

基于 Q 型聚类分析和贝叶斯判别算法 研究储层分类评价

付殿敬¹, 徐敬领², 王贵文²

1. 中国地质大学(北京)能源学院, 北京 100083
2. 中国石油大学(北京)油气资源与探测国家重点实验室, 北京 102249

摘要 储层分类评价是油藏研究的重要内容,为了使储层分类更准确、合理,本文采用 Q 型聚类分析和贝叶斯判别算法相结合的方法,进行储层分类评价研究。综合优选砂岩厚度、孔隙度、渗透率、碳酸盐含量及泥质含量等各种储层参数,应用数理统计方法将这些参数集合起来,采用 Q 型聚类分析的算法对储层进行分类;在此基础上,采用贝叶斯判别算法,建立这些储层参数与储层分类评价的定量判别关系,即建立进行储层分类评价的判别函数,依据此判别函数对非取心井的目的层进行了定量分类评价。实例证明,应用 Q 型聚类分析和贝叶斯判别算法相结合进行储层分类评价是有效的,其应用效果良好。

关键词 储层分类评价;储层参数;Q 型聚类分析;贝叶斯判别

中图分类号 TE122.2

文献标识码 A

doi 10.3981/j.issn.1000-7857.2011.03.03

Reservoir Classification and Evaluation Based on Q Cluster Analysis Combined with Bayesian Discrimination Algorithm

FU Dianjing¹, XU Jingling², WANG Guiwen²

1. School of Energy Resources, China University of Geosciences, Beijing 100083, China
2. State Key Laboratory of Petroleum Resources and Processing, China University of Petroleum, Beijing 102249, China

Abstract The reservoir classification and evaluation is an important topic of study, and in order to make reservoir classification more accurate and reasonable, in this paper, the method of Q cluster analysis combined with Bayesian discrimination algorithm is adopted. First, various reservoir parameters are optimized globally, including sandstone thickness, porosity, permeability, carbonate content, shale content and others, with the mathematical statistics method to assemble the parameters, and the algorithm of Q cluster analysis to do the reservoir classification. On this basis, Bayesian discrimination algorithm is adopted to establish the discriminant relationship between reservoir parameters and reservoir classification and evaluation, together with the discriminant function of reservoir classification and evaluation, based on which, the target beds of non-coring wells are classified and evaluated. Examples show that the method of Q cluster analysis combined with Bayes discrimination algorithm is effective to achieve reservoir classification and evaluation.

Keywords reservoir classification and evaluation; reservoir parameter; Q cluster analysis; Bayesian discrimination algorithm

0 引言

储层评价工作贯穿于油气田开发的始终,当前国内外储层评价的特点是紧紧围绕勘探开发的需要^[1-2],进行多学科结合的综合评价,其总的发展趋势是从宏观到微观、从静态

到动态、从定性向定量发展^[3-6]。国内外学者针对储层评价进行了大量的工作,1966年 Leveson 首次提出用宏观参数(孔隙度、渗透率)作为储层分类评价的标准,随后 WeKeng 和罗蛰潭等提出应用通过岩石毛管压力曲线求取的微观孔隙结构

收稿日期:2010-10-20;修回日期:2010-12-05

作者简介:付殿敬,博士研究生,研究方向为油气田开发,电子信箱:fudianjing@163.com

参数进行储层分类评价等^[7-9]。

由于储层的非均质性,单纯用宏观或微观参数对储层进行分类评价存在片面性。因此,必须选用综合方法对储层多参数进行评价。评价方法的选择必须符合研究地区的地质特征,为勘探开发决策提供依据,另外还要具备综合性及定量化。因此本文采用综合宏观、微观等参数的一种新的储层评价方法——Q型聚类和贝叶斯判别法,对储层的多个影响因素进行评价,最终得到一个综合评价指标,并据此对储层分类^[10-12]。

1 Q型聚类分析

聚类分析是一种逐级归类的方法,主要思想是根据一定的相似性指标,按照研究对象的相似程度合理地进行归并和分类,当利用其解决样品的分类问题时称为Q型聚类,解决变量的分类问题时称为R型聚类。本文主要利用聚类分析解决储层的分类评价问题,储层类型划分的流程如图1所示,根据样品的观测指标和Q型聚类算法,计算样品间的相似程度,把相似的样品归为一类,不相似的归为另一类,把关系密切的归到一个小的分类单位,把关系不密切的归到一个大的分类单位,总之,Q型聚类分析的结果是形成一个由大到小的分类谱系图或聚类图。聚类图不仅可以直观地表示研究对象之间的相似关系和分类情况,反映各类储层,而且可以定量地指示相似程度,从而为地质解释提供良好的依据。

进行储层定量分类,首先必须确定一些划分类型的定量指标,它们是反映样品(或变量)相似(或相关)程度的度量。

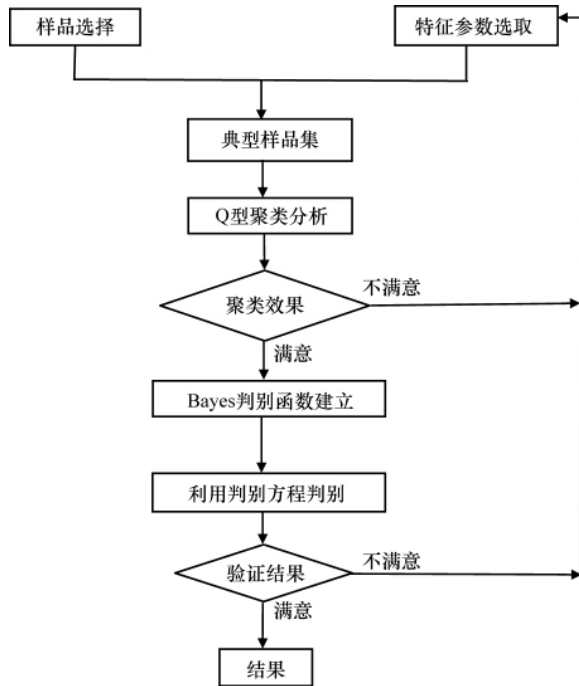


图1 储层类型划分与判别流程图

Fig. 1 Flow chart of division and discriminant of reservoir types

1.1 距离系数

距离系数是Q型聚类分析常用的分类统计量。如果把在m个变量上进行观测的N个样品看成m维空间的N个点,则任意两样品点 x_j 与 x_k 之间的相似程度可用m维空间两点间的距离表示,则距离系数定义为

$$d_{jk} = \left[\frac{1}{m} \sum_{i=1}^m (x_{ij} - x_{ik})^2 \right]^{1/2} \quad (1)$$

其中, x_{ij} 为样品j的第i个特征参数变量, x_{ik} 为样品k的第i个特征参数变量。除以m是为了得到一个与变量数无关的相对距离。

由式(1)可以看出,当各变量的单位和数量级不同时,如果直接用原始数据计算,就会强化那些绝对值大的变量而淡化绝对值小的变量,因此在计算前要对原始数据作一些处理或变换,常用归一化方法对原数据进行预处理,计算公式为

$$x'_{ij} = \frac{x_{ij} - x_{i\min}}{x_{i\max} - x_{i\min}} \quad (2)$$

其中, $x_{i\min}$ 为样品i的特征参数变量的最小值, $x_{i\max}$ 为样品i的特征参数变量的最大值。

根据式(1),对所有样品两两求距离系数 d_{jk} ,可得到距离矩阵D,

$$D = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1p} \\ d_{21} & d_{22} & \cdots & d_{2p} \\ d_{31} & d_{32} & \cdots & d_{3p} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{np} \end{bmatrix} \quad (3)$$

则该矩阵为对角线元素为0的对称矩阵,样品j与样品k越相似, d_{jk} 越趋于0。

1.2 相似系数

相似系数是描述样品之间相似程度的度量。把每个样品看作m维空间的向量,两样品 x_j 与 x_k 的相似程度定义为两向量夹角的余弦,即

$$\cos\theta_{jk} = \frac{\sum_{i=1}^m x_{ij} \cdot x_{ik}}{\sqrt{\sum_{i=1}^m x_{ij}^2} \cdot \sqrt{\sum_{i=1}^m x_{ik}^2}} \quad (4)$$

类似于距离矩阵,根据式(4)对所有样品两两求相似系数,可得到相似矩阵 $[\cos\theta_{jk}]$,则该矩阵为主对角线元素为1的对称矩阵,样品j与样品k越相似,相似系数越趋近于1,反之则趋于0。

本次分类评价在样品之间定义某种距离系数,然后按距离系数进行亲近类的逐次合并,最后以分类谱系图表示样品间的亲疏关系,从而达到分类评价的目的。

2 贝叶斯判别分析

确定了储层划分的类型后,将标准样品分组,建立判别分析的输入文件,利用贝叶斯判别分析法建立各类储层的判别函数模型。

第 i 类储层的判别函数为

$$P_i = \alpha_0 + \sum \alpha_{ij} x_{ij} \quad (i=1, 2, 3, 4; j=1, 2, \dots, 11) \quad (5)$$

其中, α_{0i} 为判别系数, x_{ij} 为 i 类储层的第 j 个特征变量, P_i 为待判储层属于第 i 类储层的判别值。

Bayes 判别分析的准则为假设样品来自各类母体是等可能的, 由贝叶斯公式计算这一样品来自第 i 类的后验概率 p_i

$$p_i = \frac{q_i p_i(y_1 y_2 \dots y_m)}{\sum_{i=1}^h q_i p_i(y_1 y_2 \dots y_m)} \quad (6)$$

其中, $p_i(y_1 y_2 \dots y_m)$ 为样品 $Y(y_1 y_2 \dots y_m)$ 属于第 i 组的概率密度, q_i 为第 i 组的先验概率。可用样品的频率作为先验概率的估计值, 即 $q_i = n_i/n$, 判别样品归属类型的准则是 $p_i(y_1 y_2 \dots y_m)$ 为最大, 当判别函数为最大, 即满足以下式

$$q_s p_s(y_1 y_2 \dots y_m) = \max_{1 \leq i \leq h} [q_i p_i(y_1 y_2 \dots y_m)] \quad (7)$$

时, 把新样品 $Y(y_1 y_2 \dots y_m)$ 划入第 S 类。

3 储层类型划分与判别

3.1 样品集的形成

利用 Q 型聚类方法进行储层分类时, 为尽可能合理、准确地划分储层, 分类样品和特征参数的选取是十分重要的。所选择的样品应尽可能包括各种岩性, 并且有岩心观察分析等资料相对应, 这样取得的样品才具有典型性、代表性, 有助

于储层的完整分类。在研究中针对研究区储层低孔低渗的特点, 结合岩心、含油性情况, 并充分考虑各项资料的分布情况, 在 8 口井的储层中选择了 25 个样品组成小层的样品集, 每个样品都反映了对应砂层单元的总储层类型。

3.2 特征参数选取

不同类型的储层有不同的特征参数值, 各参数的大小综合反映了该类储层的内在特征。在测井解释处理中共形成了多种参数, 为了选取能反映本区储层类型的参数, 对每个参数进行综合分析, 通过对应分析方法与地质经验结合进行参数选取, 最终选用孔隙度、渗透率、泥质含量、砂层厚度、碳酸盐含量 5 个参数作为储层评价的特征参数。

(1) 孔隙度、渗透率是储层研究中的经典参数, 反映了储层的总体特征, 是表征储层质量优劣的重要参数。

(2) 泥质含量从组成砂岩的胶结物中泥质含量高低的角度来说明储层性能。一般地, 泥质含量低、粒度中值较大的储层, 其孔、渗较好, 有利于油气的储集。

(3) 砂岩厚度反映了储层砂体的发育程度, 可以从侧面反映储集体的体积规模。

(4) 碳酸盐含量反映储层胶结致密程度, 从成岩作用方面反映了对储层改造程度。碳酸盐含量高则渗透率低, 储层物性较差。

综上所述, 通过细致分析, 筛选了庙南 25 个典型样品, 各样品由 5 个参数定量表征, 从而建立了扶新隆起南泉四段储层分类的样品集(表 1)。

表 1 扶新隆起南泉四段储层分类标准样品集

Table 1 Reservoir classification standard sample sets of Nanquan fourth stage in Fuxin uplift

样品编号	井号	小层	砂岩厚度/m	孔隙度/%	渗透率/($10^{-3} \mu\text{m}^2$)	碳酸盐含量/%	泥质含量/%
1	庙 138	7	4	15.68	0.82	1.6	6.68
2	庙 138	5	7	14.3	0.379	3.07	16.3
3	庙 133	4	4.4	14.53	0.73	4.03	10.88
4	庙 133	4	5	16.8	1.68	2.55	7.36
5	庙 134	10	5.4	14.65	6	2.47	5.6
6	庙 134	11	2.4	15.5	10	2.8	6.2
7	庙 134	10	1.74	12.5	1.84	6.98	15.37
8	庙 134	10	2.4	13.25	1.6	1.9	13.52
9	庙 134-1	4	3.9	14.73	2.6	1.3	10.13
10	庙 134-1	8	3	14.3	1.35	2.4	9.55
11	庙 134-1	7	1.8	13.47	0.89	2.87	16.25
12	庙 134-1	10	3	14.5	2.45	13.6	1.7
13	庙 4	10	6.1	16.88	6.26	2.05	4.5
14	庙 4	9	4.8	13.26	6.96	3.9	5
15	庙 4	9	1.8	13.47	4.33	5.83	15.97
16	庙 112	12	1.8	17	0.89	1.3	11.78
17	庙 112	3	3.5	9.26	0.2	2.76	18.59
18	庙 112	4	7	8.2	0.11	5.5	29.01
19	庙 112	2	2.5	6.35	0.02	8.65	25.92
20	庙 112	3	5	16.8	4.07	0.8	10.2
21	庙 135	7	3.6	11.02	0.5	1.16	14.94
22	庙 135	4	4.8	9.38	0.16	4.25	17
23	庙 135	10	5.4	8.73	0.08	2.3	13.42
24	庙 22-37-39	8	5.2	11.37	0.36	3.87	9.17
25	庙 22-37-39	6	1.8	8.6	0.28	6.9	32.75

3.3 储层类型划分

针对研究区块,利用 Q 型聚类分析中计算的样品距离系数,对以上所选样品集进行聚类计算,得出聚类成果图,即聚类谱系图(图 2)。由图 2 可以看出,谱系图最右边有两个分支,逐次向左,每个分支又分为 2 个次一级分支,此时聚类值恰好符合储层划分的标准,且每个分支再往下细分都是独立的,没有相互联系,因此由谱系图可以定量地将该区储层分为 4 类(表 2)。各参数在储层分类中均是定量的,该方法不仅对储层进行分类,而且实现了储层定量分类的结果,即每个参数在每类储层中都有一个量值。根据对比发现,这个储层分类标准与中国石油股份有限公司制定的分类标准具有一致性。

为了进一步验证分类的合理性,与岩心观察结果和现场储层分类结果进行对比。结果表明,第 I 类为物性好的储层,本文所判断出的具有 I 类储层的井段,与现场的划分结果是一致的;第 II 类为物性较好的储层,所判断出的 II 类储层的井段与现场划分也是一致的;第 III 类是物性较差的储层;第 IV 类是差储层。其中,每类储层划分结果与现场划分储层的结果是一致的,因此按此类型划分该区各井各小层的储层类型是可行的。

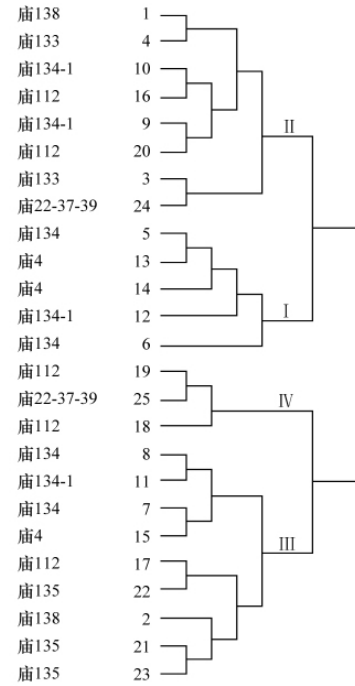


图 2 扶新隆起聚类分析谱系

Fig. 2 Hierarchical cluster analysis in Fuxin uplift

表 2 扶新隆起 4 类储层划分标准

Table 2 Four reservoir type division standards in Fuxin uplift

储层类型	孔隙度/%	渗透率/(10 ⁻³ μm ²)	泥质含量/%	砂岩厚度/m	碳酸盐含量/%
I	13.26~16.88	>5	4.5~6.2	2.4~6.1	2.05~3.09
II	11.37~17.00	0.36~5	6.68~11.78	1.8~5.2	1.3~4.03
III	8.73~13.47	0.08~4.33	13.42~18.59	1.74~5.4	1.16~6.98
IV	6.35~8.6	0.02~0.28	25.92~32.75	1.8~7.0	5.5~8.65

3.4 储层类型判别

确定了储层划分的 4 种类型后,将标准样品按上述 4 类储层类型分组,建立判别分析的输入文件,利用 Bayes 判别分析法建立各类储层的判别函数模型,判别公式如下:

$$I \text{ 类储层} = 3.643H + 4.320\phi + 4.255K - 6.194w(\text{CaCO}_3) - 9.115V_{\text{sh}} - 179.393$$

$$II \text{ 类储层} = 3.076H + 5.951\phi + 2.442K - 3.692w(\text{CaCO}_3) - 3.095V_{\text{sh}} - 69.559$$

$$III \text{ 类储层} = 2.9981H + 4.982\phi + 2.839K - 4.139w(\text{CaCO}_3) - 4.956V_{\text{sh}} - 80.398$$

$$IV \text{ 类储层} = 7.021H + 7.868\phi - 1.255K + 4.200w(\text{CaCO}_3) + 4.056V_{\text{sh}} - 97.456$$

其中, H 为砂岩厚度, ϕ 为孔隙度, K 为渗透率, $w(\text{CaCO}_3)$ 为碳酸盐含量, V_{sh} 为泥质含量。

遵循以上判别公式,编制判别分类程序,将研究区单井的每个砂岩层按砂组分类并进行储层类型判别归类,统计发现,随储层类型变化,重要特征参数呈规律性变化,如果储层

变差,则孔隙度、渗透率依次降低,泥质含量及碳酸盐含量则趋于增大。该储层分类标准与已有的岩性、物性等测井解释资料特征符合良好,反映了在精选样品集基础上,利用 Q 型聚类、贝叶斯判别分析综合划分储层类型是合理可靠的。

4 结论

(1) 聚类分析的目的是用数学方法对给定的样品找出一个合理的分类体系,以解决储层分类问题。

(2) 利用 Q 型聚类分析对关键取心井研究,确定了储层的类型、各类储层的评价标准及储层参数之间的内在关系。

(3) 应用贝叶斯判别函数建立了储层分类的综合判别函数及定量分类的标准。

(4) 采用 Q 型聚类分析和贝叶斯判别相结合的方法对低渗透储层的定量分类评价是有效的,经取心井岩心检验表明,该定量分类评价方法可行且具有较高的综合性和科学性。本文为说明问题,仅选少量样品和参数,若样品、指标数量大,可借助计算机进行计算和分类,则更能体现该方法的优点。

参考文献 (References)

- [1] 吴胜和, 熊琦华. 油气储层地质学[M]. 北京: 石油工业出版社, 1998: 113-128.
Wu Shenghe, Xiong Qihua. Reservoir geology [M]. Beijing: Petroleum Industry Press, 1998: 113-128.
- [2] 于志钧, 赵旭东. 石油数学地质[M]. 北京: 石油工业出版社, 1986: 50-60.
Yu Zhijun, Zhao Xudong. Mathematical petroleum geology [M]. Beijing: Petroleum Industry Press, 1986: 50-60.
- [3] 张旭, 颜其彬, 李祖兵. 陆相碎屑岩储层定量评价的新方法——以河南某油田为例 [J]. 天然气地球科学, 2007, 18(1): 141-148.
Zhang Xu, Yan Qibin, Li Zubing. *Natural Gas Geoscience*, 2007, 18(1): 141-148.
- [4] 吕晓光, 赵永胜, 史晓波. 储层分类方法的应用及评价[J]. 大庆石油地质与开发, 1995, 14(3): 10-15.
Lu Xiaoguang, Zhao Yongsheng, Shi Xiaobo. *Petroleum Geology & Oil Field Development in Daqing*, 1995, 14(3): 10-15.
- [5] 朱志伟, 高茂生, 朱远峰. 海岸带基本类型与分布的定量分析[J]. 地学前缘, 2008, 15(4): 315-321.
Zhu Zhiwei, Gao Maosheng, Zhu Yuanfeng. *Earth Science Frontiers*, 2008, 15(4): 315-321.
- [6] 刘宗彦, 王燕. 测井储层分类评价方法的研究[J]. 国外测井技术, 2008, 23(4): 19-22.
Liu Zongyan, Wang Yan. *World Well Logging Technology*, 2008, 23(4): 19-22.
- [7] 罗平, 邓恂康, 罗蛰潭. 克拉玛依油田八区下乌尔禾组储层评价[J]. 石油与天然气地质, 1987(1): 1-6.
Luo Ping, Deng Xunkang, Luo Zhetan. *Oil & Gas Geology*, 1987(1): 1-6.
- [8] 李世兴, 王威, 任士伟, 等. 白庙气田储层评价分类新方法[J]. 内蒙古石油化, 2008(10): 329-331.
Li Shixing, Wang Wei, Ren Shiwei, et al. *Inner Mongolia Petrochemical Industry*, 2008(10): 329-331.
- [9] 丁熊, 谭秀成, 雷一文, 等. 基于模糊聚类分析的复杂碳酸盐岩储层定量评价 [J]. 西安石油大学学报: 自然科学版, 2009, 24(3): 25-34.
Ding Xiong, Tan Xiucheng, Lei Yiwen, et al. *Journal of Xi'an Shiyou University: Natural Science Edition*, 2009, 24(3): 25-34.
- [10] 操应长, 王艳忠, 徐涛玉, 等. 东营凹陷西部沙四上亚段滩坝砂体有效储层的物性下限及控制因素 [J]. 沉积学报, 2009, 27(2): 230-237.
Cao Yingchang, Wang Yanzhong, Xu Taoyu, et al. *Acta Sedimentologica Sinica*, 2009, 27(2): 230-237.
- [11] 张连文, 郭海鹏. 贝叶斯网引论 [M]. 北京: 科学出版社, 2006.
Zhang Lianwen, Guo Haipeng. Introduction to Bayesian network[M]. Beijing: Science Press, 2006.
- [12] 高新波. 模糊聚类分析及其应用[M]. 西安: 西安电子科技大学出版社, 2004.
Gao Xinbo. Fuzzy cluster analysis and application [M]. Xi'an: Xidian University Press, 2004.

(责任编辑 刘志远)

·学术动态·



“2011 年国际电气电子工程师学会 服务、运筹、物流与信息化 国际会议(SOLI'2011)”征文

由国际电气电子工程师学会智能交通系统委员会(IEEE ITSS)主办,中国科学院自动化研究所承办的“2011 年国际电气电子工程师学会服务、运筹、物流与信息化国际会议(SOLI'2011)”将于 2011 年 7 月 10—12 日在北京市召开。会议旨在为服务科学、服务运筹、物流和信息化领域的研究者提供交流机会,探讨该领域的研究问题、挑战和未来的发展方向,分享他们的研发成果和经验。

同时召开的还有“2011 年国际电气电子工程师学会汽车电子与安全国际会议(ICVES'11)”和“2011 年国际电气电子工程师学会情报与安全信息学国际会议(ISI'11)”。

征文范围:服务设计、工程、运营和创新;服务规划和设计,服务过程工程,快件服务和快速物流,医疗保健系统,金融服务,零售及服务管理,质量和客户满意度,指标和基准,安全及与安全有关的服务和管理、应急规划、运筹学、生产工程、智能交通、工程咨询、交通规划、综合运输、服务业务;物流及供应链管理:按需投递,物流规划,货运代理和报关,会场物流管理,仓储和配送,交通管理系统,逆向物流,物流可视性和控制、采购、供应链协作、供应链过程,物流网络;物质流(MF)科学和技术:物质流基本科学(MF 数学,物理,化学,生物,等),综合物质流理论,自然世界的物质流,社会世界的物质流,经济世界的物质流,物质流元理论,物质流性质,物质流工程,物质流产业,物质流技术经济,循环物质流系统,X 方物质流(XPMF),物质流的复杂性和涌现,物质流信息与仿真技术,物质流系统和网络,物质流的财政措施;服务/事件管理与制造:需求预测,客户关系管理,事件通信和报警,服务培训,服务维持,服务质量,服务捆绑,电子商务服务市场,事件管理系统,活动赞助,基于事件的生产 and 供应链,基于事件的产品和制造业,智能制造,个性化;信息通信技术和系统(ICTS):ICTS 服务设计和管理,ICTS 服务标准,定位,组成和捆绑,过程建模,扩大和自动化,实时识别与跟踪,物流的普适计算,决策支持系统,基于代理的软件系统,RFID,数据仓库和数据/网页挖掘,商业智能,系统互操作性和集成性,信息安全,IT 项目管理,建设项目的信息管理;电子商务与知识管理:无线通信和移动商务,移动服务,电子政务,信息资源管理,IT 和企业创新管理,IT 和可持续发展企业战略,符号学,企业绩效管理,客户关系管理,信息经济学,网络文化与和谐社会,分布式计算,传感网络。

大会收录的论文将全部收录 IEEE 数据库,并被 Ei、ISTP 检索,提交的论文文字必须是英文。征文截止时间:2011 年 3 月 13 日。会议网站:www.ieeeves.org。