

## 基于知识图谱表示学习的协同过滤推荐算法

吴玺煜<sup>1</sup> 陈启买<sup>1</sup> 刘 海<sup>1</sup> 贺超波<sup>2</sup>

(1. 华南师范大学 计算机学院 广州 510631; 2. 仲恺农业工程学院 信息科学与技术学院 广州 510225)

**摘 要:** 针对协同过滤算法仅使用物品-用户评分矩阵而未考虑语义的问题,提出一种协同过滤推荐算法。使用知识图谱表示学习方法,将业界已有的语义数据嵌入一个低维的语义空间中。通过计算物品之间的语义相似性,将物品自身的语义信息融入协同过滤推荐。算法弥补了协同过滤算法没有考虑物品本身内涵知识的缺陷,在语义层面上增强了协同过滤推荐的效果。实验结果表明,该算法能够有效地提升协同过滤推荐的准确率、召回率和F值。  
**关键词:** 协同过滤; 知识图谱; 表示学习; 语义相似性; 推荐系统

中文引用格式: 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, 44(2): 226-232. 263.

英文引用格式: WU Xiyu, CHEN Qimai, LIU Hai, et al. Collaborative Filtering Recommendation Algorithm Based on Representation Learning of Knowledge Graph[J]. Computer Engineering, 2018, 44(2): 226-232. 263.

## Collaborative Filtering Recommendation Algorithm Based on Representation Learning of Knowledge Graph

WU Xiyu<sup>1</sup>, CHEN Qimai<sup>1</sup>, LIU Hai<sup>1</sup>, HE Chaobo<sup>2</sup>

(1. School of Computer, South China Normal University, Guangzhou 510631, China;

2. School of Information Science and Technology, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)

**【Abstract】** To solve the problem that collaborative filtering algorithm only uses the items-users rating matrix and does not consider semantic, a collaborative filtering recommendation algorithm is presented. Using the knowledge map to represent the learning method, this method embeds the existing semantic data into a low-dimensional semantic space. It integrates the semantic information of items into the collaborative filtering recommendation by calculating the semantic similarity between items. The shortcoming of collaborative filtering algorithm which does not consider the semantic information of items is overcome, and therefore the effect of collaborative filtering recommendation is improved on the semantic level. Experimental results show that the proposed algorithm can get higher values on precision, recall and F-measure for collaborative filtering recommendation.

**【Key words】** collaborative filtering; knowledge graph; representation learning; semantic similarity; recommendation system  
DOI: 10.3969/j.issn.1000-3428.2018.02.039

### 0 概述

协同过滤是至今应用最广泛的推荐技术,已经成功地应用于电子商务<sup>[1]</sup>、网上学习<sup>[2]</sup>和新闻媒体<sup>[3]</sup>等工业界领域,在学术界内也一直受到大量学者关注。协同过滤推荐主要根据用户对资源的历史行为信息,寻找用户或者资源的近邻集合,以此来计算用户对资源的偏好值<sup>[4]</sup>。文献[5]将Facebook用户的个性数据融合进协同过滤中,用合并多领域数据集的方式提高了推荐的准确性。文献[6]从评分上下文信息着手,提出了一种融合奇异性和扩散过

程的协同过滤模型,解决了信息过载问题。文献[7]融合了基于内容与协同过滤的推荐算法,提高了新闻推荐的性能。现有各种协同过滤算法都使用了用户隐性反馈数据(比如购买记录)、用户显性反馈数据(比如评分)等信息,准确地将用户所需要的物品推荐给对方。然而这些算法主要利用物品-用户评分矩阵这一外在信息,尚未充分考虑物品自身的内在信息。随着知识图谱技术的发展,业界已经积累了大量开放的语义数据,如Freebase<sup>[8]</sup>、DBpedia<sup>[9]</sup>,包含了物品丰富的内涵知识。将内涵知识和外在评分融合在一起,可以使得传统协同过滤推荐算法的

基金项目: 广东省自然科学基金(2016A030313441); 广东省科技计划项目(2015B010129009, 2016A030303058, 2016A090922008, 2015A020209178); 广东省高性能计算重点实验室开放课题(T191527)。

作者简介: 吴玺煜(1993—),男,硕士研究生,主研方向为推荐系统、本体理论; 陈启买,教授; 刘 海(通信作者),贺超波,副教授、博士。

收稿日期: 2016-12-30 修回日期: 2017-02-21 E-mail: wuxiyu@m.scnu.edu.cn

信息更加完备。文献[10]尝试使用了知识图谱的结构特征, 将本体融进协同过滤算法。现有研究表明, 知识图谱表示学习方法能将知识图谱嵌入到一个低维语义空间中, 可以利用连续数值的向量反映知识图谱的结构特征。这种方法可以高效地计算实体间的语义联系。文献[11]尝试将知识图谱表示学习算法与基于隐性反馈的协同过滤相结合, 把原始数据转化为偏好序列进行参数学习, 加强了协同过滤推荐算法的性能。

围绕上述背景, 本文提出基于知识图谱表示学习的协同过滤推荐算法。算法的出发点与文献[11]相近, 后者使用用户隐性反馈数据, 而本文算法是针对用户显性反馈数据, 即物品-用户评分矩阵。算法通过知识图谱表示学习算法 TransE<sup>[12]</sup> 获取物品的语义信息, 计算物品之间的语义相似性, 在推荐过程中融入物品的语义信息, 更好地为用户进行推荐。

## 1 相关理论

### 1.1 基于物品的协同过滤推荐

基于物品的协同过滤推荐<sup>[13]</sup>是一种基于最近邻的推荐系统算法, 主要基于如下假设: 用户倾向于喜欢相似的物品。在推荐系统中, 算法通过计算物品之间的相似性, 从而向用户推荐最合适的物品。整个推荐过程形式化表示如下: 假设有  $m$  个用户  $U = (U_1, U_2, \dots, U_m)$ ,  $n$  个物品  $I = (I_1, I_2, \dots, I_n)$ , 输入的数据集是一个  $m \times n$  的物品-用户评分矩阵  $R_{m \times n}$ :

$$R_{m \times n} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1j} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2j} & \cdots & R_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ R_{i1} & R_{i2} & \cdots & R_{ij} & \cdots & R_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ R_{m1} & R_{m2} & \cdots & R_{mj} & \cdots & R_{mn} \end{bmatrix} \quad (1)$$

其中, 分数  $R_{ij}$  为用户  $U_i$  对物品  $I_j$  的评分, 且该评分代表用户  $i$  对物品  $j$  的喜好程度。矩阵元素之间的相似性度量主要采用余弦相似性实现, 即通过计算两个向量之间夹角的余弦值, 度量它们之间的相似性。假设  $A$  和  $B$  为 2 个向量,  $n$  是它们的维度, 则它们的余弦相似性表示为:

$$\begin{aligned} \text{sim}_{\cos}(A, B) &= \cos(\theta) = \frac{A \cdot B}{\|A\| \times \|B\|} = \\ &= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i^2)} \times \sqrt{\sum_{i=1}^n (B_i^2)}} \end{aligned} \quad (2)$$

该式计算结果的值越大, 向量就越相似。当  $\text{sim}_{\cos}(A, B)$  值为 0 表示两个向量完全不相似, 1 则是完全相似。

此外, 针对物品-用户评分矩阵中评分相当稀疏的问题, 文献[14]提出加入相似性权重重要性的策

略, 用于避免数值过分偏倚的问题。该方法采用了贝叶斯学派里收缩的思想, 当只有少量评分用于相似性计算时, 采取策略降低两者之间的相似性权重。因此, 改进后的物品间相似性权重计算公式为:

$$w'_{ij} = \frac{|U_{ij}|}{|U_{ij}| + \beta} \times w_{ij} \quad (3)$$

其中,  $w'_{ij}$  是对原始权重进行收缩后的相似性权重, 对应于原始权重  $w_{ij}$ 。  $|U_{ij}|$  表示同时有  $|U_{ij}|$  个人对于物品  $i$  和物品  $j$  进行评分。  $\beta$  是一个收缩系数, 当  $|U_{ij}|$  很小时,  $\beta$  起到一种收缩的作用。当  $|U_{ij}| > \beta$  时, 则影响不大。

利用式(3), 基于物品的协同过滤推荐算法将物品-用户评分矩阵转换成一个物品-物品相似性矩阵, 并对这个相似性矩阵进行近邻选择。近邻选择是对  $U$  所挑选过的物品分别进行排序筛选, 分为以下 2 个步骤: 1) 对于线下已经计算好的物品-物品相似性矩阵, 进行预过滤近邻。2) 对某一个用户, 读取内存中的数据, 使用 Top- $k$  进行选择推荐。推荐系统中用户和物品往往数量巨大, 因此需要预先过滤物品间相似性度量的数量。本文采用了 2 种预过滤技术: Top- $N$  过滤和阈值过滤。对于预测的评分, 本文也使用了 Top- $k$  选择和正值选择, 使推荐效果提高。

### 1.2 知识图谱表示学习

文献[15]提出的 Word2Vec 模型, 可以将词汇嵌入到一个  $K$  维空间。这种分布式表示掀起了一股表示学习的热潮。表示学习这一古老而又新颖的话题, 又一次焕发出新的光彩。该算法的核心思想是利用类比性的优点, 更好地发现词与词之间的相关性。文献[12]将该思想应用于知识图谱领域, 提出了一种知识图谱表示学习算法 TransE。该算法的核心思想是将知识图谱中的实体与关系嵌入到一个低维的向量空间里面去, 同时将两者转化为向量表示。因此, 从直观上看, 相近的实体在空间里的向量表示也是相近的。更进一步, 在语义上有联系的实体, 在空间里也有所关联。

基于上述思想, 能够通过知识图谱表示学习, 对知识图谱中实体和关系进行语义表示学习, 通过将富有语义信息的知识图谱三元组嵌入  $n$  维语义空间并生成对应向量, 从而实现了知识图谱的数值化。通过知识图谱表示学习, 可以快速计算两个实体间的语义相似性, 并方便地将知识图谱用到其他学习任务中。知识图谱表示学习的实现主要有 2 种方法: 1) 基于张量分解的方法; 2) 翻译的方法。前者包括了 NTN<sup>[16]</sup>、RESCAL<sup>[17]</sup> 等。后者包括了 TransE、TransH<sup>[18]</sup>、TransR/CTransR<sup>[19]</sup> 等。对于 Freebase 这类关系数目众多而又非常稀疏的大规模知识库, 基于张量分解的方法效果不佳, 选择翻译方法<sup>[20]</sup>更为合适。

## 2 基于知识图谱表示学习的协同过滤推荐算法

本文提出了一种基于知识图谱表示学习的协同过滤推荐算法。其基本思想是:对于协同过滤计算出来的最近邻,系统推荐给用户;而对于该用户喜欢的物品在语义上相似的物品,系统也可以推荐给用户。相对于协同过滤推荐算法仅使用外部评分,加入内涵知识(知识图谱)会得到更好的效果,提高推荐的有效性。

利用知识图谱表示学习算法,将推荐的物品嵌入到一个低维空间,然后计算物品之间的语义相似性,生成语义相似性矩阵,最终可以得到物品的语义近邻。同时,通过调节融合比例,对语义近邻和协同过滤潜在物品按比例融合,利用丰富的语义数据一定程度上解决了推荐系统的冷启动问题。图1给出了本文算法(记为 TransE-CF)的流程图。

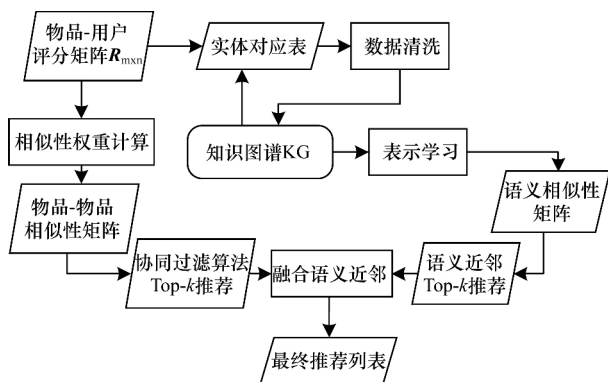


图1 TransE-CF 算法流程

### 2.1 融合语义近邻集合的协同推荐过程

#### 2.1.1 基于 TransE 算法的知识图谱表示学习

对于一个典型的知识图谱来说,可以通过有向图表示的三元组,以及三元组之间的相互链接构成一个网状的知识集合,这种三元组携带着实体自身的语义信息。其中实体作为节点,实体之间的关系作为边。

以电影领域为例,电影实体中主要包括了演员、类型、导演等主要特征。这些特征从一定程度上概括了这部电影。利用电影特征,可以得到类似图2所示的一个电影知识图谱的三元组。

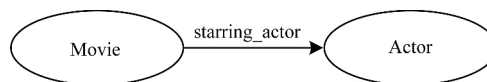


图2 知识图谱三元组

在图2中,电影实体 Movie 和演员实体 Actor 之间通过关联,构成三元组( $M, \text{starring\_actor}, A$ ),再通过三元组之间的相互链接形成知识图谱。图3为 Freebase 中节选的部分电影及其属性所成的网状结构。从图3中可以看出,在知识图谱中越相似的2个节点,在语义上也往往十分接近。因此,对于电影推荐来说,不仅可以利用电影的用户评分信息,也可以使用电影自身的语义信息。协同过滤推荐认为两物品的用户评分分布相近,则它们被判定为近邻。同样地,如果该电影在知识图谱中相近,直观上也可以被判定为近邻。

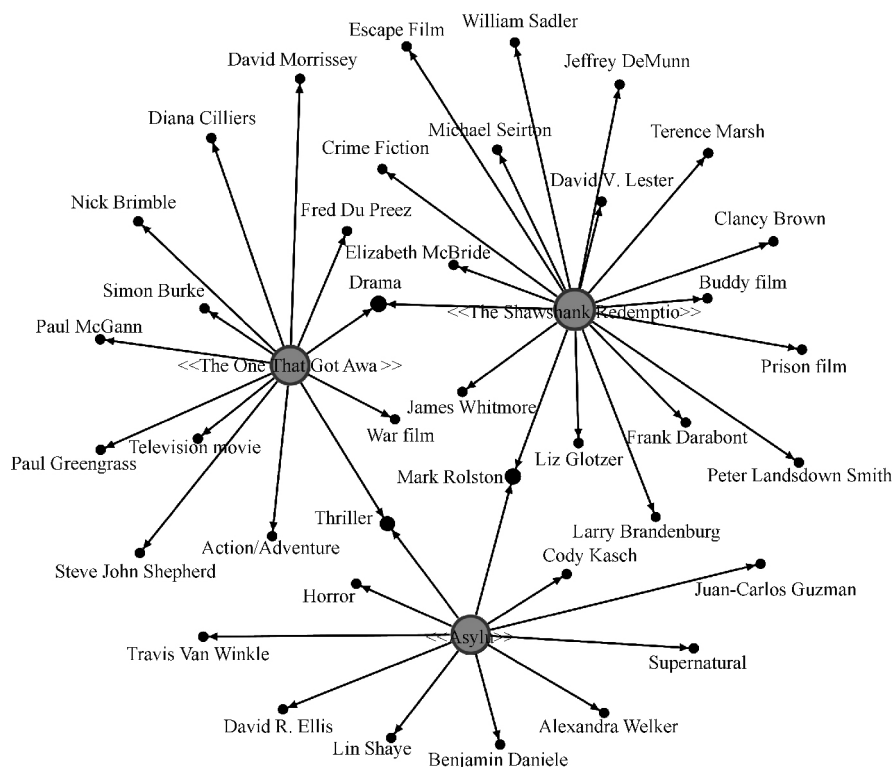


图3 电影知识图谱

基于上述思想, 本文使用在翻译方法上有较好性能的 TransE 算法, 增强协同推荐算法中物品-用户评分矩阵中物品的语义信息。按照 TransE 算法的定义, 对于知识图谱  $S$  中的一个三元组  $(h, r, t)$ , 可以用式(4)所示的损失公式进行训练:

$$L = \sum_{(h, r, t) \in S} \sum_{(h', r, t') \in S(h, r, t)} [\gamma + \|h + r - t\| - \|h' + r - t'\|]_+ \quad (4)$$

其中  $h'$  和  $t'$  是错误的三元组向量, 作为训练的负样本。该类负样本是将 TransE 算法定义的正样本, 也就是原来正确三元组的头实体或者尾实体随机替换成其他实体而得。 $\gamma$  为间距大小, 一般设  $\gamma = 1$ 。符号  $[\cdot]_+$  表示合页损失函数, 形式化描述如式(5)所示。

$$[x]_+ = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (5)$$

在整个训练过程中, TransE 采用了最大间距的思想, 拉开正样本和负样本向量之间的距离进行训练学习。通过随机梯度下降不断迭代, 使损失函数达到最优。最终的训练结果是: 1) 在知识图谱中相似的实体, 在低维空间里面的距离相近; 2) 头节点的向量加上关系向量基本等于尾节点的向量。具体过程训练见文献[12]。

### 2.1.2 语义相似性度量

上一节所述的 TransE 算法将对象表示成为一组低维实值向量, 且使得这些向量的值满足是一组连续实数的限定条件。在不同的知识图谱表示学习方法中, 向量的含义有所不同, 也就需要不同的相似性度量。式(6)给出了 TransE 算法中向量的表示形式:

$$\|h + r\| \approx t \quad (6)$$

其中  $h$  和  $t$  分别是头实体向量和尾实体向量,  $r$  是关系向量, 式子中的范数  $\|\cdot\|$  可以选择 L1 范数或者 L2 范数。在知识图谱里越接近的实体, 向量也就越相似, 如图4所示。

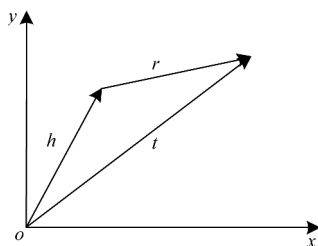


图4 TransE 模型

由于 TransE 算法选择欧氏距离计算损失函数, 因此本文采用同等范数的欧氏距离对实体进行相似性度量。其中, 选择的欧氏距离的范围为  $[0, \infty)$ , 相似性度量范围为  $[0, 1]$ 。对于  $A$  和  $B$  两个实体向量, 通过数学变换使两个值域一一对应, 最终的语义

相似性度量公式如式(7)所示。

$$\text{sim}_{\text{sem}}(A, B) = 1 - \frac{\|A - B\|}{\|A - B\| + 1} = \frac{1}{\|A - B\| + 1} \quad (7)$$

从上式中可以看出, 当  $\text{sim}_{\text{sem}}(A, B)$  的计算的数值越接近 1, 实体向量  $A$  和  $B$  就越相似, 即知识图谱里两者的关系就越紧密。相反, 如果  $\text{sim}_{\text{sem}}(A, B)$  的数值为低, 则意味着  $A$  和  $B$  关系越疏远, 即语义相似度越低。

在利用语义相似性进行推荐时, 使用和基于物品的协同过滤推荐类似的流程, 根据式(7)生成了一个语义的物品-物品相似性矩阵, 如表1所示。

表1 语义的物品-物品相似性矩阵

物品	$Item_1$	$Item_2$	...	$Item_j$	...	$Item_n$
$Item_1$	1	$a_{12}$	...	$a_{1j}$	...	$a_{1n}$
$Item_2$	$a_{21}$	1	...	$a_{2j}$	...	$a_{2n}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$Item_i$	$a_{i1}$	$a_{i2}$	...	$a_{ij}$	...	$a_{in}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$Item_n$	$a_{n1}$	$a_{n2}$	...	$a_{nj}$	...	1

在表1的相似性矩阵中,  $a_{ij}$  是序号为  $i$  的物品和序号为  $j$  的物品之间的语义相似性, 通过式(7)计算而得, 相似性矩阵元素满足条件  $a_{ij} = a_{ji}$ 。矩阵第  $i$  行所构成的集合定义为序号  $i$  物品的语义近邻元素, 简称为语义近邻。利用这种语义近邻, 将基于协同过滤推荐得到的协同过滤近邻集进行语义融合替换, 最终得到所需的推荐列表。

### 2.1.3 融合语义近邻

假定  $I$  是一个协同过滤算法将要推荐给用户的物品的集合。对每一个对象  $i \in I$  的预测评分进行排序, 生成一个对应的有序数列  $L$ 。同理, 对语义近邻集  $T$  按照式(8)进行排序生成一个待替换有序数列  $K$ 。 $K$  中元素满足  $\{j | j \in T, j \notin I\}$ 。假定数列  $L$  长度为  $l$ , 融合比例为  $p:q$ , 则需要替换的个数为  $n = \left\lceil l \times \frac{p}{p+q} \right\rceil$ , 其中  $\lceil \cdot \rceil$  为向上取整。得到以下融合过程:

#### 算法1 融合语义近邻算法

输入 协同过滤近邻集  $I$ , 语义近邻集  $T$

输出 推荐集  $C$

1. Set  $L = \text{list}(I)$ ,  $K = \text{list}(T)$

2. for  $L_i \in \{L_1, L_2, \dots, L_l\}$  do:

3.     for  $K_j \in \{K_1, K_2, \dots, K_n\}$  do:

4.          $L_i = K_j$

5.     end do

6. end do

7. Set  $C = \text{set}(L)$

利用这一过程得到了最终的推荐集  $C$ 。

## 2.2 算法描述

结合图 1 的算法流程 描述本文算法如下:

### 算法 2 TransE-CF 算法

输入 物品-用户评分矩阵  $R_{m \times n}$ , 该物品所在领域的知识图谱 KG

输出 基于知识图谱表示学习的协同过滤推荐算法 TransE-CF

1) 根据式 (3) 利用的物品-用户评分矩阵  $R_{m \times n}$  计算物品与物品之间的相似性  $w_{ij}$ , 生成物品-物品相似性矩阵, 并根据 1.1 节对协同过滤算法的潜在推荐物品进行过滤。

2) 将物品与知识图谱 KG 中的实体进行一一对应, 得到一个实体对应表。

3) 将知识图谱 KG 中的实体和关系分别转化为向量集  $E$  和  $R$ 。

4) 根据式 (7) 对第 3) 步得到的结果分别计算物品间的语义相似性  $sim_{sem}(A, B)$ , 生成语义的物品-物品相似性矩阵, 并根据 2.1.3 节对其进行过滤, 得到语义近邻集合。

5) 利用第 2) 步的结果, 选择融合比例, 根据算法 1 对第 1) 步生成的潜在推荐物品和第 4) 步所生成的语义近邻集合进行融合替换, 得到最终推荐所需的物品集合推送给用户。

## 3 实验设置及结果分析

### 3.1 数据集

本节主要针对上述已提出的 TransE-CF 算法的有效性进行实验与分析, 其中实验数据集为加州大学圣迭戈分校提供的 Amazon 的电影评分数据。该数据集为 Amazon 的商品评分及其元数据。对每条数据记录而言, 主要描述了不同用户对电影的评分, 分值从 1 分到 5 分逐渐增高, 分别对应于用户对电影的差评到好评。同时, 使用 metadata 将每条记录中电影的 Amazon 编号 asin 与知识图谱里该部电影的名字所表示的电影实体一一对应。

在知识图谱的数据集选择方面, 本文选用了 2012-11-09 版本的 Freebase 中的电影本体。该电影本体数据主要包括了电影、导演、演员等本体对象。为了将 Freebase 提供的多个本体对象进行融合, 本文用实体的名称替换了 Freebase 自身的 id, 提取出所需要的实体和关系。同时, 为了降低噪声数据的影响, 选择了 Freebase 中 1979 年以后发布的电影与 Amazon 评分进行匹配, 并筛选去除了出现次数少于 3 次的关系, 最终得到“genre”“starring\_actor”等总共 20 个知识图谱的语义关系。

同时, 为了使得 Amazon 评分数据和 Freebase 更好地匹配, 本文对电影名称进行了进一步的处理。

由于知识图谱采用的是无监督或者是半监督的关系抽取手段<sup>[21]</sup>, 因此 Freebase 中的电影名称抽取结果受限于该处理手段, 存在物品名称和本体实体名称字面无法完全匹配的问题。例如, 存在有的电影使用罗马数字 II, 而有的使用阿拉伯数字 2。此外, 例如电影 Here Comes the Groom (1934) 和电影 Here Comes the Groom 实质上是同一部电影但由于物品版本号的不同, 造成 Amazon 电影名和 Freebase 电影本体的名称匹配时无法完全对应。

针对上述问题, 本文采用了编辑距离和字符串规则匹配相结合的方式对电影名称进行了版本号删除等数据预清洗工作。通过编辑距离算法对原数据进行字符串字面相似性识别矫正, 得到了较好的清洗结果。经过最终处理得到包括 672 910 项用户数据和 41 255 部电影数据, 以及 20 个语义关系的实验数据集。本文对该评分数据集进行了划分, 其中 80% 作为训练集, 20% 作为测试集, 训练集用于构建评分矩阵, 测试集用于测试算法性能。

### 3.2 评价指标

对于推荐系统算法推荐的结果, 使用 3 个指标对其进行分析: 准确率 (Precision), 召回率 (Recall), F 值 (F-measure)。这 3 个指标可以从混淆矩阵中导出, 如表 2 所示。

表 2 混淆矩阵

系统	用户喜欢	用户不喜欢
推荐	TP	FP
未推荐	FN	TN

根据混淆矩阵, 有公式如下:

召回率 (Recall):

$$Recall = \frac{TP}{TP + FN} = 1 - \frac{FN}{T} \quad (8)$$

准确率 (Precision):

$$Precision = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + FN + FP + TN} \quad (9)$$

F 值:

$$F\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

召回率反映了被推荐系统所推荐的物品占用户真正喜欢的物品的比重。准确率反映了推荐系统的推荐水平, 能将用户喜欢的物品推荐给用户, 而用户不喜欢的物品则不推荐。F 值是准确率和召回率的加权平均, 均匀地反映了推荐效果。

### 3.3 实验结果分析

实验硬件处理器型号为 Intel (R) Core (TM) i7-6700, 内存为 8 GB; 实验软件环境为 Python 3.5。

### 3.3.1 混合比例的调整确定

本文所提出的算法中可调节的参数为两者的融合比例。在融合比例取不同值时, 所取得的推荐效果也有所不同。选取 Top- $k$  的邻居数  $k=100$ , 表示学习嵌入为 200, 实验从完全使用语义进行推荐 (0:10) 到完全使用协同过滤进行推荐 (10:0) 分别取值, 图 5~图 7 分别为准确率、召回率和 F 值曲线。对于每一组实验, 均循环 10 次并取其平均值。

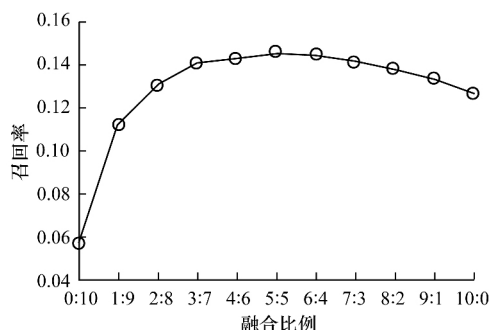


图5  $k=100$  时的召回率

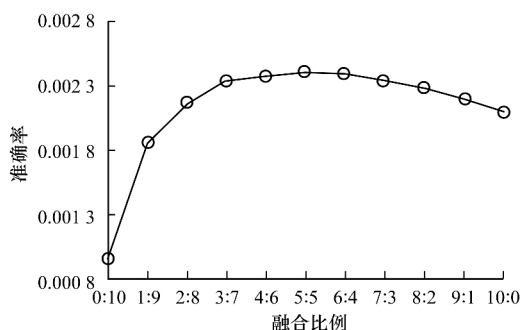


图6  $k=100$  时的准确率

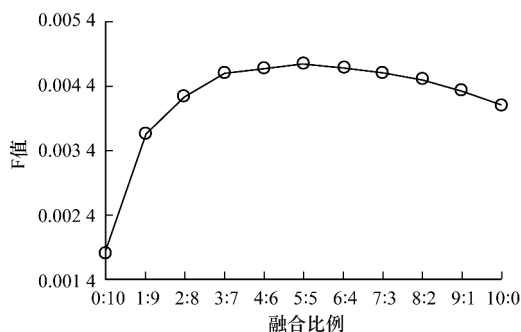


图7  $k=100$  时的F值

从图 5~图 7 可以看出, 随着融合比例的提高, 准确率、召回率和 F 值均有所提升, 并在比例为 5:5 时达到了顶峰。对于  $k=100$  的情况, 融合比例为 5:5 时效果最好。

### 3.3.2 表示学习的嵌入维度比较

知识图谱表示学习是将知识图谱嵌入一个低维的空间里, 对于不同维度会有不同的效果。选取

100 维~500 维分别进行实验, 对于每一组实验, 均循环 10 次并取其平均值。

从图 8~图 10 可以看出各个维度对推荐效果的影响。从图中得出, 在选取 200 维时取得最佳。

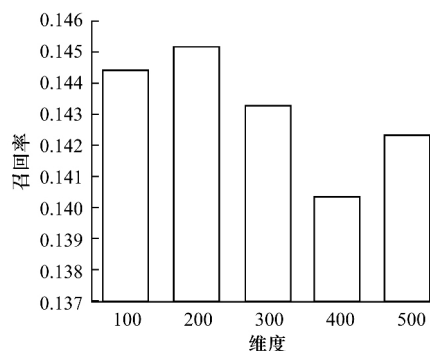


图8 不同维度的召回率

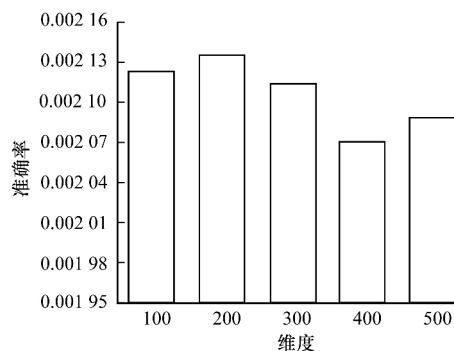


图9 不同维度的准确率

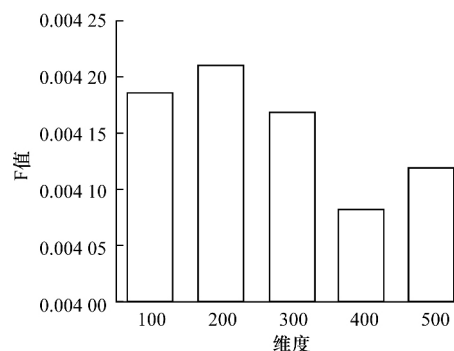
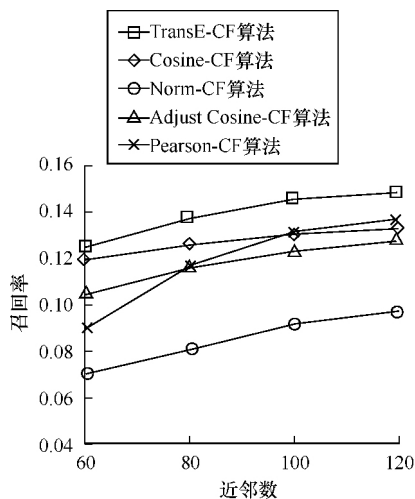
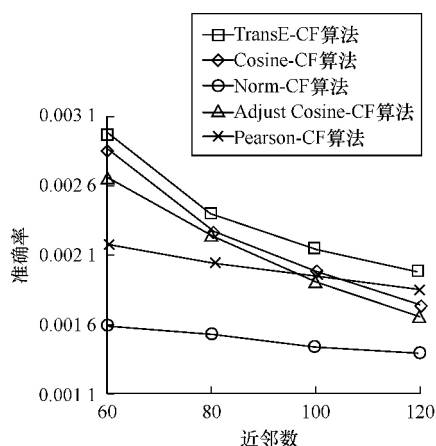
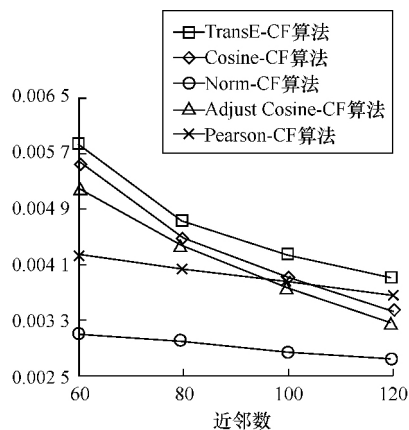


图10 不同维度的F值

### 3.3.3 算法比较

这里选取 F 值最高的比例作为代表, 与其他协同过滤推荐算法进行比较。选取嵌入维度为 200,  $k$  近邻数分别为 60、80、100、120。对于每一组实验, 均循环 10 次并取其平均值。

从图 11~图 13 可以看出, 本文提出的基于知识图谱表示学习的协同过滤推荐算法, 优于协同过滤推荐算法, 利用语义信息可以在一定程度上对协同过滤推荐算法进行改进, 在召回率、准确率和 F 值上均有所提升, 可以在语义上弥补基于物品的协同过滤算法的不足。

图 11 不同  $k$  值的召回率图 12 不同  $k$  值的准确率图 13 不同  $k$  值的 F 值

#### 4 结束语

本文提出了一种基于知识图谱表示学习的协同过滤推荐算法,既利用了物品本身内在的语义信息,又使用了物品外在的评分矩阵,能够更加全面地反映物品的属性。算法通过知识图谱表示学习将实体嵌入到低维空间里,计算实体间的语义相似性,并将其融入到推荐系统中,在语义的层面上增强了协同过滤

推荐的效果,解决了协同过滤推荐未考虑语义的问题,提高协同过滤推荐的精度;同时又利用语义的信息,在一定程度上解决了冷启动问题。实验表明,基于知识图谱表示学习的协同过滤推荐算法能够提高推荐算法的效果。下一步将尝试把该算法应用于电影本体以外内容进行推荐,并进一步优化推荐性能。

#### 参考文献

- [1] 李文海,许舒人. 基于 Hadoop 的电子商务推荐系统的设计与实现[J]. 计算机工程与设计, 2014, 35(1): 130-136.
- [2] SONG Bo, GAO Jie. The Enhancement and Application of Collaborative Filtering in e-Learning System [C]// Proceedings of International Conference in Swarm Intelligence. Berlin, Germany: Springer, 2014: 188-195.
- [3] GARCIN F, ZHOU K, FALTINGS B, et al. Personalized News Recommendation Based on Collaborative Filtering [C]// Proceedings of International Joint Conferences on Web Intelligence and Intelligent Agent Technology. Washington D. C. USA: IEEE Press, 2012: 164-173.
- [4] 文俊浩,袁培雷,曾骏,等. 基于标签主题的协同过滤推荐算法研究[J]. 计算机工程, 2017, 43(1): 247-252, 258.
- [5] FERNÁNDEZ-TOBIÁS I, CANTADOR I. Personality-aware Collaborative Filtering: An Empirical Study in Multiple Domains with Facebook Data [C]// Proceedings of International Conference on Electronic Commerce and Web Technologies. Berlin, Germany: Springer, 2014: 125-137.
- [6] 杨兴耀,于炯,吐尔根·依布拉音,等. 融合奇异性和扩散过程的协同过滤模型[J]. 软件学报, 2013, 24(8): 1868-1883.
- [7] 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.
- [8] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge [C]// Proceedings of ACM SIGMOD International Conference on Management of Data. New York, USA: ACM Press, 2008: 1247-1250.
- [9] BIZER C, LEHMANN J, KOBILAROV G, et al. DBpedia—A Crystallization Point for the Web of Data [J]. Web Semantics Science Services & Agents on the World Wide Web, 2009, 7(3): 154-165.
- [10] ZHANG Zijian, GONG Lin, XIE Jian. Ontology-based Collaborative Filtering Recommendation Algorithm [C]// Proceedings of International Conference on Brain Inspired Cognitive Systems. Berlin, Germany: Springer, 2013: 172-181.
- [11] ZHANG Fuzheng, YUAN N J, LIAN Defu, et al. Collaborative Knowledge Base Embedding for Recommender Systems [C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2016: 353-362.

(下转第 263 页)

- New York, USA: ACM Press 2009: 41-48.
- [9] BENGIO Y, COURVILLE A, VINCENT P. Representation Learning: A Review and New Perspectives [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [10] KUMAR M P, PACKER B, KOLLER D. Self-paced Learning for Latent Variable Models [C]//Proceedings of the 23th Annual Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press 2010: 1189-1197.
- [11] JIANG L, MENG D, YU S I, et al. Self-paced Learning with Diversity [C]//Proceedings of the 27th Annual Conference on Neural Information Processing Systems. Cambridge, USA: MIT Press 2014: 2078-2086.
- [12] WANG H, KLASER A, SCHMID C, et al. Dense Trajectories and Motion Boundary Descriptors for Action Recognition [J]. International Journal of Computer Vision, 2013, 103(1): 60-79.
- [13] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press 2005: 886-893.
- [14] LAPTEV I, MARSZALEK M, SCHMID C. Learning Realistic Human Actions from Movies [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press 2008: 1-8.
- [15] DALAL N, TRIGGS B, SCHMID C. Human Detection Using Oriented Histograms of Flow and Appearance [C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer 2006: 428-444.
- [16] GEHLER P V, NOWOZIN S. Let the Kernel Figure It out: Principled Learning of Preprocessing for Kernel Classifiers [C]//Proceedings of IEEE Conference on Computer Vision. Washington D. C., USA: IEEE Press 2009: 2836-2843.
- [17] 许允喜, 蒋云良, 陈方. 基于支持向量机增量学习和 LPBoost 的人体目标再识别算法 [J]. 光子学报, 2011, 40(5): 758-763.
- [18] 方育柯, 傅彦, 周俊临, 等. 基于选择性集成的最大化软间隔算法 [J]. 软件学报, 2013, 34(5): 1132-1147.
- [19] XIA L, CHEN C C, AGGARWAL J K. View Invariant Human Action Recognition Using Histograms of 3D Joints [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. Washington D. C., USA: IEEE Press 2012: 20-27.
- [20] SEIDENARI L, VARANO V, BERRETTI S, et al. Recognizing Actions from Depth Cameras as Weakly Aligned Multi-part Bag-of-poses [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops. Washington D. C., USA: IEEE Press 2013: 479-485.
- [21] 汤荣志, 段会川, 孙海涛. SVM 训练数据归一化研究 [J]. 山东师范大学学报(自然科学版), 2016, 31(4): 60-65.

编辑 吴云芳

(上接第232页)

- [12] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating Embeddings for Modeling Multi-relational Data [C]//Proceedings of International Conference on Neural Information Processing Systems. New York, USA: ACM Press 2013: 2787-2795.
- [13] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based Collaborative Filtering Recommendation Algorithms [C]//Proceedings of International Conference on World Wide Web. New York, USA: ACM Press 2001: 285-295.
- [14] BELL R, KOREN Y, VOLINSKY C. Modeling Relationships at Multiple Scales to Improve Accuracy of Large Recommender Systems [C]//Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM Press, 2007: 95-104.
- [15] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed Representations of Words and Phrases and Their Compositionality [C]//Proceedings of International Conference on Neural Information Processing Systems. New York, USA: ACM Press 2013: 3111-3119.
- [16] SOCHER R, CHEN Danqi, MANNING C D, et al. Reasoning with Neural Tensor Networks for Knowledge Base Completion [C]//Proceedings of International Conference on Neural Information Processing Systems. New York, USA: ACM Press 2013: 926-934.
- [17] NICKEL M, TRESP V, KRIEGER H P. A Three-way Model for Collective Learning on Multi-relational Data [C]//Proceedings of the 28th International Conference on Machine Learning. Berlin, Germany: Springer 2011: 809-816.
- [18] WANG Zhen, ZHANG Jianwen, FENG Jianlin, et al. Knowledge Graph Embedding by Translating on Hyperplanes [C]//Proceedings of the 28th AAAI Conference on Artificial Intelligence. Berlin, Germany: Springer 2014: 1112-1119.
- [19] LIN Yankai, LIU Zhiyuan, ZHU Xuan, et al. Learning Entity and Relation Embeddings for Knowledge Graph Completion [C]//Proceedings of the 29th AAAI Conference on Artificial Intelligence. Berlin, Germany: Springer 2015: 2181-2187.
- [20] 刘康, 张元哲, 纪国良, 等. 基于表示学习的知识库问答研究进展与展望 [J]. 自动化学报, 2016, 42(6): 807-818.
- [21] 肖明, 邱小花, 黄界, 等. 知识图谱工具比较研究 [J]. 图书馆杂志, 2013, 32(3): 61-69.

编辑 顾逸斐