S degree in computer science and technology from University of Science and Technology of China (USTC) in 2015. He is currently working toward a Ph.D. degree in the School of Computer Science and Technology, USTC. His current research interests include natural language processing and recommender system.



Hengshu ZHU was born in 1986. He is currently a senior data scientist at Baidu Inc. He received his Ph.D. degree in 2014 and B.E. degree in 2009, both in computer science from University of Science and Technology of China. His general area of research is data mining and machine learning, with a focus on developing advanced data analysis techniques for emerging applied business research.



Enhong CHEN was born in 1968. He is a professor and vice dean at the School of Computer Science, University of Science and Technology of China. His general area of research includes data mining and machine learning, social network analysis, and recommender systems.



Hui XIONG was born in 1972. He is currently a full professor at Rutgers, the State University of New Jersey. He received his B.E. degree from University of Science and Technology of China (USTC), M.S. degree from National University of Singapore (NUS), and Ph.D. degree from University of Minnesota (UMN), USA. His general area of research is data and knowledge engineering, with a focus on developing effective and efficient data analysis techniques for emerging data-intensive applications.