

# An Improved Collaborative Filtering Recommendation Model and Method Based on Social Trust

Yingliang Wu<sup>1,2</sup>, Kaimei Huang<sup>1</sup>, Huaidong Yao<sup>3</sup>

<sup>1</sup>Department of Electronic Business, School of Economics and Commerce, South China University of Technology, Guangzhou Guangdong

<sup>2</sup>Business Intelligence Research Center, Institute of Modern Services, South China University of Technology, Guangzhou Guangdong

<sup>3</sup>Haikou Branch of Shanghai Pudong Development Bank, Haikou Hainan  
Email: adylwu@scut.edu.cn, 849602425@qq.com, yaohd\_scut@126.com

Received: Apr. 16<sup>th</sup>, 2019; accepted: Apr. 30<sup>th</sup>, 2019; published: May 7<sup>th</sup>, 2019

---

## Abstract

Trust relationship in the social business environment has a profound impact on consumers' purchase behavior and decision-making, and has become an important factor to support the development of online business activities. Collaborative filtering recommendation algorithm based on user history evaluation data usually faces the problem of data sparseness; that is, the sparse rating data leads to the decline of recommendation quality. In order to solve this problem, the combination of auxiliary data has become an inevitable trend. Therefore, with the development of social media, trust-based social recommendation algorithm has been proved to be an effective solution. However, most of the current algorithms directly use the binary trust relationship of the social network to improve the recommendation quality, without considering the difference in the trust strength of the user for each friend. In order to improve the accuracy of social recommendation algorithm, this paper calculates personal reliability and mutual identify reliability based on social data, and quantifies the social attention matrix based on mutual identify reliability and alleviates the data sparsity problem based on the idea of score matrix pre-filling. The experiment and analysis results based on the real data set of public comments show that the new collaborative filtering recommendation model and algorithm proposed in this paper further improve the recommendation accuracy.

## Keywords

Social Commerce, Social Attention Matrix, Collaborative Filtering Recommendation, Pre-Filling, Personal Reliability, Mutual Identify Reliability

---

# 一种基于社会化信任的改进的协同过滤推荐模型和方法

吴应良<sup>1,2</sup>, 黄开梅<sup>1</sup>, 姚怀栋<sup>3</sup>

<sup>1</sup>华南理工大学经济与贸易学院电子商务系, 广东 广州

<sup>2</sup>华南理工大学现代服务业研究院商务智能研究中心, 广东 广州

<sup>3</sup>上海浦东发展银行海口分行, 海南 海口

Email: adylwu@scut.edu.cn, 849602425@qq.com, yaohd\_scut@126.com

收稿日期: 2019年4月16日; 录用日期: 2019年4月30日; 发布日期: 2019年5月7日

## 摘要

社会化商务环境中的信任关系深刻地影响着消费者的购买行为与决策, 成为了支撑网络商务活动开展的重要因素。基于用户历史评价数据, 协同过滤推荐算法通常面临着数据稀疏的问题, 即评分数据过于稀疏导致推荐质量下降。为了解决这一问题, 结合辅助数据成为一种必然的趋势。因此, 随着社交媒体的发展, 基于信任关系的社会化推荐算法被证明为一种有效的解决方法。然而, 目前大部分算法直接利用社交网络的二值信任关系来提高推荐质量, 没有考虑用户对每个好友信任强度的差异。为了提高社会化推荐算法的准确性, 本文以社交数据为基础, 计算用户个人可信度评分和互鉴可信度评分, 并基于可信评分对社交关注矩阵进行可信量化, 以及基于评分矩阵预填充的思想来缓解数据稀疏性问题。基于大众点评真实数据集的实验与分析结果表明, 本文提出的新的协同过滤推荐模型与算法, 进一步提高了推荐精度。

## 关键词

社会化商务, 社会关注矩阵, 协同过滤推荐, 预填充, 个人可信度评分, 互鉴可信度评分

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着 Facebook、Twitter、微博等社交媒体或平台的出现, 信息量急速增长, 用户在海量信息中很难找到真正中意的商品。与此同时, 在现实社会化商务平台中常出现“诈骗用户”、“广告用户”等一系列恶意用户, 然而, 在已有的社会化信任推荐研究中, 通常直接将社交关注关系表达为社会化信任关系, 但是并非存在社交关注关系就代表用户值得信任。其次, 在庞大的社会化网络中, 社交关注数据同样会出现极度稀疏的情况, 这将导致构造出的社会化信任关系过少。

在社会化商务环境下, 用户频繁的社交行为和交易活动产生了丰富的社交数据。用户在社交上的关注关系能在一定程度上反映了用户之间的社会化信任关系。基于这种社会化关系进行推荐, 不但能缓解

数据稀疏性对传统推荐方法造成的影响,同时也将更加符合社会化商务的口碑营销、信任建议、好友推荐购买等特征。已经有相当多的学者通过使用社会网络来提高推荐系统的准确性[1][2][3][4]。而这类推荐系统的研究一般都是基于其属于的信任网络[5][6][7],用户获得的推荐也往往是来自于信任网络中与距离较近的用户[8][9][10]。

因此,本文提出基于社会化信任改进的协同过滤推荐方法,对社交关注数据进一步进行挖掘和利用,进一步缓解传统协同过滤算法的数据稀疏性问题,最终提高推荐效果。

## 2. 研究现状

### 2.1. 国外研究现状

Massa 等[11]首次提出了基于信任的推荐系统框架,他通过将项目评分矩阵和信任关系矩阵作为系统算法的输入项并提供给评分预测模型,最终得到了预测评分矩阵,改善了传统算法数据稀疏性问题,但其采用的是未加处理的布尔型信任标注值,表达信任的能力较弱,而且由此得到的信任关系也难以抵御恶意用户的攻击。Golbeck 提出了一种 TidalTrust 算法,该算法通过广度优先搜索(BFS)的方法对信任传播方式进行了挖掘[11]。Massa 等[3]在后来对 TidalTrust 进一步改进,提出了一种设定最大信任传递距离的 MoleTrust 算法。Zarghami 等[12]提出了一种衡量信任程度的 T 指数,同时引入了信任传递机制,提高了推荐的精度和覆盖率。Jamali 等[13]将基于项目的协同过滤推荐方法与信任推荐方法相结合,提出了 TrustWalker 随机游走算法,该算法认为信任程度较高的邻居用户比信任程度低的邻居用户在与源用户相似的目标项目上的评分更可靠,由此对源用户的未评分目标项目进行预测评分。Chen 等[14]在传统协同过滤算法的基础上引入了信任的传播机制,提出了 TPCF 信任推荐模型。Bedi 等[15]提出了一种信任动态更新的协同过滤推荐算法,通过用户的评分相似度和共同评分数占比综合推出。

### 2.2. 国内研究现状

国内对信任推荐的相关研究中,出现较早的是唐文等人提出的一种公开网络环境下的信任管理模型[16]。陈婷等[17]提出一种融合社交信息的推荐新方法 Trust-PMF,融合由评分数据产生的相似度和由信任关系信息产生的信任度构建用户的偏好模型生成邻居,再综合目标用户自身的偏好和邻居用户对其评分的影响预测评分。潘一腾等[18]在社会化推荐算法 SocialMF 的基础上,提出一种基于信任关系隐含相似度的推荐算法,综合考虑了评分相似和信任关系隐含相似对每组用户之间信任度的影响,得到了更精确的信任度和推荐模型。王瑞琴等[19]借鉴社会心理学中的信任产生原理,综合考虑多种信任要素在社交信任度量中的作用,提出一种基于多元社交信任的协同过滤推荐算法 CF-CRIS。王占等[20]提出基于信任与用户兴趣变化的协同过滤改进方法,该方法将信任引入到传统协同过滤算法中,构建用户信任模型,用信任的传递特性为用户匹配更多邻居用户,从而可以在一定程度上缓解数据稀疏性等问题。吴应良等[21]提出基于社会网络分析的协同过滤推荐算法,利用网络中存在的信任关系健壮原先的推荐技术,新算法的核心是利用用户间邻接矩阵得到直接信任关系,进一步应用凝聚子群分析方法,根据成员间的可达性和捷径距离找出子群成员得到间接信任关系,摆脱单纯依赖直接信任关系,形成用户综合信任网络,并融入到协同过滤算法中。杜巍等[22]提出基于个性化情景的移动商务信任推荐模型,该模型首先通过挖掘训练集为每个用户找出对其信息需求影响最大的 K 个情景要素,在此基础上融合社会网络与信任机制,分别构建基于个性化情景的移动商务富信任信息推荐模型及基于个性化情景的移动商务稀疏信任信息推荐模型。陈梅梅等[23]提出一种基于标签簇的多构面信任关系定义的方法,在标签聚类得到的标签簇基础上,引用 TF-IDF 思想及 Pearson 相似度定义簇间和簇内信任关系,构建有利于反映不同构面信任强度的信任张量,并融入基于张量分解模型的个性化推荐算法中。薛福亮等[24]通过筛选信任用户作为相

似用户，根据选择的信任用户和目标用户形成一个项目的评分集，并对目标用户未评价过的项目进行评分估算(根据信任用户评分进行简单的评分计算)，将用户间的信任关系依据方差大小进行量化，形成一个调节因子，并将调节因子纳入用户相似性计算，形成相似性用户聚类簇，以在相似用户之间进行交叉推荐。

### 3. 基于可信评分的矩阵预处理

#### 3.1. 可信评分计算

本文从用户社交属性中抽取了影响用户可信度的 5 种重要因子：点评数(Comment)、粉丝数(Fans)、关注数(Follow)、收藏数(Collection)和互粉数(Mutual-Fans)。选择这 5 种影响因子的原因在于它们具有通用性。本文可信评分的计算将由个人可信度评分和互鉴可信度评分两部分组成。可信评分的构成，如图 1 所示。

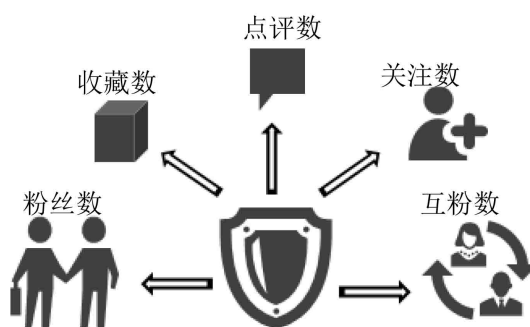


Figure 1. The impact factor of user credibility score  
图 1. 用户可信评分影响因子

##### 1) 个人可信度评分的计算

个人可信度(Personal Reliability, PR)用以衡量推荐系统中社交用户的可信赖程度，由点评数，粉丝数，关注数，收藏数以及互粉数的可信度得分共同计算得出。

##### ① 点评数可信度得分

点评数(Comment, FS)是衡量用户在推荐系统中参与活跃程度的一种重要影响因子。假设不存在恶意评论行为的情况下，用户点评数较多意味着用户参与度高，所以相对于点评数较少甚至几乎不参与点评的用户显得更为忠诚和可信。用户  $u$  的点评数可信度得分的具体计算如公式(1)所示：

$$PR\_CT(u) = \frac{mf(u)}{fs(u)} \times \left( (1 - \varepsilon_1) + \varepsilon_1 \times \sum_{v=1}^k \lg \frac{ct(v)}{|ct(v) - ct(u)|} \right) \quad (1)$$

其中， $PR\_CT(u)$ 表示用户  $u$  个人的点评数得分，而  $ct(u)$ 表示用户  $u$  的点评数。通常情况下，双向的关注关系比单向关注更具有可靠性和研究价值[25]，所以本章的五种影响因子中引入了用户  $u$  的互粉数  $mf(u)$  与粉丝数  $fs(u)$  之比，即互粉比例  $\frac{mf(u)}{fs(u)}$ ，用以强调用户在互粉数影响因子上的重要性。

$\frac{ct(v)}{|ct(v) - ct(u)|}$ 代表点评数的相对影响力，用以强调用户  $u$  相对于用户  $v$  的点评数影响程度。相对影响力越大，则说明用户  $v$  越具影响力，即用户  $v$  更值得用户  $u$  信任，否则反之。 $ct(v)$ 表示除用户  $u$  以外的  $k$  个用户对项目的点评数，用户  $v$  属于这  $k$  个用户集合，后续四种影响因子得分与此同理，不再复述。 $\varepsilon_1$ 代表点评数阻尼系数，具体计算公式在后续给出。

同理，我们可以相似得出粉丝可信度评分、关注数可信度评分、收藏数可信度评分以及互粉数可信度评分，其计算公式分别如下所示：

② 粉丝数可信度得分

$$PR\_FS(u) = \frac{mf(u)}{fs(u)} \times \left( (1 - \varepsilon_2) + \varepsilon_2 \times \sum_{v=1, v \neq u}^k \lg \frac{fs(v)}{|fs(v) - fs(u)|} \right) \quad (2)$$

③ 关注数可信度得分

$$PR\_FW(u) = \frac{mf(u)}{fs(u)} \times \left( (1 - \varepsilon_3) + \varepsilon_3 \times \sum_{v=1, v \neq u}^k \lg \frac{fw(v)}{|fw(v) - fw(u)|} \right) \quad (3)$$

④ 收藏数可信度得分

$$PR\_CN(u) = \frac{mf(u)}{fs(u)} \times \left( (1 - \varepsilon_4) + \varepsilon_4 \times \sum_{v=1, v \neq u}^k \lg \frac{cn(v)}{|cn(v) - cn(u)|} \right) \quad (4)$$

⑤ 互粉数可信度得分

$$PR\_MF(u) = \frac{mf(u)}{fs(u)} \times \left( (1 - \varepsilon_5) + \varepsilon_5 \times \sum_{v=1, v \neq u}^k \lg \frac{mf(v)}{|mf(v) - mf(u)|} \right) \quad (5)$$

本章的阻尼系数与 PageRank 算法中的阻尼系数并不相同，它可随机地由指定用户  $v$  的可信度评分除以评分集合中  $k$  个用户评分的最大值计算得出，其随机性体现在结果并不是一个固定值，而是随着不同目标用户的不同可信度评分进行变动的数值，由此计算出来的影响因子更具合理性。上述 5 个影响因子的计算公式中，权重值皆可利用用户各项指标在所有用户中的占比随机计算得出。为了让计算出来的结果限定在 0 到 1 的范围内，5 种阻尼系数  $\varepsilon$  的取值范围也在 0 到 1 之间，见公式(6)所示：

$$\begin{aligned} \varepsilon_1 &= \frac{ct(v)}{\max[ct(n)]}, \varepsilon_2 = \frac{fs(v)}{\max[fs(n)]}, \varepsilon_3 = \frac{fw(v)}{\max[fw(n)]}, \\ \varepsilon_4 &= \frac{cn(v)}{\max[cn(n)]}, \varepsilon_5 = \frac{mf(v)}{\max[mf(n)]}, n = 1, 2, \dots, k \end{aligned} \quad (6)$$

通过上述的公式(1)~(5)分别计算出每一位用户的各项影响因子后，将它们加总求和可以得到依据用户自身属性的可信评分，如公式(7)所示：

$$PR(u) = PR\_CT(u) + PR\_FS(u) + PR\_FW(u) + PR\_CN(u) + PR\_MF(u) \quad (7)$$

2) 互鉴可信度评分的计算

互鉴可信度(Mutual Identify Reliability, MI)是指相互鉴别后的可信度，即目标用户的可信度评分由其他用户的可信度相互鉴别后计算得出。

引入互鉴可信度的原因在于：在社会化商务环境下的推荐系统中，常常出现用户或商家恶意“刷分”的情况，系统中的用户同样也可能会为了追求特殊“信誉”和“地位”而疯狂地对自身社交属性进行“刷分”，由此可能会导致个人可信度评分虚高的情况。

为了使用户可信度评判过程更加公平公正，本章以用户的互鉴可信度评分作为用户最终的可信得分评判指标。互鉴可信度评分的计算如公式(8)所示：

$$MI(u) = 1 - \rho + \rho \times \sum_{v=1, v \neq u}^k \lg \frac{PR(v)}{|PR(v) - PR(u)|}, u = 1, 2, \dots, k \quad (8)$$



其中,  $\rho$  表示阻尼系数, 计算方法与其他阻尼系数相似。

### 3.2. 社交关注矩阵可信量化

在社交关注矩阵中, 如果用户 A 关注用户 B 则将 TA, B 标注为 1, 否则标注为 0, 但是在社会化商务中, 关注并不意味着信任, 每一位用户在系统中的自身属性和行为中间隐含着一些属于自己的可信赖信息, 并可以传达给其他用户。如果能在用户社交关注关系中引入目标用户的可信赖程度, 则可以构造出可能反映源用户对目标用户预期可信程度较高的信任关系矩阵。

对于这种可信赖程度的衡量, 本章已对用户社交信息中的五种重要影响因子进行了提取并评估得出了互鉴可信度评分, 可将其作为目标用户的可信赖程度衡量指标, 如表 1 所示。此时, 用户之间的信任度值由原先的 0 和 1 二值细分量化为[0, 1]区间中的值, 由此更加细致地表达用户之间更加可信的社会化信任关系, 有利于提高后续算法的推荐质量。

**Table 1.** A social attention matrix based on mutual trust

**表 1.** 互鉴可信度的社交关注矩阵

	A	B	C	D	E
A		$T_{A,B} \times MI_B$	$T_{A,C} \times MI_C$	$T_{A,D} \times MI_D$	$T_{A,E} \times MI_E$
B	$T_{B,A} \times MI_A$		$T_{B,C} \times MI_C$	$T_{B,D} \times MI_D$	$T_{B,E} \times MI_E$
C	$T_{C,A} \times MI_A$	$T_{C,B} \times MI_B$		$T_{C,D} \times MI_D$	$T_{C,E} \times MI_E$
D	$T_{D,A} \times MI_A$	$T_{D,B} \times MI_B$	$T_{D,C} \times MI_C$		$T_{D,E} \times MI_E$
E	$T_{E,A} \times MI_A$	$T_{E,B} \times MI_B$	$T_{E,C} \times MI_C$	$T_{E,D} \times MI_D$	

### 3.3. 项目评分矩阵预填充

传统协同过滤算法中存在数据稀疏性问题, 本文基于评分矩阵预填充的思想来缓解数据稀疏性问题, 利用上文得到的可信评分模型来对稀疏的评分矩阵进行预填充: 首先, 需要根据用户集中的任意两位用户之间的互鉴可信评分 MI 和公共评分项目 PI 建立起用户两两之间的预估可信度(Forecast Reliability, FR)。在建立互信度的过程中, 源用户通常会向目标用户进行咨询, 如果目标用户的反馈(项目评分)与自己预期大体相同或偏差较小时, 该目标用户的反馈建议可视作有效且会被采纳, 否则反之。通过计算两位用户之间评分差距较小的项目集合数(Undifferentiated Item, UI)占公共评分项目总数的(Public Item, PI)比例计算得到预估可信度 FR 可由公式(9)计算出:

$$FR(u, v) = \frac{|UI_{u,v}|}{|PI_{u,v}|} \tag{9}$$

其中,  $UI_{u,v}$  表示  $PI_{u,v}$  中用户评分差距小的项目集合。对于任意  $i \in PI_{u,v}$ , 若用户  $u$  对项目  $i$  的评分  $r_{u,i}$  满足  $|r_{u,v} - p_{u,i}| < \theta$ , 则  $i \in UI_{u,v}$ 。  $P_{u,i}$  表示用户  $u$  对项目  $i$  评分预测函数, 由公式(10)计算:

$$P_{u,i} = \bar{r}_u + MI(v) \times (r_{v,i} - \bar{r}_v) \tag{10}$$

然后, 根据预估可信度 FR 的大小进行降序排序, 选取排名前 20 且评分不为 0 的用户作为目标用户  $u$  的可信邻居集  $RNu$ 。下一步, 根据  $RNu$  中用户的评分, 对目标用户  $u$  尚未评分项目进行评分预测, 预测评分的计算如公式(11)所示:

$$r'_{u,i} = \bar{r}_u + \frac{\sum_{k \in RNu} FR(u, k) \times (r_{k,i} - \bar{r}_k)}{\sum_{k \in RNu} FR(u, k)} \tag{11}$$

最后，将  $r'_{u,i}$  填充到初始评分矩阵的相应位置，直到未评分项目的预测评分填充完毕，由此构造出了能提高初始评分矩阵数据密度的可信评分预估矩阵。

## 4. 基于可信评分预估矩阵的相似性计算

### 4.1. 用户相似性计算

传统的用户相似性度量可由传统的相关相似性算法计算得出，如公式(12)所示：

$$\text{SIM\_A}(u, v) = \frac{\sum_{k \in \text{PI}_{u,v}} (r_{u,k} - \bar{r}_u) \cdot (r_{v,k} - \bar{r}_v)}{\sqrt{\sum_{k \in \text{PI}_{u,v}} (r_{u,k} - \bar{r}_u)^2} \cdot \sqrt{\sum_{k \in \text{PI}_{u,v}} (r_{v,k} - \bar{r}_v)^2}} \quad (12)$$

其中  $r_{u,k}$  和  $r_{v,k}$  分别表示用户  $u$  和用户  $v$  对共同评分项目  $k$  的评分， $\bar{r}_u$  和  $\bar{r}_v$  分别表示用户  $u$  和用户  $v$  已有评分项目的平均评分。为了将结果限定在  $[0, 1]$  范围内，需要归一化处理，如公式(13)所示：

$$\text{Sim\_A}(u, v) = \frac{1 + \text{SIM\_A}(u, v)}{2} \quad (13)$$

### 4.2. 预测评分计算

对于预测评分的计算，本章算法采用传统协同过滤算法评分预测公式，如公式(14)所示：

$$P_{u,i} = \bar{r}_u + \frac{\sum_{v \in \text{NS}(u), \text{RS}(i)} \text{Sim\_A}(u, v) \times (r_{v,i} - \bar{r}_v)}{\sum_{v \in \text{NS}(u), \text{RS}(i)} \text{Sim\_A}(u, v)} \quad (14)$$

其中， $\bar{r}_u$  和  $\bar{r}_v$  分别表示用户  $u$  和  $v$  的评分均值， $\text{NS}(u)$  和  $\text{RS}(i)$  分别表示与用户  $u$  相似的 Top-N 相似邻居集和对项目  $i$  存在评分的用户集。

## 5. 算法实施步骤

图 2 展示了基于社会化关系与信任传播的协同过滤算法的算法步骤，具体步骤如下：

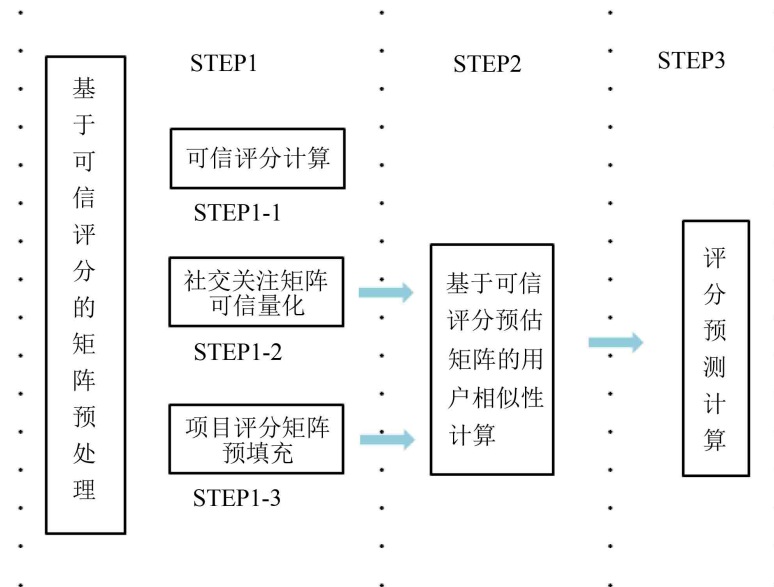


Figure 2. The implementation step diagram of the algorithm in this paper  
图 2. 本文算法实施步骤图

STEP1-1 可信评分的计算：根据推荐系统的用户数据库中，提取用户的点评数，关注数，粉丝数，收藏数，利用公式(1)至公式(7)离线计算每一位用户的五种影响因子得分和个人可信度评分，再利用公式(8)计算出每一位用户的互鉴可信度评分。

STEP1-2 社交关注矩阵可信量化：将 STEP1-1 中计算出来的互鉴可信度融入到初始社交关注关系矩阵中，如表 1 所示，得到改善可信水平后的社会化信任矩阵。

STEP1-3 项目评分矩阵预填充：将 STEP1-1 输出的互鉴可信度利用公式(9)至公式(11)，先转化为预估可信度后对用户未评分项目进行评分预测，并补充到初始用户项目评分矩阵的相应位置，最后得到数据密度改善的用户项目评分矩阵。

STEP2 基于信任传播的相似性计算：基于 STEP1-3 填充后的项目评分矩阵，利用公式(12)至公式(13)计算相似度矩阵。

STEP3 预测评分计算：根据公式(14)对用户的评分进行预测，将预测评分按照降序排序后形成推荐列表，位于前列的是目标用户最可能乐意接受的项目，以此进行实时的在线推荐。

## 6. 实验及结果分析

### 6.1. 实验数据集和评估标准

大众点评是国内领先的 Web2.0 生活信息交易平台(<http://www.dianping.com/>)，是最早在全球建立的第三方消费购物点评网站之一。大众点评具备虚拟社交社区，社区中的用户具有丰富可挖掘的用户属性，信任数据和项目评分数据，非常适用于本章的信任推荐研究。在使用八爪鱼软件采集后，本文对部分空值和重复数据进行处理，并统计出每一位用户的项目评分数量和关注其他用户的数量以及每一种项目被用户评分的数量列表，作为算法各部分的输入项。

大众点评数据集的特征统计如表 2 所示。可以看出，数据集的评分和信任数据都相当稀疏。为了让实验更具可靠性，将预处理后的数据集分别分成五组在实验中进行交叉验证，每组数据均随机划分为 80% 训练集用以构造本章推荐模型和 20% 测试集，最终结果取五组数据集的平均值，用以检验模型的性能。评价推荐算法的优劣程度可以由很多指标来评价。为了比较提出的算法，本文选择平均绝对误差 MAE、准确率 Precision、召回率 Recall 以及综合评价指标 F 值作为其评价指标。

Table 2. Characteristic statistics of experimental data set

表 2. 实验数据集特征统计

数据集	数据类型	用户数量	项目数量	评分数量	信任数量	数据稀疏等级
大众点评	评分数据(1~5)	1854	10152	154,350	-	99.17%
	信任数据(0, 1)	1854	-	-	123,847	96.4%

### 6.2. 实验结果讨论

#### 1) 可信评分影响因子充分合理性验证

在可信评分计算中，影响因子得分引入的充分与否会通过影响用户可信评分，进而影响算法推荐的效果。所以，需要对引入的五种影响因子充分性与合理性进行验证。在大众点评数据集上，测试缺失任意一种影响因子对 MAE 值的影响，最后 MAE 值取五个训练数据分组集计算的平均结果，如表 3 所示。

从表 3 可以看出，五种影响因子得分中的任意四种结合效果都没有全部包含的精度高。互粉数和点评数的缺失对算法推荐效果的影响较大。

#### 2) 社交关注和评分矩阵处理影响验证



**Table 3.** Calculating MAE results with different influence factors of credible score  
**表 3.** 不同可信评分影响因子下 MAE 的计算结果

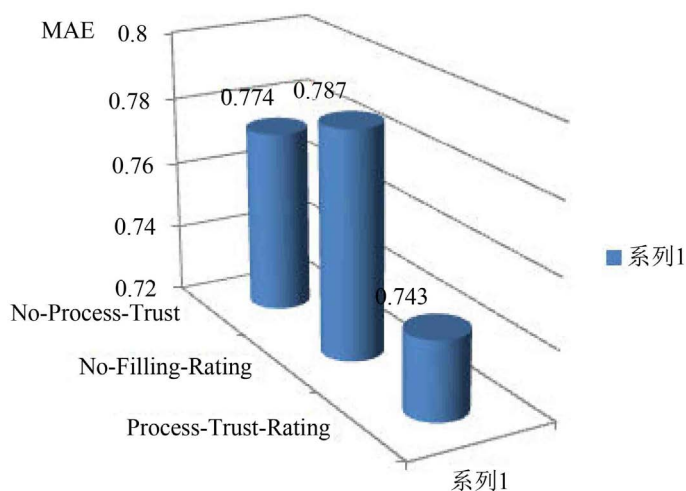
点评数得分	粉丝数得分	关注数得分	收藏数得分	互粉数得分	MAE
PR_CT	PR_FS	PR_FW	PR_CN	PR_MF	大众点评
√	√	√	√		0.792
√	√	√		√	0.751
√	√		√	√	0.778
√		√	√	√	0.776
	√	√	√	√	0.791
√	√	√	√	√	0.743

本章在基于可信评分的矩阵处理步骤中，不仅对社交关注矩阵进行了可信量化，还对评分矩阵进行了预填充，因此需要对这两部分对算法推荐效果的影响进行检验。

MAE 值取五个训练数据分组集计算的平均结果。测试分为三种情况：

- ① 未可信量化社交关注矩阵但预填充评分矩阵(No-Process-Trust)；
- ② 未预填充评分矩阵但可信量化社交关注矩阵(No-Filling-Rating)；
- ③ 既量化社交关注矩阵又预填充评分矩阵(Process-Trust-Rating)。

如图 3 所示，推荐精度最高的是本章即量化处理信任数据又预填充评分矩阵的情况 (Process-Trust-Rating)，其次是(No-Process-Trust)，最低的是(No-Filling-Rating)。可见，如果在仅提高信任数据的可信水平而未预填充评分数据的情况下，推荐效果并不可观。同样，如果仅预填充了评分数据而未对社交关注矩阵量化处理的情况下，推荐效果虽有提升，但是效果并不突出。由此说明，本章基于可信评分对社交矩阵的量化计算和对评分矩阵的预填充的处理步骤能使推荐效果更加显著。



**Figure 3.** The impact of social attention and score matrix processing on MAE  
**图 3.** 社交关注和评分矩阵处理对 MAE 的影响

## 7. 结论及展望

已有研究表明，用户之间的信任关系对改善推荐系统的推荐效果起到不可忽视的作用。在社会化商务环境下，通过社交数据反映出的社会化关系进行推荐，不但能提高传统推荐方法的推荐效果，同时也

将更加符合社会化商务的口碑营销、信任建议等特征。因此，本文通过充分挖掘社会化商务环境下的用户社交和项目评分数据，计算用户可信评分，然后基于可信评分对社交矩阵进行量化计算以及对评分矩阵进行预填充，从而提高了推荐效果。

在下一步工作中，值得我们进一步研究是：根据信任的动态特性，考虑构造一种附带时间因素的效用函数，对信任网络中的用户节点接受推荐后的效用情况进行反馈，从而模拟此信任变化过程，进一步提高推荐算法的抗恶意用户攻击性能。

## 基金项目

国家自然科学基金项目“管理科学理论和方法的综合集成研究”(70440011)，国家社会科学基金项目“分享经济下基于 TRIZ 理论的网络约租车服务创新研究”(16BGL190)，国家社会科学基金项目“基于关联数据的政府数据开放研究”(14BTQ009)。

## 参考文献

- [1] Ma, H., Yang, H.-X. and Michael, R. (2008) SoRec: Social Recommendation Using Probabilistic Matrix Factorization. In: *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, IEEE Press, New York, 931-941. <https://doi.org/10.1145/1458082.1458205>
- [2] 冯勇, 李军平, 徐红艳, 等. 基于社会网络分析的协同推荐方法改进[J]. 计算机应用, 2013, 33(3): 841-844.
- [3] Massa, P. and Avesani, P. (2007) Trust-Aware Recommender Systems. *Proceedings of the 2007 ACM Conference on Recommender Systems*, Minneapolis, 19-20 October 2007, 17-24. <https://doi.org/10.1145/1297231.1297235>
- [4] 俞琰, 邱广华. 用户兴趣变化感知的重新启动随机游走推荐算法研究[J]. 现代图书情报技术, 2012(4): 48-53.
- [5] Mohsen, J. and Martin, E. (2010) A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks. *Proceedings of the Fourth ACM Conference on Recommender Systems*, Barcelona, 26-30 September 2010, 135-142.
- [6] Yang, X.-W., Steck, H. and Liu, Y. (2012) Circle-Based Recommendation in Online Social Networks. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Beijing, 12-16 August 2012, 155-165. <https://doi.org/10.1145/2339530.2339728>
- [7] Ma, H., Zhou, D.-Y. and Liu, C. (2011) Recommender Systems with Social Regularization. *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*, Hong Kong, 9-12 February 2011, 287-297. <https://doi.org/10.1145/1935826.1935877>
- [8] 李春, 朱珍敏, 高晓芳. 基于邻居决策的协同过滤推荐算法[J]. 计算机工程, 2010, 36(13): 34-36.
- [9] 罗辛, 欧阳元新, 熊璋. 通过相似度支持度优化基于 K 近邻的协同过滤算法[J]. 计算机学报, 2010, 33(8): 1437-1445.
- [10] Massa, P. and Bhattacharjee, B. (2005) Using Trust in Recommender Systems: An Experimental Analysis. *Proceedings of Conference on Trust Management*, Paris, 23-26 May 2005, 221-235. [https://doi.org/10.1007/978-3-540-24747-0\\_17](https://doi.org/10.1007/978-3-540-24747-0_17)
- [11] Golbeck, J.A. (2005) Computing and Applying Trust in Web-Based Social Networks. Ph.D. Thesis.
- [12] Zarghami, A., Fazeli, S., Dokoohaki, N., et al. (2009) Social Trust-Aware Recommendation System: A T-Index Approach. *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technologies*, Milan, 15-18 September 2009, 85-90. <https://doi.org/10.1109/WI-IAT.2009.237>
- [13] Jamali, M. and Ester, M. (2009) Trustwalker: A Random Walk Model for Combining Trust-Based and Item-Based Recommendation. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Paris, 28 June-1 July 2009, 397-406. <https://doi.org/10.1145/1557019.1557067>
- [14] Chen, X.C., Liu, R.J. and Chang, H.Y. (2010) Research of Collaborative Filtering Recommendation Algorithm Based on Trust Propagation Model. *Computer Application and System Modeling*, Taiyuan, 22-24 October 2010, Vol. 4, 177-183.
- [15] Bedi, P. and Sharma, R. (2012) Trust Based Recommender System Using Ant Colony for Trust Computation. *Expert Systems with Applications*, **39**, 1183-1190. <https://doi.org/10.1016/j.eswa.2011.07.124>
- [16] 唐文, 陈钟. 基于模糊集合理论的主观信任管理模型研究[J]. 软件学报, 2003(8): 1401-1408.
- [17] 陈婷, 朱青, 周梦溪, 等. 社交网络环境下基于信任的推荐算法[J]. 软件学报, 2017, 28(3): 721-731.

- [18] 潘一腾, 何发智, 于海平. 一种基于信任关系隐含相似度的社会化推荐算法[J]. 计算机学报, 2018, 41(1): 65-81.
- [19] 王瑞琴, 蒋云良, 李一啸, 等. 一种基于多元社交信任的协同过滤推荐算法[J]. 计算机研究与发展, 2016, 53(6): 1389-1399.
- [20] 王占, 林岩. 基于信任与用户兴趣变化的协同过滤方法研究[J]. 情报学报, 2017, 36(2): 197-205.
- [21] 吴应良, 姚怀栋, 李成安. 一种引入间接信任关系的改进协同过滤推荐算法[J]. 现代图书情报技术, 2015(9): 38-45.
- [22] 杜巍, 高长元. 基于个性化情景的移动商务信任推荐模型研究[J]. 情报科学, 2017, 35(10): 23-29.
- [23] 陈梅梅, 薛康杰. 基于标签簇多构面信任关系的个性化推荐算法研究[J]. 数据分析与知识发现, 2017, 1(5): 94-101.
- [24] 薛福亮, 刘君玲. 基于用户间信任关系改进的协同过滤推荐方法[J]. 数据分析与知识发现, 2017, 1(7): 90-99.
- [25] 闫光辉, 刘晓飞, 王梦阳. 基于链接的微博用户可信度研究[J]. 计算机应用研究, 2015, 32(10): 2910-2913+2917.

#### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2168-5843, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [ecl@hanspub.org](mailto:ecl@hanspub.org)