三维有偏权值张量分解在授课推荐上的应用研究

姚敦红[1,2]，李石君[2]，胡亚慧[2,3]

(1. 怀化学院计算机科学与工程学院 湖南 怀化 418000；2. 武汉大学计算机学院 武汉 430072; 3. 空军预警学院四系 武汉 430010)

【摘要】为解决现今学校授课安排无推荐依据这一实际问题，首先给出了一系列形式化方法用于规约教师的专业基础、课程难度及教学评价；定义了一种加权函数计算出每组专业基础、课程难度和教学评价的综合有偏权值；构建了一种基于“教师-课程-评价-权值”四元关系的三维有偏权值张量模型，张量元素使用综合有偏权值。在此基础上，设计了一种基于Tucker分解的算法，对张量进行高阶奇异值分解(HOSVD)得到降维后的近似张量，按课程分类实现了Top_*N*授课推荐。实验结果表明，当迭代阈值达到一个合理值时，该方法能实现精准授课推荐，可作为一种新的智能化授课推荐方法应用于各类学校。

关 键 词 数据规约； 授课推荐； 张量分解； 三维有偏权值张量

中图分类号 TP391 **文献标志码** A doi:10.3969/j.issn.1001-0548.2017.05.018

A Three-Dimensional Partial Weight Tensor Model for Teaching Recommendation

YAO Dun-hong[1,2], LI Shi-jun[2], and HU Ya-hui[2,3]

(1. College of Computer Science & Engineering, Huaihua University Huaihua Hunan 418000;

2. School of Computer, Wuhan University Wuhan 430072;

3. The Fourth Department of Air Force Early Warning Academy Wuhan 430010)

Abstract To address the problem that the teaching arrangements are not on the basis of recommendation in current school, a series of formalized methods are used to specify teachers' specialty foundation, course difficulty, and teaching evaluation first. Then, a kind of weighted function is defined to calculate the comprehensive partial weight for each group of teachers' professional foundation, course difficulty, and teaching evaluation. Next, the three-dimensional tensor model with partial weight is built on the 4-tuples relation of teacher-course-evaluation-weight and the comprehensive weight is endowed to the tensor elements. Finally, on the basis of above, a new kind of decomposition algorithm based on Tucker Decomposition is designed to obtain the approximate tensor of dimensionality reduction with the higher-order singular value decomposition (HOSVD), achieving the Top-*N* recommendation of teaching arrangements. Experiment results show that our proposed method can realize precise teaching arrangements recommendations when the iterative threshold value reaches a reasonable value, which can be used as a new intelligent recommendation method applied to the teaching arrangements in all kinds of schools.

Key words data reduction; teaching recommendation; tensor decomposition; three-dimensional partial weighted tensor

推荐系统是对用户历史行为数据进行分析、预测并主动为用户给出相关推荐的系统。自文献[1]推出第一个推荐系统以来，涌现出了大量的推荐系统，特别是在电子商务、社交网络、搜索引擎等方面，如亚马逊基于兴趣的广告推荐、NEC研究院的CiteSeer搜索引擎、IBM的Websphere电商平台、阿里云推荐、京东推广、百度推广、博客挖掘、社交推荐等。这些推荐应用的实现一般是根据用户行为数据建立起的“用户-项目”二元关系挖掘分析而得。随着社会化标签的出现，又出现了“用户-产品-标签”的三元关系，使个性化推荐更趋向精准。

目前，推荐系统常用的技术有基于欧氏距离、Pearson相关系数、余弦相似性和Tanomi等最近邻启发式协同过滤推荐算法[2]；有基于上下文感知模型、潜在因子模型、贝叶斯模型、信任感知模型、聚类模型、最大熵模型[3]等协同过滤推荐算法；有以决

收稿日期：2016－03－17；修回日期：2017－05－05
基金项目：国家自然科学基金(61272109)；湖南省教育厅科学研究项目(15C1086)
作者简介：姚敦红(1972－)，男，副教授，主要从事数据挖掘、机器学习方面的研究.

策树、神经网络、向量、TF-IDF、自适应过滤、阈值设定等基于内容的推荐算法；还有其他如关联规则推荐、效用推荐、知识推理等算法，以及使用标签的图、标签的FolkRank、层叠、加权、变换、标签层次聚类[4]和张量分解的组合推荐算法等。

应用张量分解算法进行个性化推荐，在近年来也有了一些研究，文献[5-7]采用了融合某种关系或附加某种标签信息的张量分解推荐算法。文献[8-10]也有采用加权张量模型，即通过提取标注关键特征，再得出一个权值作为张量元素。

在现有研究中，还未曾涉及学校授课推荐。一直以来，学校授课安排没有一种好的推荐依据，很多是随教师意愿而为，或是强加给教师，这些方式未能使教学达到最优效果，难以提高教学质量。所以，在学校多年大量的教学数据中进行分析挖掘，找到一种实现精准授课推荐的方法，具有一定的现实意义和实用价值。

本文借鉴文献[11]的四元元组张量分解算法，优化文献[12]中提出的张量稀疏问题，设计一种基于Tucker张量分解的算法。并利用历史教学数据集进行授课推荐实验，验证该方法在授课推荐上的准确性。

1 基本概念

借鉴文献[13]对推荐系统的定义，可将授课推荐系统(teaching recommendation system)定义为：设有教师集合teacher=$\{t_1,t_2,\cdots,t_n\}$、课程集合course=$\{c_1,c_2,\cdots,c_n\}$和评价集合evaluation=$\{e_1,e_2,\cdots,e_n\}$，推荐系统目标就是使得衡量教师t、课程c与评价e之间的相关性效用函数$f(t,c,e)$最大，即 $\forall\, t\in$teacher，$f(t)=\max\{f(t,c,e)\}$。

张量是高维数组的总称[14]，一维张量是向量，二维张量是矩阵，三维或以上的张量为高阶张量[6]。张量分解即HOSVD，是对高维数据进行特征提取，或是一种低秩逼近。常见的张量分解模型有：CP模型、Tucker模型[15]。Tucker模型将N维张量分解成N个维度上的低秩特征矩阵与一个核心张量的乘积，其本质是一种高阶主成分分析。如三维张量X的Tucker分解为：

$$X\approx\hat{X}=C\times_i\boldsymbol{V}^{(i)}\times_j\boldsymbol{V}^{(j)}\times_k\boldsymbol{V}^{(k)}=\sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}c_{pqr}\boldsymbol{v}_p\circ\boldsymbol{v}_q\circ\boldsymbol{v}_r \tag{1}$$

式中，$\boldsymbol{V}^{(i)}\in\mathbb{R}^{I\times P}$，$\boldsymbol{V}^{(j)}\in\mathbb{R}^{J\times Q}$，$\boldsymbol{V}^{(k)}\in\mathbb{R}^{K\times R}$代表3个维度主成分且相互正交的低秩特征矩阵；$C\in\mathbb{R}^{P\times Q\times R}$是核心张量；运算符$o$表示向量的外积[16]。

如果P，Q，R对应小于I，J，K，则又称C为张量X的压缩张量(规模远远小于原张量的相似张量)，这在大数据集稀疏张量的应用上效果非常显著。由式(1)可知，当$\boldsymbol{V}^{(i)}$、$\boldsymbol{V}^{(j)}$和$\boldsymbol{V}^{(k)}$确定后，核心张量C就可近似由原张量X与各维特征矩阵的转置运算得到：

$$C=X\times_i\boldsymbol{V}^{(i)^{\mathrm{T}}}\times_j\boldsymbol{V}^{(j)^{\mathrm{T}}}\times_k\boldsymbol{V}^{(k)^{\mathrm{T}}} \tag{2}$$

三维张量通过Tucker分解后得到的相似张量，可采用最小化函数$\min\limits_{\hat{X}}\left\|X-\hat{X}\right\|$计算其相似程度。为便于计算，对最小化函数平方得到：

$$\begin{aligned}&\left\|X-C\times_i\boldsymbol{V}^{(i)}\times_j\boldsymbol{V}^{(j)}\times_k\boldsymbol{V}^{(k)}\right\|^2=\\&\left\|X\right\|^2-2\left\langle X,C\times_i\boldsymbol{V}^{(i)}\times_j\boldsymbol{V}^{(j)}\times_k\boldsymbol{V}^{(k)}\right\rangle+\\&\left\|C\times_i\boldsymbol{V}^{(i)}\times_j\boldsymbol{V}^{(j)}\times_k\boldsymbol{V}^{(k)}\right\|^2=\\&\left\|X\right\|^2-2\left\langle X\times_i\boldsymbol{V}^{(i)}\times_j\boldsymbol{V}^{(j)}\times_k\boldsymbol{V}^{(k)},C\right\rangle+\left\|C\right\|^2=\\&\left\|X\right\|^2-2\left\langle C,C\right\rangle+\left\|C\right\|^2=\left\|X\right\|^2-\left\|C\right\|^2=\\&\left\|X\right\|^2-\left\|X\times_i\boldsymbol{V}^{(i)^{\mathrm{T}}}\times_j\boldsymbol{V}^{(j)^{\mathrm{T}}}\times_k\boldsymbol{V}^{(k)^{\mathrm{T}}}\right\|^2\end{aligned} \tag{3}$$

根据式(3)可知，求$\min\limits_{\hat{X}}\left\|X-\hat{X}\right\|$的最优解可转化为$\left\|X\times_i\boldsymbol{V}^{(i)^{\mathrm{T}}}\times_j\boldsymbol{V}^{(j)^{\mathrm{T}}}\times_k\boldsymbol{V}^{(k)^{\mathrm{T}}}\right\|^2$最大化问题的最优解，于是分别对$\boldsymbol{V}^{(i)^{\mathrm{T}}}$，$\boldsymbol{V}^{(j)^{\mathrm{T}}}$和$\boldsymbol{V}^{(k)^{\mathrm{T}}}$做奇异值分解降维处理后，再组合可得到规模比原张量小得多的相似张量$\hat{X}$，这有利于加快推荐的速度、提高推荐的精度。

2 数据预处理

为构建用于授课推荐的有偏权值张量模型，和适应使用基于Tucker张量分解算法的要求，需对采集得到的相关教学数据进行预处理。首先从教师信息表、课程信息表及学生评教表等多个数据库表中，采用ETL方式构建一个事实星座模式的教学信息数据仓库，其结构如图1所示。图中，Course ID表示课程编号，Eva表示综合评价值，Sf(1)表示第1毕业学校因子，Sf(2)表示最后毕业学校因子，Pdb表示专业基础度。

然后采用下述定义对数据仓库中的相关属性进行规约处理：

定义 1 毕业学校因子(school factor, Sf)：用来规约教师的毕业学校，按下列规则赋值，毕业于“985工程”与“211工程”高校Sf=0.4，毕业于“211工程”高校Sf=0.3，毕业于其他一本院校Sf=0.2，毕业于二

本及以下院校Sf=0.1。

定义 2 学位系数(degree coefficient, Dc)：用于规约教师取得的学位，本文约定博士、硕士、学士和无学位的Dc分别取0.4、0.3、0.2和0.1。

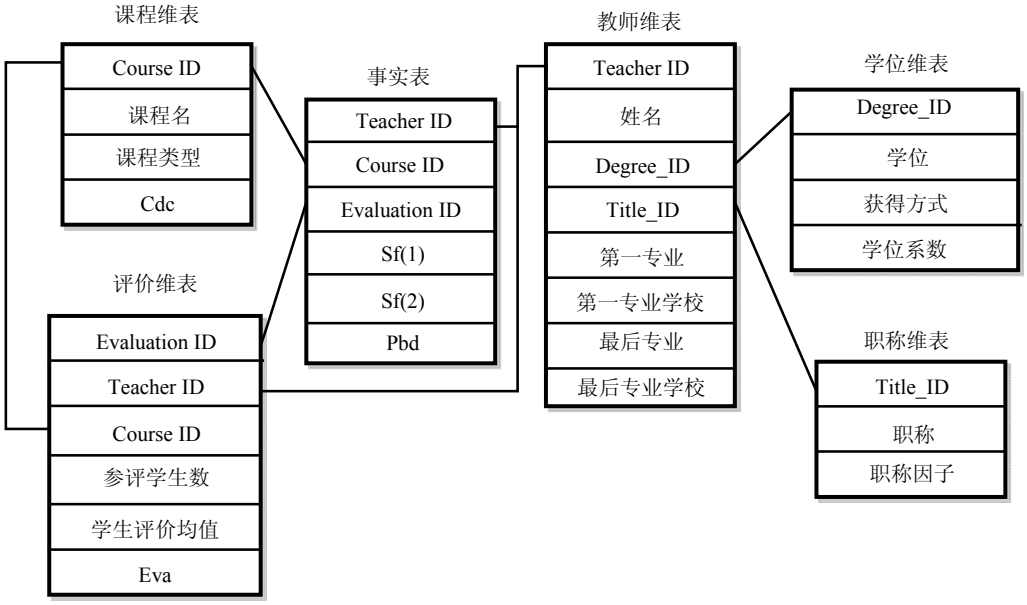

图1　事实星座模式结构图

定义 3 专业基础度(professional basic degree, Pbd) ($0.1 \leqslant \mathrm{Pbd} \leqslant 1$)：用于规约教师的专业基础，值越大表示专业基础越雄厚：

$$\mathrm{Pbd} = \sqrt{n}\sum_{i=1}^{2}(w_i \times \mathrm{Rm}_i) + \frac{1}{2}\sum_{i=1}^{2}\mathrm{Sf}_i + \mathrm{Dw} \times \mathrm{Dc} \tag{4}$$

式中，n表示教师总数；$\mathrm{Dw}(0 < \mathrm{Dw} \leqslant 1)$ 为教师的学位获得方式系数(本文约定全日制获得方式为1，非全日制获得方式小于1)；$w_i(i = 1,2,\cdots)$ 表示教师第i个毕业专业(一般取第一和最后毕业专业的权重值：

$$w_i = \mathrm{Rm}_i \Big/ \left(2\sum_{p=1}^{n}\mathrm{Rm}_p\right) \text{且} \sum_{p=1}^{n}\sum_{i=1}^{2} w_i = 1 \tag{5}$$

式中，$\mathrm{Rm}_i(i = 1,2,\cdots)$ 表示教师的第i个毕业专业与所从事的专业的相关值，其取值方式定义如下：

$$\mathrm{Rm}_i = \begin{cases} 1 & \text{全相关} \\ r & \text{相关} \end{cases} \quad 0 < r < 1 \tag{6}$$

式中，r 称为专业相关系数。

定义 4 课程难度系数(curriculum difficulty coefficient, Cdc)($0.1 \leqslant \mathrm{Cdc} \leqslant 1$)：用于规范课程难度的指标，值越大表示课程难度越大。为使课程难度系数的评定趋于公认值，邀请校内外该专业优秀毕业生及专家教师在课程难度系数网上问卷调查，问卷调查中为每一专业的每门课程给出1～10个选项，每个专业总问卷份数不少于指定的阈值(如200)。然后将每门课程的难度系数规范化至区间[$C_{\min}$，$C_{\max}$](本文中设$C_{\min}$为0.1，$C_{\max}$为1.0)上的一个难度系数，表示为：

$$\mathrm{Cdc} = \frac{\mathrm{Qr} - \mathrm{Qr}_{\min}}{\mathrm{Qr}_{\max} - \mathrm{Qr}_{\min}}(C_{\max} - C_{\min}) + C_{\min} \tag{7}$$

式中，Qr表示某门课程按专家教师问卷调查所占权重$w(0<w<1)$得到的难度值：

$$\mathrm{Qr} = \frac{w}{p}\sum_{i=1}^{p}\mathrm{Cd}_i + \frac{(1-w)}{q}\sum_{\mathrm{i}=1}^{q}\mathrm{Cd}_i \tag{8}$$

式中，p为某专业回收的教师专家问卷份数；q为回收的学生问卷份数；Cd_i为第i门课程在问卷中所给出的难度系数值。

定义 5 教师授课综合评价值(evaluation, Eva) ($0.1 \leqslant \mathrm{Eva} \leqslant 1$)：表示教师所授的某一门课程总的综合评价分，分值越高表示越受欢迎。可采用最小-最大规范化方法将Eva规范化至区间[$E_{\min}$，$E_{\max}$](本文设$E_{\min}$为0.1，$E_{\max}$为1.0)上的一个综合评价值，表示为：

$$\mathrm{Eva} = \frac{\mathrm{Stu\_sco} - \mathrm{Stu\_sco}_{\min}}{\mathrm{Stu\_sco}_{\max} - \mathrm{Stu\_sco}_{\min}}(E_{\max} - E_{\min}) + E_{\min} \tag{9}$$

式中，$\mathrm{Stu_sco}_{\min}$为某专业内所有课程中评价最低分值；$\mathrm{Stu_sco}_{\max}$为评价最高分值；Stu_sco表示某教师所授同一课程，在M个学期上学生评价分的总平均值：

$$\mathrm{Stu\_sco} = \sum_{m=1}^{M}(\mathrm{Numbers}_m \times \mathrm{Stu\_sco}_m) \Big/ \sum_{m=1}^{M}\mathrm{Numbers}_m \tag{10}$$

式中，$\mathrm{Numbers}_m$表示某门课程第$m(1 \leqslant m \leqslant M)$学期参与评价的学生数；$\mathrm{Stu_sco}_m(0 < \mathrm{Stu_sco}_m \leqslant 100)$表

示该门课程第m学期的学生平均评分值。

3 模型及算法设计

3.1 三维有偏仅值张量模型

为构建三维有偏权值张量模型，数据集按“教师(T)-课程(C)-评分(E)-权值(W)”四元关系$(t_i,c_j,e_k,w_{t_i,c_j,e_k})$构成维度分别为$T$、$C$、$E$的三维张量$X\in R^{I_t\times I_c\times I_e}$，其元素对应下标是$(t_i,c_j,e_k)$，通过对应的元素值计算得到综合有偏权值：

$$w_{t_i,c_j,e_k}=\begin{cases}\rho_1\mathrm{Pbd}_i+\rho_2\mathrm{Cdc}_j+\rho_3\mathrm{Eva}_k & \mathrm{Eva}_k\neq 0\ \sum_{m=1}^{3}\rho_m=1\\ 0 & \mathrm{Eva}_k=0\end{cases} \tag{11}$$

式(11)表示如果存在某专业基础度为Pbd_i的教师(t_i)讲授难度系数为Cdc_j的课程(c_j)且获得了评分(e_k)Eva_k，则张量对应下标(t_i,c_j,e_k)的元素值取加权计算得到w_{t_i,c_j,e_k}，否则对应元素取0。其中ρ_1，ρ_2和ρ_3分别为专业基础度、课程难度和教学评价的比重系数，可根据授课推荐偏重面不同而设置不同值，得到不同偏重性的推荐结果。这体现出授课推荐综合考虑教师专业基础与课程难度及评分值因素，是一种综合性的和有偏向性的权值。

在实际应用中，课程集与教师集均是大数集，但每位教师所教授的课程仅占课程集中几个元素。这样势必会造成三维有偏权值张量X中绝大部分元素为0，即构建的张量X是非常稀疏的。

3.2 算法设计

首先按前面的定义，将原始数据集中的数据进行规约、变换和计算，得出Pbd、Cdc与Eva；然后按式(11)计算出综合有偏权值w_{t_i,c_j,e_k}，以“教师-课程-评分-权值”方式构建加权四元元组；再以教师、课程和评分作为维度，以综合有偏权值w_{t_i,c_j,e_k}作为元素值，建立一个稀疏的三维有偏权值张量模型；最后，基于Tucker张量分解方法，采用交替最小二乘法获得降维后的近似张量，根据近似张量元素值的大小，按课程分类产生Top-N推荐列表，算法伪代码如下：

输入：迭代收敛阈值ε和最大迭代次数max-iteration;

输出：核心张量C和特征矩阵$\boldsymbol{V}^{(1)}$，$\boldsymbol{V}^{(2)}$和$\boldsymbol{V}^{(3)}$，以及按课程分类的不同Top_N的推荐结果列表;

Begin

数据预处理，按式(11)计算w_{t_i,c_j,e_k};

按教师(T)-课程(C)-评分(E)-权值(W)构建三维有偏权值张量X;

初始化$\boldsymbol{V}^{(1)}$，$\boldsymbol{V}^{(2)}$和$\boldsymbol{V}^{(3)}$;

初始化$C_0=X\times_1\boldsymbol{V}^{(1)^{\mathrm{T}}}\times_2\boldsymbol{V}^{(2)^{\mathrm{T}}}\times_3\boldsymbol{V}^{(3)^{\mathrm{T}}}$;

for(t=0; t< max-iteration; t++) {

for each $n\in[1,2,3]$ {

$\hat{X}=X$;

for each $m\in[1,n-1]$ && $m\neq n$

$\hat{X}=\hat{X}\times_m\boldsymbol{V}_{t+1}^{(m)^{\mathrm{T}}}$;

for each $m\in[n,3]$

$\hat{X}=\hat{X}\times_m\boldsymbol{V}_{t}^{(m)^{\mathrm{T}}}$;

$\left(\boldsymbol{V}_{t+1}^{(n)},\sum_{t+1}^{(n)}\boldsymbol{W}_{t+1}^{(n)}\right)=\mathrm{SVD}(uf(\hat{X},n),R)$;} // 采用SVD分解，$\boldsymbol{W}_{t+1}^{(n)}$是$\boldsymbol{V}_{t+1}^{(n)}$正交矩阵，$\sum_{t+1}^{(n)}=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_{t+1})$

$C_{t+1}=X\times_1\boldsymbol{V}_{t+1}^{(1)^{\mathrm{T}}}\times_2\boldsymbol{V}_{t+1}^{(2)^{\mathrm{T}}}\times_3\boldsymbol{V}_{t+1}^{(3)^{\mathrm{T}}}$;

if $\|C_{t+1}\|^2-\|C_t\|^2<\varepsilon$ exit; }

返回核心张量C和特征矩阵$\boldsymbol{V}^{(1)}$，$\boldsymbol{V}^{(2)}$和$\boldsymbol{V}^{(3)}$，组合成近似张量$\hat{X}$；

根据近似张量$\hat{X}$按课程分类进行Top-N授课推荐;

End

算法时间复杂度由每次迭代计算核心张量C_{t+1}的复杂度$O\left(\sum_{i=1,i\neq n}^{3}\left(I_nR_n\prod_{j=1}^{n-1}R_j\prod_{j=n+1}^{3}I_j\right)\right)$、对$\hat{X}$进行SVD计算的复杂度$O\left(I_n\prod_{j=1,j\neq n}^{3}R_jR_n^2\right)$和求近似张量$\hat{X}$的模乘运算复杂度(与求核心张量相同)构成。在算法中，因为有偏权值张量X维度大小I_n远大于分解因子维度R_n，故该算法的复杂度可以简化为$O\left(\prod_{i=1}^{3}I_n\right)$。

4 实验与结果分析

4.1 实验数据集

数据来源于某二本院校4年间728名任课教师、1 683门课程和256 632个真实评价原始记录，实验数据选用了某二级学院40名教师、128门课程以及465个评分(每位教师4年所授课程的学生评分的总平均值按式(9)计算)的记录数据。

设定不同的比重系数r、Dw和w，得到不同的实验数据集。根据定义3，不同的r和Dw对Pbd有影响，表1为r=0.7，Dw=0.4时的Pbd值。r=0.5，Dw=0.2时，

Pbd变化情况如表2所示。

表1　教师信息维表(非全日制)

Teacher ID	Dc	Dw	Ptc	Rm(1)	Sf(1)	Rm(2)	Sf(2)	Pbd
CS001	0.3	0.4	0.4	0.7	0.1	1.0	0.3	0.453 5
CS002	0.3	0.4	0.2	1.0	0.4	1.0	0.4	0.700 5
CS003	0.3	0.4	0.2	1.0	0.1	1.0	0.4	0.550 5
CS004	0.3	1.0	0.2	0.7	0.1	0.7	0.3	0.588 5
CS005	0.4	1.0	0.3	1.0	0.4	1.0	0.4	0.980 5

表2　Pbd变化情况

Teacher ID	Pbd (r=0.5,Dw=0.2)	Pbd (r=0.7,Dw=0.4)	ΔPbd	Rate/%
CS001	0.376 3	0.453 5	0.077 2	20.52
CS002	0.654 0	0.700 5	0.046 5	7.11
CS003	0.504 0	0.550 5	0.046 5	9.23
CS004	0.548 5	0.588 5	0.04	7.29
CS005	0.994 0	0.980 5	−0.013 5	−1.36

w是确认课程难度中教师专家给出的值的比重，根据定义4可以很明显的看出，w的变化对课程难度的评定也是有影响的，如表3所示。

表3　w值对课程难度的影响

Course ID	$\overline{T}$	$\overline{S}$	Qr (w=0.4)	Cdc (w=0.4)	Qr (w=0.6)	Cdc (w=0.6)	ΔCdc
60188	0.89	0.63	0.73	0.79	0.79	0.89	0.10
60254	0.81	0.65	0.71	0.75	0.74	0.81	0.06
60262	0.90	0.69	0.78	0.87	0.82	0.95	0.08
60309	0.70	0.55	0.61	0.55	0.64	0.61	0.05
60318	0.80	0.58	0.67	0.66	0.71	0.75	0.09

表中 $\overline{T}$ 和 $\overline{S}$ 分别表示课程难度调查中教师专家评分的均值和学生评分的均值，ΔCdc表示在两种不同比重系数下Cdc值的差异。

课程评价数据Eva按定义5中的式(9)和式(10)可以得到，如表4所示。

表4　学生评分

Teacher ID	Course ID	Numbers	Student-Score	Eva
CS008	60 058	93	90.97	0.66
CS007	60 064	22	93.50	0.87
CS035	60 095	289	93.40	0.86
CS021	60 185	458	91.41	0.70
CS003	60 188	679	93.14	0.84

本文根据实验数据来源的二本院校的实际情况，设定相关比重系数分别为：r=0.7，Dw=0.4，w=0.4，得出Pbd、Cdc与Eva后，授课推荐实验数据集以侧重评价为例选取，即设定 ρ_1=0.2， ρ_2=0.2和 ρ_3=0.6，得到如表5所示的实验数据集(E)。

表5　实验数据集(E)

Teacher ID	Pbd	Course ID	Cdc	Eva	w_{t_i,c_j,e_k}
CS001	0.453 5	60 308	0.57	0.81	0.690 7
CS002	0.700 5	60 384	0.54	0.83	0.746 1
CS003	0.650 5	60 337	0.11	0.87	0.674 1
CS003	0.650 5	60 339	0.55	0.85	0.7501
CS004	0.550 5	60 337	0.11	0.85	0.642 1
CS004	0.550 5	60 339	0.75	0.88	0.788 1

根据表5的实验数据，按有偏权值张量模型构建稀疏程度为90.92%的张量X，其非0值元素在三维张量模型中的分布如图2所示。

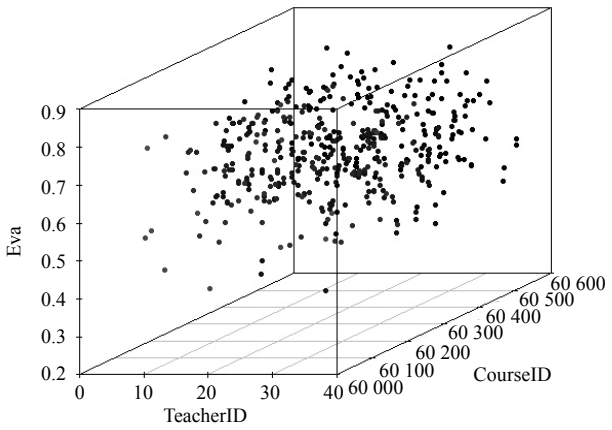

图2　稀疏有偏权值张量X非零元素分布图

4.2　实验及结果分析

实验1：推荐精度与排序准确性

为了保证每门课程在训练集和测试集中都有数据，在实验数据集E中，任选每门课程的20%作为测试集T_E，在余下的80%实验数据E-T_E中随机选取每门课程的60%、70%、80%、90%和100%作为训练集，进行授课推荐实验。在每个不同比例的训练集上，将迭代收敛阈值ε分别设为0.005、0.001、0.000 5和0.000 1。

然后采用文献[17]中的平均绝对误差(mean absolute error, MAE)[18]评价指标来衡量各推荐实验的精度，定义如下：

$$\mathrm{MAE}=\sum_{tc\in T_E}\left|r_{tc}-r_{tc}^{*}\right|\Big/\left|T_E\right| \tag{12}$$

式中，$\left|T_E\right|$ 表示测试集T_E的大小；r_{tc}表示测试集中教师t所授课程c的真实综合有偏权值 w_{t_i,c_j,e_k}；r_{tc}^{*}表示教师t所授课程c的预测值 $\hat{w}_{t_i,c_j,e_k}$。

采用P@N[19](Precision at N)来评价课程的前N个被推荐教师的相关性(实验中N仅考虑1、3、5这3种值)，该评价指标适合TOP_N推荐评测：

$$\mathrm{P@}N=\frac{\#\text{relevant items in the TOP\_}N\text{ items}}{N} \tag{13}$$

经过实验发现，任选E-T_E中60%、70%、80%、90%和100%的实验数据作为训练集实验时，不同迭代收敛阈值ε下MAE结果如图3所示：

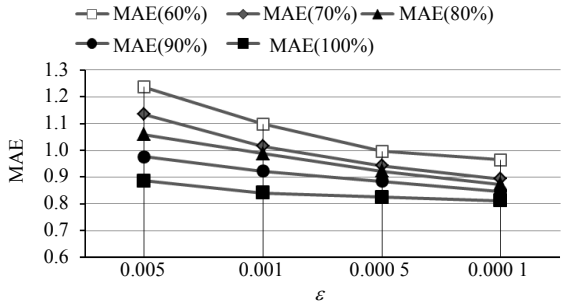

图3 不同比例训练集在不同ε下的MAE对比图

从图中可以看出，使用不同比例训练集的预测精度是不一样的，比例越高，预测精度越好；算法迭代收敛阈值ε越小，预测精度也越好。实验表明，迭代阈值小于或等于0.000 5，采用上述任一比例训练集，其平均绝对误差MAE均小于1。如果训练集大于余下的实验数据集的90%及以上，迭代阈值$\varepsilon\in$[0.000 1，0.005]，也可使MAE值小于1，在这些情况下，可认为预测精度达到要求。

固定迭代阈值ε=0.000 5，训练集任选E-T_E的60%、70%、80%、90%和100%，在取不同N时P@N排序准确性对比如图4所示：

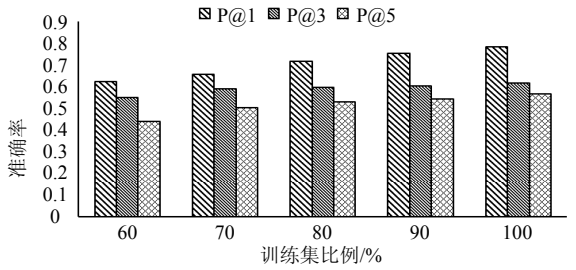

图4 不同比例训练集在不同N下的P@N对比图

从图中可以看出，训练集越大，算法排序准确性越高；N值越小，排序准确性相对来说也会越高。

通过上述实验，证明了模型的MAE均很低且P@N值较理想，这样可基于模型预测每位教师对给定课程的综合有偏权值 $\hat{w}_{t_i,c_j,e_k}$，并为给定课程推荐TOP_N个预测值 $\hat{w}_{t_i,c_j,e_k}$ 最高的授课教师。

实验2：不同比重系数下的推荐对比

用一系列对比实验检验不同比重系数下的推荐差异，在每组对比实验中，约定从各实验数据集中任选每门课程的20%作为测试集，余下的80%作为训练集，算法迭代阈值ε=0.000 5，对比在同一门课程下的Top_5的推荐差异：

1) 不同r、Dw对推荐结果的影响。固定w=0.4，ρ_1=0.6、ρ_2=0.2、ρ_3=0.2，取r=0.7、Dw=0.4，得到实验数据集E_1，取r=0.5、Dw=0.2，得到实验数据集E_2。对E_1和E_2分别作授课推荐实验，推荐结果如表6所示。

表6 不同r、Dw值下的Top_5推荐对比(Course ID=60 264)

E_1		E_2	
Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$
CS022	0.657 0	CS011	0.646 0
CS009	0.644 4	CS009	0.634 1
CS012	0.637 9	CS022	0.629 1
CS011	0.637 9	CS012	0.614 0
CS035	0.634 3	CS035	0.606 4

从表中可以看出，取不同的比重系数r、Dw得到不同的预测值 $\hat{w}_{t_i,c_j,e_k}$，同一门课程下的授课教师推荐顺序也会有所改变。从式(4)～式(6)上可以看出，因每位教师的专业相关值和学位获得方式或许不一样，即Rm$_i$与Dw值不一样，进而改变了Pbd的值；再从式(11)可知，为同一门课程推荐的授课教师的顺序也会有改变。

2) 不同比重系数w对推荐结果的影响。固定r=0.7、Dw=0.4、ρ_1=0.2、ρ_2=0.6、ρ_3=0.2，取w=0.3，得到实验数据集E_3，取w=0.7，得到实验数据集E_4。分别对E_3和E_4作授课推荐实验，推荐结果如表7所示。

表7 不同w值下的Top_5推荐对比(Course ID=60 264)

E_3		E_4	
Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$
CS012	0.465 1	CS012	0.542 0
CS022	0.459 3	CS022	0.536 2
CS035	0.456 6	CS035	0.533 5
CS004	0.453 9	CS004	0.530 7
CS011	0.448 2	CS011	0.525 1

可以看出，随着w的改变(即Cdc改变)，在其他数据和比重系数不变的情况下，预测值 $\hat{w}_{t_i,c_j,e_k}$ 也发生变化，但对同一门课程的授课教师Top_5推荐结果没有改变。究其原因，在式(8)和式(7)中，单一改变w值，仅会改变课程的Cdc值，这时会使预测值发生改变。因是对同一门课程作推荐，依据式(11)，仅Cdc值改变不会影响该门课程的授课教师推荐顺序。

3) 不同ρ_1、ρ_2、ρ_3对推荐结果的影响。固定r=0.7、Dw=0.4、w=0.4，使用前述偏重于Pbd的实验数据集E_1，取ρ_1=0.2、ρ_2=0.6、ρ_3=0.2得到偏重于Cdc实验数据集E_5，和使用前述偏重于Eva实验数据集E。分别对E_1、E_5和E作授课推荐实验，推荐结果如表8所示。

表8　不同偏重系数下的Top_5推荐对比(Course ID=60 264)

E_1		E_5		E	
Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$
CS022	0.657 0	CS012	0.538 9	CS012	0.736 0
CS009	0.644 4	CS022	0.533 1	CS004	0.717 5
CS012	0.637 9	CS035	0.530 4	CS015	0.709 2
CS011	0.637 9	CS004	0.527 6	CS035	0.705 7
CS035	0.634 3	CS011	0.522 0	CS022	0.693 9

从表中的授课教师推荐结果来看，在不同的偏重系数下，实验结果明显有不同的预测值 $\hat{w}_{t_i,c_j,e_k}$，且每组推荐顺序能体现出偏向性。这说明教务部门可以根据实际偏重需要，设定不同的偏重系数ρ_1、ρ_2和ρ_3，可得到所需的推荐结果。

4) 任意组合比重系数对推荐结果的影响。选r=0.3、Dw=0.3、w=0.5、ρ_1=0.2、ρ_2=0.4、ρ_3=0.4得到实验数据集E_6，选r=0.8、Dw=0.5、w=0.2、ρ_1=0.5、ρ_2=0.2、ρ_3=0.3得到实验数据集E_7，选r=0.6、Dw=0.7、w=0.6、ρ_1=0.2、ρ_2=0.3、ρ_3=0.5得到实验数据集E_8。分别对E_6、E_7、E_8作授课推荐实验，推荐结果对比如表9所示。

表9　任意比重系数下的Top_5推荐对比(Course ID=60 264)

E_6		E_7		E_8	
Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$	Teacher ID	$\hat{w}_{t_i,c_j,e_k}$
CS012	0.588 2	CS022	0.639 1	CS004	0.672 9
CS004	0.587 3	CS012	0.634 9	CS012	0.666 1
CS035	0.582 8	CS035	0.625 0	CS035	0.664 7
CS022	0.578 3	CS004	0.610 9	CS022	0.656 6
CS011	0.566 1	CS011	0.603 5	CS015	0.639 5

根据式(4)、式(5)、式(7)和式(8)，若取不同比重系数会得到不同的Pbd和Cdc，当然实验结果也会得到不同的预测值 $\hat{w}_{t_i,c_j,e_k}$，也会改变推荐顺序。

上述实验表明，采用文中的形式化定义规约教师专业基础度、课程难度和课程评价，取综合有偏权值作为三维加权张量模型元素，使用Tucker分解算法，可按不同侧重点精确实现授课推荐。因此，建议每所学校根据自身需求设定授课推荐依据，选取合适的比重系数，获得较理想的推荐结果，有效地提高教学质量。

5　结 束 语

从授课安排无较好的推荐依据的实际问题出发，通过归约教师专业基础、课程难度及教学评价，定义具有偏重性的加权方法，构建基于“教师–课程–评价–权值”四元关系之上的三维有偏权值张量模型，使用基于Tucker的分解算法，成功地实现了精准授课推荐，解决了一直以来授课安排无推荐依据的现状，为实现智能化精准授课推荐找到了一种新方法。如何更好地结合教师年龄、职称、专业方向等特征，更进一步精确地和多样化地实现个性化授课推荐，将是下一步研究的重点。

参 考 文 献

[1] GOLDBERG D, NICHOLS D, OKI B M, et al. Using collaborative filtering to weave an information tapestry[J]. Communications of the ACM, 1992, 35(12): 61-70.

[2] 李聪, 梁昌勇, 马丽. 基于领域最近邻的协同过滤推荐算法[J]. 计算机研究与发展, 2008, 45(9): 1532-1538.
LI Cong, LIANG Chang-yong, MA Li. A collaborative filtering recommendation algorithm based on domain nearest neighbor[J]. Journal of Computer Research and Development, 2008, 45(9): 1532-1538.

[3] 于江德, 李学钰, 樊孝忠, 等. 最大熵模型的事件分类[J]. 电子科技大学学报, 2010, 39(4): 612-616.
YU Jiang-de, LI Xue-yu, FAN Xiao-zhong, et al. Event classification based on maximum entropy model[J]. Journal of University of Electronic Science and Technology of China, 2010, 39(4): 612-616.

[4] 叶茂, 陈勇. 基于分布模型的层次聚类算法[J]. 电子科技大学学报, 2004, 33(2): 171-174.
YE Mao, CHENG Yong. Hierarchical clustering algorithm based on distribution model[J]. Journal of University of Electronic Science and Technology of China, 2004, 33(2): 171-174.

[5] 廖志芳, 李玲, 刘丽敏, 等. 三部图张量分解标签推荐算法[J]. 计算机学报, 2012, 35(12): 2625-2632.
LIAO Zhi-fang, LI Ling, LIU Li-min, et al. A tripartite decomposition of tensor for social tagging[J]. Chinese Journal of Computers, 2012, 35(12): 2625-2632.

[6] 邹本友, 李翠平, 谭力文, 等. 基于用户信任和张量分解的社会网络推荐[J]. 软件学报, 2014, 25(12): 2852-2864.
ZOU Ben-you, LI Cui-ping, TAN Li-wen, et al. Social recommendations based on user trust and tensor factorization[J]. Journal of Software, 2014, 25 (12): 2852-2864.

[7] 廖志芳, 王超群, 李小庆, 等. 张量分解的标签推荐及新用户标签推荐算法[J]. 小型微型计算机系统, 2013, 34(11): 2472-2476.
LIAO Zhi-fang, WANG Chao-qun, LI Xiao-qing, et al. Tag recommendation and new user tag recommendation algorithms based on tensor decomposition[J]. Journal of Chinese Computer Systems, 2013, 34(11): 2472-2476.

[8] 孙玲芳, 冯遵倡. 基于特征加权张量分解的标签推荐算法研究[J]. 江苏科技大学学报: 自然科学版, 2015, 29(6): 574-579.
SUN Ling-fang, FENG Zun-chang. Tag recommendation

algorithm based on feature weighting and tensor decomposition[J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2015, 29(6): 574-579.

[9] 孙玲芳, 李烁朋. 基于K-means聚类与张量分解的社会化标签推荐系统研究[J]. 江苏科技大学学报: 自然科学版, 2012, 26(6): 597-601.
SUN Ling-fang, LI Shuo-peng. Social tagging recommendation system based on K-means cluster and tensor decomposition[J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2012, 26(6): 597-601.

[10] 张昌利, 龚建国, 闫茂德. 基于复杂网络的社会化标签语义相似度分析[J]. 电子科技大学学报, 2012, 41(5): 642-648.
ZHANG Chang-li, GONG Jian-guo, YAN Mao-de. Complex network based semantic similarity measure for social tagging systems[J]. Journal of University of Electronic Science and Technology of China, 2012, 41(5): 642-648.

[11] SYMEONIDIS P, NANOPOULOS A, MANOLOPOULOS Y. A unified framework for providing recommendations in social tagging systems based on ternary semantic analysis[J]. IEEE Transactions on Knowledge & Data Engineering, 2010, 22(2): 179-192.

[12] SYMEONIDIS P, NANOPOULOS A, MANOLOPOULOS Y. Tag recommendations based on tensor dimensionality reduction[C]//Proceedings of the 2008 ACM Conference on Recommender Systems. New York: ACM, 2008: 43-50.

[13] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge & Data Engineering, 2005, 17(6): 734-749.

[14] BADER B W, KOLDA T G. Tensor decompositions and applications[J]. Siam Review, 2009, 51(3): 455-500.

[15] TUCKER L R. Some mathematical notes on three-mode factor analysis[J]. Psychometrika, 1966, 31(3): 279-311.

[16] 余刚, 王知衍, 邵璐, 等. 基于奇异值分解的个性化评论推荐[J]. 电子科技大学学报, 2015, 44(4): 605-610.
YU Gang, WANG Zhi-yan, SHAO Lu, et al. Singular value decomposition-based personalized review recommendation [J]. Journal of University of Electronic Science and Technology of China, 2015, 44(4): 605-610.

[17] 朱郁筱, 吕琳媛. 推荐系统评价指标综述[J]. 电子科技大学学报, 2012, 41(2): 163-175.
ZHU Yu-xiao, LÜ Lin-yuan. Evaluation metrics for recommender systems[J]. Journal of University of Electronic Science and Technology of China, 2012, 41(2): 163-175.

[18] BREESE J S, HECKERMAN D, KADIE C. Empirical analysis of predictive algorithms for collaborative filtering[C]//Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. Madison, USA: ACM, 1998: 43-52.

[19] WANG L, MENG X, ZHANG Y, et al. New approaches to mood-based hybrid collaborative filtering[C]//The Workshop on Context-Aware Movie Recommendation. Barcelona: ACM, 2010: 28-33.

编 辑 叶 芳