

哈长城市群高端装备制造业供应商 协同过滤推荐方法研究

杨迪 周磊 张昕 李松江
(长春理工大学, 吉林长春 130022)

摘要: 为发挥哈长城市群高端装备制造业科技服务平台优势, 完善科技服务流程, 提升企业运作效率, 本文对供应商推荐方法进行研究, 构建供应商推荐服务模块。通过信息采集、构建和降低企业特征向量, 计算企业用户相似度并进行排名, 形成供应商推荐列表, 完成供应商智能推荐, 实现供需双方精准对接。该方法能够节省企业运营时间、提升企业的综合竞争力。

关键词: 高端装备制造业; 供应商; 推荐机制; 企业特征; 服务平台; 哈长城市群

中图分类号: TP520

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2020.06.011

Research on Collaborative Filtering Recommendation Method for Supplier of High-end Equipment Manufacturing Industry in Harbin-Changchun Urban Agglomeration

YANG Di, ZHOU Lei, ZHANG Xin, LI Songjiang

(Changchun University of Science and Technology, Changchun 130022)

Abstract: In order to give full play to advantages of high-end equipment manufacturing technology service platform of Harbin-Changchun urban agglomeration, and improve the process of technology service and advancing the operational efficiency of enterprises, this paper conducts research on the supplier recommendation mechanism and builds a service module of supplier recommendation. By collecting enterprise information, constructing and reducing enterprise feature vectors, calculating and ranking similarity among enterprises, a supplier recommendation list is formed for completing supplier intelligent recommendation. The supplier recommendation module achieves accurate docking between supply and demand, saves enterprise operating time, and improves the overall competitiveness of enterprises.

Keywords: high-end equipment manufacturing, supplier, recommendation mechanism, enterprise characteristic, service platform, Harbin-Changchun urban agglomeration

0 引言

高端装备制造业又称先进装备制造业, 是

指生产制造高技术、高附加值的先进工业设施设备的行业, 在我国国民经济中占有举足轻重的地位。但高端装备制造业目前还存在企业众多、分

作者简介: 杨迪 (1990—), 女, 长春理工大学讲师, 主要方向: 智能交通 (通信作者); 周磊 (1985—), 男, 长春理工大学硕士研究生, 主要研究方向: 大数据; 张昕 (1974—), 男, 长春理工大学讲师, 主要研究方向: 计算机网络; 李松江 (1984—), 男, 长春理工大学研究生导师, 主要研究方向: 数据挖掘。

基金项目: 国家重点研发计划现代服务业重点专项项目“哈长城市群综合科技服务平台研发与应用示范”(2017YFB1401800)。

收稿时间: 2020 年 3 月 23 日。

布广、资源分布不均衡等问题。在众多资源中寻找优质资源就显得尤为重要,特别是寻找优质供应商更为困难。优质供应商能够与供需方实现精准对接,从而提高企业生产研发效率,强化企业间的合作,达到协同发展、互利共赢的目的。寻找优质供应商是企业间合作成功的重要前提,直接关系到每个企业的生产和运营效率。在为企业用户寻找供应商方面,哈长城市群高端装备制造业科技服务平台(以下简称“高端装备制造业服务平台”)发挥了重要的作用。高端装备制造业服务平台运用供应商协同过滤推荐机制,形成精确的企业用户所需供应商列表,有效地节省了企业用户选择供应商的时间,提升了企业用户使用高端装备制造业服务平台的服务体验,增加了企业用户的供应商储备数量,为企业用户未来业务调整和战略规划提供了有力的帮助。该推荐机制的实施充分发挥了平台优势,更加科学地选择和管理供应商,实现了供需双方优势互补,有效提升了企业用户与供应商的协作能力。本文将对找寻优质供应商的整体运作模式展开研究,旨在解决企业用户对供应商的甄选问题。

1 供应商推荐

1.1 供需关系

高端装备制造业中所谓的“高端”主要表现在3个方面:一是技术含量很高,集中在知识、技术密集、体系结构等,表现出多学科跨领域及高精尖技术和跨地区协作等方面;二是位于整个产业价值链高端,具有高附加值;三是在高端产业链中占据主要位置,高端装备制造业的发展水平决定产业链的整体竞争力。

崔贺琨等^[1]对我国装备制造企业制造商与供应商创新性作用机理的问卷调研中,共回收286份有效问卷。其中,机械设备制造占45.5%,汽车及船舶制造占20.6%,仪器仪表制造占12.2%,其他占21.7%。研究结果发现:制造商与供应商相互协作越密切,合作意愿越高,就越有助于提升供应商的创新能力,呈现正向关系;由制造商引导和帮助,可更好地整合供需双方知识,实现

知识共享,帮助供应商创造更多合作机会,呈现正向关系;制造商与供应商知识共享水平越高,供应商知识学习和储备越多,越有利于激发供应商的创新潜力,呈现正向影响。由此可见,供需双方理应建立稳健的、相辅相成的紧密关系。

高端装备制造业大部分企业都有着自己固定的供应商,供应商之间也会存在竞争与协作的关系。因此,出现一个需求方对应多个供应商的情况。在选择合适的供应商上,需要需求方企业根据企业自身情况进行甄选。在维护现有供应商的同时,也在拓展供应商列表,以应对各种不定因素的出现。所以,本文所讨论的供应商推荐,意指需要拓展的能够促进企业扩大再生产的供应商或是应对市场不确定因素而挑选的备用供应商。

1.2 供应商推荐流程

高端装备制造企业之间不管从经营模式、生产产品还是对技术应用上都有着较大的差异,每个企业都有着各自的特点。企业用户是高端装备制造业服务平台的使用者。高端装备制造业服务平台能够根据企业用户自身特点和企业用户的操作行为,有针对性地为企业用户推荐供应商。企业用户也可以通过高端装备制造业服务平台,按照企业用户的需求甄选合适的供应商。

在高端装备制造业服务平台推荐供应商和企业用户自行选择供应商过程中,供应商推荐问题是高端装备制造过程的一个关键问题。高端装备制造业服务平台通过构建协同过滤推荐机制,为平台中的企业用户提供有效的供应商推荐策略,提高平台整体服务水平与效率。具体推荐机制如图1所示。

2 企业信息采集

由于高端装备制造业呈“金字塔”形,以“高”“中”“低”产业为序,通过“低”段供应商提供的项目产品,由“中”段企业用户进行再次加工、组装形成“中”段项目产品,并销售给“高”段企业进行最终的整合、装配,形成高端商品。因此,在生产、制造和组装的过程中,每个企业用户既可以是需求方也可以是供应方。本

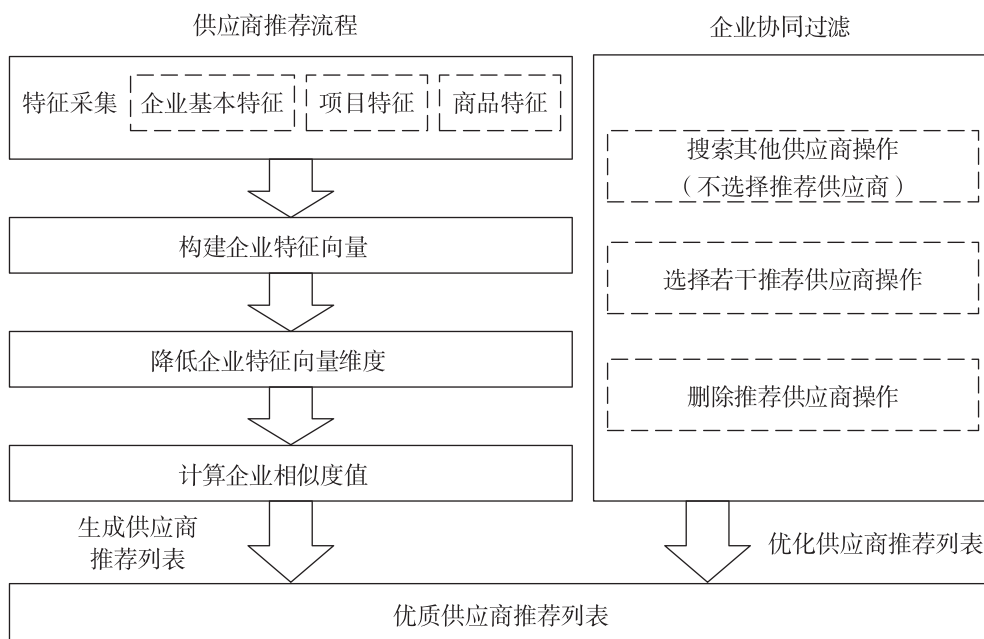


图1 供应商协同过滤推荐机制

文讨论的供应商协同过滤推荐的第一步就是对企业用户的相关信息进行系统采集。企业用户信息也可称为企业特征向量。企业的特征向量是计算企业用户相似度的关键步骤，通过高端装备制造业服务平台可以自动收集企业基本特征、项目特征和商品特征。

智能推荐过程关键是特征向量的构建。对于高端装备制造业企业通过高端装备制造业服务平台发布任务后构建的特征向量包含企业基本特征、项目特征和商品特征3个部分。

2.1 企业基本特征

在企业用户通过平台进行注册时就能够采集到企业用户的相关特征。主要包括企业用户名称、地址、联系人、经营范围、供应产品指标等信息。这是企业特征向量构建的关键要素。经营范围是高端装备制造企业基本特征中最为突出的特征向量，也是关键特征向量，体现了高端装备制造制造业的特点和企业间上下游关系。以汽车发动机行业为例，生产商需要采购下游供应商的产品，如气缸体、气缸盖罩、气缸盖、汽缸垫和油底壳等相关产品，而每种产品都是下游供应商企业列出的经营范围。

在高端装备制造业服务平台下的高端装备

制造企业属于同一集群下的企业，需要高端装备制造业服务平台通过收集企业基本特征向量的方式，为平台下的每个企业用户建立属于各自的企业基本特征向量。

2.2 项目特征

高端装备制造业的项目都具有高技术性，需求方企业为了使技术均衡，提高项目整体完成进度，可将一个项目任务分解成多个子项目任务进行发布，同时也会产生多个供应商进行协同合作，在所有的子项目完成并验收后，整体项目才算完成。在整个过程中，子项目分解数目、子项目任务完成时间、子项目设计零件耦合度、需求方企业对每个子项目完成情况的评价等，都是重要的项目特征向量。

高端装备制造业服务平台新的企业用户可申请需求方企业发布的相关项目，其中申请时间、申请类型、需求方企业回执等都可作为一类项目特征向量，并在供应商推荐列表中进行说明推荐处理。

2.3 商品特征

高端装备制造企业用户在高端装备制造业服务平台中发布的商品信息，每个商品数据属性都是有价值的特征向量，例如汽车点火系统中的火

花塞,其尺寸(螺纹直径、螺纹长度)、热值等都是重要的特征向量。火花塞又可以和汽车分电器等进行适配组合形成新的产品,它的数据信息也将形成新的相关特征向量。需求方企业对组装后的产品的相关评价也是重要的特征向量。

综上所述,采集企业信息,包括企业基本特征、项目特征和商品特征,作为基础数据;对基础数据按行业、地区和企业规模等进行划分,即可有效地对企业用户进行分类。基础数据质量和分类决定了推荐机制运行效果,即决定了推荐的供应商列表中供应商的质量,影响了企业用户的体验和高端装备制造业服务平台的服务质量。

3 供应商推荐方法

3.1 特征向量的构建

根据上述高端装备制造企业基本特征、项目特征和商品特征,构建每个企业各自的特征向量,并结合高端装备制造业协同开发、协同制造的特点,对在企业中采集的数据进行数字化编码,形成相应的特征向量。以企业用户长春某汽车空调有限公司为例,通过高端装备制造业服务平台,采集企业基本特征、项目特征、商品特征,构建企业部分特征向量(表1)。

在表1中,特征编号1—3表示企业基本特征,特征编号4—10表示商品特征,特征编号11—20表示项目特征。其中,特征编号7的空调高压管,有空调高压管长度属性,即作为特征并采集。特征编号12的离合器,是特征编号11的空调压缩机的子项目,离合器吸盘与离合器线圈是子项目的属性,即作为特征并采集。特征编号15的冷凝器,有冷凝器长度属性、冷凝器宽度属

性,即作为特征并采集。

高端装备制造业的特殊性决定了每个企业的特征向量将会有成百上千个特征。只有将企业用户的所有特征采集完整,构建企业完整的特征向量,才能完全呈现企业特点、项目优势、商品种类等。表1中只呈现了该企业特征向量中的一部分,也可称作节选的20维度企业特征向量。特征向量的构建是为推荐机制的输入准备合适的输入数据,是供应商推荐机制中不可或缺。

3.2 企业特征向量降维

对高端装备制造企业采集形成的特征向量普遍维度较高,就其产品特性而言就有很多特征向量需要表示,例如汽车空调制冷系统中的风扇以及它的车型、大小、孔距、数量、转速等都是需要采集的特征向量,同时也会出现一个零件可进行多个产品的匹配、装配使用的情况,也会造成特征向量的相似性。根据项目特征向量优先于商品特征向量,且优先于企业基本特征向量的原则,在项目特征向量中,往往包含商品特征向量和企业基本特征向量,因此应去掉无关或重复的特征向量,并通过文本分析的方法找到项目产品的相关性,再根据相关性分析的结果,去掉相关的特征向量数据,提高高端装备制造业服务平台的响应速度,使用主成分分析PCA^[2-3]进行降维和提取主成分。具体过程:一是设企业特征向量为 n 维向量,总企业样本数为 m ,全部样本组成的矩阵为 X ,并由均值求得企业样本矩阵的协方差矩阵;二是解出企业样本协方差矩阵的特征值 λ_i 和特征向量;三是将特征向量按对应特征值的大小从上到下按行排列成矩阵,设定降维的维数 k ,取前 k 行组成矩阵 P ;四是通过计算 $P \times X$ 为降

表1 企业特征向量

特征编号	企业基本特征			商品特征						
	1	2	3	4	5	6	7	8	9	10
特征名称	企业名称	地址	联系人	储液干燥器	储液干燥器尺寸	膨胀阀	空调高压管	空调高压管长度	空调低压管	空调低压管长度
特征编号	项目特征									
	11	12	13	14	15	16	17	18	19	20
特征名称	空调压缩机	离合器	离合器吸盘	离合器线圈	冷凝器	冷凝器长度	冷凝器宽度	蒸发箱	蒸发箱长度	蒸发箱宽度

维后的数据组，完成降维操作。以企业用户长春某汽车空调有限公司为例，降维后的企业特征向量如表 2 所示。

企业特征向量的降维，在既不破坏供应商关键特征信息又能突出企业特点、项目优势和商品种类的前提下，大大减少了高端装备制造业服务平台的计算压力。特征向量降维操作是节省高端装备制造业服务平台计算资源和提高高端装备制造业服务平台计算速度的重要手段，并且企业特征向量具有很大的降维空间，为后续的企业用户和供应商相似匹配计算提供了很好的加速机制。

3.3 供应商相似度计算

以长春某汽车空调有限公司和汽车行业的另外 5 家公司为例（经过行业过滤），得到每个企业的特征向量后，运用余弦相似度计算方法计算 5 家公司与长春某汽车空调有限公司的相似度情况，即通过两家企业的特征向量之间的夹角余弦来确定企业间的相似度情况。夹角余弦越小，说明两家企业相似度越高。为了方便相似度值之间进行比较，采用 0—1 规范化（如表 3 第二列），该值表示每个企业与长春某汽车空调有限公司的相似度情况，并对相似度值进行排名，即得到长春某汽车空调有限公司的供应商推荐列表，如表 3 所示。

综上所述，相似度计算和规范化是供应商推荐的重要步骤，也是供应商列表产生的重要依据。通过供应商特征向量和企业用户特征向量的相似计算，逐一核算供应商满足企业用户需求能力，高端装备制造业服务平台可以过滤掉评分较

低或不合适的供应商，完成对供应商的推荐。相似度计算效果和准确性是供应商推荐效果的重要影响因素，即严重影响优质供应商的选择。

3.4 协同过滤

通过高端装备制造业服务平台得到供应商推荐列表。经判断该列表中的供应商，企业用户做出不选择该列表中供应商、选择该列表若干供应商、删除该列表若干供应商以及搜索供应商等行为。这都将被高端装备制造业服务平台所记录，并根据此记录修正企业相似度值。如为长春某汽车空调有限公司推荐了供应商列表，可是该公司却没有选择推荐列表中的供应商，将推荐列表中供应商相似度值降低 0.05。长春某汽车空调有限公司删除推荐列表的某一供应商，将该供应商相似度值降为 0。长春某汽车空调有限公司主动搜索供应商，将所搜索到的供应商相似度值增加 0.05。

综上所述，企业相似度排名会随着企业用户使用频率、合作次数等的变化而发生变化。因此，企业用户在高端装备制造业服务平台上的操作，对供应商排名结果起到了协同过滤的作用，即企业用户对供应商推荐排名结果进行的操作，会反馈给后台，对供应商推荐列表进行优化，达到精准推荐目的。

供应商协同过滤推荐机制可有效地为企业用户推荐合适供应商列表，并在日后企业用户通过高端装备制造业服务平台不断地进行使用时，其使用数据也在不断进行反复迭代计算，这是不断完善供应商推荐列表的过程，也是不断优化企业

表 2 降维后的企业特征向量

特征编号	4	6	7	9	11	15	18
特征名称	储液干燥器	膨胀阀	空调高压管	空调低压管	空调压缩机	冷凝器	蒸发器

表 3 供应商推荐列表

企业名称	相似度值	推荐分析
长春某汽车制造有限公司	0.55	该企业相似度值在 5 家企业中最高，推荐该企业作为供应商
哈尔滨某汽车生产有限公司	0.53	该企业相似度值在 5 家企业中较高，推荐该企业作为供应商
黑龙江某汽车空调制造有限公司	0.18	该企业相似度值过低，不推荐该企业作为供应商
四平某汽车零部件生产有限公司	0.17	该企业相似度值过低，不推荐该企业作为供应商
吉林某汽车空调制造厂	0.13	该企业相似度值过低，不推荐该企业作为供应商

用户的使用体验,形成良性循环。高端装备制造业服务平台将会成为企业用户找寻合作伙伴理想的服务平台。

4 结语

高端装备制造业服务平台通过供应商协同过滤推荐机制,为企业用户提供精准的供应商列表,为企业用户节省了成本,提高了效率。本文对供应商推荐进行详细的描述,结合高端装备制造业的特点,通过企业用户主要经营范围、项目特征以及能够直观区分企业用户特征的相关商品等构建和降低企业特征向量维度,并经过企业相似度计算得到排名,形成优质供应商的推荐列表。企业用户的参与会使供应商推荐列表不断更新,达到协同过滤推荐的目的,即对供应商的选择始终保持着良好的匹配状态。这样可有效地加强对优质供应商的管理,提升企业间合作机遇和服务质量,提升企业整体产能,为企业获得更多经济效益。

参考文献

- [1] 崔贺理,李随成,乔建麒,等. 供应商创新性提升机理: 制造商—供应商社会资本作用研究[J]. 科技进步与对策, 2020, 37(13): 92–101.
- [2] AJAY G, ADRIAN B. Parameterized principal component analysis[J]. Pattern Recognition, 2018(78): 215–227.
- [3] AMMAR I S, FATEN A D, ALI H K. Data analysis and classification of autism spectrum disorder using principal component analysis[J]. Advances in Bioinformatics, 2020(2020): 1–8.
- [4] 李一野, 邓浩江. 基于改进余弦相似度的协同过滤推荐算法[J]. 计算机与现代化, 2020(1): 69–74.
- [5] 赵晓宇. 基于BP神经网络—协同过滤的云制造供应商推荐系统[J]. 价值工程, 2019, 38(20): 128–130.
- [6] 岳希, 唐聃, 舒红平, 等. 基于数据稀疏性的协同过滤推荐算法改进研究[J]. 工程科学与技术, 2020, 52(1): 198–202.
- [7] 申晋祥, 鲍美英. 基于用户聚类与项目划分的优化推荐算法[J]. 计算机系统应用, 2019, 28(6): 159–164.
- [8] 杨娟, 吴科宏. 云制造环境下设计需求与服务匹配模式算法[J]. 重庆大学学报, 2016, 39(2): 82–88.
- [9] 马婧, 吴清烈. 面向C2B个性化定制的智能推荐算法研究[J]. 工业工程, 2018, 21(5): 87–92.
- [10] 董跃华, 朱纯煜. 基于改进用户属性评分的协同过滤算法[J]. 计算机工程与设计, 2020, 41(2): 425–431.
- [11] 李顺勇, 张钰嘉, 张海玉. 基于NKL和K-means聚类的协同过滤推荐算法[J]. 河南科学, 2020(1): 6–12.
- [12] 李艳娟, 牛梦婷, 李林辉. 基于蜂群K-means聚类模型的协同过滤推荐算法[J]. 计算机工程与科学, 2019, 41(6): 1101–1109.
- [13] 邓园园, 吴美香, 潘家辉. 基于物品的改进协同过滤算法及应用[J]. 计算机系统应用, 2019, 28(1): 182–187.
- [14] SHAO G W, SANG N. Regularized max-min linear discriminant analysis[J]. Pattern Recognition, 2017(66): 353–363.
- [15] JU F J, SUN Y F, GAO J B, et al. Probabilistic linear discriminant analysis with vectorial representation for tensor data[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(10): 2938–2950.
- [16] 与财务绩效关系研究[J]. 会计之友, 2016(8): 51–54.
- [17] 黎春. 中国上市公司财务指数研究[D]. 成都: 西南财经大学, 2010: 143–156.
- [18] 柯忠义. 创业板上市公司经济绩效及影响因素: 基于贝叶斯模型平均法(BMA)的实证研究[J]. 数量经济技术经济研究, 2017(1): 146–160.
- [19] HAIR, JOSEPH F, WILLIAM C, et al. Multivariate Data Analysis[M]. 7th ed. New Jersey: Prentice Hall, 2009: 92–98.
- [20] 梁娜, 姚长青, 高影繁, 等. 政府补助对企业研发投入与企业成长性的影响[J]. 中国科技资源导刊, 2018, 50(6): 16–22, 47.
- [21] 宋建波, 高翔. 经济金融化挤出了企业实体投资吗?: 基于我国A股上市公司的经验证据[J]. 湖北大学学报哲学社会科学版, 2019, 46(5): 143–151.
- [22] 李连华, 柯东昌, 许倩芸. 经济下行中企业研发投入

(上接第34页)

[J]. Research Policy, 2012, 41(2): 334–345.

- [25] 梁娜, 姚长青, 高影繁, 等. 政府补助对企业研发投入与企业成长性的影响[J]. 中国科技资源导刊, 2018, 50(6): 16–22, 47.
- [26] 宋建波, 高翔. 经济金融化挤出了企业实体投资吗?: 基于我国A股上市公司的经验证据[J]. 湖北大学学报哲学社会科学版, 2019, 46(5): 143–151.
- [27] 李连华, 柯东昌, 许倩芸. 经济下行中企业研发投入