



计算机集成制造系统
Computer Integrated Manufacturing Systems
ISSN 1006-5911, CN 11-5946/TP

《计算机集成制造系统》网络首发论文

题目: 基于灰色关联预测与信任云混合算法的方案推荐
作者: 杨珍, 耿秀丽
收稿日期: 2018-06-07
网络首发日期: 2018-09-29
引用格式: 杨珍, 耿秀丽. 基于灰色关联预测与信任云混合算法的方案推荐. 计算机集成制造系统. <http://kns.cnki.net/kcms/detail/11.5946.TP.20180927.1410.003.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于灰色关联预测与信任云混合算法的方案推荐

杨 珍，耿秀丽

(上海理工大学 管理学院，上海 200093)

摘 要：面向用户需求的方案配置是产品或服务开发的重要内容。电子商务等领域的个性化方案推荐方法研究给方案设计带来新的思路，可以根据用户的需求推荐出合适的方案配置。针对传统推荐方法中的数据稀疏性问题，常用的方法通常受到数据量的制约，采用灰色关联预测法计算方案评分数据间的相关系数，以预测空缺的评分数据；针对面向新用户的冷启动问题，考虑用户兴趣特征相似度和基于信任云的用户对方案评分的相似性，计算用户间的综合相似度，将合适的方案推荐给新用户。最后，以汽车方案推荐为例进行方法验证，并通过与协同过滤，云模型等推荐算法进行对比，证明该方法的有效性。

关键词：方案推荐，稀疏性；冷启动；灰色关联预测；信任云

中图分类号：TH122;N94

文献标识码：A

Scheme Recommendation Based on Grey Prediction and Trust Cloud Hybrid Algorithm

Yang Zhen, Geng Xiuli

Business School, University of Shanghai for Science and Technology, Shanghai 200093, China

Abstract: The configuration of solutions for user needs is an important part of product or service development. In e-commerce and other fields, the research of personalized program recommendation methods brings new ideas to the design of the program. The appropriate scheme configuration can be recommended according to the user's needs. In terms of data sparsity of traditional recommendation method, common methods were limited to the amount of data. Therefore, the grey correlation prediction method was used to calculate the correlation coefficient between the program score data to predict the vacancy score data. In terms of the cold start problem for new users, the similarity of users' interest features and the similarity of users' ratings based on trust cloud were considered. Then the comprehensive similarity between users was calculated to recommend the appropriate solution to the new user. Finally, the method validation was carried out by taking the car scheme recommendation as an example. The effectiveness of the method was proved by comparing with the recommendation algorithms such as collaborative filtering and cloud model.

Keywords: Scheme recommendation, sparsity; cold start; grey correlation prediction; trust cloud

0 引 言

随着市场竞争愈加激烈，迅速发展的企业日益增多，顾客对产品或服务的要求也更加多元化、个性化。由于信息不对称、专业知识欠缺等原因，顾客难以在众多企业提供的多个方案中选择适合自己的方案，方案的个性推荐成为增强企业竞争力的关键因素。针对产品方案推荐的研究较少，但商品/服务的推荐已有了一定的研究。例如，文献[1-3]利用协同过滤、隐语义模型和线性回归技术将WEB 服务推荐给目标用户；文献[4-5]分别采用贝叶斯网络和用户偏好方法实现对用户的商品推荐。协同过滤作为推荐技术的常用方法，被广泛应用于各类推荐系统，如图书推荐系统、电子商务和社

收稿日期：2018-06-07；**修订日期：**2018-09-19。Received 07 June 2018; accepted 19 Sep. 2018.

基金项目：国家自然科学基金资助项目（71301104，51475290）；高等学校博士学科点专项科研基金资助项目（20133120120002，20120073110096）；上海市教育委员会科研创新资助项目（14YZ088）；上海市一流学科资助项目（S1201YLXK）；沪江基金资助项目（A14006）。

交网络等推荐系统。而常用的协同过滤算法存在数据稀疏性、冷启动和用户偏好不一等问题，已有文献针对其中的某一问题给予了解决方案。如文献[6]为减轻稀疏性对推荐结果的影响，采用矩阵聚类降维技术，引入偏好方差找出用户最近邻并推荐；文献[7]针对冷启动问题，提出采用时间权重计算的方式约束用户对项目的偏爱程度，并利用用户-项目-属性三部图关系，为新用户推荐目标项目；文献[8]针对电子商务中推荐算法对目标用户定位不准确等问题，提出利用用户兴趣偏好性挖掘相似目标用户，准确定位并推荐。文献[6-8]中的方法考虑了三类问题中的其中一类问题，不能有效地实现产品/服务的推荐。

本文综合考虑上述文献中的推荐问题，提出了采用灰色关联与信任云结合的混合推荐算法。信任云是在云模型基础上融入了用户信任度，增加方案推荐的可信度。而实际上云模型不仅适用于方案推荐过程，在方案评价中也得到广泛的应用，如文献[9-10]都采用云模型定量化方案评价信息，通过云优化的算法对方案进行分析；文献[11]针对方案评价方法未考虑定性指标中不定因素，提出利用云模型弥补层次分析法，通过云发生器实现指标的有效量化，再对方案进行评价以及优选。综上所述，云模型在方案评价中，将信息指标量化转换为数值，可以增加方案的可行性。

针对某些用户对产品或服务的评分较少而导致的数据稀疏性问题，已有研究采用余弦、皮尔森相似度将数据聚类，并预测填充空缺数据，减轻稀疏性对推荐结果的影响。但是个性化推荐算法中，余弦、皮尔森相似度因数据量剧增难以对数据快速聚类，而灰色关联度既不受限于数据量的影响，又能考虑用户项目评分的不定因素，在不减少非近邻数据的损失的情况下，能够计算项目评分的相似性，预测空缺的评分数据，弥补稀疏性。因此，本文提出采用灰色关联度代替传统推荐算法中的相似度，用于预测方案评分数据，弥补稀疏性。

针对新用户数据库中无记录，难以推荐合适项目的问题，即冷启动问题，文献[12]引入用户-标签以及项目-标签的加权相似度，对目标新用户相应的项目评分进行预测以及推荐；文献[13]提出运用粒来描述用户和产品，构建粒关联规则的指标，挖掘出用户和产品间的关联规则，并对新用户进行推荐；文献[14]提出采用动态潜在因素模型与用户属性相结合的方法，并随用户和推荐者间的交互关系，模型自动调整其系数并合理地推荐；文献[15]通过用户项目间的评分相似度，利用最近邻预测新用户的评分，解决冷启动问题。上述这些关于冷启动的解决方案，都是基于已有项目、标签、用户间相似度的改进，来预测最近邻集并给予推荐，在一定程度上可以减轻冷启动问题，但未考虑到用户主观信任的模糊性和不确定性问题。因此，本文提出将信任度引入云模型中，构建信任云，利用信任云相似度与用户特征相似度的综合值，来预测目标方案评分，减轻冷启动问题。

信任模型是为解决用户、项目间关系的信任模糊性和不确定性提出的，信任模型因应用范围的不同，构建方式有所不同。如文献[16]通过改进推荐属性的相似度，与服务推荐构建行为构成可信联盟；文献[17]通过信任模型筛选恶意评价的用户，得到信任度高的用户集以及候选服务集。由于云模型本身的不确定性和模糊性，学者们提出信任云模型，如文献[18]着重研究一维主观信任云，文献[19]提出多维信任云的模型，既能解决主观信任云的不确定性，又能提供更准确的信任数据。考虑到信任云能够解决用户主观不确定性问题，本文提出建立信任云模型，得到方案推荐信任度，来处理用户

主观性问题。考虑方案匹配过程中，用户兴趣不一问题，引入新旧用户兴趣特征相似度，既能减轻新用户引起的冷启动问题，又能解决兴趣相异的问题，如文献[19]利用兴趣信任度有效地解决面向新用户的冷启动问题。

综上所述，本文将方案推荐过程分为三个部分，第一部分主要通过灰色关联相似度预测空缺的评分数据，弥补数据稀疏性；第二部分通过建立信任云，求出用户间的信任相似度；第三部分根据用户兴趣特征的相似度，与信任相似度加权得到综合相似度，获取最近邻集并匹配最优方案推荐给目标用户。

1 基于灰色关联预测弥补数据稀疏性

灰色关联分析是灰色系统理论的重要组成部分，主要研究系统内部因素间的关联程度^[14]。灰色关联分析不同于传统的推荐方法，可以减少非近邻数据的信息损失，弥补稀疏性，提高预测精度。考虑数据稀疏性导致推荐不准确的问题，提出采用灰色关联预测空缺数据，弥补稀疏性。灰色关联分析利用少量的已知信息或者隐含信息，融入系统的不定因素，挖掘系统间各因素的特征，建立因素之间的关系模型。本文基于用户项目的协同过滤算法，利用灰色关联分析计算用户-项目间评分的关联度，得到项目间的灰色关联度，并用来预测用户对其他项目的评分。本文采用的灰色关联度计算过程如下。

将每个用户 i 对方案 $S=\{s_1, s_2, s_3, \dots, s_n\}$ 的评分 r_{ij} 表示为一个序列，则待估矩阵为（1）所示：

$$X_0(i) = \{X_0(1), X_0(2), X_0(3), \dots, X_0(m)\} \quad i = 1, 2, 3, \dots, m \quad (1)$$

为了计算各序列的关联度，首先将序列值进行标准化，即将每个序列值除以该序列值总和，从而得到数据标准值。

$$X_0(i) = \frac{1}{\text{count}_i} \sum_{j \in A} r_{ij} \quad i = 1, 2, 3, \dots, m \quad (2)$$

式中， $X_0(i)$ 表示第 i 个用户所评分项目的评分均值， count_i 表示第 i 个用户已经评分的项目数量， A 表示第 i 个用户已评分的项目编号。被比较的序列值表示如（3）所示：

$$X_j(i) = (X_j(1), X_j(2), \dots, X_j(m)) \quad (3)$$

式中， $X_0(k)$ 和 $X_j(k)$ 都是表示对应向量的第 k 个元素的值，则两者间的灰色关联系数表示为：

$$\xi_j(k) = \frac{\min_j \min_k |X_0(k) - X_j(k)| + \rho \max_j \max_k |X_0(k) - X_j(k)|}{|X_0(k) - X_j(k)| + \rho \max_j \max_k |X_0(k) - X_j(k)|} \quad (4)$$

式中， $\min_j \min_k |X_0(k) - X_j(k)|$ 表示参考向量 $X_0(k)$ 与 $X_j(k)$ 第 k 个差值的绝对最小值，未评分的不予参与计算。同理可知， $\max_j \max_k |X_0(k) - X_j(k)|$ 为两级之间的绝对最大值。当 $|X_0(k) - X_j(k)|$ 取最

小值，关系系数 ξ_j 为 1，当 $|X_0(k) - X_j(k)|$ 取最大值时，关系系数 $\xi_j(k)$ 最小，且 $\rho \in (0,1)$ 。灰色关联度 γ_j 表示比较向量和参考向量关系的紧密程度，公式如下：

$$\gamma_j = \frac{1}{m} \sum_{k=1}^m \xi_j(k) \quad (5)$$

本文在给用户推荐方案之前，考虑数据稀疏性对推荐结果的影响，基于已有数据采用灰色关联算法来预测并填充空缺数据。在协同过滤推荐算法中，评分矩阵 $R=(r_{ij})_{m \times n}$ ，评分项目对应的灰色关联度为：

$$\tau=(\gamma_1, \gamma_2, \dots, \gamma_j) \quad j=1,2,3, \dots, n \quad (6)$$

根据评分项目 s 的近邻集合 $N(s)$ ，预测目标用户 i 对 s 的评分，计算公式如下所示：

$$pred(i, s) = X_0(i) + \gamma_j \times (r_{ij} - X_0(i)) \quad (7)$$

2 基于云模型的方案匹配

通过灰色关联度以及预测公式对数据稀疏性的弥补，得到完整的用户方案评分数据。但在推荐方案的过程中，由于人为主观性因素等影响，比如有些用户偏好某种方案，就会将分值打高，而有些用户对某种方案有偏见，就会将分打的特别低，这样会对最终的相似性方案匹配结果，与实际推荐结果偏差较大。因此，本文综合考虑多方面的因素，比如，方案推荐的信任度，方案推荐的频率，用户兴趣度等来约束推荐结果的精度。以下是云模型相关定义，信任云的构建过程，以及与用户信任云与用户兴趣的综合方案匹配过程的描述。

2.1 云模型的相关定义

云模型是李德毅院士提出的处理定性描述和定量概念间不确定性的转换模型，通过结合自然语言中的模糊性和随机性，实现定性语言与定量数值间的转换，为研究数据挖掘过程中不确定信息的处理提供了新方法。本文考虑到用户方案评分信息的主观性和模糊性，提出采用云模型实现定性定量数据的转换，云模型相关定义如下。

定义 1^[18] 设 U 是一个用数值表示的定量论域， C 是 U 上的定性概念。若定量值 $x \in U$ 是定性概念 C 上的一次随机实现， x 对 C 的确定度 $\mu(x) \in [0,1]$ ，是有稳定倾向的随机数： $\mu: U \rightarrow [0,1]$ ， $\forall x \in U: x \rightarrow \mu(x)$ ，则 x 在论域 U 上的分布称为云(Cloud)，记为云 $C(x)$ ；每个 x 称为一个云滴(Drop)。

定义 2^[18] 正向云是将独立参数表示为云的数字特征，描述事务的不确定性。采用参数期望 E_x ，熵 E_n ，超熵 He 三个特征描述云，即 $C(E_x, E_n, He)$ 。当随机变量 x 满足： $x \sim N(E_x, E_n^2)$ ，其中 $E_n' \sim N(E_n, He^2)$ ，

定性概念 C 的确定度满足公式 $\mu(x) = e^{-\frac{(x-E_x)^2}{2E_n^2}}$ 。 E_x, E_n, He 三个特征指标的计算方式如 (8) 到 (10) 所示：

$$E_x = \frac{1}{N} \sum_{f=1}^n x_c \quad (8)$$

$$E_n = \sqrt{\frac{\pi}{2} \sum_{f=1}^n |x_c - E_x|} \quad (9)$$

$$He = \sqrt{S^2 - E_n^2} \quad (10)$$

式中，（8）到（10）将定性描述转换为云的特征指标，其中公式（10）中， $S^2 = \sum_{f=1}^n (x_c - \bar{x})^2$ ，本文采用正向云发生器，将特征指标转换为云滴，具体过程如下。

输入：表示定性概念的三个特征值 (E_x, E_n, He)

输出：云滴 $Drop(x_1, \mu(x_1))$ ， $Drop(x_2, \mu(x_2))$ ，…… $Drop(x_N, \mu(x_N))$

步骤 1：生成以 E_n 为期望值， He 为方差的一个正向随机数 $E_n' = N(E_n, He)$

步骤 2：生成以 E_x 为期望值， E_n' 为方差的一个正向随机数 $x = N(E_x, E_n')$

步骤 3：计算 $\mu(x_c) = e^{-\frac{(x_c - E_x)^2}{2(E_n')^2}}$

步骤 4：带有确定度 $\mu(x_c)$ 的 x_c 称为数域中的一个云滴；

步骤 5：重复 1-4 直到产生要求的 N 个云滴 $(x_c, \mu(x_c))$ ($c=1,2,3,\dots,N$)

论域 U 可以是一维或者多维的形式，云是在隶属度 $\mu(x)$ 的任意一次实现的随机变量，由无次序性的云滴组成，实现了定性概念在数量上的转变，反映了概念特征。云滴不是简单的模糊或随机，而是具有隶属度约束的随机变量，隶属度表示云滴能够代表定性概念的程度，云滴出现概率与云滴确定度成正比，符合人们的主观理解。而云模型的核心处理过程主要分两种，正向云发生器和逆向云发生器。本文采用正向云算法，将定性描述的整体向量 $C(E_x, E_n, He)$ ，转换为描述个体定量数值的云滴。以云 $(0.5, 0.05, 0.01)$ 的三个特征值及 $n=1000$ 为输入值，通过正向云发生器输出云图，如图 1 所示。

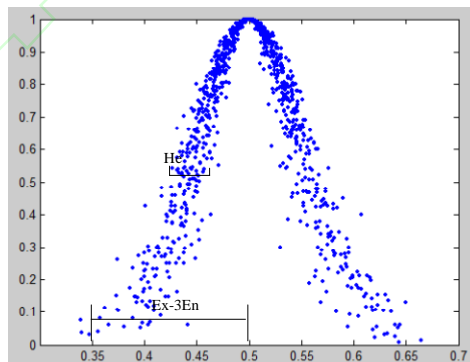


图 1 云图

其中，期望 E_x 表示云滴在论域空间分布的期望值，即最能代表定性概念的点；熵 E_n 表示定性概念的不确定量，反映云的离散程度；超熵 He 是熵的熵，描述云的厚度。

2.2 方案信任云的构建

信任关系主要分为直接信任和间接信任，由主观经验可以分成若干个等级的信任度，具有模糊性和随机性。而云模型中的云可以将定性和定量之间的关系进行转换，提出采用云模型将用户方案的定量数据转换成云滴，形成信任云。下面介绍方案的信任度及信任云的相关概念。

定义 3^[18]信任是指对推荐者的诚信，能力以及可靠性的认可程度，表现出主观概念，具有模糊随机性。由于信任程度不同，因此将其划分为几个等级，即信任等级。每个信任等级可以赋予不同的实数值 q ($q \in [0,1]$)，如 0 代表不信任，1 代表信任， q 称为信任度。

定义 4 信任度空间是等级的集合，每个等级表示为一个信任度，将等级映射到信任空间上。信任空间上的模型可以设置为如下分段函数：

$$f(x) = \begin{cases} \text{不信任} & 0 \leq x \leq 0.25 \\ \text{有点信任} & 0.25 \leq x \leq 0.5 \\ \text{很信任} & 0.5 < x < 0.75 \\ \text{完全信任} & 0.75 \leq x \leq 1 \end{cases} \quad (11)$$

定义 5 信任云是用云模型将若干个定性概念转换为若干个云滴，将信任度空间为 $T=[0,1]$ 表示为云的定量论域， $x \in T$ 是指定性概念上的一次定量信任评价， x 是由论域上的隶属度 $\mu(x)$ 决定的具有稳定倾向的随机数。信任云的定性表示是由三个数字特征构成的元组 $C(E_x, E_n, He)$ ， E_x 表示信任云在空间上分布的期望， E_n 是信任云的熵，反映了信任概念的云滴的离散程度， He 表示超熵，即熵的熵，由熵的随机模型性决定。

信任云的定性表示可以由多种方式得到，一是通过综合评价的方法，即专家经验和问卷给出信任等级的划分以及期望信任度。二是通过推荐者和被推荐者的实体信任云可以推理得到。本文是从主观概念出发，采用综合评价的方法得出信任空间中定性概念的信任云，信任空间的每个信任等级可以用单独的信任云表示，对于“不信任”和“信任”的期望值采用 0 和 1 的信任值，其他信任等级用区间平均值表示等级的期望值，也表示为信任度 θ 。熵能够反映信任概念的云滴离散程度以及能被接受信任的确定度范围。

定义 6^[19] 信任云的匹配

如果有两朵云 TC_1 和 TC_2 ，通过正向云发生器将 TC_1 信任云生成 N 个云滴 (x_i, μ_i) ，若 x_i 在云 TC_2 的隶属度为 μ_i' ，则 $\frac{1}{N} \sum \mu_i'$ 称为 TC_1 和 TC_2 的相似度，记为 φ ，假设 TC_1 云为综合信任云， TC_2 为相似云，匹配过程如下图 2 所示。

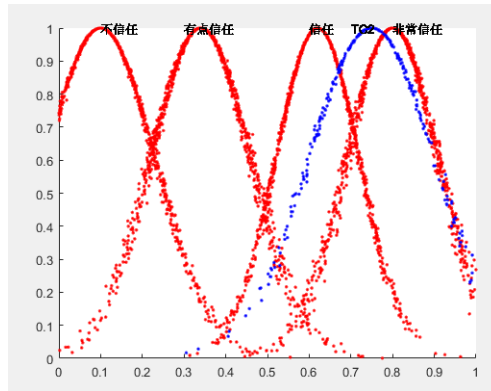


图 2 相似云匹配图

以某用户为例，与其他用户的方案信任云相似度分别为 $\varphi_1, \varphi_2, \varphi_3 \dots \varphi_s$ ，将这些相似度的平均值作为某用户最终的相似度 x ，而 x 在 $f(x)$ 中的某个区间内，该区间的信任度对应的权重可以用该区间的中间值作为权重。

2.3 基于用户兴趣和信任度的方案匹配模型

为用户推荐方案前，方案选择不仅考虑信任云推荐的相似度，而且会考虑用户兴趣等原因。如果某个用户偏好某项方案，导致推荐结果与实际情况偏差太多，则推荐结果不精确。考虑到上述问题，在方案匹配前引入用户兴趣度，约束用户方案的推荐效率。用户兴趣值的获取是以用户需求、兴趣或者偏好为基础，得到评分值。因此，在用户兴趣的方案匹配之前，需要建立用户兴趣模型。对于新用户，主要通过用户的咨询情况得到兴趣度，对于旧用户，主要通过实际反馈的信息，并将其转化用户的兴趣分值。根据专家的评分，将用户兴趣相似度以及对应的权重定义如表 1 所示。

表 1 用户兴趣度表

相似度	兴趣程度	权重(w)
0.8-1	非常满意	0.9
0.6-0.8	满意	0.7
0.4-0.6	一般	0.5
0.2-0.4	不喜欢	0.3
0-0.2	讨厌	0.1

对于某一方案，用户 A 在内容上对用户 B 感兴趣的方案信息相似，则推断用户 A 对此方案也感兴趣。本文的用户兴趣特征用 $f = \{f_1, f_2, f_3, \dots, f_q\}$ 表示，其中 $f_k, k = 1, 2, 3, \dots, q$ 是一种特征值，每个用户 $U = \{u_1, u_2, u_3, \dots, u_p\}$ ，其中 u_t 表示用户， $t = 1, 2, 3, \dots, p$ ，用户对特征评分用 r_{tk} 表示。针对余弦相似性未考虑用户评分标准的不同，即用户偏向打高分还是低分的问题，修正的余弦相似性通过减去用户对项目的评分均值来改善余弦相似性存在的缺陷，提出采用修正的余弦相似度给予解决，对用户兴趣特征的相关性进行计算，如公式（12）所示。

$$sim(u,v) = \frac{\sum_{t \in U} (r_{tk} - \bar{r}_t) \times (r_{sk} - \bar{r}_s)}{\sqrt{\sum_{t \in U} (r_{tk} - \bar{r}_t)^2} \times \sqrt{\sum_{s \in V} (r_{sk} - \bar{r}_s)^2}} \quad (12)$$

考虑推荐者评分的主观性以及兴趣偏好的影响，将基于信任云的用户相似度 $\varphi = \frac{1}{N} \sum \mu_i'$ ，信任云依赖度 θ ，与用户兴趣相似度 $sim(u,v)$ 以及用户对方案的评分关联度 γ_j 加权求和，作为综合相似度，如公式 (13)，进行方案匹配，将最近邻用户的相应方案进行推荐。

$$sim' = (1-w)\theta\varphi + w \times sim(u,v) \quad (13)$$

式中， w 表示为表 1 中对应的相似度权重， sim' 表示信任云与用户兴趣特征的综合相似度。为了避免用户评分的主观性，本文引入了用户兴趣特征值，即将用户对方案特征的偏好程度，根据专家给的评分等级进行量化，并给予不同的权重值，用户方案匹配以及推荐的步骤：

步骤 1: 建立用户-兴趣特征，用户-方案间的关系矩阵，针对用户评分的稀疏性，采用灰色关联度进行预测并填充空缺数据。

步骤 2: 对用户推荐方案的方式，一方面考虑旧用户的推荐，即在新用户和旧用户之间产生信任关系，通过信任云的相似度进行方案匹配。另一方面，考虑新用户的兴趣特征相似度，找寻最近邻用户的方案。

步骤 3: 将基于信任云的用户相似度与用户特征相似度加权求和，得到综合相似度。通过相似度进行方案匹配并排序，选择相似性大的用户对应方案给予推荐。

3 案例分析

某汽车企业想要提高销售效率，希望给用户推荐相关产品的购买方案。通过整理以往顾客购买产品历史记录信息，本文用户对象是各种企业，企业里人群种类的不同，购买产品的需求也就不一样。本文选取 10 个用户 $U = \{u_1, u_2, u_3, \dots, u_{10}\}$ ，以及相关产品类型分为轿车 (s_1)，SUV (s_2)，MPV (s_3)，跑车 (s_4)，皮卡 (s_5)，集卡 (s_6)，微面 (s_7) 7 类；用户对产品的兴趣特征，其实是产品功能以及非功能特征，包括：变速箱 (f_1)，制动性能 (f_2)，燃油经济性 (f_3)，安全性 (f_4)，轮胎性能 (f_5)，汽车通过性 (f_6)，售后服务 (f_7)，动力性 (f_8)，汽车噪音 (f_8)。如下表 2 是已有用户对汽车的评分，其中 1-6 分别表示对方案的满意度，依次是非常差，差，一般，还行，好，非常好；表 3 是汽车特征以及对应的量化分值，表 4 是用户对兴趣特征的评分，其中 u_{11} 表示新用户。

表 2 用户对汽车的评分（填充前）

	s_1	s_2	s_3	s_4	s_5	s_6	s_7
u_1	2	1	3	4	5	6	3
u_2	3	3	6	4	?	4	6
u_3	4	1	3	1	3	5	3
u_4	1	6	5	2	6	4	4
u_5	5	2	?	6	4	1	5

u_6	2	6	2	4	1	2	1
u_7	4	1	3	1	5	6	4
u_8	6	5	1	3	6	4	5
u_9	1	?	4	4	3	2	3
u_{10}	4	3	3	1	2	5	5

表 3 汽车特征表

f_1	手自一体(4)	自动(3)	手动(2)	
f_2	制动效能(4)	抗衰退性(3)	方向稳定性(2)	
f_3	耗油少(4)	耗油较少(3)	耗油一般(2)	耗油较多(1)
f_4	ABS+EBD(4)	安全气囊(3)	分盘式制动器(2)	鼓式制动器(1)
f_5	防抱死好(4)	防抱死中(3)	防抱死差(1)	
f_6	转弯半径(3)	越障能力(2)	离去角(1)	
f_7	维修(4)	保养(3)	定期检查(2)	折旧(1)
f_8	加速时间(4)	最高速度(3)	爬坡能力(2)	高(1)
f_9	低(4)	中(3)	较高(2)	高(1)

表 4 用户的兴趣特征的评分

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}
f_1	4	4	2	4	2	3	1	3	1	4	3
f_2	2	4	2	3	3	4	3	2	2	2	3
f_3	3	4	1	4	2	2	1	3	4	1	1
f_4	2	1	4	4	3	2	2	1	2	3	2
f_5	3	?	4	3	3	1	4	3	1	2	2
f_6	2	3	3	2	1	2	2	4	3	1	1
f_7	3	4	1	?	2	3	4	2	2	2	2
f_8	2	3	3	4	4	2	1	2	2	4	3
f_9	1	3	2	3	2	4	2	1	2	4	2

首先，采用灰色关联度，预测表 2 和表 4 的空缺数据，以公式 (4) 和公式 (5) 来计算相似度。如表 2 中用户 u_5 对汽车 s_3 的评分可以用公式 (4) 计算相关系数，其中 ρ 取值为 0.5， u_2 和 u_1 的灰色关联系数 $\xi_1=1$ 或者 0.47，根据公式 (5) 得到关联度 $\gamma_1 \approx 0.73$ 。同理可得， u_2 和 u_3 的灰色关联度 $\gamma_2 \approx 0.78$ ，依次类推， u_2 与 $u_4, u_5, u_6, u_7, u_8, u_9$ 的关联度为 $\gamma_3 \approx 0.67, \gamma_4 \approx 0.78, \gamma_5 \approx 0.67, \gamma_6 \approx 0.78, \gamma_7 \approx 0.75, \gamma_8 \approx 0.83, \gamma_9 \approx 0.78$ 。因此，选取用户 u_9 对产品 s 的评分相关度，根据公式 (7) 得到预测的分值， $pred(u_2, s_5) \approx 4.58$ 。同理， u_5 对 s_3 的预测评分，通过公式 (4) 和 (5) 计算，得到 u_9 与 u_5

的关联度最大为 0.83，预测值 $pred(u_5,s_3) \approx 3.85$ ； u_9 对 s_2 的预测评分，利用公式（4）和（5）计算得到， u_8 与 u_9 的关联度最大为 0.83，预测值 $pred(u_9,s_2) \approx 4.81$ ，表 2 填数据填充完整后的结果如表 5 所示。

表 5 用户对汽车的评分（填充后）							
	s_1	s_2	s_3	s_4	s_5	s_6	s_7
u_1	2	1	3	4	5	6	3
u_2	3	3	6	4	4.58	4	6
u_3	4	1	3	1	3	5	3
u_4	1	6	5	2	6	4	4
u_5	5	2	3.85	6	4	1	5
u_6	2	6	2	4	1	2	1
u_7	4	1	3	1	5	6	4
u_8	6	5	1	3	6	4	5
u_9	1	4.81	4	4	3	2	3
u_{10}	4	3	3	1	2	5	5

表 4 的数据填充方式，也是采用公式（4）-（6）进行计算，得到 u_5 和 u_2 的灰色关联度为 0.83， $pred(u_2,f_5) \approx 3.46$ ，同理，得到 u_6 与 u_4 的灰色关联度为 0.83， $pred(u_4,f_7) \approx 2.93$ ，表 3 的填充后的数据如表 6 所示。

表 6 用户的兴趣特征的评分（填充后）											
	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}
f_1	4	4	2	4	2	3	1	3	1	4	3
f_2	2	4	2	3	3	4	3	2	2	2	3
f_3	3	4	1	4	2	2	1	3	4	1	1
f_4	2	1	4	4	3	2	2	1	2	3	2
f_5	3	3.46	4	3	3	1	4	3	1	2	2
f_6	2	3	3	2	1	2	2	4	3	1	1
f_7	3	4	1	2.93	2	3	4	2	2	2	2
f_8	2	3	3	4	4	2	1	2	2	4	3
f_9	1	3	2	3	2	4	2	1	2	4	2

其次，根据专家打分建立方案信任云，如图 2 所示。本文将云的方案信任云与图 2 中的信任云进行匹配，得出相似度 φ ，相似度的大小决定信任依赖度 θ 所属区间。为了求出与信任云的相似度，需要将表 5 的方案评分缩小 10 倍，得到表 7 的单位化的结果，利用云模型的正向云发生器，每个用户对应的信任云特征值 $C(E_x, E_n, He)$ ，与信任云的相似度，信任度如下表 8 所示。

表 7 用户对汽车的评分（单位化）

	s_1	s_2	s_3	s_4	s_5	s_6	s_7
u_1	0.2	0.1	0.3	0.4	0.5	0.6	0.3
u_2	0.3	0.3	0.6	0.4	0.458	0.4	0.6
u_3	0.4	0.1	0.3	0.1	0.3	0.5	0.3
u_4	0.1	0.6	0.5	0.2	0.6	0.4	0.4
u_5	0.5	0.2	0.385	0.6	0.4	0.1	0.5
u_6	0.2	0.6	0.2	0.4	0.1	0.2	0.1
u_7	0.4	0.1	0.3	0.1	0.5	0.6	0.4
u_8	0.6	0.5	0.1	0.3	0.6	0.4	0.5
u_9	0.1	0.481	0.4	0.4	0.3	0.2	0.3
u_{10}	0.4	0.3	0.3	0.1	0.2	0.5	0.5

表 8 信任云特征值

	Ex	En	He	φ	θ	$\varphi\theta$
u_1	0.343	0.168	0.320	0.713	0.375	0.267
u_2	0.437	0.124	0.128	0.6995	0.625	0.437
u_3	0.286	0.133	0.611	0.718	0.125	0.089
u_4	4	0.179	0.679	0.715	0.375	0.268
u_5	0.384	0.167	0.584	0.702	0.375	0.263
u_6	0.257	0.174	0.511	0.71	0.125	0.089
u_7	0.343	0.189	0.191	0.713	0.375	0.267
u_8	0.429	0.174	0.462	0.723	0.625	0.452
u_9	0.312	0.124	0.399	0.711	0.125	0.089
u_{10}	0.329	0.148	0.193	0.694	0.125	0.087

为约束方案信任云以及用户评分的主观性，以及解决冷启动问题，即对新用户难以推荐方案的情况下，本文基于用户兴趣评分，如表 6 所示，即根据调查新用户的兴趣特征，与已有的旧用户兴趣特征相匹配，得到兴趣相似度，如表 9 所示，根据与表 8 的信任云相似度加权综合，得到新相似度并推荐。

表 9 用户特征相似度

sim	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}
u_1	0.49	0.22	0.32	0.13	0.31	0.01	0.52	0.22	0.14	0.10

u_2	0.7	0.16	0.24	0.28	0.15	0.47	0.04	0.25	0.12
u_3		0.04	0.4	0.57	0.10	0.04	0.41	0.16	0.08
u_4			0.54	0.12	0.55	0.35	0.06	0.48	0.38
u_5				0.17	0.01	0.47	0.37	0.43	0.64
u_6					0.01	0.45	0.07	0.32	0.32
u_7						0.07	0.36	0.35	0.03
u_8							0.23	0.57	0.37
u_9								0.60	0.71
u_{10}									0.7

以新用户 u_{11} 为例，采用公式 (13) 计算与其他用户的综合相似度，并进行匹配，其中的权重值参考表 1，求得 u_{11} 与用户的相似度如表 10 所示。

表 10 新用户相似度

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
u_{11}	0.10	0.12	0.08	0.38	0.64	0.32	0.03	0.37	0.71	0.7
$\varphi\theta$	0.267	0.437	0.089	0.268	0.263	0.089	0.267	0.452	0.089	0.087
w	0.1	0.1	0.1	0.3	0.7	0.3	0.1	0.3	0.7	0.7
sim'	0.2503	0.4053	0.0081	0.3016	0.5269	0.1583	0.2433	0.4274	0.5237	0.5161

由表 10 可以看出 u_{11} 和用户的综合相似度，明显与 u_5 和 u_9 以及 u_{10} 的相似度最高，因此，将用户 u_5 、 u_9 和 u_{10} 对应方案给予推荐。

3.1 结果对比分析

本文采用灰色关联度与信任云的混合算法，为企业用户推荐合适的汽车方案。首先利用灰色关联度，对稀疏的评分数据进行预测填充，弥补稀疏性；其次根据已有方案评分数据，计算用户的信任云相似度；获取用户兴趣值，计算用户兴趣特征的相似度；最后采用信任云与用户兴趣的综合相似度，匹配与新用户合适的方案并给予推荐。由于本文整体方法执行过程比较繁琐，未将其与其他推荐算法进行比较。在同等数量用户的评分数据中，采用所提方法的第二个步骤，即，方案匹配算法，与其他推荐算法进行比较。其中，方案匹配算法考虑用户特征各异，与传统信任云的计算方式不同，输入用户特征相似度和信任度的加权综合值，输出方案推荐结果。本文通过比较不同算法的执行时间，来突出方案匹配算法的优势。在配置 win10 系统，RAM 2GB，Inter CORE i5，matlab 9.0 平台下，将本文采用的方案匹配算法与基于灰色关联的协调过滤推荐方法，以及基于项目的协同过滤方案匹配推荐算法以及基于信任云的方案匹配推荐方法进行对比，结果如下图 3 所示。

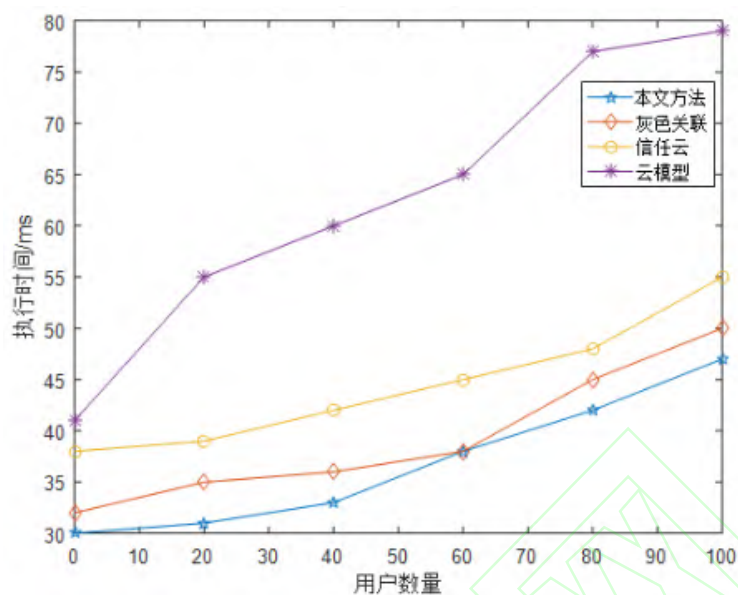


图3 所有算法的方案匹配执行效率对比图

如上图所示，基于同等数量的用户对应的评分数据中，本文采用的方案匹配算法与其他推荐算法相比，方案匹配算法的执行时间最低，效率最高。基于灰色关联的推荐效率仅次于方案匹配算法，而信任云在匹配前的相似度计算比较繁琐，因此基于信任云的推荐算法执行时间较长。由此得出，上述四种算法相对比中，本文采用的方案匹配算法，执行速度最快，推荐效率明显最高。

4 结束语

产品或服务都是当今制造业重点研究方向，甚至产品与服务的结合也是制造业的发展趋势。因此，要想提供好的产品或服务，就要配置出满足顾客需求的产品或服务方案，推荐给新客户，从而提高企业效益。如何准确定位用户需求，将合适的配置方案推荐给新用户，推荐方法显得尤为重要。而传统推荐方法存在的数据稀疏，冷启动等问题，造成推荐结果不准确。因此，本文提出灰色关联预测与信任云的综合算法，对海量产品数据进行挖掘并推荐方案。即通过灰色关联预测空缺数据，减轻数据稀疏性，通过信任云和用户兴趣特征的综合相似度解决冷启动以及推荐效率等问题，其主要特点如下：

(1) 考虑到评分数据稀疏性以及基于余弦，皮尔森相似度的聚类会随着数据量变化影响推荐结果的精确性，提出采用灰色关联预测，计算方案评分的相似度，以填充空缺数据。相对于传统推荐算法而言，灰色关联预测减少了非近邻的信息损失，减轻稀疏性。

(2) 考虑冷启动问题，本文考虑用户兴趣特征相似度和基于信任云的用户相似度，计算加权的综合相似度，进行方案匹配并推荐给新用户。

通过汽车方案推荐的案例，验证所提方法的有效性。但是本文灰色关联相似度并未进行优化，因此，下一步考虑对灰色关联度进行改进，并用于方案推荐。

参考文献：

[1] 胡堰,彭启民,胡晓惠.一种基于隐语义概率模型的个性化 Web 服务推荐方法[J]. 计算机研究与发展, 2014,

51(8):1781-1793.

- [2] 张佩云,陈恩红,黄波.基于社会网络面向个性化需求的可信服务推荐[J]. 通信学报, 2013, 34(12):49-59.
- [3] 刘树栋,孟祥武.一种基于移动用户位置的网络服务推荐方法[J]. 软件学报, 2014(11):2556-2574.
- [4] 张光前,白雪.基于消费性格的新商品推荐方法[J]. 管理科学, 2015(2):60-68.
- [5] 谢意,陈德人,干红华.基于浏览偏好挖掘的实时商品推荐方法[J]. 计算机应用, 2011, 31(1):89-92.
- [6] 严隽薇,黄勋,刘敏,等.基于本体用户兴趣模型的个性化推荐算法[J]. 计算机集成制造系统, 2010, 16(12):2757-2762.
- [7] 于洪,李俊华.一种解决新项目冷启动问题的推荐算法[J]. 软件学报, 2015, 26(6):1395-1408.
- [8] 贺桂和.基于用户偏好挖掘的电子商务协同过滤推荐算法研究[J]. 情报科学, 2013(12):38-42.
- [9] 耿秀丽,董雪琦.粗糙信息公理与云模型集成的方案评价方法[J]. 计算机集成制造系统, 2017, 23(3):661-669.
- [10] 弓晓敏,于长锐.基于云模型的改进TODIM方案评价方法[J]. 系统工程与电子技术, 2018, 40(7):1539-1547.
- [11] 罗圆,姚令侃.基于云模型的山区铁路选线设计方案评价模型[J]. 工程设计学报, 2015(4):372-380.
- [12] 肖扬,王道平,杨岑.基于三部图网络结构的知识推荐算法[J]. 计算机应用研究, 2015(2):386-390.
- [13] 巫文佳,何旭.基于粒关联规则的冷启动推荐方法[J]. 计算机科学, 2014, 41(3):71-75.
- [14] Lin R, Wang W. Research on the dynamic latent factor model combining attributes of user and item[C].IEEE International Conference on Information and Automation. IEEE, 2017.
- [15] Sang A, Vishwakarma S K. A ranking based recommender system for cold start & data sparsity problem[C].Tenth International Conference on Contemporary Computing. IEEE Computer Society, 2017:1-3.
- [16] 王海艳,杨文彬,王随昌,等.基于可信联盟的服务推荐方法[J].计算机学报, 2014, 37(2):301-311.
- [17] 吴文明,刘茜萍.基于可信相似用户的服务推荐方法[J]. 计算机工程, 2016, 42(11):57-63.
- [18] 蔡红云,杜瑞忠,田俊峰,基于多维信任云的信任模型研究[J]. 计算机应用, 2012, 32(1):5-7.
- [19] 邵淳亮,谢怡,孙知信.一种基于用户兴趣的位置服务推荐算法[J]. 计算机技术与发展, 2017, 27(9):64-69.

作者简介:

杨 珍(1994-), 女, 江苏泰州人, 硕士研究生, 研究方向: 信息系统与决策支持, E-mail: jenny_0513@163.com;
+耿秀丽(1984-), 女, 山东东营人, 副教授, 博士, 研究方向: 产品服务系统、服务科学、产品/服务方案设计技术等。