

基于云理论和信息融合理论的系统性能评估方法研究^{*}

A Method of System Performance Evaluation Based on the Cloud Theory and the Information Fusion Theory

秦 旻, 卢 强, 黄士坦

QIN Yi, LU Qiang, HUANG Shi-tan

(西安微电子技术研究所, 陕西 西安 710054)

(Xi'an Institute of Microelectronics Technology, Xi'an 710054, China)

摘 要:现代系统结构复杂,其性能评估需要从多个层次综合评判形成合理可信的最终评估结果,评估过程大量存在多源不确定信息。不确定性的存在增加了评估过程中定性与定量综合集成的难度。云理论和信息融合理论在不确定信息和多源信息的处理上各有优点,在功能上具有一定的互补性。两者相结合能够增强对信息的处理能力,挖掘出数据中更有价值的隐含信息。本文创新地将二者结合起来,提出了新的系统性能评估模型和方法,较好地融合了影响系统性能的多源信息,充分利用和集成了信息的模糊性和随机性,完成定性和定量之间的映射。这种方法能有效地处理各多源信息的不确定性,具有科学性,同时最终得到的评价结果是语言评判值,具有直观性。该方法是对信息融合理论和云理论联合使用的探索。

Abstract: The structure of modern systems is very complex, and the performance evaluation is the comprehensive evaluation of multiple levels and multiple objectives. During the process of evaluation, uncertainty exists widely. In order to use and process the uncertainty and improve the creditability of system evaluation, this paper combines the cloud theory and the information fusion theory and put forward a novel evaluation method of system performance. The cloud theory and the information fusion theory have their own advantages in processing uncertain information. The union of them can excavate more valuable implicit information. The method promoted in the paper is more scientific and more visual. Firstly, the paper introduces the basic theory. Secondly, it studies the mapping method between nature variable and quantity variable. Thirdly, it constructs a system performance evaluation model. Finally, the validity of the method is verified by an example. The method in this paper is new exploration of the combined use of the cloud theory and the information fusion theory.

关键词:云理论;信息融合;云模型

Key words: cloud theory; information fusion; cloud model

doi:10.3969/j.issn.1007-130X.2012.02.033

中图分类号:TP302

文献标识码:A

1 引言

现代系统产品中包含有大量的多种多样的元

器件,且这些器件以各种各样的形式互连,工作过程中这些元件呈现不同的状态,当然会发生不同的故障。系统和器件之间的状态转移是相互渗透、相互转化的,在实践中会出现亦此亦彼的状态,存在

^{*} 收稿日期:2010-12-31;修订日期:2011-03-28
通讯地址:710054 陕西省西安市西安微电子技术研究所
Address: Xi'an Institute of Microelectronics Technology, Xi'an, Shaanxi 710054, P. R. China

着很大的随机性和模糊性。而可靠性本身是一个定性的概念,复杂系统的评估往往需要从多个层次综合评判形成合理可信的最终评估结果。在不同的评估方式中,可能采用定量的精确的数值描述变量,也可能采用“好、一般”这样的定性的模糊语言描述变量。复杂系统的评估尤其存在着这种不确定性,主要包括被评估对象本身行为的不确定性,测试数据的不确定性,专家意见自然语言描述的模糊性和随机性等。

云理论和信息融合理论在不确定信息和多源信息的处理上各有优点,在功能上具有一定的互补性。两者结合能够增强对信息的处理能力,挖掘出数据中更有价值的隐含信息。本文提出了一种基于云理论和信息融合理论的系统性能评估方法,能够使系统性能评估更加有效客观,是信息融合理论和云理论联合使用的探索。

2 基本理论的再理解

云模型是李德毅院士提出的云理论的一个组成部分,是在对概率理论和模糊集合理论交叉渗透的基础上,采用特定的算法,构造定性概念与定量表示之间的转换模型。

云模型是用语言值表示的某个定性概念与其定量表示之间的不确定转换模型,或者简单地说云模型是定性和定量间转换的不确定性模型。

设 U 表示论域,用精确数值量的集合 $U = \{x\}$ 表示,关于论域 U 上对应的定性概念 A ,对于任意论域中的元素 x ,都存在一个有稳定倾向的随机数 $y \in [0, 1]$,称为 x 对应 A 的隶属度,隶属度在论域上的分布称为隶属云,简称为云^[2]。云由许许多多云滴组成,云的整体形状反映了定性概念的重要特性,云滴则是对定性概念的定量描述。云滴的产生过程,表示定性概念和定量值之间的不确定性映射。根据论域 U 的维数,云又可以分为一维云、二维云、多维云等^[2]。

图 1 描述了自然语言“约 20”的隶属云。

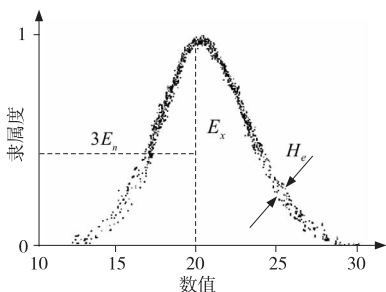


图 1 自然语言“约 20”的隶属云

从图 1 中可以看出:所有在论域上的与元素 x 到区间 $[0, 1]$ 的映射都是一对多的转换。 x 对于定性概念 A 的隶属度是一个概率分布而并不是一个固定值,所以产生了云。每个云滴都是定性向定量的一次转换。云的“厚度”是不均匀的,越靠近中心值的隶属度随机性越小。

云的数字特征反映了定性知识的定量特性,用期望值 E_x (Expected Value)、熵 E_n (Entropy)、超熵 H_e (Hyper Entropy)三个数值来表征^[3]。

期望值 E_x :是定性概念在论域中的中心值,是最能代表这个定性概念的值^[3]。

熵 E_n :是定性概念模糊度的度量,反映了在论域中可被定性概念 A 所接受的数值范围,体现了亦此亦彼性的裕度。熵越大,概念所接受的数值范围也越大,概念越模糊^[3]。

超熵 H_e :也称作熵的熵,反映了云滴的离散程度。超熵越大,云滴离散度越大,隶属度的随机性越大,云的“厚度”也越大^[3]。

云模型的三个数字特征值勾画出由许许多多的云滴构成的云,把模糊性(定性概念的亦此亦彼性)和随机性(隶属度的随机性)完全集成到一起,构成定性和定量相互间的映射。

由此总结出云的基本特征:在云模型中每一元素的隶属度都是遵循某一概率分布的随机数,不是一个唯一的值。这反映了定性变量与定量变量转换的不确定性。尽管每个元素的映射结果都在变化,但这种变化是有一定范围的,不会影响到云的整体形状。反应定性变量到定量变量映射真正含义的就是云的整体形状。

正态云模型是表征语言值最重要的工具。若论域中的元素对于定性概念 A 的隶属度满足:

$$y(x) = \exp\left[-\frac{(x - E_x)^2}{2E_m^2}\right]$$

则称 x 在论域 U 上的分布为正态云^[4]。正态云的生成算法步骤如下^[4]:

(1) $x = G(E_x, E_n)$ 。生成以 E_x 为期望值、 E_n 为标准差的正态随机数 x 。

(2) $E_m = G(E_n, H_e)$ 。生成以 E_n 为期望值、 H_e 为标准差的正态随机数 E_m 。

(3)将整态随机数代入式(1)计算 y_i ,令 (x, y) 为云滴。

给定正态云的三个数字特征值 (E_x, E_n, H_e) ,可以用上面的算法生成任意个数云滴组成的正态云。该算法生成的云自然地具有不均匀厚度的特性。三个数字特征值能够很好地描述整个云的形

态。

给定符合某一正态云分布规律的一组云滴作为样本 (x_i, y_i) , 产生云所描述的定性概念的 3 个数字特征 (E_x, E_n, H_e) , 即从定量到定性的转换, 其软件或硬件实现称为逆向云发生器^[5]。逆向云发生器算法如下^[5]:

- (1) $E_x = \text{mean}(x_i)$;
- (2) $E_n = \text{stdev}(x_i)$;
- (3) $E_m = \frac{-(x_i - E_x)^2}{2 \ln(\mu_i)}$, $H_e = \text{stdev}(E_m)$ 。

其中, mean 、 stdev 分别为求均值和标准差的函数。

以上的逆向云发生器算法是一种统计方法, 求出的数字特征值是一种估计值, 当云滴数较少时, 误差可能比较大, 随着云滴数增加, 误差将减小。

信息融合的基本原理就是模拟人脑综合处理信息的过程, 充分利用多源信息, 通过对这些多源信息的合理支配和使用, 将这些多源信息在时间和空间上的冗余和互补特性按照有效的准则加以组合, 最后获得对被处理对象在时间和空间上的一致性理解。

融合的概念最早出现于 20 世纪 70 年代的美国, 并最早应用于 C³I(command control communication and information, 简称 C³I)系统。其实在我国古代“瞎子摸象”就是信息融合的一个生动的实例。

目前比较受到公认的几种定义方式如下:

(1) Waltz 给出的定义是: 对来自多源的信息和数据进行检测、关联、估计和综合等多级多方面的处理, 以得到精确的状态和身份估计, 以及完整、及时的态势评估和威胁估计。

(2) JDL 当前的最新定义是: 信息融合是组合数据或信息以估计和预测实体状态的过程。

(3) Li X R 最近给出的定义是: 信息融合是为了某一目的对多个实体包含的信息的组合。

从以上学者的定义可以看出, 信息融合的定义越来越简单, 但是所涵盖的内容却越来越丰富。

3 系统评估云模型的建立

基于云理论和信息融合理论的系统性能评估, 就是分解影响系统性能的各级指标并获取各指标信息, 再运用云理论的有关知识来分别刻画定性指标和定量指标, 得到各指标的云模型, 最后根据信息融合理论, 综合评估系统的性能。主要由三个模块(如图 2 所示)构成:

- (1) 变量云化;
- (2) 评估指标的云模型;
- (3) 系统云模型融合。

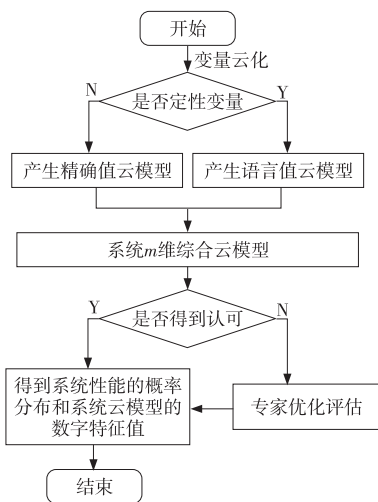


图 2 云模型处理流程图

3.1 变量云化并获得各评估指标云模型

根据云的定义, 对于一个论域 U 的定量变量, 由过程 P 来构造 U 上对应的定性概念的集合 A , 集合中的概念由不同的云刻画。反之, 过程 P 对表示定性概念的定性变量进行云描述, 包括获取云数字特征 (E_x, E_n, H_e) 和云的形状。我们称这个可逆的过程为变量云化过程。

在系统性能指标体系中, 既有精确数值型表示的, 又有语言值表示的。精确数值可以表示为熵和超熵都为 0 的值, 其云的数字特征表示为 $(E_x, 0, 0)$, 语言值的云数字特征表示为 (E_x, E_n, H_e) 。

提取 n 组系统状态组成决策矩阵, 那么 n 个精确数值表示的 1 个性能指标用 1 个云模型表示, 其数字特征如公式(2)^[6]:

$$\begin{cases} E_x = (E_{x1} + E_{x2} + \dots + E_{xn})/n \\ E_n = [\max(E_{x1}, E_{x2}, \dots, E_{xn}) - \min(E_{x1}, E_{x2}, \dots, E_{xn})]/6 \end{cases} \quad (2)$$

n 个语言值表示的一个性能指标用一个云模型表示, 其数字特征如公式(3)^[6]:

$$\begin{cases} E_x = \frac{(E_{x1} \cdot E_{n1} + E_{x2} \cdot E_{n2} + \dots + E_{xn} \cdot E_{nm})}{(E_{n1} + E_{n2} + \dots + E_{nm})} \\ E_n = E_{n1} + E_{n2} + \dots + E_{nm} \end{cases} \quad (3)$$

利用上述正态云生成算法, 可以分别生成定性变量(语言值描述)的云模型和定量变量(精确值描述)的云模型。

3.2 系统云模型融合

将定性变量和定量变量的云化结果置于统一

坐标系中,形成云簇。这些云簇反应了定性变量和定量变量的激活区间。

若系统具有 m 个性能指标,将表示 m 个性能指标的 m 个云模型置于同一个坐标系,形成反应该系统性能的 m 维综合云簇。当系统状态发生变化时,该云簇也会发生变化。

3.3 得到系统的云模型

根据系统在某一状态下的云簇,任意时刻,定性变量和定量变量都可以通过激活相应区间,这样,通过多时间点的多个输入,可以得到定性变量的采样点,将这些点看作云滴,通过逆向云发生器,还原出云模型的数字特征,可得到系统的云模型。

表 1 指标状态表

状态	可抵抗性	可识别性	可恢复性
1	一般	好	一般
2	一般	好	一般
3	很差	差	好
4	差	很差	差

$$B = \begin{pmatrix} 0.5 & 0.6 & 0.5 \\ 0.5 & 0.6 & 0.5 \\ 0.3 & 0.4 & 0.6 \\ 0.4 & 0.3 & 0.4 \end{pmatrix}$$

从决策矩阵中求得各指标的云模型的 E_x 和 E_n ,如表 2 所示。

表 2 各指标的云模型数字值

指标	可抵抗性	可识别性	可恢复性
E_x	0.425	0.475	0.5
E_n	0.13	0.008	0.1

4 实例应用

以某分布式信息融合系统的性能评估为例,该系统中存在大量的多源不确定信息,基于云模型的评估方法是在不确定的环境中,综合考虑多种因素的影响,并做出较为合理的评估方法。

影响该系统性能的因素如图 3 所示。

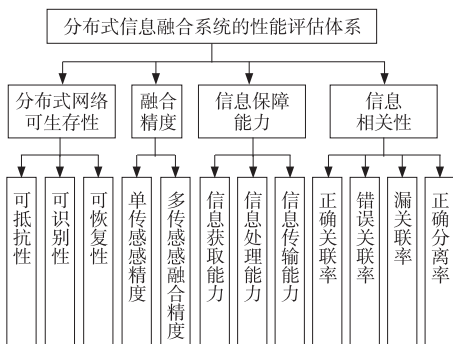


图 3 系统性能指标分级图

从图 3 可以看出,在分布式信息融合系统的性能评估体系中,分布式网络可生存性属定性信息,需要用语言值表示。融合精度、信息保证能力和信息相关性,属定量信息,需要用精确数值表示。

按照文中上述内容,对于定量的精确数值表示的信息,提取 n 组指标,利用式(2),用一个云模型表示;对于定性的语言值表示的信息,提取 n 组指标,利用式(3),用一个云模型表示。最后得到系统的综合云模型。

下面说明计算过程,文中的数据为设计值,旨在说明算法过程,无实际意义。对于定型变量,从系统中抽取分布式网络可生存性的 3 个指标的 4 组值,如表 1 所示。

设评语与评定值的对应关系:一般-0.5;好-0.6;差-0.4;很差-0.3,则可得到决策矩阵为:

得到分布式网络可生存性的三个指标的云模型如图 4 所示。

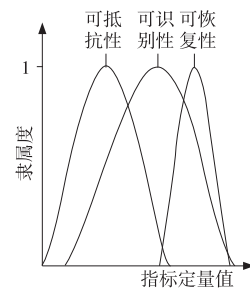


图 4 分布式网络可生存性的 3 个指标的云模型

同样,得到系统中其它各指标的云模型的 E_x 和 E_n ,如表 3 所示。

表 3 各指标的云模型数字值

指标	E_x	E_n
单传感器精度	0.525	0
多传感器融合精度	0.527	0
信息获取能力	0.575	0
信息处理能力	0.600	0
信息传输能力	0.625	0
正确关联率	0.635	0
错误关联率	0.645	0
漏关联率	0.650	0
正确分离率	0.700	0

将各个云化结果置于统一坐标系,得到云簇,如图 5 所示。通过多时间点的多个输入,得到各性能指标的采样点,还原出系统性能云模型,如图 6 所示。

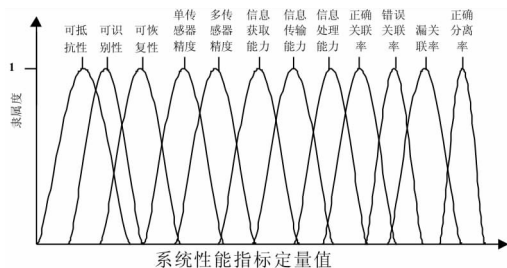


图5 系统云簇

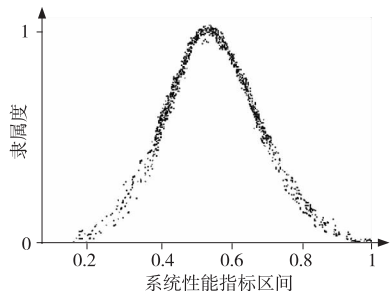


图6 系统性能云模型

从图5和图6的对比可以得到如下结论:虽然系统的信息处理能力、信息传输能力、关联率和分离率性能很不错,但是由于可抵抗性和可识别性性能一般,导致整个系统性能的评定值为0.5,对应的评语为一般。如果提高系统在可抵抗性和可识别性方面的能力,将会大大提高系统性能。

利用上述方法和模型所提供的评估结论,相对于传统的定量评估方法,该方法考虑的因素更多、更全面,得出的结论更直观、更接近实际。

5 结束语

不确定性的存在增加了评估过程中定性与定量综合集成的难度,云模型在不确定信息处理时,不强调精确的函数表示而是利用三个数字特征。充分利用这种不确定性,建立定性和定量之间的映射和变换,在处理系统性能评估过程中出现的不确定性具有优势。而在系统性能评估中采用信息融合技术,是根据信息理论中关于信息量的可加性,充分利用各种时空条件下多种信息源的信息,运用融合的方法,先对系统性能的各分指标进行分析,然后再综合获得系统性能更加完整准确的信息。

本文创新性地将二者结合起来,提出了新的系统性能评估方法,较好地融合了影响系统性能的多源信息,充分利用和集成了信息的模糊性和随机

性,完成了定性和定量之间的映射。这种方法不仅充分利用了影响系统性能的多源信息,又有效地处理了多源信息中的不确定性,具有科学性,同时最终得到的评估结果是语言值,具有直观性。该方法是对信息融合理论和云理论联合使用的探索。

由于文中涉及到多源信息的联合识别,比传统的评估方法增加了时间开销,因此如何改进算法降低时间复杂度是下一步的工作。

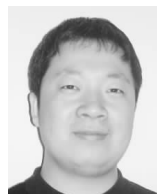
参考文献:

- [1] 刘常显,李德毅,潘莉莉. 基于云模型的不确定性知识表示[J]. 计算机工程与应用, 2004, 40(2): 32-35.
- [2] 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[J]. 系统工程与电子技术, 2008, 30(4): 772-776.
- [3] 王媛. 基于云模型的网络信息系统可生存性评估方法:[硕士学位论文][D]. 南京:南京理工大学, 2008.
- [4] 罗自强,张光卫,李德毅. 一维正态云的概率分析[J]. 信息与控制, 2007, 36(4): 741-745.
- [5] 刘婷,陈晓云. 基于逆云模型的支持向量机多类分类方法[J]. 福州大学学报, 2008, 34(3): 431-436.
- [6] 郭强,毕义明. 基于云模型的导弹信息作战指挥效能评估[J]. 指挥控制与仿真, 2008, 30(4): 61-64.



秦昶(1978-),女,陕西西安人,博士,工程师,研究方向为信息融合、计算机系统结构和计算机可靠性。E-mail: sad-dyqin1120@163.com

QIN Yi, born in 1978, PhD, engineer, her research interests include information fusion, computer architecture, and computer reliability.



卢强(1977-),男,山东莒南人,硕士,高级工程师,研究方向为计算机系统结构。E-mail: Luqiang_mail@sohu.com

LU Qiang, born in 1977, MS, senior engineer, his research interest includes computer architecture.



黄士坦(1937-),男,陕西西安人,研究员,研究方向为计算机系统结构。E-mail: shitanhuang@163.com

HUANG Shi-tan, born in 1937, research fellow, his research interest includes computer architecture.