

正态云模型研究回顾与展望

杨洁¹⁾ 王国胤¹⁾ 刘群¹⁾ 郭毅可²⁾ 刘悦²⁾ 涂文燕³⁾ 刘玉超⁴⁾

¹⁾(重庆邮电大学计算智能重庆市重点实验室 重庆 400065)

²⁾(上海大学计算机工程与科学学院 上海 200444)

³⁾(解放军理工大学指挥信息系统学院 南京 210007)

⁴⁾(中国指挥与控制学会 北京 100089)

摘要 不确定性信息的表达和处理是人工智能的一个重要研究问题,目前有多种理论模型从不同的角度研究不确定性问题,包括模糊集、粗糙集、概率论、证据理论等.1995年,李德毅院士在概率论和模糊集理论两者的基础上,提出了一种处理不确定性问题的双向认知模型——云模型,即通过正向云变换和逆向云变换算法进而实现定性概念与定量数值的双向转换.经过20多年的研究与发展,该模型逐渐得到完善,并在不确定性信息处理方面得到了广泛应用.该文回顾了正态云模型理论的研究现状和进展,并在此基础上分析了其存在的挑战和问题:(1)双向认知计算方面,虽然多步逆向云能够实现稳定的双向认知变换,但是它是基于单个云概念的,尚未有多粒度云模型双向认知变换的研究成果;(2)云模型相似性度量方面,由于不同的领域问题需要不同的评价标准,需要开展针对特定问题的云模型相似性度量研究;(3)粒计算机制方面,高斯云变换能够实现由细到粗的粒度变换以及多粒度概念的自适应生成,解决了云模型的变粒度问题,但是没有体现不同云概念以及不同粒度层次之间的关联;(4)多维云模型方面,目前在这方面的研究工作相对较少,缺乏较有效的多维云表示方法.该文针对以上问题,围绕当前的研究热点——大数据存在的挑战,进一步提出了大数据的云模型研究框架,深入探讨了未来的研究方向,并指出未来的工作需要以大数据为中心,结合粒计算、机器学习以及统计学的思想,进一步完善云模型的理论机制,该文的工作对于大数据和云模型理论的研究提供了重要的参考价值.

关键词 云模型;双向认知;粒计算;不确定性;大数据

中图法分类号 TP18 **DOI号** 10.11897/SP.J.1016.2018.00724

Retrospect and Prospect of Research of Normal Cloud Model

YANG Jie¹⁾ WANG Guo-Yin¹⁾ LIU Qun¹⁾ GUO Yi-Ke²⁾ LIU Yue²⁾ TAN Wen-Yan³⁾ LIU Yu-Chao⁴⁾

¹⁾(Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065)

²⁾(School of Computer Engineering and Science, Shanghai University, Shanghai 200444)

³⁾(Institute of Command Information System, PLA University of Science and Technology, Nanjing 210007)

⁴⁾(Chinese Institute of Command and Control, Beijing 100089)

Abstract The expressing and processing of uncertain information is one of the key basic issue in the era of artificial intelligence. Currently, there have been many theories, such as fuzzy set, rough set, probability statistics, evidence theory, etc., to deal with uncertainty information from different perspectives. The cloud model theory as a new cognition model for uncertainty proposed by Prof D. Y. Li in 1995 based on probability theory and fuzzy set theory, which provides a way to realize the bidirectional cognitive transformation between qualitative concept and quantitative

收稿日期:2016-10-26;在线出版日期:2017-08-19.本课题得到国家重点研发计划(2016YFB1000905)、国家自然科学基金(61572091)、重庆市研究生科研创新项目(CYB16106)资助.杨洁,男,1987年生,博士研究生,副教授,主要研究方向为云模型、粒计算、粗糙集、机器学习. E-mail: 530966074@qq.com. 王国胤(通信作者),男,1970年生,博士,教授,中国计算机学会(CCF)杰出会员,主要研究领域为粗糙集、粒计算、认知计算、智能信息处理、数据挖掘. E-mail: wanggy@ieee.org. 刘群,女,1969年生,博士,教授,主要研究领域为非线性理论、粒计算. 郭毅可,男,1963年生,博士,教授,主要研究领域为机器学习、数据挖掘、并行计算. 刘悦,女,1975年生,博士,副教授,主要研究方向为机器学习、数据挖掘. 涂文燕,女,1971年生,博士,副教授,主要研究方向为机器学习、数据挖掘. 刘玉超,男,1980年生,博士,高级工程师,主要研究方向为云模型、粒计算、智能驾驶.

data by forward cloud transformation algorithm (FCT) and backward cloud transformation algorithm (BCT). After more than 20 years of research and development, Cloud model has been extensively studied and gradually improved, and it has been widely applied in uncertain information processing in the past decades. Through reviewing the fundamental results and state-of-art research, we discuss some key problems and challenges for normal cloud model, including: (1) Although stable bidirectional cognition transformation between qualitative concept and quantitative data can be realized by multi-step backward cloud transformation, it is still only a cognition transformation based on single cloud concept, and there are still lack of results involving cognition transformation on multi-granularity cloud concepts. (2) For the similarity measure of cloud model, there exist different evaluation criterions for the similarity measure of cloud models in many fields, therefore it is necessary to develop the similarity measure of cloud model for specific application problems. (3) For granular computing mechanism of cloud model, the Gaussian Cloud Transformation (GCT) can realized the granularity transformation from fine to coarse and generate the multi-granularity concepts adaptively, which solve the problem of variable granularity. However, it has no ability to reflect the relationship between different cloud concepts or different granularity. (4) Owing to lack of studies on multi-dimensional cloud model, there are still no valid and reliable methods for its representation. Based on the above problems, in this paper, we point out the existing challenge problems in big data which is a current research issue, and we propose research framework for cloud model based big data. Future perspectives and potential research topics are also discussed. Finally, we point out the future work of cloud model should combine with granular computing, machine learning and statistics. These works are significant for the development of big data and cloud model.

Keywords cloud model; bidirectional cognitive; granular computing; uncertainty; big data

1 引言

不确定性知识的表达和处理是人工智能中最重要的问题之一. 如何有效地从不确定性信息和数据中获取知识是不确定性人工智能的一个重要研究课题. 概念的不确定性具有随机性、模糊性以及不完备性等特征, 其中随机性和模糊性是最基本的两种不确定性, 具有很强的关联性, 在人类认知过程中尤为重要^[1-2]. 在不确定性的研究中已涌现出许多理论模型, 从不同的角度处理不确定问题, 例如, 概率论^[3]、模糊集^[4]、粗糙集^[5]、证据理论^[6]等. 这些理论模型有着比较严谨的约束条件, 但在处理问题时也具有一定的局限性. 例如, 模糊理论通过隶属函数刻画了自然语言中概念的亦此亦彼性, 但是隶属函数的选择具有主观性, 忽略了隶属函数本身具有的不确定性. 传统的贝叶斯推理^[7]基于概率表示事件的不确定性. 对于形如 if A then D ($A \rightarrow D$) 的规则, 事件概率由属于 A 事件的样本频度确定, 然而概率表示侧

重于描述事件 A 出现的不确定性, 即预测的不确定性, 无法描述 A 事件相关概念的语义不确定性, 例如规则“年老的驾驶员很少闯红灯”, 概率只能描述闯红灯的驾驶员为老人的频度, 而无法描述究竟多大年龄的人可认为是年老的驾驶员.

大多数知识表示方法的研究主要聚焦在从定量数据中提取定性概念, 也就是从概念外延转换到概念内涵. 1995 年, 李德毅院士基于概率测度空间提出了定性概念与定量数据双向转换的认知模型——云模型^[8], 通过引入 3 个数字特征: 期望 Ex (Expected Value), 熵 En (Entropy) 和超熵 He (Hyper Entropy), 并结合特定生成算法, 构造出服从泛正态分布的随机变量——云滴, 将认知中的模糊性纳入到概率框架中进行统一描述. 其中, 期望 Ex 代表定性概念的基本确定性, 是云滴在论域空间分布中的数学期望, 即最能够代表定性概念的点; 熵 En 代表定性概念的不确定性度量, 由概念的随机性和模糊性共同决定; 超熵 He 代表熵的不确定性, 可以将定性概念的随机性约束弱化为某种“泛正态分布”, 反映定性概念所对应

的随机变量偏离正态分布的程度. 再者, 云模型通过正向云变换 (Forward Cloud Transformation, FCT) 和逆向云变换 (Backward Cloud Transformation, BCT) 实现定性概念及其定量表示之间的相互映射, 提供了从数据到知识以及从知识到数据的双向认知通道. 因此, 云模型不仅反映了定性概念自身所具有的不确定性, 同时揭示了客观事物的随机性和模糊性的关联^[9]. 以上特性决定了云模型可以作为表示概念的基本模型.

云模型理论发展二十余年来, 在理论研究上不断得到完善, 云模型发生器、云规则发生器、逆向云算法、云变换和云模型的粒计算等理论相继被提出. 特别是逆向云算法和云模型的粒计算方法的提出, 突破了云模型理论存在的许多瓶颈, 为以后的研究奠定了坚实的基础. 2012 年王国胤教授和李德毅院士等合作出版了相关专著《云模型与粒计算》^[10], 对云模型与粒计算的结合进行了探讨. 2014 年李德毅院士再次出版了《不确定性人工智能》^[11], 进一步深化了云模型理论研究. 同时, 云模型在智能控制、数据挖掘、系统评测等方面也得到了较好的效果^[12-18].

不同的云模型具有不同的分布形状, 其中正态云模型是最重要的一种云模型, 具有普适性^[19-20], 并在许多领域中的应用都取得了成功^[21-23]. 正态云模型生成的云滴分布不仅具有“中间大、两头小”的特征, 可以用于研究近似服从高斯分布的现象, 而且具有尖峰肥尾的特征, 可以描述二八定律所反映的幂律现象, 具有非常重要的数学意义. 王国胤教授在文献^[24]中通过正态云生成云滴的特点, 提出了一般正态云模型, 建立了正态云模型与正态分布之间的联系. 在云模型随机变量符合高斯分布的基本假设下, 我们可以建立云模型与统计理论中多个概念的对应关系, 如表 1 所示. 这样的对应关系是开拓正态云模型的理论基础.

表 1 云模型和统计学习的对应关系表	
云模型	统计学习
云滴	随机变量
云发生器	分布随机数产生器
逆向云发生器	点估计
云模型	测度空间
云变换	混合分布模型
期望	后验估计期望
熵	似然分布中的方差
超熵	超参数
概念树	特征结构及组合表示

从认知科学的角度来说, 要掌握一门新的学科, 就必须学习思维和计算方法, 建立其体系结构, 只有

这样才能从已知推到未知. 本文将在总结已有研究成果的基础上, 帮助读者建立起一个完整的云模型体系结构, 同时对当前云模型研究存在的问题进行深入分析并提出一些观点和看法.

本文第 2 节介绍正态云模型的相关定义; 第 3 节对正态云模型的相关研究工作及其存在问题进行综述和分析; 第 4 节对云模型在众多领域的应用成果进行梳理; 第 5 节是对云模型进一步研究方向进行展望; 最后部分是总结.

2 相关定义

定义 1^[11]. 设定性概念 C 是定量论域 U 上的概念, 若 $x \in U$ 是概念 C 的一次随机实现, x 对 C 的确定度 $\mu(x) \in [0, 1]$ 是有稳定分布的随机数: $\mu(x): U \rightarrow [0, 1] \forall x \in U$, 则 x 在论域 U 上的分布称为云模型. 其中, 定量数值 x 体现了表示概念的定量值的随机性, 而 $\mu(x)$ 反映了定量数值 x 隶属于定性概念 C 的确定程度.

从定义 1 可知, 云模型不仅能够使用隶属度函数 $\mu(x)$ 描述概念的模糊性, 而且能够描述隶属度函数 $\mu(x)$ 的随机性和云滴的随机性. 由此可见, 云模型与二型模糊集的不同之处在于, 云模型既能描述不确定概念的模糊性又能描述随机性.

定义 2^[11]. 设定性概念 C 是定量论域 U 上的概念, 且 C 包含 3 个数字特征: (Ex, En, He) , 若 $x \in U$ 是概念 C 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: $\mu(x): U \rightarrow [0, 1] \forall x \in U$, 其中满足 $x = R_N(Ex, |y|)$, $y = R_N(En, He)$, 且隶属度函数满足

$$\mu(x) = \exp\left\{-\frac{(x-Ex)^2}{2y^2}\right\} \quad (1)$$

那么所有云滴构成随机变量 X 的分布称为正态云模型.

在定义 2 中, 如果 $He=0$, 那么云滴分布将退化为正态分布, 如果 $He=0, En=0$, 那么 x 将等于常量 Ex 且 $\mu(x)=1$. 即确定性是不确定性的特殊情形. 当 He 很大时, 云滴分布呈重尾性^[25].

从定义 2 可知, 云滴的概率密度函数如下:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy \\ &= \frac{1}{2\pi He} \int_{-\infty}^{+\infty} \frac{1}{y} e^{-\frac{(x-Ex)^2}{2y^2} - \frac{(y-En)^2}{2He^2}} dy \end{aligned} \quad (2)$$

由式(2)可知, 此时的概率密度函数无解析解.

通过对公式进行积分变换可以得到正态云模型云滴的期望、方差、三阶矩、四阶矩, 分别如下:

$$EX = Ex \quad (3)$$

$$DX = En^2 + He^2 \quad (4)$$

$$E(X - EX)^3 = 0 \quad (5)$$

$$E(X - EX)^4 = 3(3He^4 + 6He^2En^2 + En^4) \quad (6)$$

3 研究现状

目前, 围绕云模型理论的研究大致可以分为以下几个方面: 双向认知计算模型、云模型相似性度量、云模型的粒计算机制、多维云模型。通过 20 年的发展, 这几个方面的理论逐步得到完善, 但是仍存在不足之处, 下面将围绕这四个方面的内容展开讨论。

3.1 双向认知计算模型

从粒计算的角度来说, 人们不仅能够在不同粒度层次上求解问题, 而且可以很快地从一个粒度世界跳到另一个粒度世界。这种处理不同粒度世界的能力, 正是对人类问题求解的强有力的表现^[26]。云模型就提供了这样一种实现双向认知的理论模型。双向认知计算模型是云模型理论中比较重要的模型之一, 其中, 正向云变换(FCT)和逆向云变换(BCT)是实现双向认知变换的重要工具。如图 1 所示, 通过 FCT 和 BCT, 云模型可以实现内涵与外延的双向转换。

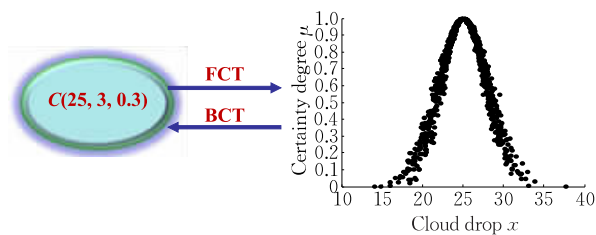


图 1 云模型双向认知转换图

对于正态云模型来说, 由于人类对特定概念的认知总体是一致的, 所以用于实现概念内涵到外延转化的正向云算法通常是固定的; 而用于实现概念外延到内涵转化的逆向云算法各有不同, 不同的逆向云算法所导致的误差和稳定性也不同, 这恰好体现了不同人的思维能力。因此, 对于双向认知计算模型来说, 目前的研究工作基本集中于对逆向云算法的研究。

逆向云算法通过对数据样本集合估计事物的整体特征, 可以模拟人类在认知过程中从具体到抽象, 部分到整体的思维过程, 对于人工智能领域有重要应用价值。从数学的角度来说, 逆向云算法本质上是一个参数估计的过程, 即利用统计学方法对数据样本进行参数估计, 从而得到内涵的数字特征参数。在

逆向云变换中, 难免会有误差产生。如果逆向云算法的估计误差越大, 那么产生的“概念漂移”越明显, 即不能实现概念内涵与外延之间稳定的双向认知映射。因此, 要建立稳定的双向认知计算模型就必须对逆向云算法进行改进, 尽可能的减少估计误差, 以解决内涵与外延变换过程中存在的“概念漂移”问题。

在逆向云算法的研究问题上, 已有了许多相关成果。这些研究成果大致可以分为两类: 有确定度逆向云算法^[27-29]和无确定度逆向云算法^[30-32]。由于有确定度逆向云算法在计算参数 En 和 He 时是需要利用数据样本的确定度, 而在实际应用中代表性概念的确定度却难以获得, 所以有确定度的逆向云算法缺乏实用性。而无确定度的逆向云算法利用矩估计方法对云模型三个特征参数进行估计, 可以克服云模型的概率密度函数积分表达式为非显式解析式而求解复杂的缺点, 对云模型的研究更具有普遍性。从实现概念内涵的数字特征参数中 En 和 He 的估计方式(直接或间接)来看, 无确定度的逆向云算法又可以细分为单步式和多步式的逆向云变换算法。其中, 单步式逆向云算法主要有如算法 1 所描述的基于一阶绝对中心矩的逆向云算法(Backward Cloud Algorithm based on the First Order Absolute Central Moment, SBCT-1stM)^[30]和算法 2 描述的基于四阶绝对中心矩的逆向云算法(Backward Cloud Algorithm based on the Four Order Absolute Central Moment, SBCT-4thM)^[31]。

算法 1. SBCT-1stM.

输入: N 个云滴 $x_i (i=1, 2, \dots, N)$

输出: 数值特征估计值(期望 $\hat{Ex}$, 熵 $\hat{En}$, 超熵 $\hat{He}$)

$$1. \hat{Ex} = \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i;$$

$$2. \hat{En} = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |x_i - \hat{Ex}|;$$

$$3. \hat{He} = \sqrt{S^2 - \hat{En}^2}, \text{ 其中 } S^2 \text{ 为样本方差,}$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$$

SBCT-1stM 算法是首个无云滴确定度的逆向云算法, 既简单又实用, 而且容易推广到高维。但是在实际计算中如果样本的方差 S^2 偏小或者熵的估计值 $\hat{En}$ 偏大, 可能出现 $S^2 - \hat{En}^2 < 0$ 的情形, 此时超熵为虚数, 意味着是一次无意义的计算, 导致估计参数过程不稳定。文献[32]通过删除离期望最近样本点对 SBCT-1stM 算法进行了改进, 保证了超熵的估计为正实数。但是, 在实际应用中, 人类在学习知识的过程中是对定量的样本集合进行抽象, 形成固有的

概念,样本集合是客观存在的. 因此,文献[32]中强制去掉样本的方法并不符合人类认知过程.

算法 2. SBCT-4thM.

输入: N 个云滴 $x_i (i=1,2,\cdots,N)$

输出: 数值特征估计值(期望 $\hat{E}x$, 熵 $\hat{E}n$, 超熵 $\hat{H}e$)

1. $\hat{E}x = \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$
2. $\hat{E}n = \sqrt[4]{\frac{9(S^2)^2 - \bar{\mu}_4}{6}}$; 其中 S^2 为样本方差,
 $S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$, $\bar{\mu}_4$ 为四阶中心距,
 $\bar{\mu}_4 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^4$;
3. $\hat{H}e = \sqrt{S^2 - \sqrt[4]{\frac{9(S^2)^2 - \bar{\mu}_4}{6}}}$

同 SBCT-1stM 算法一样,SBCT-4thM 算法在实际计算中并不能保证每次逆向云变换中 $S^2 - \hat{E}n^2 \geq 0$, 即 $S^2 - \sqrt[4]{\frac{9(S^2)^2 - \bar{\mu}_4}{6}} \geq 0$, 所以超熵的估计值有可能为虚数,从而导致在实现概念外延到内涵的转化过程中缺乏稳定性. SBCT-1stM 算法和 SBCT-4thM 算法都是通过不同阶数的矩估计直接得到云模型的数字特征参数,即单步式估计,忽略了云滴是通过两次随机过程产生的特点. 基于对以上算法的缺点分析以及结合云模型的特点,许昌林和王国胤^[33-34]提出了如算法 3 所示的多步式的逆向云变换算法(Multiple Backward Cloud Transformation based on Sampling with Replacement, MBCT-SR).

算法 3. MBCT-SR.

输入: N 个云滴 $x_i (i=1,2,\cdots,N)$

输出: 数值特征估计值(期望 $\hat{E}x$, 熵 $\hat{E}n$, 超熵 $\hat{H}e$)

1. $\hat{E}x = \bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$

2. 随机抽样分组

FOR($i=1, i \leq m, i++$)
FOR($j=1, j \leq r, j++$)
对 N 个云滴 $x_i (i=1,2,\cdots,N)$ 样本进行随机抽样
ENDFOR
 $X_i = \{X_{i1}, X_{i2}, \cdots, X_{ir}\}, \bar{X}_i = \frac{1}{r} \sum_{j=1}^r x_{ij}$
 $DX_i = Y_i^2 = \frac{1}{r-1} \sum_{j=1}^r (X_{ij} - \bar{X}_i)^2$
ENDFOR
 $Y^2 = \{Y_1^2, Y_2^2, \cdots, Y_m^2\}$

3. $\hat{E}n^2 = \frac{1}{2} \sqrt{4(EY^2)^2 - 2DY^2}, \hat{H}e^2 = EY^2 - \hat{E}n^2$

MBCT-SR 算法保证超熵 He 的估计不会出现 $He^2 < 0$ 的情形,具有较强的稳定性,能实现较稳定的双向认知计算模型.

表 2 为几个初始概念分别采用以上几种逆向云算法进行 50 次双向认知转换后形成的概念. 由表 2 可知,当初始为一个清晰概念时,例如 $C(25, 3, 0.1)$, 在 SBCT-1stM 与 SBCT-4thM 算法得到的结果中 He 发生漂移较严重,而此时 MBCT-SR 算法表现较稳定,当初始为一个不确定概念时,SBCT-1stM 与 SBCT-4thM 算法得到的结果发生了一些漂移,而 MBCT-SR 算法基本接近初始概念,当初始为一个模糊概念时,SBCT-1stM 得到的结果发生漂移较严重,而 SBCT-4thM 算法与 MBCT-SR 算法得到的结果较好. 因此,MBCT-SR 算法的提出为实现稳定的双向认知计算模型提供了必要的基础.

但是,目前的双向认知研究成果只能实现单层次、单粒度的双向转换,需要进一步研究多粒度、多层次的双向转换,从而实现多粒度的知识表示以及对数据对象的全局的认知. 本文第 5 节中将会对此方面的研究方向进行进一步探讨和展望.

表 2 50 次双向认知后概念的三个特征参数^[34]

$C(Ex, En, He)$	SBCT-1stM	SBCT-4thM	MBCT-SR
$C(25, 3, 0.1)$	(24.92, 2.96, 0.34)	(24.92, 2.98, 0.30)	(24.92, 3.01, 0.09)
$C(25, 3, 0.55)$	(24.93, 2.96, 0.36)	(24.93, 3.01, 0.31)	(24.92, 3.03, 0.54)
$C(25, 1, 0.8)$	(25.03, 1.25, 0.41)	(25.03, 0.96, 0.83)	(24.92, 1.02, 0.78)

3.2 云模型相似性度量

云模型不仅可以用于数据挖掘过程中的不确定性表示,而且为挖掘过程的表示提供更符合人类认知的定性分析方法. 定量数据可以通过云模型实现定性概念的转换,而建立在定性概念基础上的数据挖掘任务需要进行相似性度量,例如在系统性能评估中,运用云模型相似性度量给出的结果更有说服

力;在网络舆情检测与控制中,运用云模型相似性度量可以更加准确地衡量概念是否产生了舆情. 再者,人类的认知既具有差异性,又具有相似性. 对于同一个人而言,随着知识和经验的增加,不同时期对同一概念的认知也不尽相同;对于不同人而言,由于先天和后天等各种因素的影响,对于同一概念的认知也不相同. 那么如何对这些差异性进行度量才能

判断原始概念与新认知概念之间的相似性呢?通过 3.1 节的分析可知,概念漂移的大小本身也是检验双向认知是否稳定的指标.因此,云模型相似性度量是一个值得研究的问题.

目前,云模型的相似性计算方法有值系数和形系数两种度量形式.前者通常算其距离(欧氏距离、海明距离等)来进行度量,后者主要计算曲线的相似系数(向量间的夹角余弦及其推广).张仕斌等人^[35]提出了基于交互式的云滴确定度算法来度量两个云之间的相似性.虽然文献^[35]体现了云模型不确定性的特性,但这种方法的实验结果波动较大,精度依赖于云滴数量和实验次数,时间复杂度较高,而且当两个云概念完全相同时,相似度却不为 1,故不满足度量准则中的自反性^[36].廖列法等人^[37]将云模型的数字特征作为特征向量利用欧氏距离来度量云概念间的相似性,忽略了云模型三个数值特征的权重问题.张光卫等人^[38]提出的夹角余弦法(Likeness Comparing Method based on Cloud Model,LICM)法,将云模型的数字特征作为特征向量,通过计算它们之间的夹角余弦来度量云概念间的相似性,虽然在协同过滤系统中取得了较好的效果,但是当熵和超熵远远小于期望时,余弦值差异并不明显,削弱了不确定性.李海林等人^[39]提出的期望曲线法(Expectation based on Cloud Model,ECM)和最大边界曲线法(Max boundary based on Cloud Model,MCM)克服了云相似度计算基于随机选取云滴和特征向量的带来的时间复杂度过高、结果不稳定以及云模型数值特征过于显著等问题,这两种算法分别通过云模型的期望曲线和边界曲线以云模型相交面积作为衡量依据来计算云模型的相似度.但 ECM 法在计算云相似度时忽略了超熵 He 的作用.MCM 法虽然考虑了超熵 He ,但是当超熵 He 很大时,它不能很好地刻画云概念相似度.在 ECM 法基础上提出的超熵约束的含熵期望曲线法(Restricted hyper-entropy expection method based Cloud Model,RHECM)^[40],

将基于含熵期望曲线的“与区域”和“或区域”的面积比作为云概念相似性度量的基准,但超熵系数 k 需要人为指定,带有一定的主观性.查翔等人^[41]提出了基于概念跃升的云模型相似性计算方法(Concept skipping indirect approach of Cloud Model,CCM).这种方法以表征云相似性的期望曲线相交区域为切入点,通过概念跃升在更高层次概念上计算其面积,实现原相交区域代表概念比重的等价转换,本质上借鉴了多粒度的思想,将细粒度的子云分别与粗粒度的综合云进行比较计算子云在综合云中的比例,从而间接得出两个子云之间的相似性.但无论是 MCM 法、ECM 法还是 RHECM 法都是以云模型之间的相交面积代表概念之间的公共成分,当待比较的云模型中某一云模型越清晰时,它们的相交面积越小,得到相似度也就更接近于 0,显然与人的认知不符.例如,利用 ECM 或其他几种基于面积的算法来计算 $C_1(23,0.01,0.5)$ 与 $C_2(23,5,0.4)$ 的相似度则接近 0,显然不符合人的认知.许昌林等人^[42]基于信息论中 KL 散度结合云模型最大边界曲线来刻画云概念的相似性,由于 KL 散度法(Kullback Leibler Divergence based on Cloud Model,KLDCM)可以较好地刻画两个概率分布的差异,相对而言,KLDCM 具有较高的普适性,可以较好的反映云概念之间的相似性.Yang 等人^[43]提出一种基于地球移动距离的云模型相似性度量方法(The Earth Movers Distance based Cloud Model,EMDCM),实现单粒度上多对多的云模型相似性计算,而且易推广到多粒度云模型的计算.

表 3 给出了以上所述的几种云模型相似性度量方法的性能对比,由于 ECM、MCM、RHECM、CCM、KLD、EMDCM 这几种度量方法都是基于期望曲线或边界曲线计算相似度的,而期望曲线由云模型的数字特征计算得到,因此以上方法的稳定性较强,且时间复杂度较低.

表 3 云概念相似性度量算法一览表

	LICM	ECM	MCM	RHECM	CCM	KLDCM	EMDCM
算法效率	高	高	高	高	高	高	高
值影响因素	显著特征值	期望,熵	超熵	熵,超熵	期望,熵	特征值	特征值
稳定性	低	较高	高	高	高	高	高

但是,目前的云模型相似性度量仅是从概念内涵(数字特征值)和概念外延(云滴、交叉面积)上进行研究,而没有考虑具体的领域.只有在相同领域问

题下进行度量才具有统一的评价标准,即云模型的三个特征参数具有固定的权重.除此之外,没有考虑属性量纲也是云模型相似性计算存在的一个问题.

本文第 5 节将对此问题进行详细的阐述和分析。

例 1. 现通过甲、乙、丙三人 50 次考试测试的数据统计来评价三位选手的水平. 假设甲、乙、丙的 50 次测试数据分别通过逆向云算法转化为定性概念 $C_1(79,1,0.3)$, $C_2(80,10,0.8)$, $C_3(75,1,0.1)$.

说明甲发挥较稳定,其成绩在 79 分左右波动,且每次测试的分数较集中;乙发挥很不稳定,其成绩在 80 分左右波动,且每次测试的分数很离散;丙发挥非常稳定,其成绩在 75 分左右波动,且每次测试的分数很集中. 由此可知,虽然乙的 50 次测试的平均分最高,但是发挥非常不稳定. 假设现在要比较三者能力的相似性 S , 如果从知识储备的角度来说, $S(\text{甲}, \text{乙}) > S(\text{甲}, \text{丙}) > S(\text{乙}, \text{丙})$; 从心理素质及考试能力的角度来, $S(\text{甲}, \text{丙}) > S(\text{甲}, \text{乙}) > S(\text{乙}, \text{丙})$; 从综合能力的角度来说, $S(\text{甲}, \text{丙}) > S(\text{甲}, \text{乙}) > S(\text{乙}, \text{丙})$. 但是, 如果没有给定测试的领域问题以及选拔的标准, 就无法评价出甲、乙、丙三人谁能力最强, 哪两个选手最相似. 因为不同的领域问题, 评价选手的尺度和角度是不一样的, 比如传统的考试, 看重的是平均分; 选拔射击选手, 则更看重的是心理素质等.

3.3 云模型的粒计算机制

粒度原本是一个物理学的概念,是指物质微粒大小的平均度量^[10]. 在这里被作为对概念中包含信息量的度量,从不同概念层次分析和处理论域空间中的数据. 如图 2 所示的整个粒空间结构中, $Layer_i$ 表示第 i 层, 其中, $Layer_k$ 表示最细粒度的粒层, $Layer_k$ 中各个小圆点代表最细粒度数据, 具有同类关系的数据构成信息粒.

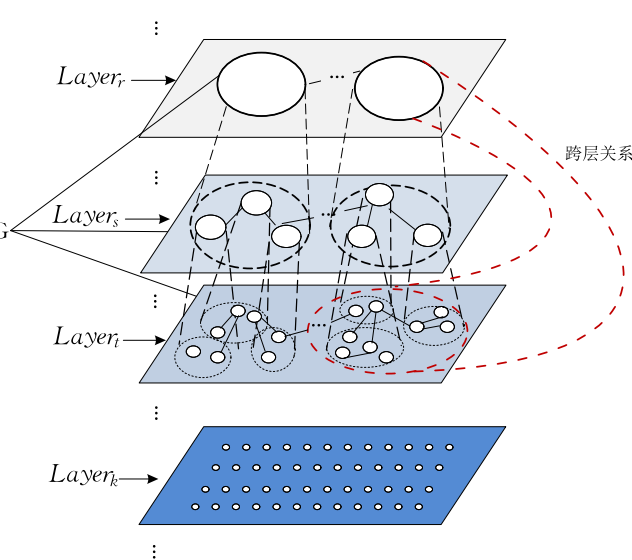


图 2 信息粒 IG、粒层 Layer 及粒结构 GS 示意图

模糊集^[5]、粗糙集^[6]、商空间^[26]是几种常见的粒计算理论模型. 这三种粒化模型分别从不同的角度描述人类从不同粒度解决问题的能力, 表 4 列出了包含云模型在内的几种粒计算模型的一些共性.

表 4 四种粒化模型的共同特征

粒化模型	信息粒	粒结构
模糊集	模糊信息粒	If-then 规则
粗糙集	等价类	层次粗糙集
商空间	商集	商结构
云模型	由特征参数生成的云	云概念树

与其他粒化模型对信息的硬划分不同, 由于定性概念的外延通常是变化的、不确定的, 并涉及到随机性和模糊性, 云模型将模糊理论与概率论相结合, 通过其期望、熵和超熵三个特征参数度量一个定性概念的信息粒度, 可以实现对信息的软划分. 其中, 熵作为概念粒度的基本确定性衡量, 超熵作为粒度的不确定性度量, 构建的概念含混度用来衡量概念的共识程度, 为一个基本概念的代表, 计算和度量提供了基础. 这种粒化方式考虑了数据的概率分布, 获取的知识具有更强的泛化能力. 而且, 通过云模型进行粒化不需要先验知识, 它可以从原始数据中统计规律, 实现定量到定性概念的转换.

概念层次在数据挖掘中有着重要的作用. 通过自动生成概念层次, 可有效地提高数据挖掘的效率, 在不同层次上发现知识^[44]. 由于数据库中大量存在各种数值型属性, 一些传统的概念层次生成算法^[45-47]通过对论域进行硬划分从而生成概念树, 不能反映定性概念的模糊性、随机性以及不同层次概念之间的多隶属关系. 由于基于云模型的粒化方式考虑了数据的概率分布, 获取的知识具有更强的泛化能力. 结合虚拟泛概念树的跃升原理, 建立多层次、多粒度概念的外延空间, 我们称这样一个过程为云变换. 云变换本质上是一种聚类算法, 相对于传统的确定性聚类算法而言, 云变换的优势在于: 通过云模型来描述概念, 能够更好的反映概念的不确定性. 云综合和云分解是云变换的两种形式. 目前的研究主要集中于云综合, 即两个或者多个云概念通过某种合并方式跃升为一个较高层次的云概念, 从粒度变换的角度来说, 就是细粒度到粗粒度的变换过程.

文献^[44]首次提出了基于云模型的泛概念树生成算法, 提出了峰值法云变换, 生成原子概念集, 再通过提出的概念提升的方法——“软或云”, 对原子概念集进行概念跃升, 最终生成泛概念树. 这种综合方法忽视了云的雾化现象, 当综合后概念的数据样

本不满足正态分布,甚至偏离正态分布程度较大时,那么综合后的云模型则会出现雾化现象.文献[48]提出一种基于截断熵的云综合算法,也叫“积分云”综合法,通过将每两个概念分别综合的方法进行云综合,但是在合并时没有受到其幅度系数的影响.云综合必须针对特定的领域,如在一定的数据集中进行概念跃升,仅对两个独立的云概念进行综合是没有意义的,而在一定的领域中不同的子概念又具有不同的幅度.文献[49-50]针对文献[48]没有考虑幅度系数对合并的影响,提出了一种新的概念跃升算法——“幅度云”,但文献[49]的方法与云滴数量有关,不符合人的认知.文献[50]是基于截断熵的幅度云算法,是文献[48]的改进,能够较准确的从数据库中提取定性概念,从而构建概念层次.文献[51]提出了“格贴近度”作为条件来判别综合后的云是否会出现雾化现象.上面三种云综合方法中,“软或云”综合方法是直接利用概念内涵对两云概念进行合并,缺乏理论基础.“积分云”法利用期望曲线和概念内涵对云概念进行综合,具有一定的几何意义.由于“幅度云”法利用逆向云变换估计合并后概念的数字特征,所以对逆向云算法的精度要求较高,且生成的云模型特征参数直接与正向云变换产生的云滴数量有关.

文献[52]从二阶正云模型的期望曲线以及内外包络曲线出发,根据不同区域内的云滴群对云概念的贡献程度,结合云模型的“3En”规则,提出一种更具有解释性的云概念综合方法,但是缺点是没有考虑幅度系数对合并的影响.综上所述,无论哪种概念合并方法都不可能解决所有概念的跃升问题,因为语境、主题等因素的不同会导致概念的具体含义发生变化.因此,概念合并方法是否有效与具体的应用领域有关.

文献[53]根据不确定性概念在不同粒度层次上的不同表现形式,提出了自适应高斯云变换(Adaptive Gaussian Cloud Transformation, AGCT),建立了一种多粒度的推理模型与方法.图3为利用AGCT对院士群体从5个概念到2个概念的变粒度认知过程,概念间相互交叠,总体上是一棵泛概念树.基于高斯混合模型,AGCT通过对原始数据分布进行拟合,实现了细粒度概念到粗粒度概念的变换,获得不同粒度层次的云概念,可以实现不同粒度概念之间的柔性切换,构建泛概念树,解决了粒计算中的变粒度问题,有着广阔的应用前景.

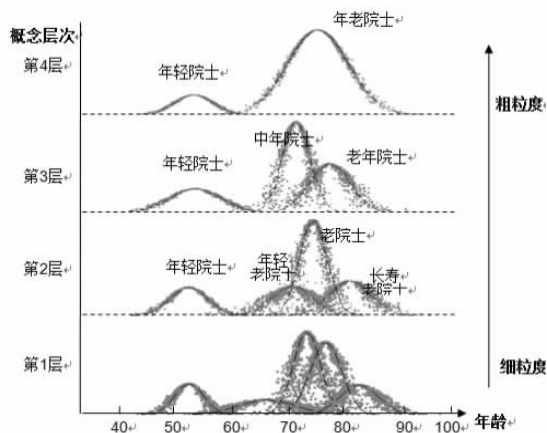


图3 通过AGCT对院士群体聚类形成的多粒度概念^[53]

以上文献所提出的云变换算法都需要首先求数据的频率直方图,即频率分布函数,而频率分布函数的求解是一个时间复杂度较高的过程.再者,目前的云变换研究成果主要集中于云综合的研究,只能实现细粒度到粗粒度的变换,而且基本都是通过不断的迭代实现变粒度,直到达到满足约束条件为止,增加了时间复杂度.反之,如果能实现从粗粒度到细粒度的云变换,增加粒度变换的灵活性,可以从很大程度上减少时间复杂度.

根据运用“粒”求解问题时对粒度层次的使用特征,基于粒计算的研究工作又可以归纳为3个基本模式^[54]:粒度空间优化、粒度层次切换和多粒度联合问题求解.图4为基于云变换的粒度空间优化示意图.优化空间解的有效性可以通过一个二元组来定义,图中用 $SolutionEffectiveness(GM(R), T)$ 来表示.其中 R 是计算的结果, $GM(R)$ 是该结果的粒度度量, T 是时限需求.若将用户需要的结果粒度表示为 $GM(R_u)$, 用户的时限约束表示为 T_u , 粒度空间寻优就是需要找到某个粒层 $Layer_i$, 使得 $GM(R_i = Solve(Layer_i)) \leq GM(R_u)$, 并且 $T_i \leq T_u$, 即求得映射 $Layer_i$ 上同时满足了解的粒度要求和时限约束的解,然后在上面进行智能计算与分析.

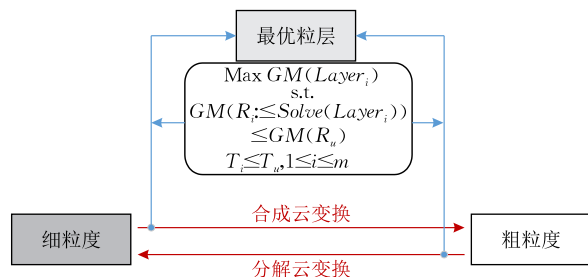


图4 利用云变换实现粒度空间优化示意图

例 2. 图 5 所示,展示了一个云模型粒度空间优化的例子,通过实际约束条件调节概念含混度,可以实现不同粒度层次上对遥感图像进行图像分割,使其不同的概念层次具有不同的分割结果.根据当前的时限约束条件,从多个分割结果中选择最符合人类认知的结果.

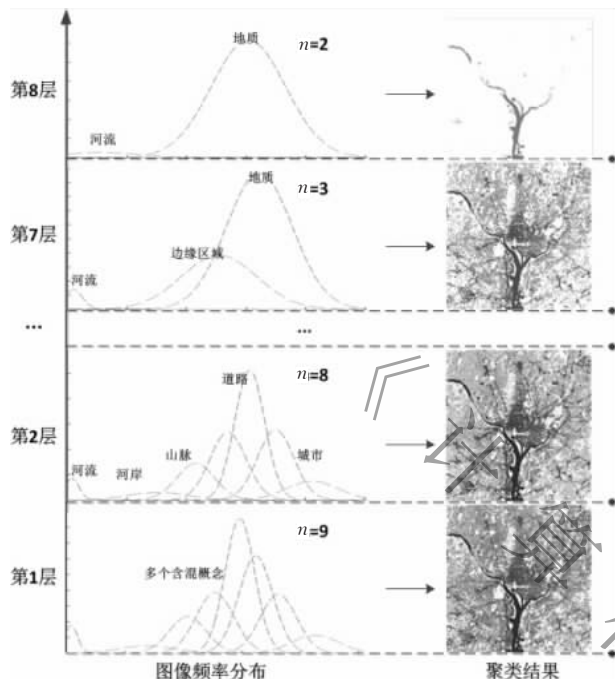


图 5 粒度空间优化例子

云模型的粒计算机机制从概率统计出发,开拓了计算研究领域的新方向.目前,基于云模型的粒计算机机制在粒度空间优化机制方面取得了一些成果^[51,55-56],即在粒结构中选择某一粒度层次进行问题求解.而在另外两种粒计算机机制方面的研究成果甚少.如果能充分将其它粒计算的方法和机制运用到云模型上面,则可以在很大程度上促进云模型在大数据和其他方面的应用研究.在第 5 节中将会对此方面的研究方向进行进一步探讨和展望.

3.4 多维云模型

基于一维云模型的规则发生器只能表示单输入、单输出的情形,而现实生活中任何一次推理过程的完成都是由多种相关联的因素共同决定的,尤其在处理的大数据处理中,我们常常要面对的是多维的属性数据.由于维数间存在着耦合关系,使得对高维数据空间的划分更为困难.所以,一维云模型在某种程度上存在局限性.图 6 为利用 X 条件云和 Y 条件云来构造的二维云规则发生器^[57],确定一个输入值后,就能由规则发生器推出相应的有规律的结果,最终达到预测的效果.

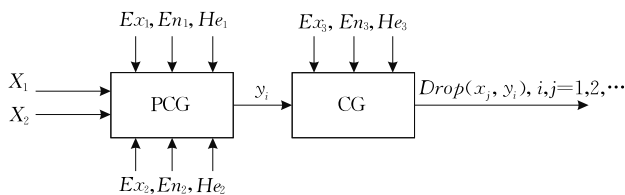


图 6 二维云模型规则发生器^[57]

在多维云的研究方面,杨朝晖和李德毅^[57]在一维云模型的基础上首次提出了如图 7 所示的二维正态云模型,并给出了二维云的数字特征和数学模型,从而解决了双输入、单输出的复杂性规则表示.张国英等人^[58]在二维云的基础上提出基于核空间的高维云模型,并将其应用于对地理信息系统内选定区域进行定性评价,从而选出合适的观察所的位置.王成全^[59]提出了一种二维尺度云变换算法,通过对定量的二维数据进行划分,从中提取出定性规则.该方法不仅能够处理双输入单输出的系统样本数据,而且无需任何先验知识.文献[60]提出了高维正向云模型生成算法(Multidimension Forward Normal Cloud Transformation, MFNCT),并结合基于香农熵和层次分析法设计的权重公式用于水质富营养化程度评价,相比之前的评价方法取得了更好的效果.总体来说,当前仍然缺乏多维云模型方面的相关研究,除了多维云模型的表示之外,多维云模型的双向认知计算、相似性度量、粒计算机机制都是未来需要研究的内容.

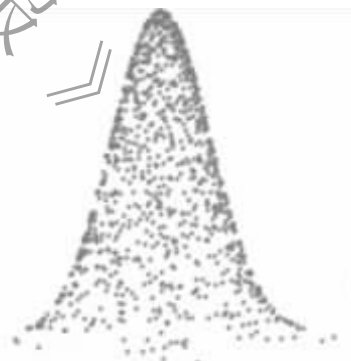


图 7 二维云模型示意图

4 基于云模型理论的应用研究

目前,云模型的应用研究主要基于云模型的定性概念的不确定知识表示,定性概念到定量数值双向认知转换和多层次泛概念树等理论方面.这些研究成果成功应用于许多领域,其中包括智能控制、数据挖掘、系统评估、信号处理、图像处理、知识建模等

领域,有的已经获得了商业价值。

4.1 智能控制

在智能控制领域,基于云模型的定性知识表示与推理的方法已被成功应用于实现三级倒立摆^[16]、无人驾驶智能车运动控制^[13]等非线性控制系统的稳定性控制难题。三级倒立摆作为一个多变量、非线性、不稳定且具有不确定因素的控制系统,不能采用精确数学方法进行形式化描述,但人类却可以通过自然语言表述的定性规则进行稳定控制。对此类复杂的被控系统,引入云模型实现定性规则中语言值概念的不确定性表示,通过云控制与云推理,可以实现被控系统的有效控制。李德毅院士^[16]基于云控制器有效地实现了单电机控制的一、二、三级倒立摆的多种不同动平衡姿态。该控制器不需要利用被控对象的精确数学模型,通过语言原子和云模型转换到语言控制规则器中,就能很好地实现对倒立摆的控制。图8所示的是基于云模型对三级倒立摆实现智能控制的实验结果。其中,图8(a)为三级倒立摆四种动平衡模式下小车的位移和时间关系曲线,可以看出在不同模式下小车的摆动幅度和周期的差别;图8(b)为三级倒立摆不同动平衡姿态之间的动态切换关系曲线,控制系统不仅可以实现不同平衡模

式之间的动态切换,而且具有较强的鲁棒性。

三级倒立摆的研究成果具有普遍意义,容易推广到其他控制对象上,比如,云控制与云推理被推广应用于无人驾驶智能车的不确定性控制,实现跟车行驶、换道行驶、路口通行等多种模式下对车辆速度和转角的不确定性控制。采用云推理和云控制进行定性控制时,首先要对人类驾驶员的定性经验和知识进行云模型表示,建立云控制器的控制规则集合,然后通过云推理与云控制,生成带有不确定性的车辆速度与转角的控制输出值,从而控制车辆在道路环境中安全、稳定行驶。

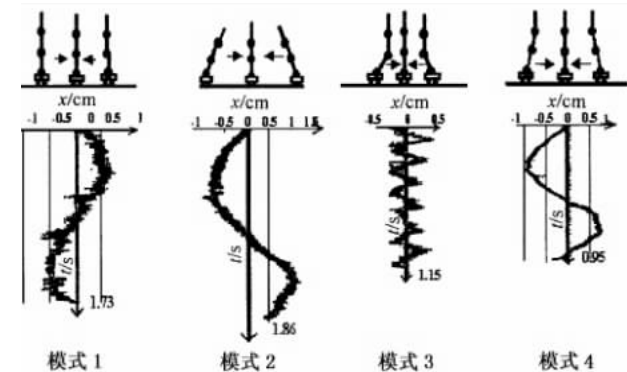
李众等人^[61-63]基于混合维云模型的推理方法以及结合复杂控制规则的简化方法,利用云模型能够很好的模拟人类逻辑推理的能力,通过云模型将自然语言表达的定性控制经验,转换到语言控制规则器中,针对不同的控制对象,提出了相应的云控制器,取得了很好的效果。

总体来说,基于云模型的智能控制方面的研究取得了不错的成果,但是这方面的研究还有许多不完善的地方,如何设计出鲁棒性更强的动态云规则发生器有待于进一步研究。

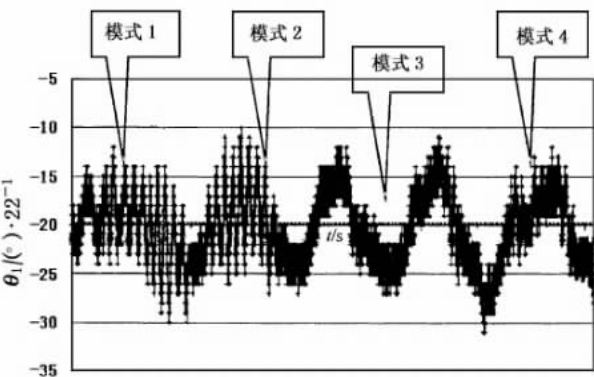
4.2 数据挖掘

虽然目前有许多数据挖掘方法,比如回归分析、神经网络、聚类分析等,但这些方法忽视了数据挖掘中的不确定性以及知识本身的不确定性,不能完全实现对数据进行挖掘的智能化。由于融合云模型的数据挖掘方法考虑了模糊性和随机性,因此其结果更加合理。

在关联规则挖掘方面,由于云模型能够体现人类思维中的不确定性,同时又能较好地反映属性值的分布特征,可以实现定量到定性的转换。基于以上特点,文献[64-68]提出了相关的基于云模型的关联规则挖掘算法。王德鲁等人^[67]利用基于云模型改进的关联规则来分析行为,用于对企业转型战略风险进行预警,取得了较好的效果。因此,可以实现概念的软划分的云模型关联规则算法,更加贴近人的思维方式。陆建江等人^[69]利用云模型对数据属性论域进行划分,并提出了基于正态云的关联规则,由于云模型对数据属性论域进行了软划分,所以利用已挖掘出来的规则和预测结果更容易让人理解。但是以上文献中的概念定义都是主观的,如何根据数据库中已有的数据进行更加合理的动态划分仍需要进一步研究。



(a) 三级倒立摆四种动平衡模式下小车的位移和时间关系曲线



(b) 三级倒立摆不同动平衡姿态之间的动态切换关系曲线

图8 基于云模型对三级倒立摆控制的实验结果^[16]

在趋势预测方面,虽然传统的数据分析法和机

器学习方法各有优点,但是它们属于精确预测,没有考虑粒度问题,使得预测的结果并不令人满意.杨朝晖等人^[57]提出了二维云模型及其规则发生器,对保险库数据进行了预测,验证了二维云模型在预测领域的可行性.蒋嵘等人^[70]以云模型为知识表示,综合两种不同粒度的预测知识实现了时间序列的预测.苗良等人^[71]提出了一种基于云模型的农业害虫的定性预测模型,并应用此模型对麦芽管蚜的发生期进行了预测.Deng 等人^[72]结合云模型和模糊预测模型,提出了一种多粒度的水质连续时间序列预测模型,与当前的水质预测模型相比,取得了更高的预测精度.此外,云模型还被应用于网络状态^[73]、石油钻井事故^[74]和矿井涌水量^[75-76]等的预测,均取得较好的效果.总体来说,以上文献用云模型来预测基本都是按照以下的步骤进行:首先,通过聚类或者专家经验划定影响预测对象的几个定性概念,并对每个概念进行程度划分,例如,年龄可分为老年、中年、青年和少年 4 种程度;然后,通过逆向云算法得出每个概念所对应的云模型 3 个数字特征;最后,利用条件云发生器,构建规则发生器,输入待预测对象的属性数据,通过激活相应的规则,得出预测结果.

在分类方面,文献^[77-80]针对分类对象的不确定性,基于云模型提出了相关的分类模型,克服了确定性分类模型的硬划分缺点,可以更加客观地描述对象行为的不确定性和随机性特征.其中,刘禹^[80]提出基于云模型的分类算法并研究了算法的效率、可解释性、伸缩性与健壮性,选用 29 个标准分类测试集数据,从平均正确率、W/T/L 指标和 Rank 指标等角度与 17 个优秀分类器进行了对比分析,验证了基于云模型的分类算法具有更好的分类性能.

4.3 系统评估

在复杂系统评估过程中,由于待评估对象本身行为以及相关测试数据等因素的不确定性,尤其是专家意见通常采用自然语言描述,具有一定的随机性和模糊性.不确定性的广泛存在一方面增加了测试与评估过程中定性定量综合集成的难度,另一方面这种不确定性也正是专家智慧的体现^[81].因此,充分利用这种不确定性,建立定性概念与定量描述之间的双向映射关系,是提高系统评估结果可信度的重要手段之一.由于云模型可以实现定性定量之间的双向转换,能够克服对某一定性概念进行评价时缺乏客观性的缺点,从而实现对评价客体的有效评估.

目前,有许多研究者将云模型应用于复杂系统

的综合评估中.宋远骏和李德毅^[82]首次将云模型应用于电子产品可靠性的评价中,解决了多变的工作条件下电子产品的可靠性评估问题,并将此方法推广到计算机可靠性评价中^[83].在土地资源方面,胡石元等人^[84]将云模型与关联度分析方法相结合应用于土地评价因素的权重挖掘,并将实验结果与其他方法进行对比,具有较高的准确度.在网络方面,黄海生等人^[85]基于主观信任的主观性、随机性和模糊性,将云模型引入到网络安全领域,给出了信任等级云的定量描述,提出一种主观信任评估模型.王守信等人^[86]提出一种基于云模型的主观信任评价方法,并以亚马逊网站的客户对商品的历史评分信息为例,为客户合理完成交易对象选择提供指导,从而做出正确的决策.虽然利用云模型来代表信任度比较理想,但是它在上下文感知的信任评价以及动态更新方面效果不明显.文献^[87]针对这些缺点引入了贝叶斯网络这种不确定的推理工具,建立了更有效的动态信任评估模型.在通信方面,徐晓斌等人^[88]提出了一种适用于无线传感网络的轻量云模型,并给出了信任管理中常见的直接信任、间接信任以及推荐信任的描述和计算方法,实现了对无线传感网络中节点信任度的不确定性的全面描述和评估.最近,高洪波等人^[89]提出了基于云模型的智能驾驶车辆变粒度测评系统,由于结合了云模型自身的优势,整个综合评判方法更具有说服力,评判结果更符合人的认知.因此,基于云模型的多因素评价策略将成为未来的一个有潜力的研究方向.

4.4 信号处理

在信号处理方面,李兴生等人^[90]通过云模型对信号聚类的结果进行建模,最后基于云模型的不确定性推理得到接收信号的调制方式.史军涛等人^[91]针对通信干扰系统效能评估的复杂性,提出了基于云模型的通信干扰系统性能评估模型.由于此方法融合了主客观信息,有效地结合了模糊性和随机性,可操作性强,为通信干扰系统效能评估提供了一种有效方法.陈金等人^[92]提出了基于云模型的相位拟合方法,利用短时频率估计值建立正态云,通过隶属度进行分类识别,从而淡化了非彼即此的硬划分,并通过实际电台实验验证了该算法具有较高的识别率.

4.5 其他

目前,云模型理论不仅在智能控制、数据挖掘中的预测、系统的综合评估等领域取得许多成果,而且在图像处理^[55,93-94]、数据预处理^[95]、知识建模^[96-99]、

网络安全^[100-102]也得到了广泛应用. 这些研究成果充分显示, 由于云模型具有处理不确定性和双向认知的优势, 可以为解决许多领域的一些算法瓶颈提供帮助. 随着云模型理论的日渐成熟, 其应用领域也将会越来越广泛.

5 云模型研究展望

互联网时代, 随着社交网络、电子商务与移动通信等技术的蓬勃发展, 人类社会进入以 PB 级数据信息为特征的大数据时代^[103-104]. 目前, 围绕大数据的相关研究大多是面向方法与工具, 聚焦于研究各种发现新知识的工具与算法, 而忽略了纷繁复杂的数据表象中蕴涵的知识不确定性以及认知中的过程不确定性. 事实上, 现代科学研究表明, 不确定性是客观世界的基石. 作为客观世界量化特性反映与本质规律描述的数据与知识, 不确定性是其本质的重要体现, 在从数据到知识的认知学习过程中, 不确定性是无时不在的基本特征.

然而, 目前主流的机器学习研究对不确定性这一物理世界与认知过程的基本特性作为大数据研究的基础问题的研究还是很不够的. 以联结主义为基础的模式识别的主要研究方向为深度学习^[105], 相关方法在计算机视觉、语音识别等传统模式识别领域表现出优异的学习性能. 但其认知原理本质难以解释, 针对特定数据与特定问题如何进行模型与参数学习, 大都是凭经验猜测加实验探索, 结构与参数选择具有太多的随意性或试凑性, 是一种不带语义的学习模型(数据缺乏语义解释, 深度神经网络内部结构是“黑箱”), 而且这一学习过程中的不确定性并没有系统的研究, 而以统计学习为基础的数据挖掘的主要研究方向是学习算法及其优化. 虽然不确定性研究在统计模型评价与选择中有很大应用, 贝叶斯理论已经被广泛应用于统计学习中的不确定性描述和推理. 但是, 以贝叶斯为基础的不确定性描述, 只是描述预测的不确定性, 而不能描述概念本身语义的不确定性.

综上所述, 机器学习的一个重大挑战在于如何广泛、有效应用于以不确定性为基本特性的现实世界, 并进一步发展成为能够仿照人类的学习模式从客观世界中获得的量化数据来得到定性的感官认知的机器智能. 为此, 有必要建立一套系统的以不确定性的数据为对象的具有语义解释的大数据机器学习理论与方法. 其核心问题是: 如何从分布在高维数据

空间的大数据中找出聚合的、不同粒度的概念及概念之间的复杂关联, 实现概念及其相互关联的不确定性描述, 以及如何实现从数据到知识的不确定性认知过程的计算建模. 除此之外, 如何从认知计算的角度应对大数据的 3V (Volume, Velocity, Variety) 特性^[106-107], 也是值得研究的问题.

云模型以刻画物理世界与认知过程的不确定性为基本出发点, 将概率统计的工具与模糊数学相结合, 建立了描述与推演数据(物理世界观察)与语义(认知的理解)之间关系, 从而创建了定性概念(语义)与定量观察(数据)之间的相互关联. 再者, 云模型用期望 Ex , 熵 En , 超熵 He 三个数字特征来表达一个概念. 期望表示最能够代表定性概念的点, 熵反映定性概念的不确定性度量, 而超熵是熵的不确定性度量, 由概念的不确定性和模糊性共同决定. 因此, 通过这三个特征参数形成的隶属度分布的统计特征可以用于表达语义不确定性. 云模型的提出恰恰为知识的不确定性给出了一个具有概率语义解释的、可以学习与推演的描述模型. 对于大数据的研究具有重大的意义. 但是, 以往的云模型研究主要是以云模型为中心, 围绕完善云模型自身理论而展开. 因此, 未来的工作需要以大数据为中心, 结合粒计算、机器学习以及统计学的思想, 进一步完善云模型的理论机制. 本节将围绕大数据为研究背景, 分别对云模型理论和应用两个方面进行展望, 希望为云模型下一步的研究方向提供有价值的参考.

5.1 云模型的理论展望

现存的机器学习计算模式从最细粒度出发对数据处理进行处理, 较少考虑人脑思考问题时对粒度的准确把握以及人类自身的认知特点. 这种直接处理最细粒度原始数据的方法显然不符合人类的认知规律. 针对粒计算在处理复杂大数据挖掘中的可行性, Skowron 等人^[108]指出利用粒计算框架应对大数据带来的挑战是一个理想的策略. 梁吉业等人^[109]认为粒计算有望为大数据挖掘提供一条极具前途的崭新途径. Xu 教授和 Pedrycz 教授^[110]分析了粒计算框架下的模糊集研究, 指出通过扩展模糊集以及结合其他理论的, 能够为大数据处理提供新的方向. 因此, 我们可以借鉴人类智能的多粒度思维与问题求解模式, 在第 3 节和第 4 节的综述分析的基础上, 针对云模型当前研究存在的问题, 以大数据为中心, 对云模型理论进行改进和扩展.

图 9 为本文提出的基于大数据的云模型研究框架. 其中, 图中左边深色框为云模型的四个理论. 目

前,较成熟的一维或二维云模型在某种程度上存在局限性.再者,云变换方法主要是针对一维属性数据,缺乏对于多维属性数据处理的有效方法.如果不能实现多维云模型的基本模型和相关的变粒度变换机制,那么利用云模型进行数据挖掘是没有意义的.所以多维云的表示、双向认知、相似性度量及其粒计算机机制将是实现大数据挖掘的基础.图9中浅色方框中的内容为本文提出的云模型未来研究的四个方向.下面将围绕这四个方面对大数据的云模型研究框架进行探讨.

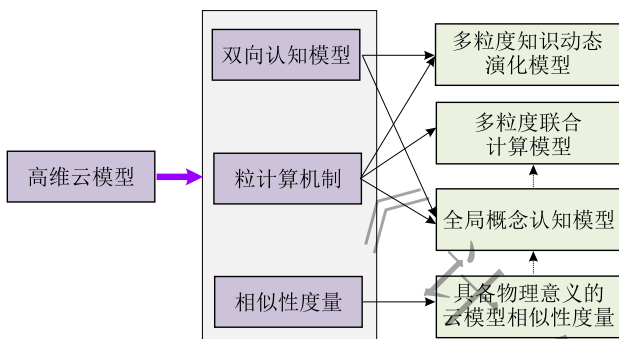


图9 基于大数据的云模型研究框架

(1) 具备物理意义的云模型相似性度量

如何从分布在高维数据空间的大数据中找出聚合的、不同粒度的概念及概念之间的复杂关联,知识之间语义的相似性是关键,而云模型可以为知识的不确定性给出一个具有概率语义解释的描述.因此,有必要进一步研究云模型的相似性度量.通过3.2节对常见的云模型相似性度量方法的分析可知,当前的云模型相似性度量研究还缺乏可靠的理论基础,单纯的根据比较对象进行相似性度量,缺乏说服力和实际意义.云模型的期望、熵和超熵的量纲不同是普遍存在的客观事实,简单地对具有不同量纲的云模型三个特征参数进行数学运算,不具备清晰的物理意义,不能作为有效的云模型相似性度量依据.比如,对于时间序列来说,其前后的数据具有相互依赖性,要利用云模型进行相似性度量首先需要选择合适的时间切割算法对时间序列进行分割,然后进行度量;对于那种只与数据分布有关,而与其他属性无关的领域,例如图片的灰度值分布,只与灰度值的大小有关,而与其在图片中位置无关,此时可直接利用云模型进行粒化后进行度量.因此,云模型相似性计算必须是针对特定的领域问题,根据待比较对象的特定属性,建立特定的相似性度量模型.

(2) 全局概念认知模型

在人类认知过程中,不但有大尺度优先,差异性

优先,还有结构优先.尤其是网络时代,各节点之间还有连接,连接构成的组织关系比节点自身还重要^[56].对于大数据同样如此,数据空间中概念及其相互关联的不确定性描述,对于有效地形式化描述大数据内在的不确定性尤为重要.

作为不确定性描述的另一种工具,数据场^[111]把现代物理学中对客观世界的认知理论引申到对主观世界的认知中来,描述认知从数据到概念、再到知识的形成过程.由于数据场可以形式化描述原始的、混乱的、不成形的概念间的复杂关联,因此,将云模型和数据场相结合,能够建立问题领域空间中的每个粒度中的云粒子之间的横向关系以及层次间的多粒度认知结构,实现如图所10所示的数据→概念→概念树→概念森林的全局认知模型,形成领域问题空间的全局概念视图(知识)或概念森林描述.因此,云模型与数据场相结合是未来不确定性研究的一个必然趋势.

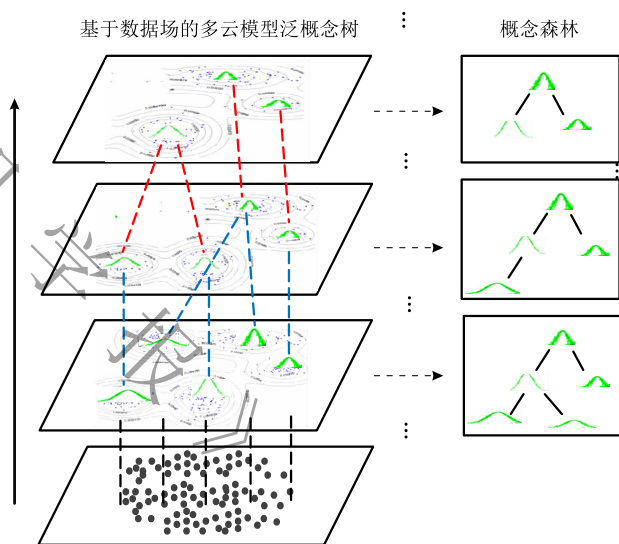


图10 全局概念认知模型示意图

(3) 多粒度联合计算模型

求解不同的问题,需要不同粒度的空间描述.对于多维度跨领域大数据复杂任务的求解,需要联合不同领域、不同层次的多粒度信息知识,从而进一步实现大数据下的智能预测和优化调度管理.再者,当前的机器学习方法注重提升模型的分类、预测性能,从数据分布角度处理学习过程中的不确定性,但学习过程中蕴含的概念、决策的语义不确定性缺少表示与处理方法.因此,如何实现对大数据复杂任务到知识的不确定性认知过程的计算建模,是需要解决的关键问题之一.

2006年由Hinton等人^[105]在《Science》上提出

了深度学习的概念,在对具有超过三层结构的深度置信神经网络(Deep Belief Network)进行了大量数据分析基础上,突破性地发现了深层网在计算能力上卓越的性能及其表示数据语义信息的结构化和层次化能力.因此,可以借鉴深度学习的思想,在全球认知模型对概念及其相互关联的不确定性描述的基础上,构建如图 11 所示的云模型的多粒度联合计算模型,其实质在面向问题/领域的粒结构上利用多层求解器求解问题.首先,将待求解任务分解为多个粒度上的子任务,每个子任务对应一个学习器,多个逐层求解器对应的不同粒度的数据构成一个粒结构.因此,将深度学习的思想与云模型粒计算机制相结合,可以产生泛化能力更强的复杂问题处理模型,这同时也是未来值得研究的一个方向.

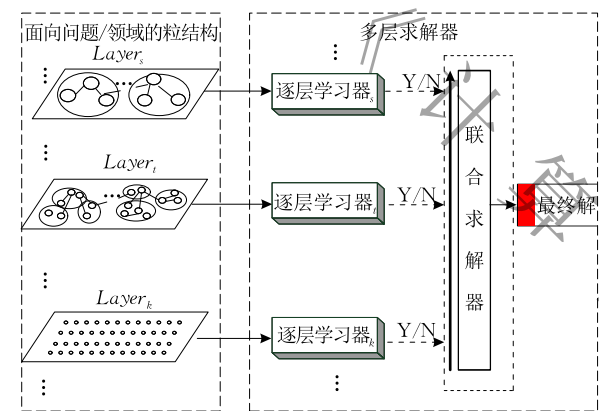


图 11 多粒度联合计算示意图

(4) 多粒度知识动态演化模型

面对大数据知识信息的汪洋大海,借鉴人脑的选择遗忘和联想记忆机理,记忆值得保留的数据与遗忘垃圾数据,将成为未来大数据处理发展的重要方向.信息多层次抽象和概念具体化是记忆构造和重塑的基础.而云模型的正向/逆向云算法恰好提供了内涵(抽象特征)与外延(具体数据)之间的双向映射计算.其中,正向云发生器实现从定性概念内涵(期望、熵、超熵)到其外延(样本集合)的映射,将抽象概念具体化,进而保留有价值数据;通过逆向云发生器实现从定性概念外延到其内涵的变换,将具体数据多层次抽象,记忆数据特征和遗忘垃圾数据.结合粒计算机制,建立基于双向认知变换的多粒度信息遗忘与联想记忆模型,从而改变传统大数据的机械存储模式,提高大数据复杂任务求解的效率和精度.再者,各种应用中的数据往往在动态地演化(数据增加、变化和老化),如何在数据的动态演化中及时有效地更新知识是动态数据挖掘的一个重要研究

问题.通过探究数据动态演化时,大数据复杂任务与其子数结构之间的联动机制,子数据结构的动态演化机制,结合基于双向认知变换的信息选择遗忘和联想记忆机理,记忆高价值信息和遗忘老化数据,建立融合选择遗忘、联想记忆机理的多粒度知识动态演化模型,实现动态数据的高效知识发现.

5.2 云模型的应用展望

通过第 4 节的分析可知,目前云模型在很多领域得到了广泛应用,充分体现了云模型在处理不确定性信息方面的优势.但是就目前的应用而言,并没有充分结合云模型理论各个方面的优势.因此,本节将围绕云模型在处理不确定性、双向认知转换和多粒度概念提取方面的优势,以 5.1 节中对云模型理论的展望为基础,对云模型的应用前景进行相关的探讨.

(1) 无人驾驶认知系统

当前,随着传感器、人工智能、移动互联网、自动控制、地图测绘等技术的发展,无人驾驶系统被广泛应用到各行各业.然而无人驾驶系统的研发是一个复杂系统工程,其核心技术是利用先进的传感器技术,实时获取周边不确定性驾驶环境的感知数据,通过环境感知理解与智能决策控制,实现自动驾驶行为的安全、稳定控制.例如,Google 无人车会将传感器数据发送给 Google 数据中心,通过感知数据分析与处理,生成详尽的前方道路地图,用于指导驾驶行为.在此意义上,以无人车为代表的无人驾驶系统研究可以看作是大数据机器学习的典型案例,而如何构建具有复杂环境认知能力的驾驶大数据机器学习系统是其面临的核心挑战.

以驾驶大数据为背景,基于 5.1 节中所述的全局概念认知模型应用于无人驾驶系统的不确定性表示与推理,以驾驶态势认知建模为研究目标,建立涵盖“感知—认知—控制—行为”的强化学习框架,结合强化学习、深度学习以及基于云模型的语义知识表示与推理,实现大数据驱动的驾驶认知过程形式化建模与全局态势模型的自主机器学习,为无人驾驶智能车的研究提供一种新方法与新思路.

(2) 带语义的深度学习

众所周知,深度学习采用的深层神经网络结构,通过各隐藏层的神经元逐层将细粒度特征抽象成粗粒度特征,利用高层抽象特征进行问题求解^[112-113].深度神经网络的结构(每层神经元的数量以及神经网络的层数等),通常利用经验和尝试法来设置,模

型的参数设置需要大量的经验和时间,而且缺乏理论指导,难度很大;再者,也缺乏度量输入变量对输出变量贡献率的尺度.因此,如何设计和优化深度神经网络的结构,仍然是尚未解决的问题.通过 5.1 节的分析可知,由于云模型可以实现对语义的不确定性表达,而深度神经网络的学习机制实质上采用了 5.1 节中所述的多粒度联合计算机制.因此,可利用两者的联系研究深度神经网络的结构设计,从而构建带语义的深度神经网络,增加深度神经网络的可解释性,同时也是机器学习、大数据挖掘未来的重要研究方向之一.

(3) 新型数据存储模式

传统的数据存储方式已经不再能够适应数据信息呈指数性增长,随着更新周期不断缩短,按时清理(删除)的信息存储机制不能满足大数据信息存储的需求.人类大脑的信息记忆存储具有带遗忘机制的长时记忆机制,当对记忆信息进行提取时,通常有两种形式:回忆和再认知.这两种形式提取信息都需要运用一定的策略,即依靠一定的线索和选择一定的中介进行联想回忆.因此,可利用 5.1 节中所述的多粒度动态演化机制,模拟人类大脑对记忆信息的储存和提取机制,为大数据存储方式的变革提供支持.

(4) 复杂网络的建模

具有自组织、吸引子世界和无标度等特点的复杂网络越来越受到人们的关注.无论是生活中的电力网、互联网、交通网,还是医学上的活性脑图、感知细胞网络等,都属于复杂网络.复杂网络拓扑结构的不确定性本身是复杂网络研究的一个基本问题.例如,在活性脑图的研究中,脑系统常常被建模为脑网络.其中,脑网络的边表示脑区与脑区之间的关系.这些边被成为脑连接.脑区的功能连接是一个数据驱动的模式^[114],它是指脑区观测信号之间的相关性.虽然功能连接的生理学意义有限,但是功能连接分析可以扩展到数百上千个节点,可以有效地刻画脑结构和脑结构的特征和变化;因此功能连接常常用作于脑疾病研究和脑认知研究的特征,有着重要且广泛的应用价值.但是,传统的以贝叶斯理论为基础的相关性与因果性分析,缺乏对脑图及脑神经系统不确定性的全局性把握,而以数据与认知的不确定性描述与推演为基础的云模型,恰恰为建模和分析活性脑图提供了一条可行的道路.因此,如何利用云模型来对复杂网络拓扑结构进行建模,是云模型的一个应用拓展.

6 总 结

云模型作为一种不确定性问题的刻画模型,从提出至今,理论逐渐成熟完善,取得了许多研究成果.作为一种处理不确定性的理论,云模型在不确定人工智能领域发挥比较重要的作用.云模型结合随机性与模糊性,可以更好的刻画对象的不确定性.通过双向认知计算模型,实现定性概念与定量数据的双向变换;通过云模型的粒计算机制,自适应抽象概念,从而形成泛概念树,反映不确定性概念在不同粒度层次上的不同表现形式.本文综述分析了正态云模型在双向认知计算模型、相似性度量、多粒度云模型和多维云等理论的研究成果及其在智能控制、预测和系统评估等应用领域的研究工作.并在此基础上,分析了各个研究方面存在的问题.以当前研究热点——大数据为背景,综合考虑云模型理论的优势和不足,指出借鉴人类智能的多粒度思维与问题求解模式改进云模型理论是今后的主要研究方向.并在此基础上提出了基于大数据的云模型研究框架,旨在通过改进云模型理论应对目前大数据中存在的挑战问题.以上这些工作将为关注云模型、粒计算和大数据处理的研究者提供一些有价值的参考和帮助,进一步推动我国在这一领域的研究工作.

参 考 文 献

- [1] Dubois D, Prade H. Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, 1991, 42(1): 87-101
- [2] Quattrone A, Vitetta A. Random and fuzzy utility models for road route choice. *Transportation Research Part E Logistics and Transportation Review*, 2011, 47(6): 1126-1139
- [3] Lush G J. Probability theory. *Nature*, 1978, 272: 107
- [4] Zadeh L A. Fuzzy sets. *Information and Control*, 1965, 8(65): 338-353
- [5] Pawlak Z. Rough sets. *International Journal of Computer and Information Sciences*, 1982, 11(5): 341-356
- [6] Yager R R. On the dempster-shafer framework and new combination rules. *Information Sciences*, 1987, 41(2): 93-137
- [7] Kahneman D, Tversky A. Subjective probability: A judgement of representativeness. *Cognitive Psychology*, 1972, 3(3): 430-454
- [8] Li De-Yi, Meng Hai-Jun. Membership and membership cloud generator. *Computer Research and Development*, 1995, 32(6): 15-20(in Chinese)

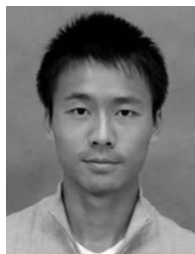
- (李德毅, 孟海军. 隶属云和隶属云发生器. 计算机研究与发展, 1995, 32(6): 15-20)
- [9] Li D Y, Liu C Y, Gan W Y. A new cognitive model: Cloud model. International Journal of Intelligent Systems, 2009, 24(3): 357-375
- [10] Wang Guo-Yin, Li De-Yi, Yao Yi-Yu, et al. Cloud Model and Granular Computing. Beijing: Science Press, 2012 (in Chinese)
(王国胤, 李德毅, 姚一豫等. 云模型与粒计算. 北京: 科学出版社, 2012)
- [11] Li De-Yi, Du Yi. Artificial Intelligence. 2nd Edition. Beijing: National Defense Industry Press, 2014(in Chinese)
(李德毅, 杜鹃. 不确定性人工智能. 第2版. 北京: 国防工业出版社, 2014)
- [12] Ding Hao, Wang Dong. Evaluation method of eutrophication degree of water body based on cloud model. Journal of Environmental Science, 2013, 33(1): 251-257(in Chinese)
(丁昊, 王栋. 基于云模型的水体富营养化程度评价方法. 环境科学学报, 2013, 33(1): 251-257)
- [13] Gao H B, Zhang X Y, Liu Y C, et al. Cloud model approach for lateral control of intelligent vehicle systems. Scientific Programming, 2016, 24(12): 1-12
- [14] Du Yi, Li De-Yi. Data source generating method of test data mining algorithm. Journal of Computer Research and Development, 2000, 37(7): 776-782(in Chinese)
(杜鹃, 李德毅. 一种测试数据挖掘算法的数据源生成方法. 计算机研究与发展, 2000, 37(7): 776-782)
- [15] Qin Kun, Li De-Yi, Xu Kai. Research on image segmentation method based on cloud model. Mapping Information and Engineering, 2006, 31(5): 3-5(in Chinese)
(秦昆, 李德毅, 许凯. 基于云模型的图像分割方法研究. 测绘信息与工程, 2006, 31(5): 3-5)
- [16] Li De-Yi. Cloud control method and dynamic balance model of three stage inverted pendulum. Chinese Engineering Science, 1999, 1(2): 41-46(in Chinese)
(李德毅. 三级倒立摆的云控制方法及动平衡模式. 中国工程科学, 1999, 1(2): 41-46)
- [17] Liu Lin-Lan, Gu Xiao-Le, Liu Song, et al. Prediction of link quality in wireless sensor networks based on cloud model. Journal of Software, 2015, 28(10): 2413-2420(in Chinese)
(刘琳岚, 谷小乐, 刘松等. 基于云模型的无线传感器网络链路质量的预测. 软件学报, 2015, 28(10): 2413-2420)
- [18] Gao H B, Jiang J, Zhang L, et al. Cloud model: Detect unsupervised communities in social tagging network// Proceedings of the International Conference on Information Science and Cloud Computing. Guangzhou, China, 2013: 317-323
- [19] Li De-Yi, Liu Chang-Yu. On the universality of the normal cloud model. Engineering Science, 2004, 6(8): 28-34 (in Chinese)
(李德毅, 刘常昱. 论正态云模型的普适性. 中国工程科学, 2004, 6(8): 28-34)
- [20] Wang Guo-Yin, Xu Chang-Lin, Zhang Qing-Hua, et al. p -order normal cloud model recursive definition and analysis of bidirectional cognitive computing. Chinese Journal of Computers, 2013, 36(11): 2316-2329(in Chinese)
(王国胤, 许昌林, 张清华等. 双向认知计算的 p 阶正态云模型递归定义及分析. 计算机学报, 2013, 36(11): 2316-2329)
- [21] Zhu Yong-Li, Gong Zheng, Wu Zhong-Li, et al. Application of normal cloud model in condition assessment of power transformer. Journal of North China Electric Power University: Natural Science Edition, 2010, 37(5): 27-31(in Chinese)
(朱永利, 宫政, 武中利等. 正态云模型在电力变压器状态评估中的应用. 华北电力大学学报(自然科学版), 2010, 37(5): 27-31)
- [22] Sun Jing, Yang Jing-Yu. Research on watermarking of one-dimensional normal cloud model. Journal of Wuhan University of Technology, 2010, 32(16): 48-51(in Chinese)
(孙菁, 杨静宇. 一维正态云模型水印研究. 武汉理工大学学报, 2010, 32(16): 48-51)
- [23] Gong Yan-Bing, Zhang Ji-Guo. A comprehensive evaluation of population development modernization based on normal cloud model and entropy weight. China Population and Resources and Environment, 2012, 22(1): 138-143(in Chinese)
(龚艳冰, 张继国. 基于正态云模型和熵权的人口发展现代化程度综合评价. 中国人口资源与环境, 2012, 22(1): 138-143)
- [24] Wang G Y, Xu C L, Li D Y. Generic normal cloud model. Information Sciences, 2014, 280(280): 1-15
- [25] Li De-Yi, Liu Chang-Yu, Gan Wen-Yan. The heavy tailed property of normal cloud model. Chinese Engineering Science, 2011, 13(4): 20-23(in Chinese)
(李德毅, 刘常昱, 淦文燕. 正态云模型的重尾性质证明. 中国工程科学, 2011, 13(4): 20-23)
- [26] Zhang Yan-Ping, Zhang Ling, Wu Tao. The description method of the world with different grain size: Quotient space method. Chinese Journal of Computers, 2004, 27(3): 328-333(in Chinese)
(张燕平, 张铃, 吴涛. 不同粒度世界的描述法——商空间法. 计算机学报, 2004, 27(3): 328-333)
- [27] Lv Hui-Jun, Wang Ye, Li De-Yi, et al. Application of reverse cloud in qualitative evaluation. Chinese Journal of Computers, 2003, 26(8): 1009-1014(in Chinese)
(吕辉军, 王晔, 李德毅等. 逆向云在定性评价中的应用. 计算机学报, 2003, 26(8): 1009-1014)
- [28] Luo Zi-Qiang, Zhang Guang-Wei, Li De-Yi. Statistical analysis of the probability of one dimensional normal cloud. Information and Control, 2007, 36(4): 471-475(in Chinese)
(罗自强, 张光卫, 李德毅. 一维正态云的概率统计分析. 信息与控制, 2007, 36(4): 471-475)
- [29] Wang Hui, Qin Shu, Liu Shao-Ying, et al. Improved reverse cloud algorithm based on curve fitting. CAAI Transactions on Intelligent Systems, 2014, 9(5): 590-594(in Chinese)
(王辉, 秦术, 刘少英等. 曲线拟合的逆向云改进算法. 智能系统学报, 2014, 9(5): 590-594)

- [30] Liu Chang-Yu, Feng Mang, Dai Xiao-Jun, et al. A new cloud based reverse cloud algorithm based on cloud X information. *Journal of System Simulation*, 2004, 16(11): 2417-2420(in Chinese)
(刘常昱, 冯芒, 戴晓军等. 基于云 X 信息的逆向云新算法. *系统仿真学报*, 2004, 16(11): 2417-2420)
- [31] Wang Li-Xin. Basic mathematical properties of normal cloud and cloud filter. *Personal Communication*, 2011(in Chinese)
(王立新. 正态云的基本数学性质及云滤波器. 个人通信. 2011)
- [32] Chen Hao, Li Bing, Liu Chang-Yu. A kind of non determined reverse cloud algorithm. *Journal of Chinese Computer Systems*, 2015, 36(3): 544-549(in Chinese)
(陈昊, 李兵, 刘常昱. 一种无确定度的逆向云算法. *小型微型计算机系统*, 2015, 36(3): 544-549)
- [33] Xu C L, Wang G Y, Zhang Q H. A new multi-step backward cloud transformation algorithm based on normal cloud model. *Fundamenta Informaticae*, 2014, 133(1): 55-85
- [34] Wang G Y, Xu C L. Cloud model: A bidirectional cognition model between concept's extension and intension//*Proceedings of the Advanced Machine Learning Technologies and Applications*. Cairo, Egypt, 2012: 391-400
- [35] Zhang Shi-Bin, Xu Chun-Xiang. Research on trust evaluation method based on cloud model. *Chinese Journal of Computers*, 2013, 36(2): 422-431(in Chinese)
(张仕斌, 许春香. 基于云模型的信任评估方法研究. *计算机学报*, 2013, 36(2): 422-431)
- [36] Rubner Y, Tomasi C, Guibas L J. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 2000, 40(2): 99-121
- [37] Liao Lie-Fa, Li Chen, Meng Xiang-Mao. Cloud model collaborative filtering algorithm based on similarity of Euclidean space. *Computer Engineering and Science*, 2015, 37(10): 1977-1982(in Chinese)
(廖列法, 黎晨, 孟祥茂. 基于欧氏空间相似度的云模型协同过滤算法. *计算机工程与科学*, 2015, 37(10): 1977-1982)
- [38] Zhang Guang-Wei, Li De-Yi, Li Peng, et al. Collaborative filtering recommendation algorithm based on cloud model. *Journal of Software*, 2007, 18(10): 2403-2411(in Chinese)
(张光卫, 李德毅, 李鹏等. 基于云模型的协同过滤推荐算法. *软件学报*, 2007, 18(10): 2403-2411)
- [39] Li Hai-Lin, Guo Chong-Hui, Qiu Wang-Ren. Similarity calculation method of the normal cloud model. *Acta Electronica Sinica*, 2011, 39(11): 2561-2567(in Chinese)
(李海林, 郭崇慧, 邱望仁. 正态云模型相似度计算方法. *电子学报*, 2011, 39(11): 2561-2567)
- [40] Yan Yan, Tang Zhen-Min. A cloud model correlation measurement method based on the entropy expectation curve. *Journal of Huazhong University of Science and Technology: Natural Science Edition*, 2012, 40(10): 95-100 (in Chinese)
(阎岩, 唐振民. 基于含熵期望曲线的云模型相关性度量方法. *华中科技大学学报: 自然科学版*, 2012, 40(10): 95-100)
- [41] Zha Xiang, Ni Shi-Hong, Xie Chuan, et al. Indirect calculation method of the concept of cloud similarity. *System Engineering and Electronic Technology*, 2015, 37(7): 1676-1682(in Chinese)
(查翔, 倪世宏, 谢川等. 云相似度的概念跃升间接计算方法. *系统工程与电子技术*, 2015, 37(7): 1676-1682)
- [42] Xu Chang-Lin, Wang Guo-Yin. Measurement and analysis of the drift of the concept of normal cloud. *Computer Science*, 2014, 41(7): 9-14(in Chinese)
(许昌林, 王国胤. 正态云概念的漂移性度量及分析. *计算机科学*, 2014, 41(7): 9-14)
- [43] Yang J, Wang G Y, Li X K. Multi-granularity similarity measure of cloud concept//*Proceedings of the International Joint Conference on Rough Sets*. Santiago, Chile, 2016: 318-330
- [44] Jiang Rong, Li De-Yi, Fan Jian-Hua. An automatic generation method for the generalized concept tree of numerical data. *Chinese Journal of Computers*, 2000, 23(5): 470-476 (in Chinese)
(蒋嵘, 李德毅, 范建华. 数值型数据的泛概念树的自动生成方法. *计算机学报*, 2000, 23(5): 470-476)
- [45] Cheung D, Fu A, Han J. Knowledge discovery in databases: A rule-based attribute-oriented approach//*Proceedings of the International Symposium on Methodologies for Intelligent Systems*. Charlotte, USA, 1994: 164-173
- [46] Lu Y. Concept Hierarchy in Data Mining: Specification, Generation and Implementation [Ph. D. dissertation]. Simon Fraser University, Vancouver, Canada, 1997
- [47] Han J, Fu Y. Discovery of multiple-level association rules from large databases//*Proceeding of the 21st VLDB Conference*. Zurich, Switzerland, 1995: 420-431
- [48] Li De-Yi, Du Yi. *The Uncertainty of Artificial Intelligence*. Beijing: National Defense Industry Press, 2005(in Chinese)
(李德毅, 杜鹃. 不确定性人工智能. 北京: 国防工业出版社, 2005)
- [49] Qin Kun, Xu Min. Remote sensing image segmentation method based on cloud model and FCM clustering. *Geo-Information Science*, 2008, 10(3): 302-307(in Chinese)
(秦昆, 徐敏. 基于云模型和 FCM 聚类的遥感图像分割方法. *地球信息科学学报*, 2008, 10(3): 302-307)
- [50] Meng Hui, Wang Shu-Liang, Li De-Yi. Concept extraction and concept hierarchy construction based on cloud transformation. *Journal of Jilin University Engineering and Technology Edition*, 2010, 40(3): 782-787(in Chinese)
(孟晖, 王树良, 李德毅. 基于云变换的概念提取及概念层次构建方法. *吉林大学学报: 工学版*, 2010, 40(3): 782-787)
- [51] Xu Kai, Qin Kun, Liu Xiu-Guo, et al. A method for the construction of pan concept level and its application in remote sensing image classification. *Geomatics and Information Science of Wuhan University*, 2013, 38(9): 1078-1082(in Chinese)

- (许凯, 秦昆, 刘修国等. 泛概念层次构建方法及其在遥感图像分类中的应用. 武汉大学学报: 信息科学版, 2013, 38(9): 1078-1082)
- [52] Xu Chang-Lin. Research on the Two-Way Cognitive Computing Method Based on Cloud Model [Ph. D. dissertation]. School of Information Science and Technology, Southwest Jiaotong University, Chengdu, 2014(in Chinese)
- (许昌林. 基于云模型的双向认知计算方法研究[博士学位论文]. 西南交通大学信息科学与技术学院, 成都, 2014)
- [53] Liu Y C, Li D Y, He W, et al. Granular computing based on Gaussian cloud transformation. Fundamenta Informaticae, 2013, 127(14): 385-398
- [54] Xu Ji, Wang Guo-Yin, Yu Hong. Big data processing based on granular computing. Chinese Journal of Computers, 2015, 38(8): 1497-1517(in Chinese)
- (徐计, 王国胤, 于洪. 基于粒计算的大数据处理. 计算机学报, 2015, 38(8): 1497-1517)
- [55] Yao Hong, Wang Guo-Yin, Zhang Qin-Hua. Color image segmentation method based on rough set and cloud model. Journal of Chinese Computer Systems, 2013, 34(11): 2615-2620(in Chinese)
- (姚红, 王国胤, 张清华. 基于粗糙集和云模型的彩色图像分割方法. 小型微型计算机系统, 2013, 34(11): 2615-2620)
- [56] Liu Yu-Chao. Research on Particle Computation Method Based on Cloud Model [Ph. D. dissertation]. School of Information Science and Technology, Tsinghua University, Beijing, 2012(in Chinese)
- (刘玉超. 基于云模型的粒计算方法研究[博士学位论文]. 清华大学信息科学技术学院, 北京, 2012)
- [57] Yang Zhao-Hui, Li De-Yi. Planar model and its application in prediction. Chinese Journal of Computers, 1998, 21(11): 961-969(in Chinese)
- (杨朝晖, 李德毅. 二维云模型及其在预测中的应用. 计算机学报, 1998, 21(11): 961-969)
- [58] Zhang Guo-Ying, Sha Yun, Liu Xu-Hong, et al. High-dimensional cloud model and its application in multi attribute evaluation. Transaction of Beijing Institute of Technology, 2004, 24(12): 1065-1069(in Chinese)
- (张国英, 沙云, 刘旭红等. 高维云模型及其在多属性评价中的应用. 北京理工大学学报, 2004, 24(12): 1065-1069)
- [59] Wang Cheng-Quan. Research on rule extraction algorithm based on two dimensional cloud model. Computer Engineering and Applications, 2010, 46(26): 46-48(in Chinese)
- (王成全. 基于二维云模型的规则提取算法研究. 计算机工程与应用, 2010, 46(26): 46-48)
- [60] Wang Dong, Liu Deng-Feng, Ding-Hao, et al. A multidimension cloud model-based approach for water quality assessment. Environmental Research, 2016, 149: 113-121
- [61] Li Zhong, Yang Yi-Dong. Controllable pitch propeller control based on qualitative reasoning of variant dimension cloud model. Journal of Nanjing University of Aeronautics and Astronautics, 2003, 35(2): 162-167(in Chinese)
- (李众, 杨一东. 基于混合维云模型定性推理的调距桨螺距控制. 南京航空航天大学学报, 2003, 35(2): 162-167)
- [62] Li Zhong, Liu Yan. A single rule reasoning map for the two dimensional normal cloud model. CAAI Transactions on Intelligent Systems, 2010, 5(5): 464-470(in Chinese)
- (李众, 刘艳. 二维正态云模型的单规则推理映射. 智能系统学报, 2010, 5(5): 464-470)
- [63] Li Zhong, Gao Jian, Zhang Ri-Xun. Study on the approximation of one dimensional cloud model system. Acta Automation Sinica, 2011, 37(10): 1190-1199(in Chinese)
- (李众, 高健, 张日勋. 一维云模型系统的逼近性研究. 自动化学报, 2011, 37(10): 1190-1196)
- [64] Li De-Yi, Di Kai-Chang, Li De-Ren, et al. Mining association rules by using language cloud model. Journal of Software, 2000, 11(2): 143-158(in Chinese)
- (李德毅, 邸凯昌, 李德仁等. 用语言云模型发掘关联规则. 软件学报, 2000, 11(2): 143-158)
- [65] Du Yi, Song Zi-Lin, Li De-Yi. Association rules mining method based on cloud model. Journal of PLA University of Science and Technology: Natural Science Edition, 2000, 1(1): 29-34(in Chinese)
- (杜懿, 宋自林, 李德毅. 基于云模型的关联规则挖掘方法. 解放军理工大学学报: 自然科学版, 2000, 1(1): 29-34)
- [66] Tian Yong-Qing, Yang Bin, Li Zhi, et al. Extraction of association rules based on cloud model in a relational database. Journal of Shanghai Jiaotong University, 2003, 37(4): 512-515(in Chinese)
- (田永青, 杨斌, 李志等. 一种关系数据库中基于云模型关联规则的提取. 上海交通大学学报, 2003, 37(4): 512-515)
- [67] Wang De-Lu, Song Xue-Feng. Strategic risk early warning of enterprise transformation based on cloud model association rules. Chinese Journal of Management Science, 2009, 17(2): 152-159(in Chinese)
- (王德鲁, 宋学锋. 基于云模型关联规则的企业转型战略风险预警. 中国管理科学, 2009, 17(2): 152-159)
- [68] Lu Jian-Jiang. Research on algorithms of mining association rules with weighted items. Journal of Computer Research and Development, 2002, 39(10): 1281-1286(in Chinese)
- (陆建江. 加权关联规则挖掘算法的研究. 计算机研究与发展, 2002, 39(10): 1281-1286)
- [69] Lu Jian-Jiang, Qian Zu-Ping, Song Zi-Lin. Application of the normal cloud association rules in prediction. Journal of Computer Research and Development, 2000, 37(11): 1317-1320(in Chinese)
- (陆建江, 钱祖平, 宋自林. 正态云关联规则在预测中的应用. 计算机研究与发展, 2000, 37(11): 1317-1320)
- [70] Jiang Rong, Li De-Yi, Chen Hui. Time series forecasting based on cloud model. Journal of PLA University of Science and Technology: Natural Science Edition, 2000, 1(5): 13-18(in Chinese)
- (蒋嵘, 李德毅, 陈晖. 基于云模型的时间序列预测. 解放军理工大学学报: 自然科学版, 2000, 1(5): 13-18)

- [71] Miao Liang, Yue Xun, Zhang Xin-Yi, et al. Qualitative forecasting model based on cloud model and its application in the prediction of wheat aphid. *Scientia Agricultura Sinica*, 2002, 35(6): 654-659(in Chinese)
(苗良, 岳训, 张新谊等. 基于云模型的定性预报模型及其在麦长管蚜预报中的应用. *中国农业科学*, 2002, 35(6): 654-659)
- [72] Deng W H, Wang G Y, Zhang X R. A novel hybrid water quality time series prediction method based on cloud model and fuzzy forecasting. *Chemometrics and Intelligent Laboratory Systems*, 2015, 149: 39-49
- [73] Jia Shun-Ping, Mao Bao-Hua. Research on CFCM: Car following model using cloud model theory. *Journal of Transportation Systems Engineering and Information Technology*, 2007, 7(6): 67-73(in Chinese)
(贾顺平, 毛保华. 基于云推理的车辆跟驰模型 CFCM 研究. *交通运输系统工程与信息*, 2007, 7(6): 67-73)
- [74] Li Fei, Zhang Shi-Bin. Research on state evaluation and prediction model of network system based on cloud model. *Journal of Sichuan University: Engineering Science Edition*, 2010, 42(6): 99-104(in Chinese)
(李飞, 张仕斌. 基于云模型的网络系统状态评估与预测模型研究. *四川大学学报: 工程科学版*, 2010, 42(6): 99-104)
- [75] Liu Qi, Du Yuan-Dong. The normal cloud model of fuzzy inference prediction method for petroleum drilling accident// *Proceedings of the International Conference on Information and Control Engineering*. Wuhan, China, 2011: 1816-1821
- [76] Xie Dao-Wen, Shi Shi-Liang. Based on the cloud theory and weighted Markov model of mine water inrush pre. *Journal of Central South University of Technology: Natural Science Edition*, 2012, 43(6): 2308-2315(in Chinese)
(谢道文, 施式亮. 基于云理论与加权马尔可夫模型的矿井涌水量预. *中南大学学报: 自然科学版*, 2012, 43(6): 2308-2315)
- [77] Yan Chang-Shun, Li Yi-Jun. Research on dynamic customer segmentation and classification model based on cloud model. *Journal of Harbin Institute of Technology*, 2007, 39(2): 299-302(in Chinese)
(阎长顺, 李一军. 基于云模型的动态客户细分分类模型研究. *哈尔滨工业大学学报*, 2007, 39(2): 299-302)
- [78] Wang Wei, Ma Dong-Hui, Su Jing-Yu, et al. Site classification method based on two dimensional multi rule cloud model for qualitative reasoning. *Journal of Beijing University of Technology*, 2009, 35(10): 1362-1372(in Chinese)
(王威, 马东辉, 苏经宇等. 基于二维多规则云模型定性推理的场地分类方法. *北京工业大学学报*, 2009, 35(10): 1362-1372)
- [79] Qin Li, Li Bing. Novel method of uncertain data modeling and classification based on cloud model. *Computer Science*, 2014, 41(8): 233-240(in Chinese)
(秦丽, 李兵. 一种基于云模型的不确定性数据的建模与分类方法. *计算机科学*, 2014, 41(8): 233-240)
- [80] Liu Yu. Classification Algorithm Based on Cloud Model [Ph. D. dissertation]. School of Computing, Beihang University, Beijing, 2010(in Chinese)
(刘禹. 基于云模型的分类算法[博士学位论文]. 北京航空航天大学计算机学院, 北京, 2010)
- [81] Du Xiang-Yu, Yin Quan-Jun, Huang Ke-Di, et al. Qualitative and quantitative transformation method based on cloud model and its application. *Journal of System Engineering and Electronic*, 2008, 30(4): 772-776(in Chinese)
(杜湘瑜, 尹全军, 黄柯棣等. 基于云模型的定性定量转换方法及其应用. *系统工程与电子技术*, 2008, 30(4): 772-776)
- [82] Song Yuan-Jun, Li De-Yi. Evaluation method of reliability of electronic products by cloud model. *Acta Electronica Sinica*, 2000, 28(12): 74-76(in Chinese)
(宋远毅, 李德毅. 电子产品可靠性的云模型评价方法. *电子学报*, 2000, 28(12): 74-76)
- [83] Song Yuan-Jun, Yang Xiao-Zong, Li De-Yi, et al. Evaluation of computer reliability evaluation of computer reliability considering environmental factors. *Journal of Computer Research and Development*, 2001, 38(5): 631-636(in Chinese)
(宋远毅, 杨孝宗, 李德毅等. 考虑环境因素的计算机可靠性云模型评价. *计算机研究与发展*, 2001, 38(5): 631-636)
- [84] Hu Shi-Yuan, Li De-Ren, Liu Yao-Lin, et al. Land evaluation factors weight mining based on cloud model and correlation analysis method. *Journal of Wuhan University: Information Science Edition*, 2006, 31(5): 423-427(in Chinese)
(胡石元, 李德仁, 刘耀林等. 基于云模型和关联度分析法的土地评价因素权重挖掘. *武汉大学学报: 信息科学版*, 2006, 31(5): 423-427)
- [85] Huang Hai-Sheng, Wang Ru-Chuan. A study on the subjective trust evaluation model based on the membership cloud theory. *Journal on Communications*, 2008, 29(4): 13-19(in Chinese)
(黄海生, 王汝传. 基于隶属云理论的主观信任评估模型研究. *通信学报*, 2008, 29(4): 13-19)
- [86] Wang Shou-Xin, Zhang Li, Li He-Song. A method of subjective trust evaluation based on cloud model. *Journal of Software*, 2010, 21(6): 1341-1352(in Chinese)
(王守信, 张莉, 李鹤松. 一种基于云模型的主观信任评价方法. *软件学报*, 2010, 21(6): 1341-1352)
- [87] Jin B, Wang Y, Liu Z, et al. A trust model based on cloud model and Bayesian networks. *Procedia Environmental Sciences*, 2011, 11(1): 452-459
- [88] Xu Xiao-Bin, Zhang Guang-Wei, Wang Shang-Guang, et al. A method of uncertain trust representation based on the lightweight cloud model WSN. *Journal on Communications*, 2014, 35(2): 63-69(in Chinese)
(徐晓斌, 张光卫, 王尚广等. 基于轻量云模型的 WSN 不确定性信任表示方法. *通信学报*, 2014, 35(2): 63-69)
- [89] Gao Hong-Bo, Zhang Xin-Yu, Zhang Tian-Lei, et al. Research on variable particle size measurement of intelligent vehicle based on cloud model. *Acta Electronica Sinica*, 2016, 44(2): 365-373(in Chinese)
(高洪波, 张新钰, 张天雷等. 基于云模型的智能驾驶车辆变粒度测评研究. *电子学报*, 2016, 44(2): 365-373)
- [90] Li Xing-Sheng, Du Yu-Xuan, Li De-Yi. A method for modulation recognition of PSK/QAM signals based on cloud model. *Measurement and Control Technology*, 2003, 22(3): 15-19(in Chinese)

- (李兴生, 杜蔚轩, 李德毅. 一种基于云模型的 PSK/QAM 信号调制识别方法. 测控技术, 2003, 22(3): 15-19)
- [91] Shi Jun-Tao, Zhou Ming, Zhang Zhen-Kun, et al. Effectiveness evaluation of communication jamming system based on cloud model. Journal of Air Force Early Warning College, 2012, 26(4): 266-268(in Chinese)
(史军涛, 周铭, 张振坤等. 基于云模型的通信干扰系统效能评估. 空军预警学院学报, 2012, 26(4): 266-268)
- [92] Chen Jin, Ni Wei-Min, Qian Zu-Pin, et al. Study on the cloud model algorithm for radio identification. Journal of Signal Processing, 2012, 28(11): 1529-1534(in Chinese)
(陈金, 倪为民, 钱祖平等. 电台识别的云模型算法研究. 信号处理, 2012, 28(11): 1529-1534)
- [93] Xu Kai, Qin Kun, Liu Xiu-Guo, et al. A hybrid model of Gauss cloud transform algorithm and its application in image segmentation. Journal of Wuhan University: Information Science Edition, 2013, 38(10): 1163-1166(in Chinese)
(许凯, 秦昆, 刘修国等. 高斯混合模型云变换算法及其在图像分割中的应用. 武汉大学学报: 信息科学版, 2013, 38(10): 1163-1166)
- [94] Qin K, Li D, Wu T, et al. Comparative study of image thresholding using type-2 fuzzy sets and cloud model. International Journal of Computational Intelligence Systems, 2012, 3(1): 61-73
- [95] Li Xing-Sheng, Li De-Yi. A method of discretization of continuous attributes in decision table based on cloud model. Pattern Recognition and Artificial Intelligence, 2003, 16(1): 33-38(in Chinese)
(李兴生, 李德毅. 一种基于云模型的决策表连续属性离散化方法. 模式识别与人工智能, 2003, 16(1): 33-38)
- [96] Wu L, Zuo C, Zhang H, et al. Bimodal fruit fly optimization algorithm based on cloud model learning. Soft Computing, 2015, 21(7): 1877-1873
- [97] Li Z, Wang Y, Olivier K K S, et al. The cloud-based framework for ant colony optimization//Proceedings of the 1st ACM/SIGEVO Summit on Genetic and Evolutionary Computation. Shanghai, China, 2009: 279-286
- [98] Liu Yu, Li De-Yi, Zhang Guang-Wei, et al. Atomization characteristics of cloud model and its application in evolutionary algorithms. Acta Electronica Sinica, 2009, 37(8): 1651-1658(in Chinese)
(刘禹, 李德毅, 张光卫等. 云模型雾化特性及在进化算法中的应用. 电子学报, 2009, 37(08): 1651-1658)
- [99] Zhang Guang-Wei, He Rui, Liu Yu, et al. An evolutionary algorithm based on cloud model. Chinese Journal of Computers, 2008, 31(7): 1082-1091(in Chinese)
(张光卫, 何锐, 刘禹等. 基于云模型的进化算法. 计算机学报, 2008, 31(7): 1082-1091)
- [100] Zhao Wei-Wei, Li De-Yi. An intrusion detection method based on cloud model. Computer Engineering and Application, 2003, 39(26): 158-160(in Chinese)
(赵卫伟, 李德毅. 基于云模型的入侵检测方法. 计算机工程与应用, 2003, 39(26): 158-160)
- [101] Zhou L H, Liu Y H, Chen G L. A feature selection algorithm to intrusion detection based on cloud model and multi-objective particle swarm optimization//Proceedings of the International Symposium on Computational Intelligence and Design. Hangzhou, China, 2011: 182-185
- [102] Wang Shang-Guang, Sun Qi-Bo, Zhang Guang-Wei, et al. The uncertainty of the cloud model based on the QoS aware Skyline service selection. Journal of Software, 2012, 23(6): 1397-1412(in Chinese)
(王尚广, 孙其博, 张光卫等. 基于云模型的不确定性 QoS 感知的 Skyline 服务选择. 软件学报, 2012, 23(6): 1397-1412)
- [103] Chen C L P, Zhang C Y. Data-intensive applications, challenges, techniques and technologies: A survey on big data. Information Sciences, 2014, 275(11): 314-347
- [104] Du Z. Granularities and Inconsistencies in big data analysis. International Journal of Software Engineering and Knowledge Engineering, 2013, 23(6): 887-893
- [105] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. Science, 2006, 313: 504-507
- [106] Chen J C, Chen Y G, Du X Y, et al. Big data challenge: A data management perspective. Frontiers of Computer Science, 2013, 7(2): 157-164
- [107] Clifford L. Big data: How do your data grow? Nature, 2008: 28-29
- [108] Skowron A, Szczuka M. Toward interactive computations: A rough-granular approach. Studies in Computational Intelligence, 2010, 263: 23-42
- [109] Liang Ji-Ye, Qian Yu-Hua, Li De-Yu, Hu Qing-Hua. Theory and method of granular computing theory for big data mining. China Science Series E: Information Sciences, 2015, 45(11): 1355-1369(in Chinese)
(梁吉业, 钱宇华, 李德玉, 胡清华. 大数据挖掘的粒计算理论与方法. 中国科学 E 辑: 信息科学, 2015, 45(11): 1355-1369)
- [110] Wang H, Xu Z S, Pedrycz W. An overview on the roles of fuzzy set techniques in big data processing: Trends, challenges and opportunities. Knowledge-Based Systems, 2016, 118: 15-30
- [111] Gan Wen-Yan, Li De-Yi, Wang Jian-Min. Hierarchical clustering method based on data fields. Acta Electronica Sinica, 2006, 34(2): 258-262(in Chinese)
(淦文燕, 李德毅, 王建民. 一种基于数据场的层次聚类方法. 电子学报, 2006, 34(2): 258-262)
- [112] Erhan D, Szegedy C, Toshev A, et al. Scalable object detection using deep neural networks//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 2147-2154
- [113] Lecun Y, Bengio Y, Hinton G E. Deep learning. Nature, 2015, 521(7553): 436-444
- [114] Buckner R L, Krienen F M, et al. Opportunities and limitations of intrinsic functional connectivity MRI. Nature Neuroscience, 2013, 16(7): 832-837



YANG Jie, born in 1987, Ph.D. candidate, associate professor. His current research interests include cloud model, granular computing, rough set, machine learning.

WANG Guo-Yin, born in 1970, Ph.D., professor. His research interests include rough set, granular computing, cognitive computing, intelligent information processing, data mining.

LIU Qun, born in 1969, Ph.D., professor. Her research

interests include nonlinear theory, granular computing.

GUO Yi-Ke, born in 1962, Ph.D., professor. His research interests include machine learning, data mining, parallel computing.

LIU Yue, born in 1975, Ph.D., associate professor. Her research interests include machine learning, data mining.

GAN Wen-Yan, born in 1971, Ph.D., associate professor. Her research interests include machine learning, data mining.

LIU Yu-Chao, born in 1980, Ph.D., senior engineer. His research interests include cloud model, granular computing, intelligent drive.

Background

In the processes of human perception, people are able to receive, store and deal with various uncertainty information. It is essential to study the fuzziness and randomness of natural languages in order to make the computers have understanding and judgement abilities similar to those of humans. For the purpose, combining with probability theory and fuzzy sets theory, professor Li De-Yi proposed the cloud models a new cognition model of uncertainty in 1995, which provide a way to simulate the human cognition process for concepts by computer. In recent years, cloud model has received much attention and produced applications in uncertain information

processing.

In this paper, we survey the main theoretical development and applications of the cloud model firstly. Based on the related key problems, we discuss the further improvement respectively. Finally, we point out the existing challenge problems, and explore the research direction in future.

The work is partially supported by the National Key Research and Development Program (No. 2016YFB1000905), the National Natural Science Foundation of China (No. 61572091), and the Chongqing Postgraduate Scientific Research and Innovation Projects (No. CYB16106).