计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331,CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目：融合循环知识图谱和协同过滤电影推荐算法
作者：李浩，张亚钏，康雁，杨兵，卜荣景，李晋源
网络首发日期：2019-10-12
引用格式：李浩，张亚钏，康雁，杨兵，卜荣景，李晋源．融合循环知识图谱和协同过滤电影推荐算法．计算机工程与应用.
http://kns.cnki.net/kcms/detail/11.2127.TP.20191012.0936.010.html

融合循环知识图谱和协同过滤电影推荐算法

李　浩，张亚钏，康　雁，杨　兵，卜荣景，李晋源

云南大学 软件学院，昆明 650091

摘　要：推荐系统对筛选有效信息和提高信息获取效率具有重大的意义。传统的推荐系统会面临数据稀松和冷启动等问题。本文利用外部评分和物品内涵知识相结合，提出一种基于循环知识图谱和协同过滤的电影推荐模型——RKGE-CF。在充分考虑物品、用户、评分之间的相关性后，利用基于物品和用户的协同过滤进行 Top-K 推荐；将物品的外部附加数据和用户偏好数据加入知识图谱，提取实体相互之间的依赖关系，构建用户和物品之间的交互信息，以便揭示实体与关系之间的语义，帮助理解用户兴趣；将多种推荐结果按不同方法融合进行对比；模型训练时使用多组不同的负样本作为对比，以优化模型；最后利用真实电影数 Movielens 和 IMDB 映射连接成新数据集进行测试。实验结果证明该模型对于推荐效果的准确率有显著的提升，同时能更好的解释推荐背后的原因。

关键词：知识图谱；协同过滤；推荐系统；可解释性推荐
文献标志码：A　**中图分类号：**TP311.5　doi：10.3778/j.issn.1002-8331.1907-0131

李浩，张亚钏，康雁，等. 融合循环知识图谱和协同过滤电影推荐算法. 计算机工程与应用
LI Hao, ZHANG Yachuan, KANG Yan, et al. Fusion recurrent knowledge graph and collaborative filtering movie recommendation algorithm. Computer Engineering and Application

Fusion Recurrent Knowledge Graph and Collaborative Filtering Movie Recommendation Algorithm

LI Hao, ZHANG Yachuan, KANG Yan, YANG Bing, BU Rongjing, LI Jinyuan

School of Software, Yunnan University, Kunming 650091, China

Abstract: Recommendation system has significance in screening the useful information and improving the efficiency of information acquisition. However, sparse data and cold start are encountered in traditional recommendation systems. Therefore, this paper proposes a method combining external rating and item connotation knowledge and a movie recommendation model basing on knowledge graph and collaborative filtering——RKGE-CF. After the full consideration on correlation between items, users and ratings, Top-k recommendation is adopted which using collaborative filtering based on items and users. Adding the items external additional data and user preference data to the knowledge map, extracting the dependencies between entities and building the interactive information between users and items. Then, the model can reveal the semantics between entities and relations, improve the understanding of users' interests to make recommendations. Using different algorithm to fuse several recommendation

基金项目：国家自然科学基金（No.61762092）；云南省软件工程重点实验室开放基金项目（No.2017SE204）。
作者简介：李浩(1970-)，男，博士，教授，主要研究方向为机器学习;张亚钏(1996-)，女，硕士生，主要研究方向为推荐系统与深度学习；康雁(1972-)，女，副教授，硕士生导师，主要研究方向为机器学习；杨兵(1995-)，男，硕士生，主要研究方向为城市大数据计算与深度学习；卜荣景(1996-)，女，硕士生，主要研究方向为推荐系统与深度学习；李晋源（1994-)，男，硕士生，主要研究方向为城市大数据计算与深度学习。

results and compare it. In the training of the model, we used multiple groups of different negative samples for comparison, to optimize the model. Finally, a new dataset is obtained for testing by mapping real movie Movielens and IMDB. Experimental results show that this model improve the accuracy of recommendation effect, and explain the reasons behind the recommendation.

Key words：knowledge graph; collaborative filtering; recommendation system; explainable recommendation

1 引言

随着信息化社会的推广和普及，互联网技术的迅速发展使得信息以爆炸式增长的态势呈现在用户面前，用户难以从信息过载难题下获得对自己真正有用的那部分信息，因此如何有效地为用户筛选信息是大数据时代的一个课题。推荐系统研究的主要问题就是如何从这些过载的信息中找到每个用户感兴趣的内容，并把这些内容推送给用户。

协同过滤算法是推荐领域应用广泛的算法。传统的推荐算法不需要预先获得用户或物品的特征数据，仅依赖于用户的历史行为数据对用户进行建模，从而为用户进行推荐。该算法多数采用最近邻技术，利用用户历史喜好信息计算用户之间的距离，然后利用目标用户的最近邻居对商品评价的加权评分值来预测目标用户对特定商品的喜好程度。但常常面临着数据稀松和推荐结果挖法解释等问题。

因此学者考虑利用知识图谱来完善基于内容的推荐系统中对用户和物品的特征描述从而提升推荐效果。辅助信息可以丰富对用户和物品的描述、增强推荐算法的挖掘能力，从而有效地解决稀疏性和冷启动问题，提高推荐结果的精确性、多样性和可解释性，所以如何根据具体推荐场景的特点将各种辅助数据有效地融入推荐算法成为推荐系统研究领域的热点和难点。并且混合方法可以弥补基于内容推荐在多样性的不足[1]。

为了有效的推荐，本文依据混合推荐的基本思路，结合深度学习，在循环网络的基础上结合协同过滤和知识图谱，提出一个高效的推荐模型：RKGE-CF（Recurrent Knowledge Graph Embedding based on Collaborative Filtering）。主要内容包括：1）采用了循环知识图嵌入，自动学习实体和实体之间路径的语义表示，以表征用户对物品的偏好，得到更好的推荐结果。2）在知识图谱中同时加入外部评分，作为学习权重，更好的表达用户的偏好程度。3)改进了协同过滤算法，在传统的相似度计算中加入惩罚因子，以消除热门物品和不活跃用户对结果的影响。4）利用的不同的融合方法将内涵知识与外部评分结果融合，得到最优的融合推荐结果。5）在公开数据集 MovieLens 和 IMDB 上进行测试，测试结果在多维度上进行比较，不断调节参数，达到最优推荐性能。实验结果表明，本文所提出的框架在一定程度提高了推荐的准确性。

2 相关工作

2.1 基于协同过滤的推荐

传统的推荐系统算法可以分为协同过滤推荐、基于内容推荐和混合推荐三种。Sarwar 等[2]提出一种基于物品的预测算法，建立物品相似度的预计算模型，提高推荐系统修改的在线可伸缩性。Fletcher 等[3]利用基于个性化的协同过滤为用户提供个性化新歌推荐。Hernando[4]等提出一种基于将评价矩阵分解成两个非负矩阵的协同过滤算法预测用户口味的新技术。Liu[5]等提出一种采用关联挖掘技术从论文上下文中计算出用于协同过滤的引用论文之间的相似性。基于内容的推荐能够很好的解决用户行为数据稀疏和新用户的冷启动问题，通过使用向量空间模型、线性分类、线性回归等方法对用户兴趣特征和物品特征进行建模，为用户推荐与他感兴趣的内容相似的物品。江周峰[6]等提出一种结合社会化标签的基于内容的推荐算法，可以较好的识别模糊标签。Shu J[7]等提出一种基于卷积神经网络的基于内容的推荐算法，文本信息被直接用于进行基于内容的推荐而无需标记。混合推荐指将多种推荐技术进行混合相互弥补缺点，从而获得更好的推荐效果。Chu[8]等将视觉信息视为中间体，整合基于内容的推荐和协同过滤，具有很高的实用性。

Subramaniam[9]等提出一种基于贝叶斯算法的非个性化推荐，在计算最小网页加载时间的电影预测和推荐因子方面被证明是有效的。

2.2 基于深度学习的推荐

深度学习通过学习一种深层次非线性网络结构，表征用户和物品相关的海量数据。源异构数据中进行自动特征学习并将不同数据映射到一个相同的隐空间，从而获取用户和物品的深层次统一特征表示，将深度学习用在推荐系统上具有更好的抗噪性和有效性。Rumelhart[10]等人提出的自编码器(Autoencoder，AE)通过对用户和物品的相关信息进行隐层特征表示，应用于推荐系统中用户对物品的偏好预测。Smolensky P[11]等人提出的受限玻尔兹曼机(Restricted Boltzmann Machine，RBM)通过重构学习用户评分矩阵对推荐系统中的未知评分进行预测。Hinton[12][13]等人提出的深度信念网络(Deep Belief Network，DBN)采用贪婪追逐算法训练多层非线性变量连接组成的生成式模型，从而从无标记数据中获取更深层次的特征表示，多数应用于音乐数据的推荐。随着卷积神经网络(Convolutional Neural Network，CNN)[14][15]的出现避免了前面所提及的复杂特征提取和重构学习从而获取物品的低维空间表示，减少了推荐模型中的参数数量，成为推荐系统研究的热点。由于 CNN 未考虑到建模数据之间的序列影响，循环神经网络(Recurrent Neural Network，RNN)应运而生，并由此研究出更加有效建模长期依赖关系的长短时记忆网络(Long Short-Term Memory，LSTM)[16][17]和门控循环单元(Gated Recurrent Unit，GRU)[18]，广泛应用于结合社交网络的推荐。

2.3 基于知识图谱的推荐

知识图谱作为一种新兴类型的辅助数据源引起了越来越多学者的关注，现有的将知识图谱引入推荐系统的工作分为以 LibFM 为代表的通用的基于特征的推荐算法[19]和以 PER、MetaGraph 为代表的基于路径的推荐算法[20]，前者将知识图谱弱化为物品属性，统一地把用户和物品的属性作为推荐算法的输入，然而该方法无法高效地利用知识图谱的全部信息；后者将知识图谱视为一个异构信息网络，然后构造物品之间的基于 meta-path 或 meta-graph 的特征，充分且直观地利用知识图谱的网络结构，不过工作量大。吴玺煜[21]等人使用知识图谱表示学习方法，将语义数据嵌入到低维空间，并将物品语义信息融入协同过滤推荐。Zhang 等[22]人分别用网络嵌入、多层降噪自动编码机、层叠卷积自编码器获取结构化知识的向量化表示、文本知识特征、图片知识特征，紧接着将这三类特征融合进协同集成学习框架实现个性化推荐，实验证明基于深度学习的知识图谱推荐算法在推荐效果上优于基于协同过滤的传统推荐模型。

现有的方法局限于考虑物品外在的物品-用户评分矩阵信息，忽视了物品自身的信息。本文所提出的模型考虑到语义问题，将实体嵌入到低维空间里，还保持图中原有的结构和语义信息，通过知识图谱语义网络引入额外的一些辅助信息作为输入，丰富实体之间的语义关联，使推荐结果更急精确。此外，知识图谱发散不同的关系连接种类和历史记录，提升了推荐结果的多样性和可解释性。

3 RKGE-CF 架构

本节将分块介绍 RKGE-CF 的具体内部结构。首先采用了循环知识图的电影推荐模型，去自动学习实体和实体之间路径的语义表示，以表征用户对物品的偏好,在知识图谱的基础上结合循环神经网络进行学习，形成循环知识图谱。考虑到电影实体之间关系序列长度等问题，本文利用循环知识图谱，较为方便的学习实体关系的语义，能够对不同长度的序列进行建模，特别适用于建模路径，捕获实体和实体对之间的整个路径的语义的能力较好。对于多条路径与不同长度可能连接实体,网络能捕获所有可能的关系。

在加入一批递归的循环神经网络后，可以链接相同的实体对的路径，也就是完成了实体之间的关联。在对实体对的路径的语义进行建模，将路径无缝的融合为到推荐模型中，使得每个实体和关系可以通过学习得到对应的低维向量。既保持图中原有结构或语义信息，同时还方便链接相同语义实体的路径，再将这些路径融合到推荐中，提高推荐精确度。因此一组好的实体向量可以充分且完全地表示实体之间的相互关系，利用循环知识图谱特征学习可以很方便地将数据特征引入各种推荐系统算法中。

然后加入了协同过滤的推荐，包括基于物品的协同过滤和基于用户的协同过滤。对于协同过滤的推荐，系统会执行最近邻搜索，计算相关的相似度

之后得到推荐结果。循环知识图谱可以学习到推荐关系中的内涵知识，协同过滤可以很好地使用外部评分，我们提出的方法将内涵知识和外部评分进行组合，有效地提高推荐的效率。

循环知识图谱嵌入模型框架如图 1 所示，主要由上下两部分组成。上部分是循环知识图谱嵌入，下部分是结合基于用户的协同过滤和基于物品的协同过滤。

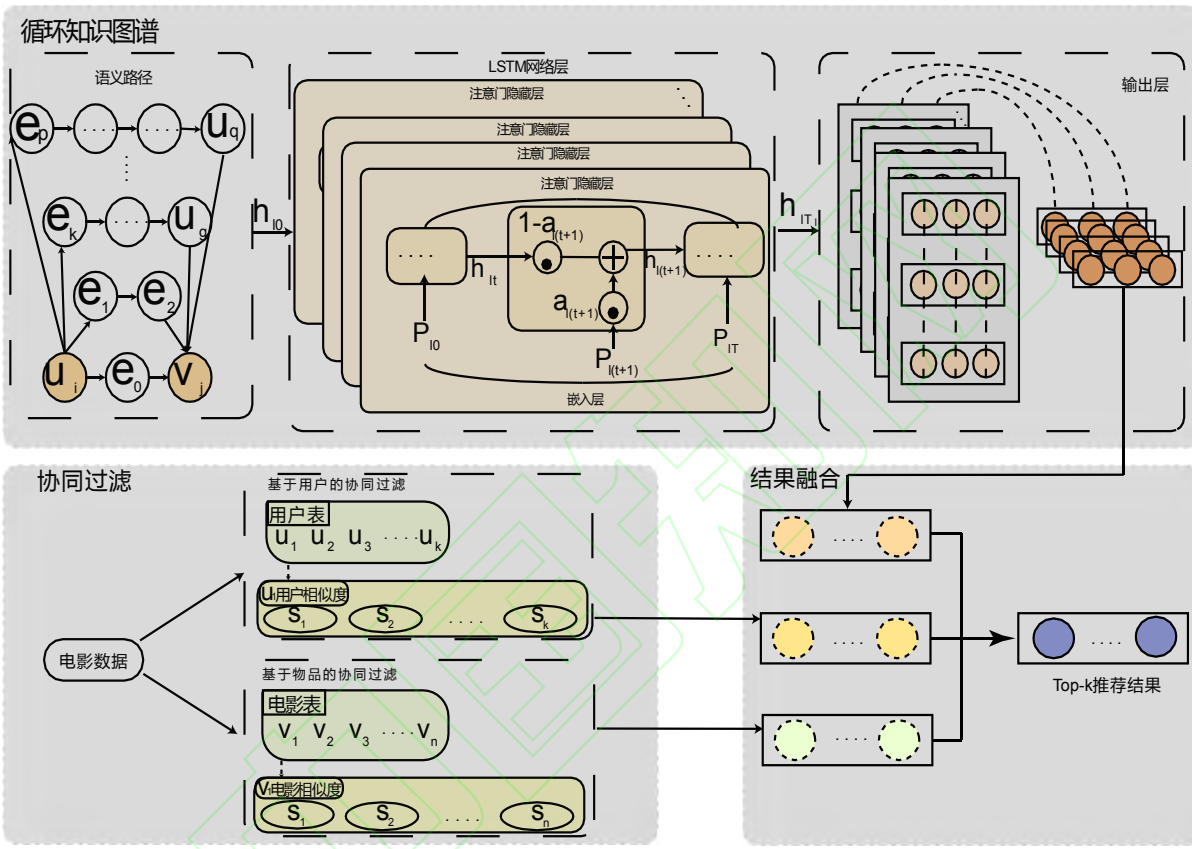

图 1　RCKG-CF 框架

3.1　循环知识图谱（RKGE）

在本节将具体介绍循环知识图谱的详细结构信息，包括语义路径、LSTM 网络层和输出层。同时本文使用真实数据集 Movielens 1M 和相应的 IMDB 数据集进行循环知识图谱的构建。

3.1.1　语义路径

知识图谱是一种特殊网络，其中每个节点代表现实世界中的实体，而节点间的边表示实体之间的关系。知识图谱一般用三元组形式表示内涵知识，每个三元组包括一个头实体、一个尾实体以及它们之间的关系，这是知识图谱的基本表示形式。

本文实验中使用电影相关的数据集，用户实体对应观看过的电影，电影实体包含演员、导演和电影类型等信息。

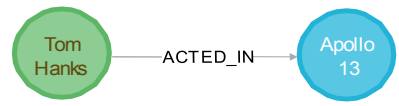

图 2　三元组

如图 2 所示，将电影特征放入知识图谱就可得到电影知识图谱三元组是基础的三元组，表示该导演指导了这部电影。蓝色表示为电影实体，绿色表示为人物实体，箭头描述的是人物与电影之间的关系，意味该人物参演或指导了该电影。将类似的多个三元组相互连接便形成知识图谱，如图 3。

在 RKGE-CF 模型中，包含一批 LSTM 结构，每个 LSTM 学习指定路径的语义表示。实体对 (u_i, v_j) 的路径长度是动态的，对于长度为 T 的任意路径p_l可表示为：

$$p_l = e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} e_2 \ldots \xrightarrow{r_T} e_T \qquad (1)$$

其中 $e_0 = u_i, e_T = v_j$。LSTM 通过学习每个实体的语义表示和整个路径的单个表示来对路径进行编码。为了充分利用知识图谱中的实体关系，我们首先挖掘出实体间具有不同语义的路径，然后将这些路径无缝地融合到循环网络批处理中进行有效推荐。为了提高模型的效率，我们用长度约束枚举的路径，即只使用长度小于阈值的路径。

因此可以根据知识图谱中内容，挖掘出实体间不同语义的路径。再抽取出的关系路径后，我们可以根据关键路径推断出用户的偏好关系，便于推荐，如下图 4。

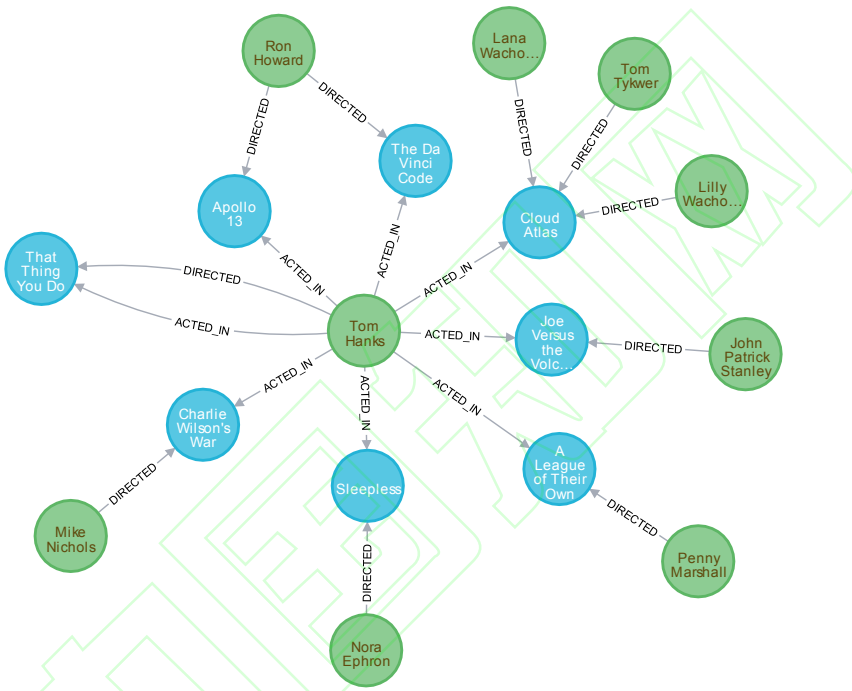

图 3　电影知识图谱

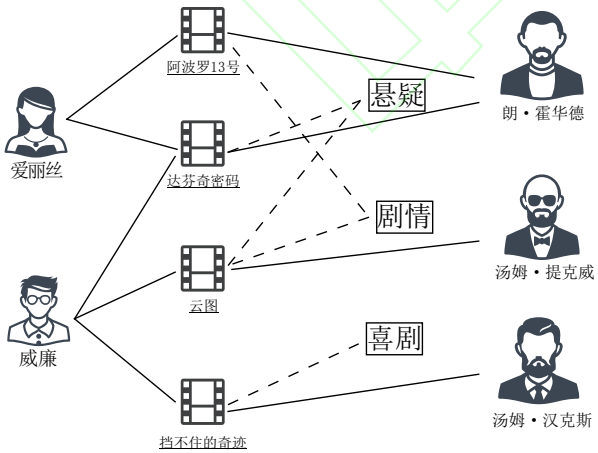

图 4　语义路径推理

以威廉和阿波罗 13 号的偏好关系为例，可得到以下路径：

1）威廉 $\xrightarrow{\text{打过评分}}$ 云图 $\xrightarrow{\text{电影类型}}$ 剧情 $\xrightarrow{\text{同一类型}}$ 阿波罗 13 号

2）威廉 $\xrightarrow{\text{打过评分}}$ 达芬奇密码 $\xrightarrow{\text{影片导演}}$ 朗·霍华德 $\xrightarrow{\text{同导演影片}}$ 阿波罗 13 号

以上捕捉到的路径 1）描述的是属于同一种电影类型之间的特征关系；路径 2）描述由同一位导演指导的其他电影。因此我们可以推断出威廉可能会喜欢的电影是《阿波罗 13 号》。所以我们基于以上的关系路径，结合相关偏好信息进行合理推测。但是，连接相同实体对于不同的路径通常具有不同的语义关系，意味着在描述用户的偏好和品位方面具有不同的重要性，某些路径可能比其他路径更能描述用户的偏好。为了充分利用知识图谱中的路径进行推荐，不仅需要捕获不同路径的语义，还需要捕获它们在描述用户对物品的偏好的显著性。然后将这些路径无缝地融合到 LSTM 批处理中进行有效推荐。

由于知识图谱的体积大、复杂度高，存在大量

连接实体对的路径，这些路径可能包含不同顺序、不同长度的不同实体类型和关系类型。为了提高模型的效率，我们只使用长度小于阈值的路径 ····。较短的路径表示了两个实体之间的近邻关系密切，如果使用较长的路径，两个实体之间就存在大量的噪声关系，并且它们之间的近邻关系就越微弱，遥远的两个实体会在一定程度上失去语义意义。

3.1.2 LSTM 网络层

在上部分中，我们将模型中用户-物品实体对看成序列，序列中的元素是路径中的实体，并使用由嵌入层和注意门隐藏层组成的 LSTM 对路径进行编码。该体系结构包含一批 LSTM， LSTM 通过学习每个实体的语义表示和整个路径的单个表示来对路径进行编码。

在嵌入层对 p_l 中的每个实体 e_t 学习一个分布式表示 p_{lt} ，该 p_{lt} 将 e_t 映射到一个低维向量并捕获该实体的语义，然后将此新表示作为输入提供给隐藏层，以学习编码整个路径的单个表示。注意门控隐藏层为了学习路径表示，考虑路径中实体的嵌入和这些实体的顺序，采用基于流的方法对路径的开始实体到结束实体的序列进行编码，最终得到整个路径的表示 h_{lT} 。

我们用 a_{lt} 表示步骤 t 处的注意门，它是 $[0,1]$ 之间的标量值。 t 时刻的隐藏状态可表示为：

$$h_{lt} = (1-a_{lt}) \cdot h_{l(t-1)} + a_{lt} \cdot h'_{lt} \quad (2)$$

其中注意门 a_{lt} 平衡了前一个隐藏状态 $h_{l(t-1)}$ 和当前候选隐藏状态 h'_{lt} 的输入贡献。通过充分考虑当前时间步长的输入，进一步给出了当前候选隐藏状态：

$$h'_{lt} = \sigma(\mathrm{W} \cdot \mathrm{h}_{\mathrm{l(t-1)}} + \mathrm{H} \cdot \mathrm{p}_{\mathrm{lt}} + \mathrm{b}) \quad (3)$$

其中 W ， H 分别是前一步和当前步的线性变换参数， b 是偏置项， σ 是 sigmoid 激活函数。最后，根据当前时间步长的输入观测值和相邻观测值在两个方向上的信息，建立了注意门的模型：

$$a_{lt} = \sigma(M^T \cdot (\overrightarrow{h_{lt}}; \overleftarrow{h_{lt}}) + b') \quad (4)$$

其中 σ 是 sigmoid 激活函数，用于将注意门的范围控制在 $[0,1]$ 之间； M 为权重向量， b' 为注意层的偏置项；“;”表示量之间的连接。 $\overrightarrow{h_{lt}}$ 总结从开始到步骤t的路径， $\overleftarrow{h_{lt}}$ 总结从结束到步骤t的路径，由下式给出：

$$\overrightarrow{h_{lt}} = \sigma(\overrightarrow{M} \cdot \overleftarrow{p_{lt}} + \overrightarrow{H} \cdot \overrightarrow{h_{l(t-1)}} + \overrightarrow{b}) \quad (5)$$

$$\overleftarrow{h_{lt}} = \sigma(\overleftarrow{M} \cdot \overleftarrow{p_{lt}} + \overleftarrow{H} \cdot \overleftarrow{h_{l(t-1)}} + \overleftarrow{b}) \quad (6)$$

通过将 u_i 和 v_j 之间的限定路径同时合并到相应的注意门控网络中，得到所有 u_i 和 v_j 的实体关系。由于 u_i 和 v_j 之间有多条路径连接，不同的路径在建模它们之间的关系时会有不同的影响程度。因此，我们通过池化操作来区分不同向量的最重要特性，max-pooling 层可形式化表示为：

$$h[j] = \max_{1 \le i \le S} h_{iTl}[j] \quad (7)$$

若 u_i 和 v_j 之间的路径为s条，其通过 LSTM 学习后，最后隐藏状态为 $h_{1T1}, h_{2T2}, h_{3T3}...., h_{STS}$ ，其中 TS 为最后一步。我们通过池化层获得所有路径上最显著的特性。然后采用全连接层，进一步量化 u_i 和 v_j 的关系（接近度）。完成模型训练后，通过根据接近度评分对物品进行排序，并向 u_i 推荐得分最高的前 K 个物品。

3.2 协同过滤（CF）

模型的下半部分，主要描述的是协同过滤算法的实现过程。我们同时使用了基于用户的协同过滤和基于物品的协同过滤算法。分别得到两个不同的 Top-k 推荐结果，在与之前循环知识图谱得到的结果融合，可以得到最终的 Top-k 结果。

首先把每个的用户或物品当作向量，然后计算其他所有的用户或物品与其他的相似度，有了两两之间相似度之后，系统也推荐给用户。循环知识图谱可以学习到推荐关系中的内涵知识，协同过滤可以很好地使用外部评分，我们提出的方法将内涵知识和外部评分进行组合，有效地提高推荐的效率。

基于用户的协同过滤中，根据用户的历史行为计算用户与其他用户之间的相似度时，计算如下式：

$$W_{uv} = \frac{\sum_{i \in N(u) \cap N(V)} \frac{1}{\lg 1 + |N(i)|}}{\sqrt{|N(u)||N(v)|}} \quad (8)$$

其中， $N(u)$ ， $N(v)$ 分别表示用户 u ， v 过

正反馈的电影集合。如果用户对于冷门的物品采取过相似的行为，更能表达出两者之间的相似度。所以特别加入惩罚因子 $\frac{1}{\lg 1+|N(i)|}$，以此来惩罚用户之间共同电影列表中热门电影相似度的影响。

基于物品的协同过滤中，计算物品与物品之间的相似度时，通过下式：

$$W_{ij}=\frac{\sum_{i\in N(i)\cap N(j)}\frac{1}{\lg 1+|N(u)|}}{\sqrt{|N(i)||N(j)|}} \qquad (9)$$

其中，$N(i)$，$N(j)$分别表示喜欢电影 i 和喜欢电影 j 的用户数。因为活跃用户对物品相似度计算的贡献会小于不活跃的用户，所以也加入惩罚因子 $\frac{1}{\lg 1+|N(u)|}$，以此降低影响。

4 循环知识图谱和协同过滤融合算法

4.1 融合流程

由上一节可知，循环知识图嵌入采用了一种新的递归网络架构，该架构包含一批递归网络，用于对链接相同实体对的路径的语义建模，这些路径无缝地融合到推荐中，并选择一条好的推荐路径进行 Top-K 推荐。同时，利用协同过滤收集用户行为以获得其对物品的显式或隐式信息，以基于物品协同过滤和基于用户协同过滤分别给出 Top-K 推荐。模型最后将三种方法给出的 Top-K 推荐结果进行融合，得到最终的推荐列表。图 5 描述了循环知识图谱与协同过滤融合的流程，将两部分的推荐结果融合，能更好的提高推荐效率。

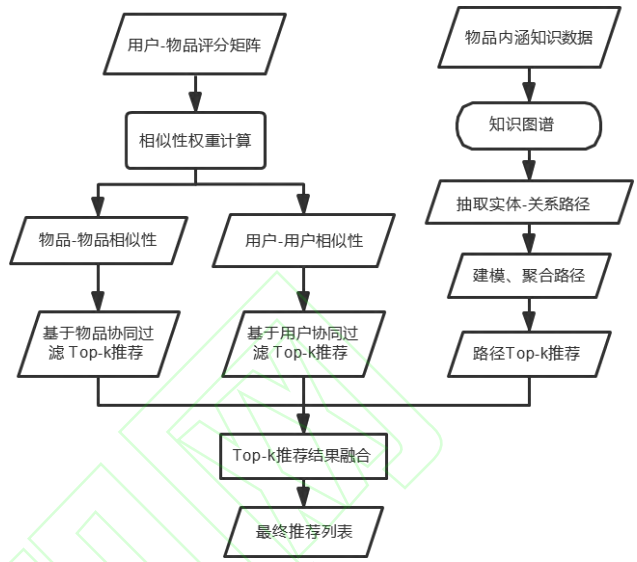

图 5 知识图谱与协同过滤融合

4.2 融合算法

为了实现循环知识图谱对协同过滤算法的支撑，本文针对实体内涵知识和外部评分的情况，提出了两种不同的结果融合算法。根据循环知识图谱得出的推荐列表和协同过滤算法得到的推荐列表，我们经过融合算法可以抽取出相同或排名靠前的结果进行融合，得到新的推荐结果。由此得到的推荐结果，一方面可以提高推荐的有效性，另一方面可以解释每个推荐结果的来源，使得协同过滤和循环知识图谱结果得以相互弥补。

在下列两个算法中$\mathbb{L}$和$\mathbb{E}$是基于用户和基于物品的协同过滤要推荐给用户的物品集合，对于集合当中的每一个对象$\{L_0,...,L_n\}$和$\{E_0,...,E_n\}$按照预测评分进行排序，也就是说$\mathbb{L}$和$\mathbb{E}$是两个有序数列。内涵知识近邻集$\mathbb{T}$也根据预测评分排序得到一个有序数列。

算法 1（图 6）概述了内涵知识与外部评分的第一种融合方式，我们将该融合方式记为循环抽取融合(Loop Extraction Fusion，LEF)。基于用户的协同过滤、基于物品的协同过滤和基于循环知识图谱生成的物品集合通过遍历，依次将三个集合中的物品放入 Top-K 推荐集合$\mathbb{T}$中，在放入推荐集合$\mathbb{T}$的过程中，要保证放入的对象不存在于$\mathbb{T}$中，也就是要保证推荐集合$\mathbb{T}$中对象的唯一性。

算法 1： 融合算法 LEF

输入： 基于用户的协同过滤近邻集:Set $\mathbb{L} = \{L_0, \dots, L_n\}$;
基于物品的协同过滤近邻集:Set $\mathbb{E} = \{E_0, \dots, E_n\}$;
内涵知识近邻集:Set $\mathbb{T} = \{T_0, \dots, T_n\}$。
输出： Top-K 推荐集 $\mathbb{C} = \{C_0, \dots, C_k\}$.

```
for i (0 ≤ i ≤ n) do
    If L_i ∉ ℂ:
        ℂ.append (L_i);
        If Len(ℂ) == k: break;
    If E_i ∉ ℂ:
        ℂ.append(E_i);
        If Len(ℂ) == k: break;
    If T_i ∉ ℂ:
        ℂ.append(T_i);
        If Len(ℂ) == k: break;
end do
输出 Top-k 推荐集Set ℂ
```

图 6　融合算法 1

算法 2（图 7）描述了内涵知识与外部评分的第二种融合方式，我们将该融合方式记为循环比较融合(Loop Comparison Fusion，LCF)。基于用户的协同过滤、基于物品的协同过滤和基于循环知识图谱生成的物品集合通过遍历，分别判定每个集合中当前对象是否存在于另外两个集合，若存在，则将当前集合的当前对象放入 Top-K 推荐集合 $\mathbb{T}$ 中，在放入推荐集合 $\mathbb{T}$ 的过程中，也要保证放入的对象不存在于 $\mathbb{T}$ 中，保证推荐集合$\mathbb{T}$中对象的唯一性。

算法 2： 融合算法 LCF

输入： 基于用户的协同过滤近邻集: $\mathbb{L} = \{L_0, \dots, L_n\}$;
基于物品的协同过滤近邻集: $\mathbb{E} = \{E_0, \dots, E_n\}$;
内涵知识近邻集: $\mathbb{T} = \{T_0, \dots, T_n\}$.
输出： Top-K 推荐集 $\mathbb{C} = \{C_0, \dots, C_k\}$.

```
for i (0 ≤ i ≤ n) do
    If L_i ∉ ℂ && (L_i in 𝔼 || L_i in 𝕋):
        ℂ.append(L_i);
        If Len(ℂ) == k: break
    If E_i ∉ ℂ && (E_i in 𝕃 || E_i in 𝕋)
        ℂ.append(E_i);
        If Len(ℂ) == k: break;
    If T_i ∉ ℂ && (T_i in 𝕃 || T_i in 𝔼):
        ℂ.append(E_i);
    If Len(ℂ) == k: break;
end do;
输出 Top-k 推荐集Set ℂ
```

图 7　融合算法 2

5　实验及结果分析

5.1　实验设置

5.1.1　数据集

为测试模型的有效性，我们利用了真实的数据集 MovieLens 中的 IM-1M 来进行验证。该数据集在 Movielens 1M 和相应的 IMDB 数据集的基础上进行构建的,数据集详细信息如表 1。在前期的循环知识图谱构建和后期的测试中，我们都使用了该数据集进行实验。其中 Movielens 1M 包含电影元数据信息和用户属性信息，也包括多个用户对多部电影的评分数据，每个用户至少有 20 个评分记录。将 Movielens 1M 数据集与 IMDB 数据集映射链接，得到我们实验数据。(数据集下载地址分别为：http://groplens.org/datatsets/movielens/http://www.imdb.com/)。

表 1　数据集信息

数据集		IM-1M
用户–物品数据	#用户	6040
	#物品	3382
	#评分	756684
	数据密度	3.704%
知识图谱	#实体	18920
	#实体类型	11
	#关系	800261
	#关系类型	10
	图密度	0.447%

在实验数据集中在实验数据集中每个用户对应多个已观看的电影，并且将用户对电影评分小于等于阈值 r 的数据作为负反馈，再对模型进行训练。$r = \{0,1,2,3,4\}$ 以获得不同的外部评分对于内涵知识的影响。其中，0 表示用户未对该电影有过评分行为。

5.1.2　实验环境

实验在 GPU 服务器上运行，详细信息见表 2。

表 2　实验环境

OS	Ubuntu 16.04
Memory	128GB
CPU	64 Intel(R) Xeon(R) Gold 6130 CPU @
GPU	2.10GHz
Number of	GTX1080(11G)
GPU-cards	4
CUDA version	9.0

cuDNN version	7.3.1
python version	3.6.5
pytorch version	1.1.0

5.2 评价指标

在模型的损失函数部分，我们将推荐问题作为二分类问题进行处理，所以在给定的训练集中将通过优化以下参数进行学习，公式如下：

$$L = \frac{1}{|D|}\sum_{r_{ij}\in D} BCELoss(\tilde{r}_{ij}, r_{ij}) \quad (10)$$

其中，D 表示训练集，$\tilde{r}_{ij}, r_{ij}$ 分别表示观测到的分数和估计得到的分数，$BCELoss$ 表示二元交叉熵损失函数（binary cross entropy）。

根据公式和训练模型，我们可以很容易地进行端到端训练。在递归层中采用时间反向传播算法对参数进行更新，在其他部分采用普通的反向传播对参数进行更新。我们为每个用户随机抽取未评分的物品作为负面实例，其数量与他的评级物品相同。连接用户及其负面实例的路径也被用来帮助平衡模型学习。

在推荐系统的评价指标中，我们使用准确率（Precision）和 MRR（Mean Reciprocal Rank）来评价模型的推荐能力。

准确率描述的是推荐系统中给出的最终推荐列表中有多少比例的用户是发生过的用户-物品评分记录,准确率（Precision）公式如下：

$$Precision = \frac{\sum_{u\in U}|R(u)\cap T(u)|}{\sum_{u\in U}|R(u)|} \quad (11)$$

其中，$R(u)$ 是根据用户在训练集上的行为给用户做出的推荐列表，$T(u)$ 是用户在训练集上。同时 $Precision@K(K=1,5,10,15)$ 表示的是评测推荐系统的准确率，并且选取不同的推荐列表长度K，计算出多组准确率，以便对比。

MRR 是平均倒数排名，表示最终推荐列表在被评价系统给出结果中的排序取倒数作为准确度，再对所有的数据取平均，公式如下：

$$MRR = \frac{1}{m}\sum_{i=1}^{m}\left(\sum_{v_j\in test(u_i)}\frac{1}{rank(u_i, v_j)}\right) \quad (12)$$

其中，m 表示用户个数，v_j 是在最终的推荐列表中正确的推荐物品，$test(u_i)$ 是 u_i 的测试数据集中物品集合，$rank(u_i, v_j)$ 是 u_i 的推荐列表中 v_j 的位置。推荐列表中第一个在推荐列表结果中物品所在的排列位置。本文实验中计算的是 $K=10$ 时的 MRR 数值，进行对比。

5.3 实验对比

为了验证是经验结果的有效性，我们将和 9 种的算法在上述数据集种进行实验对比，包括最新的协同过滤与知识图谱相结合的算法 CKE 和 RKGE，以证明该模型具有良好的性能。分别介绍如下：

MostPop:向所有用户推荐热门物品，但不属于个性化推荐算法。

BPRMF:基于矩阵因子分解的贝叶斯后验优化的个性化得分排序算法,本身不优化用户对物品的评分，只是借评分来优化用户对物品的排序。

NCF:神经协同过滤算法，是一种基于神经网络的推荐方法。主要用于解决在含有隐式反馈的基础上进行推荐的协同过滤问题。

LIBFM:基于潜在特征因子的一种经典的矩阵分解模型，其中将图谱中的物品属性当作原始特征放入该模型。

HeteRs:提出了一种基于图的推荐方法，其中利用马尔可夫链整合知识图谱。

HeteRec:使用潜在因子模型混合元路径的个性化推荐方法。

GraphLF:基于图形的方法个性化的 Pagerank 方法，再通过逻辑推理来发现用户偏好。

CKE:最近提出了一种基于协同过滤结合知识图谱嵌入的方法，在知识图谱的帮助下更好的学习物品的潜在信息。

RKGE:利用知识图谱嵌入和一组递归网络结构，自动学习实体之间的路径及语义关系，从而更好的描述物品对用户的偏好信息。

5.4 实验结果分析

我们分别使用了融合方法 1(LEF)和融合方法 2(LCF)在 MovieLens 数据集上进行了 Top-1、Top-5、Top-10、Top-15 和 MRR 的推荐，曲线展示了在 RKGE 的基础上面分别加入 userCF 和 itemCF、同时加入 userCF 和 itemCF(下列文中统一称为为 CF)，以及在同时加入 userCF 和 itemCF 基础上对 RKGE 加入不同评分的变化情况。在下面所有的数据 $RKGE(r0, r1, r2, r3, r4)$ 代表了不同的外部评分

对于内涵知识的影响，其中 $r0$ 表示未添加外部评分。例如，$RKGE(r2)$ 表示对电影评分小于等于 2 的数据作为负反馈，再对模型进行训练。

从表 3 中可以看出 LEF 的推荐性能优秀，当 Top-K 推荐的 K 值比较大时，模型也能够保持相对较好的性能，同时加入 CF 的推荐结果会明显优于单独加入 userCF 或 itemCF；在选择了评分小于 1 作负反馈时，综合推荐结果略优于其他分数，不同 Precision 下的比较曲线见图 8。

表 3　融合算法 1(LEF)精确度数据

	precision@1	precision@5	precision@10	precision@15
RKGE+userCF	0.1622	0.1533	0.1386	0.1272
RKGE+itemCF	0.1622	0.1565	0.1386	0.1262
RKGE(r4)+CF	0.1760	0.1637	0.1470	0.1331
RKGE(r3)+CF	0.1782	0.1644	0.1469	0.1332
RKGE(r2)+CF	0.1803	0.1637	0.1471	0.1333
RKGE(r1)+CF	0.1930	0.1637	0.1492	0.1344
RKGE(r0)+CF	0.1898	0.1654	0.1484	0.1330

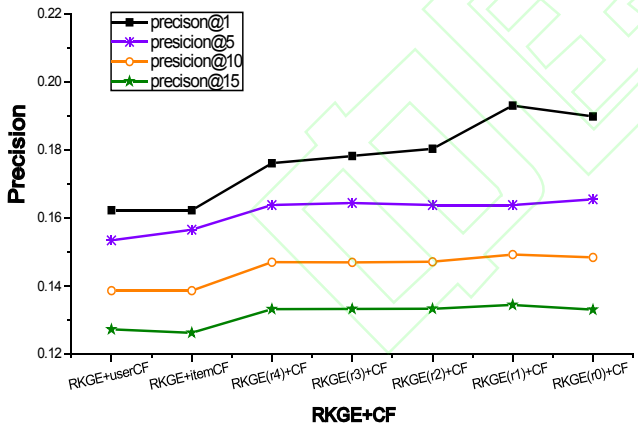

图 8　融合算法 1(LEF)精确度对比

从表 4 中可以看出 LCF 整体的性能都要优于 LEF，也可以从图 9 中看出，随着推荐个数的增加性能不会再有所增加；同时在选择了评分小于 1 作负反馈时，综合推荐结果略优于其他分数。

表 4　融合算法 2(LCF)精确度数据

	precision@1	precision@5	precision@10	precision@15
RKGE+userCF	0.2057	0.1484	0.1321	0.1281
RKGE+itemCF	0.2057	0.1543	0.1333	0.1326
RKGE(r4)+CF	0.2134	0.1904	0.1703	0.1703
RKGE(r3)+CF	0.2068	0.1905	0.1718	0.1715
RKGE(r2)+CF	0.2068	0.1892	0.1725	0.1718
RKGE(r1)+CF	0.2089	0.1881	0.1729	0.1722
RKGE(r0)+CF	0.2057	0.1885	0.1730	0.1718

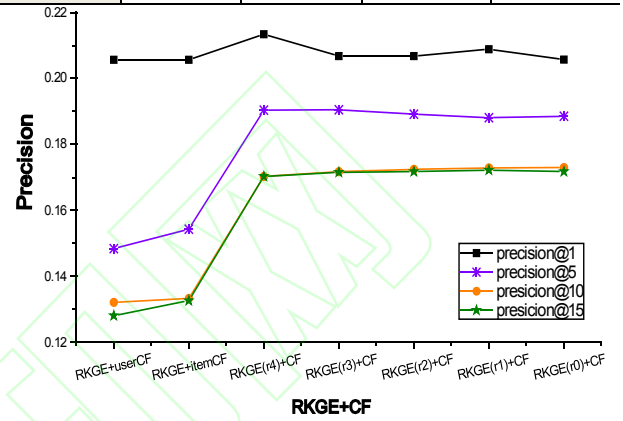

图 9　融合算法 2(LCF)精确度对比

表 5 对比了 LEF 和 LCF 两种融合方法在不同维度下的 MRR 值。图 10 展示了 LEF 和 LCF 在 MRR 上的对比曲线。只加入 userCF 或 itemCF 时 LCF 优于 LEF。但是同时加入 CF 后，LEF 会明显优于 LCF，这也说明了 EF 更适用于多推荐列表的融合。两种融合方法，都是评分 1 以下作为负反馈时效果最好。

表 5　两种融合方法 MRR 数据

	LEF	LCF
RKGE+userCF	0.3533	0.4129
RKGE+itemCF	0.3415	0.4056
RKGE(r4)+CF	0.4748	0.4365
RKGE(r3)+CF	0.4856	0.4349
RKGE(r2)+CF	0.4930	0.4352
RKGE(r1)+CF	0.4944	0.4375
RKGE(r0)+CF	0.4925	0.4299

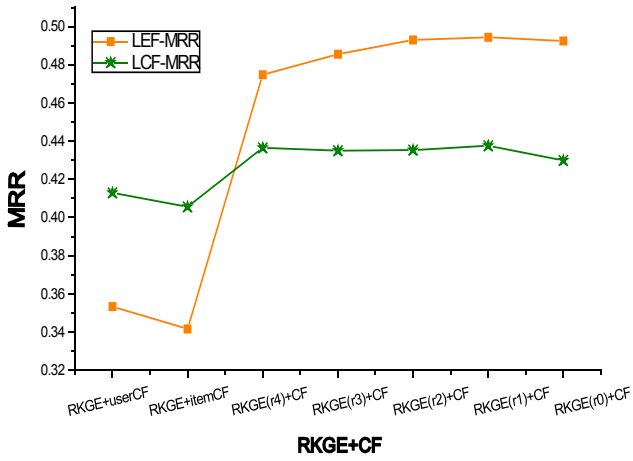

图 10　两种融合算法 MRR 对比

最后表 6 将本文提出的模型与 MostPop、BPRMF、LIBMF、NCF、HeteRS、HeteRec、GraphLF、CKE、RKGE 进行了对比实验，在图 11 在不同的 Top-1、Top-5 和 Top-10 上都可以看出本文提出的模型远远优于其他模型，并且当 K 值较大时也能保持优秀的推荐性能。

表 6　十种方法精确度对比

	precision@1	precision@5	precision@10
MostPop	0.0118	0.0064	0.0081
BPRMF	0.0409	0.0438	0.0441
LIBMF	0.0459	0.0525	0.0456
NCF	0.045	0.0482	0.0485
HeteRS	0.0689	0.0538	0.0475
HeteRec	0.0764	0.0579	0.0488
GraphLF	0.1069	0.036	0.0581
CKE	0.0954	0.0781	0.0682
RKGE	0.1396	0.1092	0.0861
RKGE-CF	**0.2089**	**0.1881**	**0.1729**

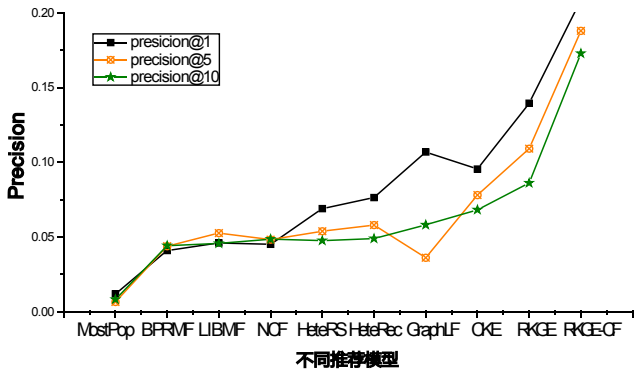

图 11　十种方法精确度对比

6　结束语

本文提出了一种基于循环知识图谱嵌入的混合推荐模型改模型既可以通过协同过滤发现用户的现有兴趣，也可以通过知识图谱挖掘用户的潜在兴趣，将两种结果融合，得到个性化的推荐结果。模型将循环神经网路、知识图谱和协同过滤相结合，模型可以自动学习实体之间的路径关系，推断出偏好关系。同时在知识图谱中加入外部评分，作为学习权重，更好的表达用户的偏好程度。最后利用的不同的融合方法将内涵知识与外部评分结果融合，得到最优的融合推荐结果。结果表明，本文所提出的框架在推荐的准确性、MRR 对比现有的模型取得了更好的效果。该模型也有一些待优化的部分，例如在融合方法能不能更好的调节比例，这也为接下来的工作提出了新的思路。

此外，本文所提出的方法，同样适合于音乐、图书等推荐场景。但是不同的产品领域相对于电影推荐会存在着评分刻度差异、领域之间相关性不同、情感差异等问题。为了实现迁移学习，可以尝试在其他场景中，提取用户和物品的标签，通过添加神经网络对用户、物品和评分等内容特征进行学习，得到用户内涵知识或特征等信息，然后再迁移到目标任务中。

参考文献：

[1] 杨武, 唐瑞, 卢玲. 基于内容的推荐与协同过滤融合的新闻推荐方法[J]. 计算机应用, 2016, 36(2):414-418.

[2] Sarwar B , Karypis G , Konstan J , et al. Item-based collaborative filtering recommendation algorithms[C]//ACM Press the tenth international conference - Hong Kong, Hong Kong (2001.05.01-2001.05.05)Proceedings of the tenth international conference on World Wide Web, - WWW \"01 - 2001:285-295.

[3] Kenneth K. Fletcher, Xiaoqing Frank Liu. A Collaborative Filtering Method for Personalized Preference-Based Service Recommendation[J]. ICWS 2015: 400-407.

[4] Antonio Hernando, Jesús Bobadilla, Fernando Ortega. A non-negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model[J]. Knowl. -Based Syst, 2016, 97: 188-202.

[5] Liu, Haifeng, et al. Context-based collaborative filtering for citation recommendation[J]. IEEE Access 3 (2015): 1695-1703.

[6] 江周峰, 杨俊, 鄂海红. 结合社会化标签的基于内容的推荐算法[J]. 软件, 2015(1):1-5.

[7] Shu J, Shen X, Hai L, et al. A content-based recommendation algorithm for learning resources[J]. Multimedia Systems, 2017(1):1-11.

[8] Chu W T, Tsai Y L. A hybrid recommendation system considering visual information for predicting favorite restaurants[J]. World Wide Web, 2017.

[9] Subramaniam, Rajan, Roger Lee, and Tokuro Matsuo. Movie Master: Hybrid Movie Recommendation[C]//2017 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2017.

[10] Rumelhart D E. Learning Representations by Back-Propagating Errors[J]. Nature, 1986, 23.

[11] Smolensky P. Information processing in dynamical systems: Foundations of harmony theory[M]//Parallel distributed processing: explorations in the microstructure of cognition, vol. 1. 1986.

[12] Hinton G E. Deep belief networks[J]. Scholarpedia, 2009, 4(6): 5947.

[13] MOHAMED, AbdelRahman, DAHL, et al. Acoustic Modeling Using Deep Belief Networks[J]. IEEE Transactions on Audio Speech & Language Processing, 2011, 20(1): 14-22.

[14] Kim D, Park C, Oh J, et al. Convolutional Matrix Factorization for Document Context-Aware Recommendation[C]//Acm Conference on Recommender Systems. ACM, 2016.

[15] Zhang, Shuai, et al. Deep learning-based recommender system: A survey and new perspectives[J]. ACM Computing Surveys (CSUR) 52.1 (2019): 5.

[16] Y. Li, T. Liu, J. Jiang, and L. Zhang. Hashtag Recommendation with Topical Attention-Based LSTM[C]//COLING, page 3019-3029. ACL, (2016)

[17] Zhou, Yuwen, et al. Personalized learning full-path recommendation model based on LSTM neural networks[J]. Information Sciences, 2018, 444: 135-152.

[18] Liu, Juntao, Caihua Wu, and Junwei Wang. Gated recurrent units based neural network for time heterogeneous feedback recommendation[J]. Information Sciences 423 (2018): 50-65.

[19] Rendle, Steffen. Factorization machines with libfm[C]// ACM Transactions on Intelligent Systems and Technology 3(3) ·May 2012:57.

[20] Davidson, James, et al. The YouTube video recommendation system[C]//ACM, Proceedings of the 2010 ACM Conference on Recommender Systems, RecSys 2010, Barcelona, Spain, 2010 DOI: 10.1145/1864708.186477.

[21] 吴玺煜, 陈启买, 刘海, 等. 基于知识图谱表示学习的协同过滤推荐算法[J]. 计算机工程, 2018.

[22] Zhang F, Yuan N J, Lian D, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]//the 22nd ACM SIGKDD International Conference. ACM, 2016.