

强化人机交互的在线会场*

罗 斌, 唐锡晋, 张 楠

(中国科学院数学与系统科学研究院, 北京 100190)

摘要: 随着信息技术的发展与应用, 以及学术会议规模的逐渐扩大和频率的不断增长, 更多增强型的信息支持技术被引入到学术会议管理系统中, 大大提高会议组织效率, 促进更广泛的学术交流和深入的科学合作. 在线会场提供计算机支持的会议事务处理, 同时展现相关学科的不同观点, 为知识创造场的涌现提供支持. 本文介绍了在线会场(OLCB)概念、系统框架及 CorMap 分析技术, 主要工作在改进 CorMap 分析人机交互方式增强在线会场功能, 最后将增强型的在线会场应用于 MCS'2007 并得出结论.

关键词: 在线会场; 会议挖掘; CorMap 技术; 知识共享

Enhancing Human-Computer Interaction in On-line Conferencing Ba

LUO Bin, TANG Xi-jin, ZHANG Nan

(Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: With the development in information technologies and growth in the scale and frequency of conference, more and more enhancing information support technology are happening to conference transaction management system, which greatly enhance the efficiency of the conference transaction processing and promote wide academic exchanges and deeper cooperation between researchers. The On-line Conferencing Ba is an computer support conference management system, and at the same time, it can exhibit the different perspectives of the relevant disciplines, and also support emergence of a ba for knowledge creation. This paper introduce the concept and the framework of OLCB and the CorMap tool. We focus on task that enhancing the OLCB by providing Human-computer Interaction to CorMap analysis of the conference data. At last, we using the revised system on MCS2007 and report the results.

Key words: on-line conference Ba; conference mining; CorMap; knowledge sharing

1 引言

据国际会议协会 2008 年 4 月 14 日对其拥有的能反映一定水平的会议信息数据库抽样调查统计, 2007 年举办的能定期召开并至少在三个国家或地区内循环的由国际协会组织的会议达到 6500 场, 同比 2006 年增加 800 场^[1]. 随着会议规模的逐渐扩大和频率的不断增长, 组织会议是一项很有挑战性的工作, 研究人员将更多增强型的信息技术应用于会议数据挖掘从而形成虚拟会场, 为组织者提供更好的帮助, 为与会者和其他研究人员提供更多更直

* 基金项目: 国家自然科学基金(70571078); 创新群体项目(70221001)

观的信息。这就促进了会议管理系统的改进,使其不仅局限于对会议事务如论文提交,会议信息发布等功能,将信息技术应用于论文分配选择评审专家、使关联度更大的专家在同一个讨论组是最典型的改进^[2,3],将数据挖掘技术对会议数据进行深入分析,并将挖掘结果推送给与会者和其他研究人员,以启发人们新的思想,提供知识共享,促进潜在合作,成为会议系统最有价值的辅助功能^[4,5]。

在线会场(On-line Conferencing Ba)由唐锡晋等 2006 年提出^[6],通过组织的国际会议尝试所研究的定性综合集成方法与技术中而沉淀出来的概念。在线会场不仅仅是一种技术系统,在完成会议信息发布、提交论文、在线论文指派及评审等会务管理的基础上,对会议数据进行一系列挖掘工作并将其可视化结果推送给与会者,为与会者提供在线研讨,形成一种创造力支持系统。本文介绍在线会场的概念和系统框架,结合已有会议挖掘工作^[7-9],进一步改进 OLCB 人机交互方式,增强在线交互功能,体现人机结合、以人为本,从定性到定量的综合集成思想,从而更大程度激发与会者的创造力,使隐性知识以不同的形式表达,促进灵感的涌现,为与会者提供更好的服务。

2 在线会场(On-line Conferencing Ba)

“场”(Ba)的存在形式不拘一格,可以是物理的研讨室,也可以是基于 Web 的虚拟环境,甚至是一种气氛等等。日本学者 Nonaka 将“场”的概念引入他所提出的关于知识转化和创造的 SECI 模型^[10],在“场”中体现出知识是动态的,有关联的,基于个体与群体思维活动的。学术会议是一个信息和知识交互的场所,会议的组织者实际上提供一个物理的“场”来进行动态知识共享及新思想的涌现。“在线会场”引用了 Nonaka 知识创造理论中“场”的概念,是为这些活动的组织而设计并提供辅助的跨空间和时间虚拟平台,目的在于提供一般性会议信息和会议事务处理,形成学术交流和信息共享,产生各种想法和灵感的场所,通过这样的“场”使与会成员都参与知识的共享和创造。

在线会场与一般会议管理系统功能相似,但又不满足于论文提交、论文指派及评审等会议事务处理,同时提供虚拟会场,提供所有分组和分组报告幻灯片,录音,建立一个讨论区,提供对所有参会论文的作者和关键词对应分析(CorMap 分析)和作者及关键词网络分析(iView 分析),其框架结构如图 1 所示。中科院数学与系统科学研究院综合集成与知识科学研究小组已研制的探测非结构问题结构的定性综合集成技术——CorMap 和 iView,在不同领域的数据挖掘中得到应用^[9,11]。由 CorMap 和 iView 两种技术对会议数据挖掘之后,将可视化结果以图片形式提供给在线会场,形成在线会场的 CorMap 分析和 iView 分析。

在线会场通过第七届知识与系统科学研讨会(KSS' 2006)和第七届国际综合集成与复杂系统研讨会(MCS'

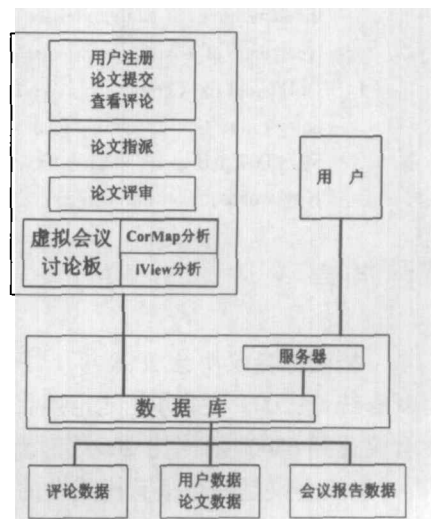


图 1 在线会场的系统框架^[5]

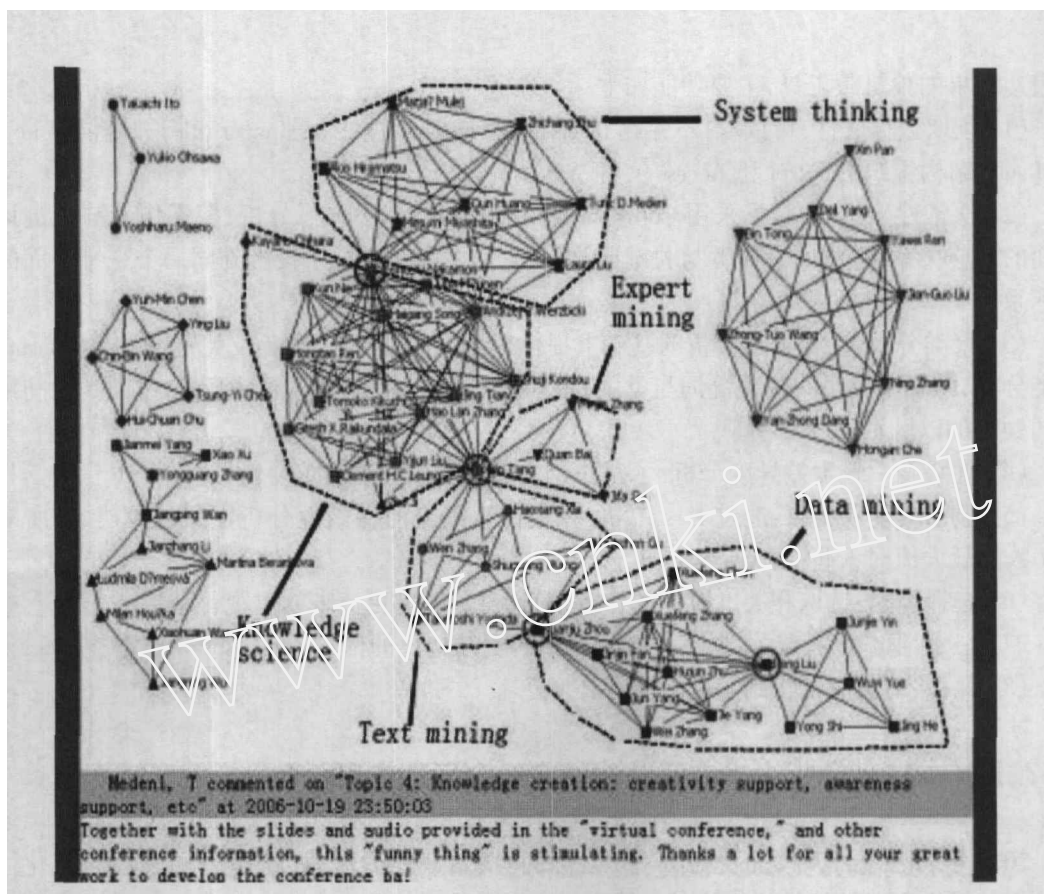


图2 KSS' 2006 iView 分析之作者网络^[5]

2007)的应用,为与会者提供了基于会议数据的不同挖掘结果,主要是通过结果可视化输出有助于参会者和相关研究人员初步了解“谁是该领域的主要人员”,“有哪些研究兴趣小组”,“各研究人员关注点是什么”等问题,了解该领域研究的一些主要变迁和现状,使人们生成一个该领域研究的概念框架,为知识创造营造知识的转换和提升的环境.图2所示即为对KSS' 2006的iView分析之作者网络页面.

在线会场不完全是一种技术支持系统,更旨在一种创造力支持系统.“场”的理念要求知识是动态的,相互关联的,需要人的参与,在线会场的人机交互功能是整个系统的关键点.目前在线会场提供静态挖掘结果并将其推送给用户,交互功能体现在用户可对会议挖掘结果进行评论,用户不能对会议挖掘过程进行操作,对于有兴趣的研究人员无法实现深入挖掘,引发更多的思考和关注点.为此,本文的主要工作是改进在线会场的CorMap分析人机交互手段,实现以人为为主的人机交互方式,尝试实证互动挖掘,增强在线会场功能.

3 CorMap 分析工具介绍

CorMap和iView分析工具由综合集成与知识科学研究小组独立研制、开发的定性建模与可视化表达工具^[12].CorMap采用对应分析、聚类分析等技术,iView则采用图论、复杂网络等技术.CorMap和iView分析工具目前在名老中医思想挖掘,社会风险关键词抽取,KSS2006会议挖掘等领域得到应用^[5,9,11,13].CorMap分析工具是能对定性数据进行建模、挖掘和可视

化表达结果的单机版工具,主要功能如下:

(1) 读取数据文件:可读取按主题 thesis,作者 author,论文标题 paper title,关键词集 keyword set 结构组织的 Excel 和 Access 文件.

(2) 探索性分析:对数据采用 Correspondence Analysis 算法进行分析,探索作者与关键词之间的对应关系,并将所有作者与关键词、以及它们之间对应关系投影到同一个二维平面的坐标系统中呈现出可视化结果.

(3) 聚类分析:在探索性分析的基础上提供对会议数据进行作者及关键词的 K - means 聚类分析,计算出类之间各关键词与类中心关键词(用大一号字体标注)距离,不同类之间用不同颜色加以区别.

(4) 优势度一致度差异度分析:了解作者之间的关联程度及重要性.

(5) 显示最佳聚类值:通过 K - Means 算法计算出当前探索性分析下最佳聚类, K 所对应 J 数值最大值时 K 即为最佳聚类值.

(6) 选择专家:选取相关性较高的作者进行分析,获取更清晰的可视化结果.

(7) 显示/隐藏作者或关键词:便于用户了解更结构化的信息.

(8) 保存:对所有分析结果进行保存.

4 改进 CorMap 分析人机交互功能的在线会场及应用

2007 年 MCS'2007 举办时,提出两年之际的 OLCB 仍然只能通过静态图片进行可视化挖掘结果的推送和对挖掘结果评论的在线交互.依照“场”的理念,本文通过将 CorMap 分析依托在 Web 技术上,实现 CorMap 分析的人机交互,用户对会议挖掘进行控制,促进深入挖掘.随着用户对挖掘的参与,可进一步丰富留言板的内容,引发更多的联想与思考,扩展大家的思考空间,增强 OLCB 在线交互的功能.

4.1 改进在线会场的 CorMap 分析人机交互功能

笔者通过运用 Java Applet 技术,在网页 HTML 文件中嵌入 Applet 小程序,实现人机交互,以人为主的会议数据挖掘分析,并将结果可视化输出.在本次改进中,由 HTML 向 Applet 传入数据,由程序调用数据和算法进行分析,从而实现对会议数据 CorMap 分析的人机交互以及可视化结果的动态输出.改进后在线会场的 CorMap 分析主页面如图 3 所示,包括以下功能:显示数据信息(DataSet),探索性分析(CorMapping),聚类分析(Keyword Cluster),最佳聚类值(Fittest K),选择作者(Select Authors),显示/隐藏作者(Show/Hide Author),显示/隐藏关键词(Show/Hide Keywords).

增强人机交互后在线会场的 CorMap 分析在二维平面上显示作者、关键词及其对应关系,无背景的是作者,有白色背景是关键词,鼠标停留于某个论文作者时会显示该作者的论文标题,停留于某个关键词时会显示对应于探索性分析中心的坐标值.具体结果如图 4 所示,图中显示了 CorMap 对 MCS2007 进行探索性分析的结果,会议所有数据(包含作者,论文标题,论文关键词)以及显示了作者(C. H. Tian)的论文标题.

4.2 改进后在线会场 CorMap 分析的应用

图 4 是实现人机交互功能后对 MCS'2007 会议进行 CorMap 分析的结果,可看出视图相

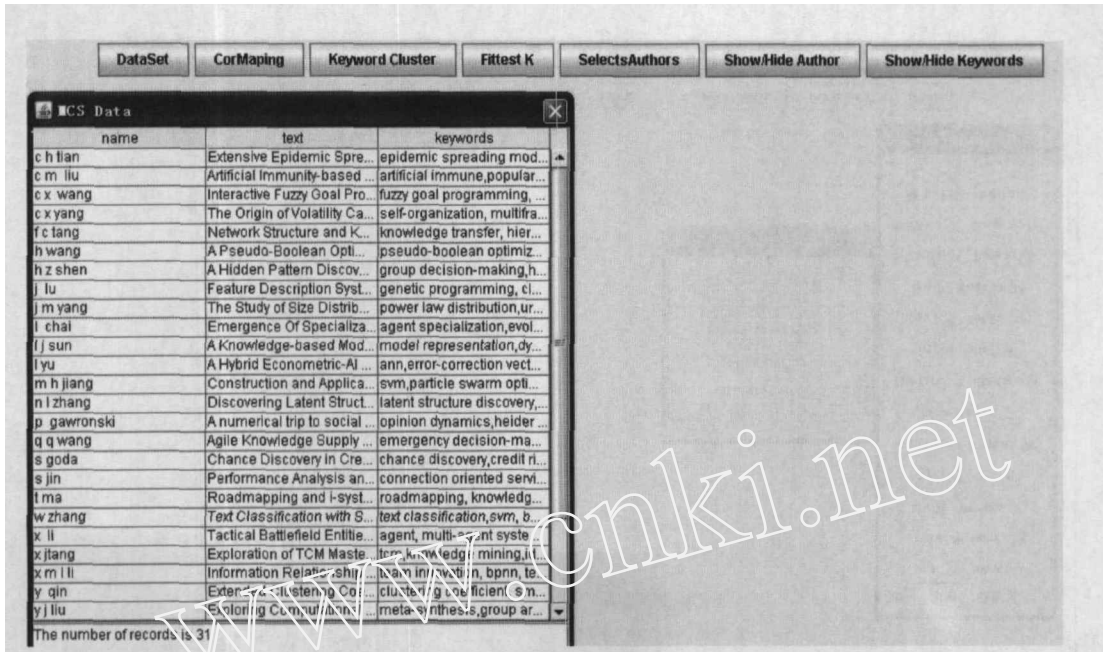


图3 在线会场 CorMap 分析页面

对分散,能反映会议数据信息,可动态选择显示更多的会议信息,提供相对稳定的数据挖掘结果.图5是利用改进的在线会场对 MCS'2007 不选择作者(P Gamronki)进行探索性分析基础上进行均值为3的聚类分析,其中最佳聚类为8类(见图中对话框中J值).

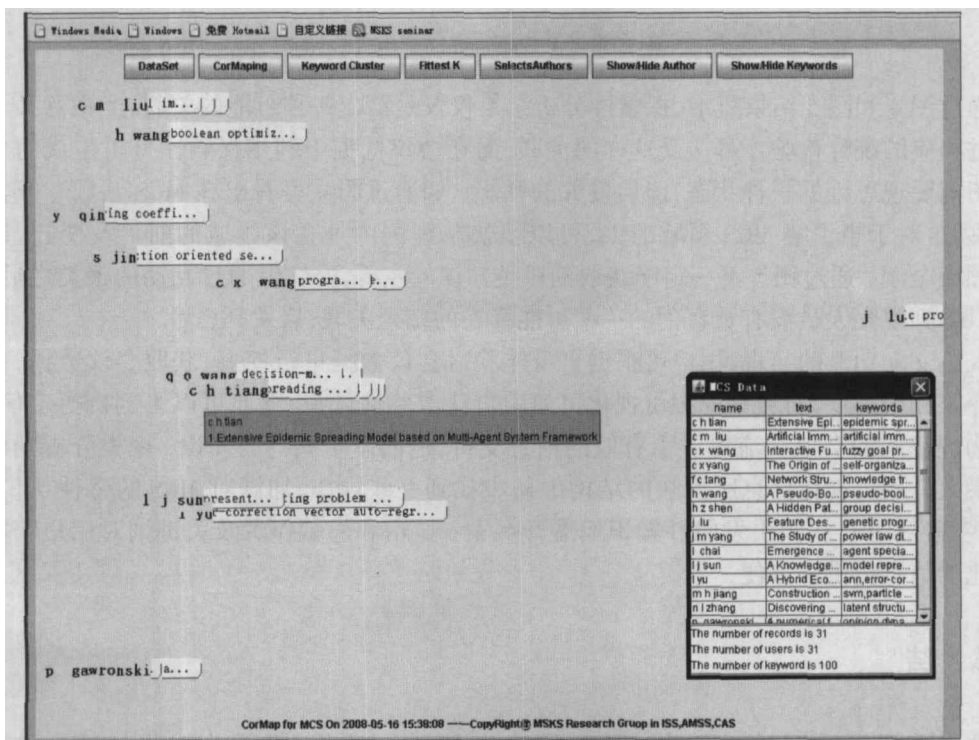


图4 改进后在线会场对 MCS'2007 CorMap 分析结果

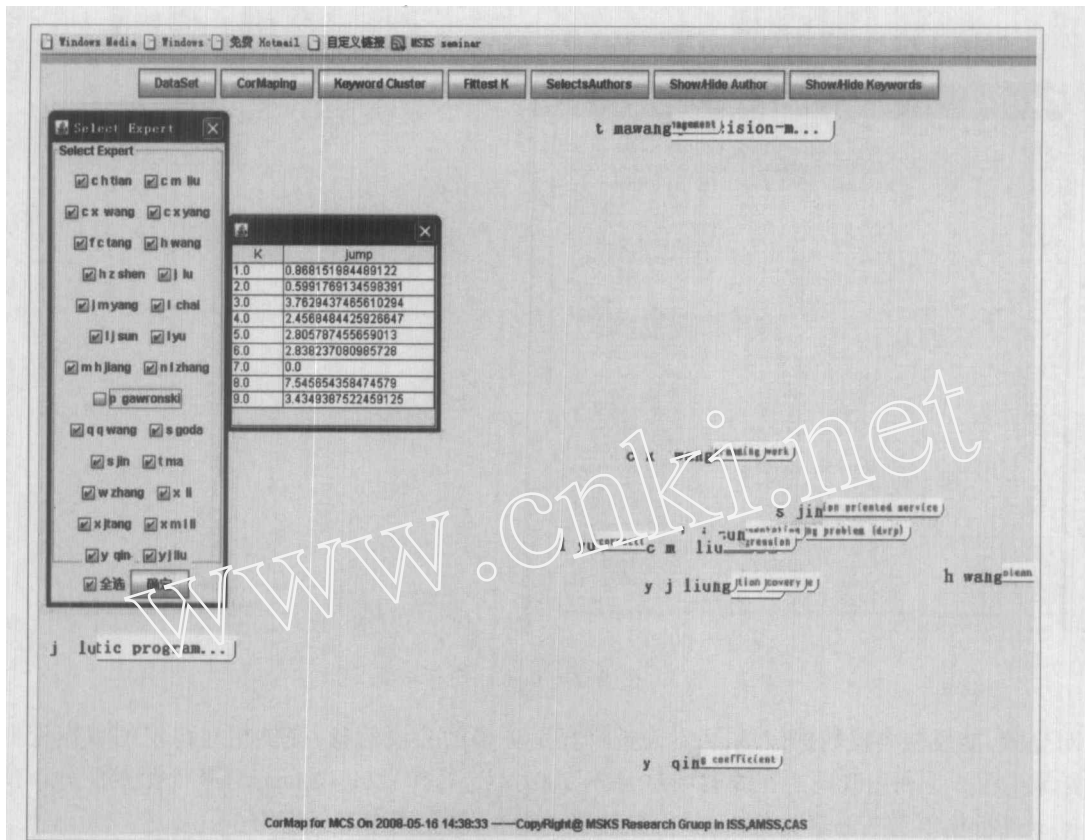


图5 不含作者 P Gamronki 的聚类分析结果

通过图4和图5结果显示,探索性分析结果仅仅是给出一种可能的关联,严格意义下只有通过相应的统计检验才可接受这样的关联,而在会议挖掘中利用探索性分析生成的可视化视图则展现思维的一种状态,能启发人的联想。如通过图4可看出在MCS领域专家关注点在哪,这对于想了解MCS领域的人,可以很快掌握国内外在该领域的前沿及变迁,以及MCS领域全貌。通过图5不一定能获取到该类知识的主要含义,但通过互动的聚类之后,至少可以看到观察人员没有觉察到的一些可能存在的知识关联,启发其思考。

人机交互功能的实现,用户可根据自身体验对会议数据进行挖掘,获取会议数据,自行选择专家和聚类度,对数据挖掘可视化可使用户获取感性认识,继而可以去剖析更多信息,为在物理空间的会议交流提供了有效的信息支持,进而产生共鸣的知识。聚类分析则可帮助观察人员从可视化信息所反映的结构中初步找到一些规律,加深对MCS的感性认识,获取一些理性认识,从而引发隐性知识而展开探索。显示和隐藏作者及关键词对用户获取更结构化的信息提供了方便。

5 总结

改进人机交互功能的在线会场可实现CorMap分析的人机交互和在线动态支持,体现了人机结合、以人为本,从定性到定量的综合集成思想,目的在于加入用户体验,方便用户进行深入挖掘获取更多的信息,以启迪用户的思想,促进与会者之间,研究人员之间的交流和信

息共享. 动态的人机交互过程使虚拟会场更接近于物理会场, 信息共享程度的提高和用户的自身体验将促进用户了解领域知识的概念框架, 为知识的创造提供更人性化的平台, 从而激发更深层次的思想. 改进后在线会场 CorMap 分析只能用于对已有会议数据, 下一步工作是实现 iView 分析的人机交互功能和在线研讨的挖掘分析, 并将在线会场的分析技术应用到 BBS 和博客等日常网络交流中.

参考文献:

- [1] <http://www.iccaworld.com/npps/story.cfm?ID=1575>(2008-5-8)
- [2] Pesenhofer A, Mayer R, Rauber A. Improving Scientific Conferences by enhancing Conference Management System with information mining capabilities. Proceedings of the 1st International Conference on Digital Information Management (ICDIM 2006), December 6-8, 2006, Bangalore, India, pp359-366.
- [3] 张正文, 唐锡晋. 论文分配的支持方法研究. 第九届全国青年管理科学与系统科学学术会议论文集(蓝海林等主编). 广州: 华南理工大学出版社, 2007. pp10-13.
- [4] Matsuo Y, Tomobe H, Hasida K, Ishizuka M. POLYFRONT: An Advanced Social Network Extraction System from the Web. Proc. 15th International World Wide Web Conference (WWW2006), 2006. 5
- [5] Tang Xijin, Zhang Nan, Wang Zheng. Augmented Support for Knowledge Sharing By Academic Conference - Online Conference Ba. in proceedings of IEEE WiCOM'2007 (the Management Track of IEEE International Conference on Wireless Communications, Networking and Mobile Computing, 2007), 2007:6400 - 6403.
- [6] 唐锡晋(编). 综合集成与复杂系统(2005-2006). No. MSKS-2006-03, .北京: 中国科学院数学与系统科学研究院, 2006年10月.
- [7] 刘怡君, 唐锡晋, 李增惠. 对香山科学会议跨学科研讨的一种初步分析. 管理科学与系统科学新进展(刘思锋等主编). 第8届全国青年管理科学与系统科学会议论文集. 河海大学出版社, 2005:35-40.
- [8] 唐锡晋, 刘怡君, 张文. Augmented Analytical Exploitation of a Scientific Forum. In Iwata, S. et al. eds. Communications and Discoveries from Multidisciplinary Data, Studies in Computational Intelligence Vol. 123, Springer-Verlag, 2008 (the first paper in Chapter 3)
- [9] 唐锡晋, 张正文. How knowledge science is studied - A vision from conference mining of the relevant knowledge science symposia. International Journal of Knowledge and Systems Sciences, 2007, 4(4): 51-60.
- [10] Nonaka I, Konno N, Toyama R. Emergence of 'Ba'. In: Nonaka I, Nishiguchi T eds. Knowledge Emergence. New York: Oxford University Press, 2001.
- [11] 唐锡晋, 张楠, 王正. Exploration of TCM masters knowledge mining. Journal of Systems Science and Complexity, 2008, 21(1):34-45.
- [12] 唐锡晋. Towards meta-synthetic support to unstructured problem solving. International Journal of Information Technology & Decision Making, 2007, 6(3):491-508.
- [13] 唐锡晋(编). 综合集成与知识科学(2007-2008). No. MSKS-2008-05, 北京: 中国科学院数学与系统科学研究院, 2008.