

DRVisSys: 基于属性相关性分析的可视化推荐系统

吴小全^{1,2}, 李 晖^{1,2}, 陈 梅^{1,2}, 戴震宇^{1,2}

WU Xiaoquan^{1,2}, LI Hui^{1,2}, CHEN Mei^{1,2}, DAI Zhenyu^{1,2}

1. 贵州大学 计算机科学与技术学院, 贵阳 550025

2. 贵州大学 贵州省先进计算与医疗信息服务工程实验室, 贵阳 550025

1.College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

2.Guizhou Engineering Lab of Advanced Computing and Medical Information Services, Guizhou University, Guiyang 550025, China

WU Xiaoquan, LI Hui, CHEN Mei, et al. DRVisSys: visualization recommendation system based on attribute correlation analysis. *Computer Engineering and Applications*, 2018, 54(7): 251-256.

Abstract: Visualization is the most effective way to unleash the insights of data. However, discovering the relevant subgroup of large multidimensional data sets and mapping it into appropriate visualizations often require large amount of sophisticated iterative analysis. In order to make it easy to obtain better visualization automatically, this paper proposes DRVisSys, a system that can produce suitable visualizations according to Canonical Correlation Analysis(CCA) based attributes relevance. Furthermore, considering that the weight of attribute is different in the real world, the system uses cascaded Hidden Markov Model(HMM) to calculate the weight of attributes, which regards weight as one of criteria of non-trivial set of attributes. Besides, the system can update the visual recommendation model based on users' feedback in order to make the recommended visualization meet the needs of users. The experiment results show that DRVisSys can quickly analyze the data and recommend the appropriate visualization to users.

Key words: data visualization; Canonical Correlation Analysis(CCA); non-trivial set of attributes; cascaded Hidden Markov Model(HMM); visual recommendation model

摘 要: 数据可视化通常是展示数据价值最有效的方式。针对大规模复杂多维数据,对相关数据子集进行分析并将分析结果自动映射成合适的可视化展现模式,是一项需要大量迭代计算的复杂技术工作。设计并实现了DRVisSys系统,该系统根据属性关联分析技术推荐出合适的可视化展现模式;其对于非平凡属性组合的选择,采用典型关联算法计算出更优的属性集。考虑到各属性权重在实际生活中是有区别的,采用层叠隐马尔可夫算法计算各属性权重,将属性权重作为非平凡属性组的评测标准之一。为使得推荐出的可视化展现模式能更好地满足用户需要,DR-VisSys系统能根据用户反馈,更新可视化推荐模型。实验结果表明,DRVisSys能够快速进行数据分析并为用户推荐出合适的可视化展现模式。

关键词: 数据可视化;典型关联算法;非平凡属性组;层叠隐马尔可夫模型;可视化推荐模型

文献标志码: A **中图分类号:** TP311 **doi:** 10.3778/j.issn.1002-8331.1710-0104

1 引言

可视化通常是数据分析中一个关键步骤。加拿大维多利亚大学的研究人员^[1]提出了可视化分析任务,包括三个主要步骤:数据属性子集选择、可视化展现模式

推荐、数据和图表映射。文献[1]证明,计算出和目标具有强相关性的属性子集并为其确定合适的可视化展现模式是一项困难的技术工作。尤其是当数据和维度数量增加时,现有的可视化分析系统^[2-3]通常会面临如下

基金项目: 国家自然科学基金(No.61462012, No.61562010); 贵州大学研究生创新基金(No. 研理工 2017078, No. 研理工 2017082)。

作者简介: 吴小全(1990—),男,硕士研究生,研究领域为大数据管理与分析;李晖(1982—),通讯作者,男,博士,副教授,研究领域为大规模数据管理与分析、云服务, E-mail: cse.HuiLi@gzu.edu.cn;陈梅(1964—),女,教授,研究领域为大规模数据管理与分析;戴震宇(1985—),男,实验师,研究领域为大规模数据管理与分析。

收稿日期: 2017-10-16 **修回日期:** 2017-12-06 **文章编号:** 1002-8331(2018)07-0251-06

问题:

(1)多维属性子集选择的搜索空间不够大。现有数据分析工具在对多维数据集进行非平凡属性组分析时,需耗费大量时间进行迭代计算并构建可视化展示模式,且这个过程通常需要人为干预和反馈,未能包含对所有重要的属性组合的探索和评估。

(2)对分析人员的技术门槛要求较高。由于通常要求分析人员能够针对初步分析结果做进一步的关联分析和探索,分析人员需具备良好的统计知识和编程技术素养,进一步增加了应用难度。

(3)分析过程的自动化程度不够高。缺乏基于用户行为反馈的推荐模型及其展现模式的自动化更新机制。

针对解决上述问题,法国里尔第二大学的研究人员设计并实现了基于知识库技术实现可视化推荐系统 VizAssist^[4]。该系统为了推荐出合适的可视化图表,在属性选择上充分考虑到属性的数据类型和属性的权重,优先选择权重高的属性通过知识库内部已制定规则构建可视化展现模式。然而在设计知识库时,它依然存在规则有限的缺陷,导致部分不符合该规则的属性无法构建可视化展现模式。此外,VizAssist通过人为设置属性的权重和属性的数据类型来决定推荐的属性列,这使得推荐出的数据可视化展示模式过于依赖人为初始化的设定,导致人为干预成为数据分析的主导地位,无法实现自动探索和推荐。Voyager^[5]是华盛顿大学的研究人员研发的可视化推荐系统,其为数据分析添加合适的统计函数,通过人机交互进行数据探索,并将探索出的数据自动生成可视化图表,从而较为有效地支撑用户进行交互式数据导航。

现有的可视化推荐系统通常是通过参数设置来推荐出部分有价值的可视化展现模式,在一定程度上提高了推荐的准确率。但大多数推荐仅仅利用静态的初始化设置而忽略了数据本身动态组合的特性。数据属性之间的关联性是数据价值的潜在表现。此外,现有可视化推荐系统通常引入了过多的人为判断操作,在处理大规模复杂多维数据时,将会增加交互难度。

针对目前可视化分析领域存在的上述问题,本文基于属性相关性分析策略,研究非平凡可视化推荐的度量标准,引入典型关联算法用于解决非平凡属性组的选择问题。设计并实现了数据分析可视化推荐系统 DRVisSys,主要技术工作如下:

(1)在非平凡属性组选择环节中引入属性关联分析技术。为了从整体考虑属性之间的相关性,应用典型关联分析技术计算属性之间的关联系数,用典型关联系数大小作为非平凡属性组合的评测标准^[6-7]。

(2)通过自动构建大量 SQL 语句来生成所有属性组及属性组对应的数据子集。并针对大量 SQL 语句执行

过程中资源消耗较高的问题,提出并行执行查询和共享公共数据子集的技术。

(3)引入知识库技术来构建可视化展示模式。只需将选择出的非平凡属性组及对应的数据子集作为输入,便可自动通过知识库构建出可视化展现模式。

(4)设计并实现了基于用户反馈的可视化推荐模型的持续更新方法,可根据用户行为迭代出更合适的展示模式^[8]。

2 预备知识

2.1 非平凡属性组评测

麻省理工学院的研究人员,在2015年提出通过计算原数据集与参考数据集之间的偏差来确定非平凡属性组的方法^[9]。该方法假设偏差值越大,属性组成为非平凡属性组的可能性越大。但部分特殊值可能会严重影响整体偏差值,从而导致非平凡属性组的分析有偏差。为解决此问题,本文通过分析属性之间相关关系来确定非平凡属性组。为了能从整体考虑属性之间相关关系,本文在计算过程中还引入了典型关联分析技术。

计算属性组的典型关联系数时存在下述问题:当属性组里存在一个或者多个非数值属性列时,无法计算非数值属性列的关联系数。本文根据 SQL92 标准将数值类型属性归类为度量属性,其他属性归为维度属性。在计算维度和度量的相关性时,首先按维度属性对数据集进行分组,然后在分组内计算度量属性的 SUM、AVG、COUNT 聚合函数值,选择组内任意属性构建属性组并计算其相关性。

为了解决当前可视化推荐系统对多维属性子集选择的搜索空间不够全的问题,DRVisSys 对整个属性子集空间进行探索^[10]。对任意数据集进行探索并生成候选属性组时,候选属性组的数量由数据集的维度属性和度量属性决定。假设数据集中有 m_1 个维度属性, m_2 个度量属性,在生成候选属性组时,根据各维度属性进行分组,计算组内度量属性的聚合函数值(SUM、AVG、COUNT)。然后任意从该组中选择属性构建属性组,则共计会产生 $\{m_1 * (C_{m_2}^2 + C_{m_2}^3 + \dots + C_{m_2}^{m_2}) = m_1 * (2^{m_2} - m_2)\}$ 个属性组。

2.2 基于可扩展 VKB 技术的知识库

基于 VKB 技术的知识库起源于 1980 年的 BHARAT 系统^[11-12]。2015 年,VizAssist 系统的研究人员提出基于 VKB 技术的知识库构建方法,用于简化复杂可视化模型的构建。然而 VKB 技术受限于有限的规则模型,无法分析当前知识库内现有规则之外的可视化模型。本文提出基于可扩展 VKB 技术的知识库模型,为扩展知识库规则定义一套扩展机制,可通过扩展 API 添加新的

根据用户行为反馈生成的新的可视化图表;C区是用于显示推荐出的所有可视化图表。

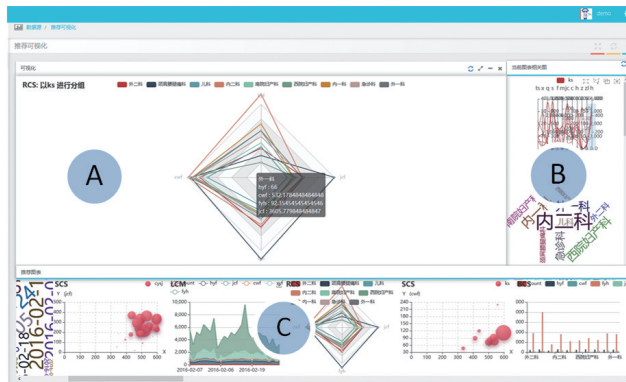


图2 DRVisSys推荐出的可视化展示图表

从图2可以看出,DRVisSys根据不同属性集推荐出各种图表对数据进行展示。以(A)中的雷达图生成为例,该图表的构建流程简述如下:

(1)构建雷达图规则模型。雷达图是用来显示多项指标数据的数值占比情况。雷达图的推荐规则如下:属性维度数目3~10;任意两项指标相关系数偏差不得超过0.1;数据量不超过100条。

(2)为数据集构建可视化展现模式。当前雷达图属性列有 $\{ks, cwf, fyh, jcf, hyf\}$,将字符型属性 ks 作为维度属性,其他作为度量属性。通过对任意属性列的关联系数计算,得知其关联系数的偏差都不超过0.1。上述计算结果符合雷达图规则,因而将该属性组映射成雷达图。

DRVisSys支持用户交互。系统可根据用户行为反馈,更新推荐模型。DRVisSys对当前用户选中可视化展示模式的偏好行为进行分析,将相应的属性添加权重并重新计算出非平凡属性组,更新当前推荐模型,构建新的可视化展现模式。相应的结果,均在B区列出。

4 查询优化及主要算法设计

本章主要介绍DRVisSys系统的两种查询优化技术,分别是并行执行查询技术和共享公共表达式技术^[13-15]。此外,为了过滤大量的非平凡属性组,DRVisSys引入了典型关联算法并采用奇异值分解法求解出典型关联系数,取Top K个属性组定义为非平凡属性组。

4.1 查询优化

4.1.1 查询的并行执行

为了增加多维属性子集选择的搜索空间,DRVisSys系统对数据集中的所有组合进行探索。对于数据集中有 m_1 个维度属性 m_2 个度量属性,生成的候选属性组的数量是 $m_1 \times (2^{m_2} - m_2)$ 。每个属性组所对应的数据子集是通过执行一条为该属性组特别构建的SQL查询语句而得到。因此,获取各候选属性组对应的数据子集需要

构建并执行大量SQL语句。考虑到候选属性组的数量,此过程的时间开销极大。DRVisSys通过采用多线程并行执行上述SQL任务的方法来提高查询执行效率。为了减少并行执行过程中数据库连接增加的问题,DRVisSys引入了数据库连接池技术来实现数据库连接共享,进一步减少数据库连接延迟。

4.1.2 查询重写和执行优化

在构建候选属性组合时,属性组之间存在相同属性,即与之相对应的SQL之间存在公共的表达式。为了减少查询的执行开销,DRVisSys系统通过基于共享公共表达式(及其对应的数据子集)的查询重写技术实现多个SQL查询语句的合并。具体优化步骤如下:

(1)合并同一属性的不同聚合查询。在同一分组属性下,对所有度量属性添加不同聚合函数(如:SUM、AVG、COUNT)。若为每个聚合函数操作构建一个查询语句,会造成资源浪费。DRVisSys将类似于如下查询 $(a, m, f_1), (a, m, f_2), \dots, (a, m, f_i)$ 合并,重写成新的查询语句 $(a, m, \{f_1, f_2, \dots, f_i\})$ 。

(2)合并同一分组属性的不同查询。针对同一分组属性下不同属性的聚合查询,本文采用group by分组属性区分不同聚合查询的结果集,将多个聚合查询重写成一个聚合查询。DRVisSys系统将类似如下的查询语句 $(a, m_1, f), (a, m_2, f), \dots, (a, m_i, f)$ 合并,重写形成新的查询语句 $(a, \{m_1, m_2, \dots, m_i\}, f)$ 。

(3)添加分组属性,合并具有公共子表达式的不同查询。此外,存在一种特例:其中一条查询语句的结果集包含另外一条查询语句的结果集。此时,不能利用原有分组属性区分查询结果集,需为其添加新的分组属性。例如:合并如下所示的 Q_1 和 Q_2 两个查询语句时,需要为其新增分组属性 G_1 和 G_2 。

$Q_1 = \text{SELECT } A, F(m) \text{ FROM } D \text{ WHERE } X = \text{'贵州'} \text{ GROUP BY } A$

$Q_2 = \text{SELECT } A, F(m) \text{ FROM } D \text{ GROUP BY } A$

将上述两个查询 Q_1 和 Q_2 合并,重写成新查询 Q_3 :

$Q_3 = \text{SELECT } A, F(m), \text{CASE IF } X = \text{'贵州'} \text{ THEN } 1 \text{ ELSE } 0 \text{ END AS } G_1, 1 \text{ AS } G_2 \text{ FROM } D \text{ GROUP BY } A, G_1, G_2$

4.2 非平凡属性组选择算法

为了从大量属性组合中过滤掉大部分平凡的属性组,本文引入CCA^[16]作为非平凡属性组的评估算法。通过CCA算法可以计算出属性之间的关联系数,将其作为非平凡属性组的评估标准。但在求解典型关联系数时,为了避免属性组出现的小样本数据可能导致无法计算出关联系数的问题,DRVisSys系统采用奇异值分解法求解典型关联系数。非平凡属性组选择算法的步骤如下。

(1)计算推荐系数。通过计算属性组之间典型关联系数作为推荐的系数,下面是CCA计算相关系数过程:

- ① 设定属性组中存在各包含 n 个样本的 U 和 V , 而且 U 和 V 的维度大于1。
- ② 计算出样本 U 的方差 S_{uu} , 样本 V 的方差 S_{vv} , 以及 U 和 V 样本的协方差 S_{uv} 。
- ③ 构建矩阵 $M=S_{uu}^{-1/2}S_{uv}S_{vv}^{-1/2}$ 。
- ④ 对矩阵 M 进行奇异值分解,计算出最大的奇异值,即为所需的最大关联系数 ρ 。

(2)加权推荐系数。考虑到部分维度属性在应用中有着特殊的实际意义,例如:日期、姓名和部门等,本文采用层叠隐马尔可夫算法^[17]识别出此类重要属性并为其添加权重值。该算法为每种命名实体的识别方法构造独有的识别模型,例如:人名识别的HMM、地名识别的HMM和机构识别的HMM等等。当对某一个属性实体进行识别时,如果属于重要实体,设置该属性的权重值 $W=1$; 否则,设置权重值 $W=0$ 。

在选择非平凡属性组时考虑到部分属性的权重值,本文引入式(1),将权重值与关联系数相结合,重新计算推荐系数。

$$\rho = \begin{cases} \rho, & W=0 \\ \rho + \frac{\rho}{\rho + W}, & W=1, \rho < 0.6 \\ 1, & W=1, \rho > 0.6 \end{cases} \quad (1)$$

(3)取Top K 个属性组。将加权后的推荐系数进行排序,取前 K 个属性组合作为非平凡属性组合。推荐引擎默认取 $K=15$,但当候选组合数超过15组且推荐系数 $\rho>0.9$,引擎会自动添加推荐系数大于0.9的属性组作为非平凡属性组,其中 K 和 ρ 的值可根据需求自定义。

4.3 计算非平凡属性组的案例

为进一步说明非平凡属性组的选择算法流程,本节以“病人治疗费用表”为例进行说明,该表共23维属性,其中有10维维度属性,13维度量属性。根据2.1节中的方法构建属性组合及对应数据子集,然后通过推荐算法计算出平凡属性组,步骤如下:

(1)利用典型关联算法计算属性组的关联系数,并对其进行排序。本节取三个属性组合进行讲解:属性组 $V_1=\{ks, cwf, hyf, fyh, jcf\}$, 计算典型关联系数 $\rho=0.94$; 属性组 $V_2=\{cyqtzd, hyf, cwf, fyh, jcf, xyf\}$, $\rho=0.96$; 属性组 $V_3=\{ks, zjf, ts\}$, $\rho=0.46$ 。

(2)使用层叠马尔可夫算法识别重要属性并对属性进行加权。针对属性组 V_1 ,通过算法识别得出 ks (科室)在命名实体中属于部门实体,属于重要属性且 $\rho>0.6$,根据式(1)重新计算出推荐系数 $\rho=1$;针对属性组 V_2 识别,不属于重要属性,推荐系数不需修改;针对

属性组 V_3 , 识别出 ks 属于重要属性且 $\rho<0.6$,根据式(1)重新计算出推荐系数 $\rho=0.775$ 。同理,其他属性组以此方法计算推荐系数。

(3)对重新计算的推荐系数进行排序,取Top K 个属性组作为非平凡属性组, $K=15$ 。

5 实验结果与分析

对DRVisSys系统性能进行评测和分析。实验过程中选用不同数据量、不同维度、不同记录数的数据集进行评估,评估指标是DRVisSys推荐出可视化展现模式所需的时间开销。测试数据包含20个不同数据集,具体数据集描述如表1所示。

表1 实验数据集描述

属性	100 MB		500 MB		10 万行		100 万行	
	Dim	Mea	Dim	Mea	Dim	Mea	Dim	Mea
10	5	5	5	5	5	5	5	5
20	10	10	10	10	10	10	10	10
50	40	10	40	10	40	10	40	10
100	80	20	80	20	80	20	80	20
500	480	20	480	20	480	20	480	20

所有实验运行在两个节点上,分别是分析数据集的数据节点和推荐引擎的执行节点,其中数据节点数量可根据实验要求进行配置,执行引擎节点是单节点。节点的参数配置如表2。

表2 实验节点参数配置

节点	CPU	内存	磁盘
数据节点	4x Intel®-Xeon-CPU		
	E5-2630@2.6 GHz		
执行引擎	8x Intel®i7-4790		
	CPU@3.6 GHz		

如图3、图4是系统在不同数据集下的性能测试结果。图3是10万条数据和100万条数据在不同维度情况下的性能测试结果。实验结果表明,在数据集记录数相同时,随着属性维度的增加,系统执行所需时间不断增加。图3中,当维度低于50时,10万条记录和100万条记录执行时间差较小,执行时间在3~25 s;当属性维度高于50维时,执行时间都偏高。这是由于:(1)随着属性维度的增加,候选属性组合数呈指数增加。(2)随着

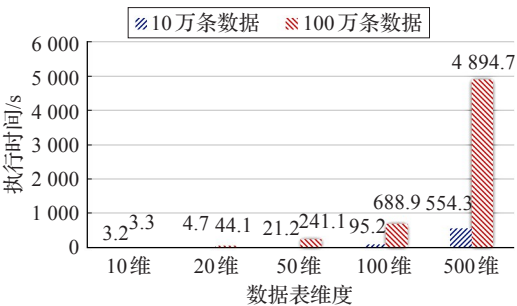


图3 10万条和100万条数据执行时间

候选属性组合数量的增长,使构建与候选属性组相对应的数据子集的CPU开销、磁盘I/O和网络传输开销均显著增加。(3)在计算出候选属性组的推荐系数时,所需要的CPU开销也随着候选属性组合数量的显著增长而增加。

为验证数据量对系统效率的影响,对100 MB和500 MB的数据集在不同维度情况下进行测试,实验结果如图4所示。维度相同的情况下,500 MB数据执行所需时间明显长于100 MB。数据量相同的情况下,维度越高,所需执行时间越长。

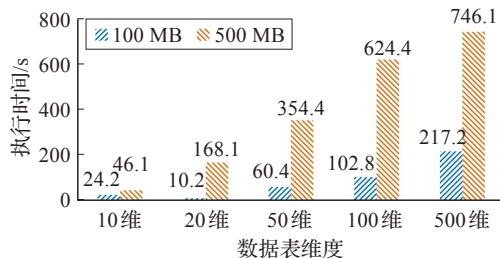


图4 100 MB和500 MB数据执行时间

图3、图4实验结果表明,该系统针对属性维度低于50的数据集,通常能够在几十秒内推荐出合适的可视化展示模式;系统针对高维度、数量大的数据集,亦能相对高效地完成推荐。

6 用户评价实验

上一章验证DRVisSys系统的性能,本章邀请用户使用真实数据进行实验,通过获取用户对该系统推荐图表的反馈来检验推荐效果。为此,该实验选择10个真实数据集(包含医疗数据、金融数据、NBA球员比赛数据)。实验招募16名计算机学院的在校硕士生参与实验,参与者有从事数据分析的人员和非数据分析方向的人员。对推荐满意度进行对比分析,本文选择VizAssist系统做实验对比。实验步骤如下:

(1)利用DRVisSys系统和VizAssist系统对10个真实数据集进行实验。两个系统在对不同数据集进行实验时,同时对同一数据集的推荐图表进行对比分析,比较各系统的推荐效果。

(2)16位参与者为系统推荐的前5个图表打分(满分为10)。参与者在评分前,需要对当前数据集进行了解,从而能更好地判断当前推荐是否有价值、有意义,推荐图表是否符合用户需求。

统计前5个推荐可视化图表平均评分情况,结果如表3所示。

表3 可视化推荐图表评分统计表

系统	推荐图表				
	图表1	图表2	图表3	图表4	图表5
DRVisSys	9.4	9.1	9.2	8.6	8.4
VizAssist	7.5	7.1	7.1	6.5	6.4

从表3可以看出,DRVisSys系统的推荐效果明显好于VizAssist系统。DRVisSys系统采用典型关联算法整体分析属性之间相关关系,并且采用层叠马尔可夫算法对重要实体属性加权。VizAssist系统通过属性权重确定推荐属性,并采用交互式遗传算法修改属性权重,重新生成推荐图表。从推荐效果上看,本文中采用的算法推荐效果更好,同时也说明属性之间关联性是数据价值的潜在表现。

7 结束语

对大规模复杂多维数据进行分析时,现有的可视化分析工具需要耗费大量时间进行反复迭代计算,并且存在对多维属性子集探索空间不够大,对分析人员技术门槛高和分析过程自动化程度不够高等问题。本文设计并实现了基于属性相关性分析的可视化推荐系统DRVisSys,运用典型关联分析算法完成非平凡属性组的选择,并基于VKB技术的知识库构建合适的可视化展示模式。同时,针对执行效率问题,引入并行执行查询和基于公共表达式的查询重写技术。实验结果表明,DRVisSys系统能快速为不同维度的数据集推荐出合适可视化展现模式。

参考文献:

- [1] Grammel L, Tory M, Storey M A. How information visualization novices construct visualizations[J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 16(6): 943-952.
- [2] Mackinlay J, Hanrahan P, Stolte C. Show me: Automatic presentation for visual analysis[J]. IEEE Transactions on Visualization and Computer Graphics, 2007, 13(6): 1137-1144.
- [3] Dan M. Tableau your data!: Fast and easy visual analysis with tableau software[M]. [S.l.]: Wiley Publishing, 2013.
- [4] Bouali F, Guettala A, Venturini G. VizAssist: An interactive user assistant for visual data mining[J]. The Visual Computer, 2016, 32(11): 1447-1463.
- [5] Wongsuphasawat K, Moritz D, Anand A, et al. Voyager: Exploratory analysis via faceted browsing of visualization recommendations[J]. IEEE Transactions on Visualization and Computer Graphics, 2016, 22(1): 649-658.
- [6] Naimi A I, Westreich D J. Big data: A revolution that will transform how we live, work, and think[J]. American Journal of Epidemiology, 2014, 177(17): 181-183.
- [7] 梁吉业, 冯晨娇, 宋鹏. 大数据相关分析综述[J]. 计算机学报, 2016, 39(1): 1-18.
- [8] Gotz D, Wen Z. Behavior-driven visualization recommendation[C]// Proceedings of International Conference on Intelligent User Interfaces, Sanibel Island, Florida, USA, 2009: 315-324.

(下转263页)

- [39] Herrem F, Herrera-Viedma E, Verdegay J L. Rational consensus model in group decision making using linguistic assessments[J]. Fuzzy Sets and Systems, 1997, 88(1): 31-39.
- [40] Xu Z S. Induced uncertain linguistic OWA operators applied to group decision making[J]. Information Fusion, 2006, 7(2): 231-238.
- [41] Xu Z S. Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment[J]. Information Sciences, 2004, 168(1/4): 171-184.
- [42] Yager R R. OWA aggregation over a continuous interval argument with applications to decision making[J]. IEEE Transactions on Systems Man & Cybernetics: Part B Cybernetics, 2004, 34(5): 1952-1963.
- [43] Xu Z S. Deviation measures of linguistic preference relations in group decision making[J]. Omega, 2005, 33(3): 249-254.
- [44] Herrera F, Herrera-Viedma E, Mart'inez L. A fusion approach for managing multi-granularity linguistic term sets in decision making[J]. Fuzzy Sets & Systems, 2000, 114(1): 43-58.
- [45] 王坚强, 刘淘. 基于综合云的不确定语言多准则群决策方法[J]. 控制与决策, 2012, 27(8): 1185-1190.

(上接169页)

- [6] 韩俊英, 刘成忠. 应用反向学习策略的果蝇优化算法[J]. 计算机应用与软件, 2014, 31(4): 157-160.
- [7] 韩俊英, 刘成忠. 基于历史认知的果蝇优化算法[J]. 计算机科学与探索, 2014, 8(3): 368-375.
- [8] 宁剑平, 王冰, 李洪儒, 等. 递减步长的果蝇优化算法及应用[J]. 深圳大学学报: 理工版, 2014, 31(4): 367-373.
- [9] 钟伟民, 牛进伟, 梁毅, 等. 多策略果蝇优化算法及其应用[J]. 化工学报, 2015, 66(12): 4888-4894.
- [10] 张前图, 房立清, 赵玉龙. 具有 Levy 飞行特征的双子群果蝇优化算法[J]. 计算机应用, 2015, 35(5): 1348-1352.
- [11] 常鹏, 李树荣, 葛玉磊, 等. 迭代步进值自适应调整的果蝇优化算法[J]. 计算机工程与应用, 2016, 52(3): 32-36.
- [12] Dai Hongde, Zhao Guorong, Lu Jianhua, et al. Comment and improvement on "A new fruit fly optimization algorithm: Taking the financial distress model as an example"[J]. Knowledge-Based System, 2014, 59: 159-160.
- [13] Shan Dan, Cao Guohua, Dong Hongjiang. LGMS-FOA: An improved fruit fly optimization algorithm for solving optimization problems[J/OL]. (2013). Mathematical Problems in Engineering. <http://dx.doi.org/10.1155/2013/108768>.
- [14] 王李进, 尹义龙, 钟一文. 逐维改进的布谷鸟搜索算法[J]. 软件学报, 2013, 24(11): 2687-2698.
- [15] Alatas B. Chaotic bee colony algorithm for global numerical optimization[J]. Expert System with Application, 2010, 37: 5682-5687.

(上接256页)

- [9] Vartak M, Rahman S, Madden S, et al. SEEDB: efficient data-driven visualization recommendations to support visual analytics[J]. Proceedings of the VLDB Endowment, 2015, 8(13): 2182-2193.
- [10] Broeksema B, Telea A C, Baudel T. Visual analysis of multi-dimensional categorical data sets[C]//Proceedings of Computer Graphics Forum, 2013: 158-169.
- [11] Gnanamgari S. Information presentation through default displays[M]. [S.l.]: University of Pennsylvania, 1981.
- [12] Senay H, Ignatius E. A knowledge-based system for visualization design[M]. [S.l.]: IEEE Computer Society Press, 1994.
- [13] Giannikis G, Makreshanski D, Alonso G, et al. Shared workload optimization[J]. Proceedings of the VLDB Endowment, 2014, 7(6): 429-440.
- [14] Dalvi N N, Sanghai S K, Roy P, et al. Pipelining in multi-query optimization[J]. Journal of Computer and System Sciences, 2003, 66(4): 728-762.
- [15] Sellis T K. Multiple-query optimization[J]. ACM Transactions on Database Systems (TODS), 1988, 13(1): 23-52.
- [16] Dan K. A singularly valuable decomposition: The SVD of a matrix[J]. College Mathematics Journal, 1996, 27(1): 2-23.
- [17] 俞鸿魁, 张华平, 刘群, 等. 基于层叠隐马尔可夫模型的中文明命名实体识别[J]. 通信学报, 2006, 27(2): 86-94.

word版下载: <http://www.ixueshu.com>

免费论文查重: <http://www.paperyy.com>

3亿免费文献下载: <http://www.ixueshu.com>

超值论文自动降重: http://www.paperyy.com/reduce_repetition

PPT免费模版下载: <http://ppt.ixueshu.com>

阅读此文的还阅读了:

1. [一种基于ElasticSearch的推荐系统架构](#)
2. [基于特征的兴趣推荐系统研究](#)
3. [基于用户兴趣的微博推荐系统研究](#)
4. [基于学习资源的推荐系统研究](#)
5. [DRVisSys: 基于属性相关性分析的可视化推荐系统](#)
6. [基于GIS的气象信息集成与可视化系统](#)
7. [基于hadoop平台的混合推荐系统](#)
8. [基于社交网络的群体推荐系统](#)
9. [基于TurnTool的排课系统教室使用属性可视化的实现](#)
10. [基于长尾理论对推荐系统的思考](#)
11. [基于PageRank的科技论文推荐系统](#)
12. [基于PDM系统的可视化工艺规程的探讨](#)
13. [基于协同推荐的高校个性化图书推荐系统](#)
14. [基于PDM系统的可视化工艺规程的探讨](#)
15. [基于PDM系统的可视化工艺规程的探讨](#)
16. [推荐系统的推荐解释研究](#)
17. [基于网购的推荐模式系统研究](#)
18. [基于属性聚类的项目评分预测推荐算法研究](#)
19. [基于地震属性分析的储层预测方法研究](#)
20. [基于数据分析的游戏玩家成长路线推荐系统](#)
21. [基于个性化混合推荐算法的网络推荐系统](#)
22. [基于非对称属性的SVD推荐算法的研究](#)
23. [基于用户属性和活跃性的协同过滤推荐算法](#)
24. [基于哈希模型的Web推荐系统](#)
25. [基于Android系统的餐饮推荐系统的开发](#)

26. 基于系统论视角的产品属性系统的研究

27. 基于机器学习的推荐系统

28. 基于商品三重属性的“系统价值论”

29. 基于共词分析的循证矫正可视化研究

30. 基于分类算法的潜在好友推荐系统

31. 基于spark的推荐系统设计

32. 一种基于混合推荐模式的图书推荐系统

33. 基于DSmTrust信任模型的推荐系统

34. 基于网页分析的可视化系统

35. 基于属性偏好自学习的推荐方法

36. 基于漏洞属性分析的软件安全评估方法

37. 基于社区协同推荐系统的攻击研究

38. 可视化数据分析的艺术

39. 基于用户的推荐系统技术应用及建构

40. 基于Fisher线性判别分析的情景感知推荐方法

41. 基于属性分析的供应链研究途径

42. 基于Iot图书推荐系统研究

43. 基于场景感知的服务系统推荐算法

44. 基于学习资源的推荐系统研究

45. 一种基于资源多属性分类的群组推荐模型

46. 基于多属性评分的酒店推荐系统

47. 基于属性偏置矩阵分解冷启动推荐

48. 基于网络结构的推荐系统

49. 基于Web的物流推荐系统分析

50. 基于L系统的树木生成与可视化研究