

社会多媒体计算

田永鸿 黄铁军 高文
北京大学

关键词：社会多媒体计算 多媒体社会网络

社会网络和在线内容分享服务（如FaceBook、Youtube等）的迅猛发展，特别是近一两年来微博客的盛行，逐渐产生了将多媒体与社会网络相结合的新型媒体形式，我们称为社会多媒体（social multimedia）。它拥有改变人们相互交流、合作和通信方式的巨大潜能，甚至能够影响我们对生活与社

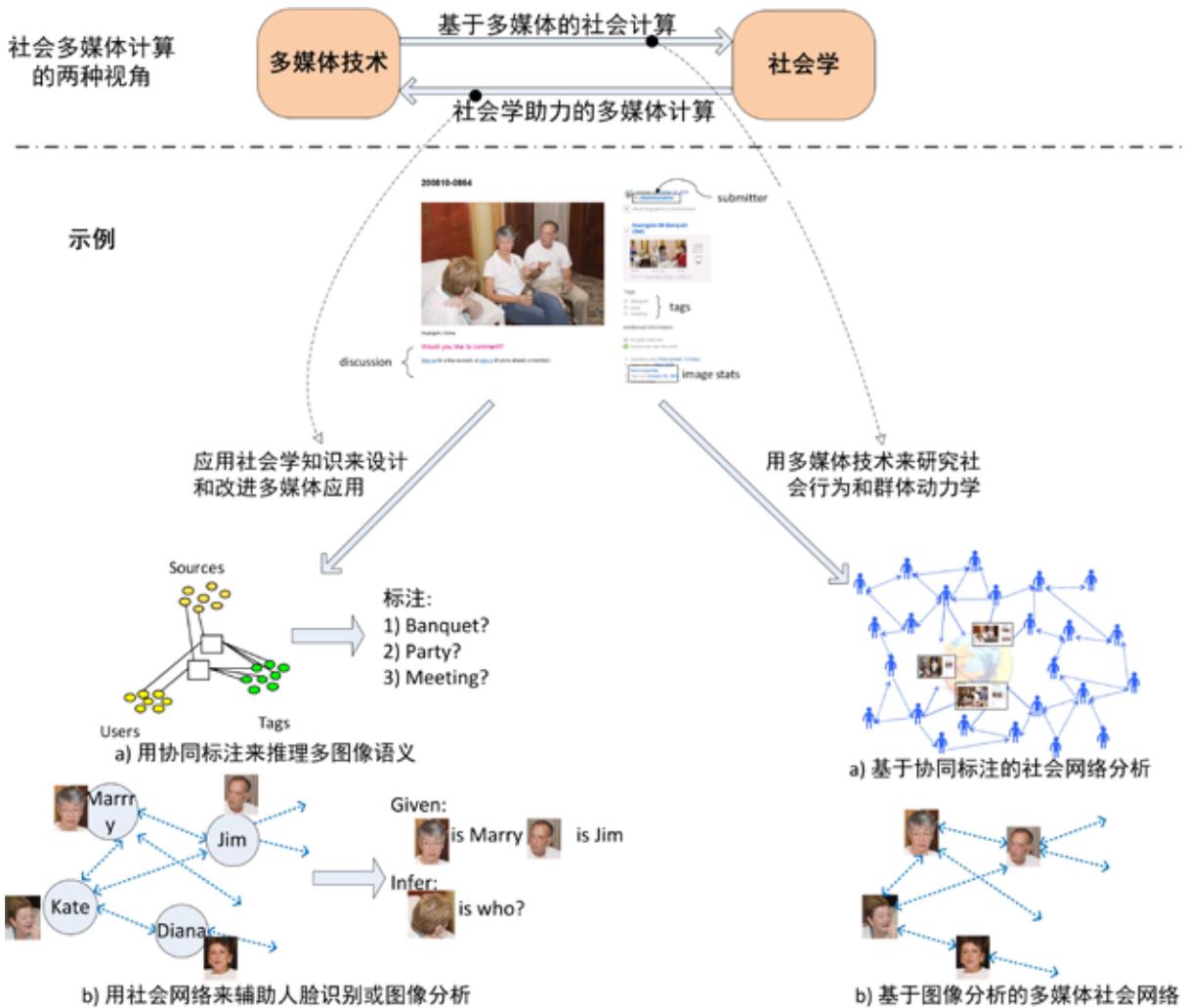


图1 社会多媒体计算的两种视角

会的理解。社会多媒体的这种将信息“爆炸式”地传播和多媒体信息真实视觉刺激结合起来的特性，已在社会政治、经济生活中展示出巨大的威力。

社会多媒体对计算技术本身的挑战和机遇是什么？早在2006年，多媒体国际会议ACM Multimedia就举办了一次专家论坛来讨论“多媒体和Web 2.0：噱头、挑

战和合作”；而在2010年多媒体检索国际会议ACM MIR也专门讨论了“Twitter、Flickr、Youtube等互联网‘噱头’对多媒体信息检索研究的影响”。另一方面，我国学者王飞跃早在2004年就提出了“社会计算”（social computing）的概念和方法^[1]；而美国哈佛大学的大卫·拉泽尔（David Lazer）等人于2009年2月在《科学》（Science）杂志上发表论文提出了“计算社会学”（computational social sciences）这一新兴研究领域^[2]。

然而，多媒体与社会学之间的碰撞与对话，已经不仅仅局限于社会计算或计算社会学的范畴。一方面，以图像、视频等为主要手段或工具的在线社会交互，与基于文本的交互方式相比有了很大不同，其信息传播规律甚至有了根本性的变化；另一方面，由于其固有的挑战性，多媒体分析与理解的研究仍是一个未完全解决的领域难题，而社会多媒体则为解决这一难题提供了新的角度与思路。正基于此，我们用“社会多媒体计算”（social multimedia computing）来指代这一社会学与多媒体技术深度融合后产生的跨学科研究领域。相关的方向性论文于2010年8月作为封面文章发表在IEEE《计算机》（Computer）杂志上^[3]。

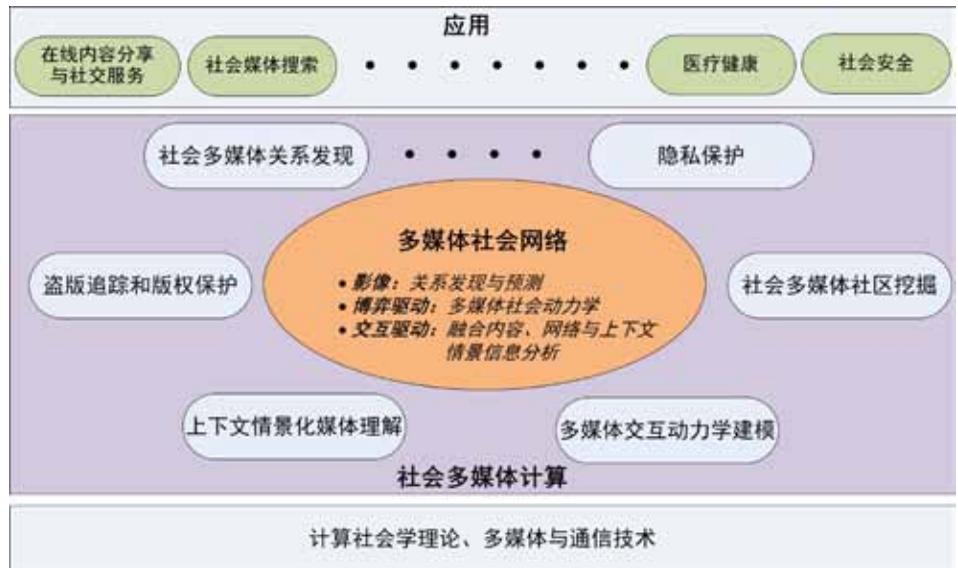


图2 社会多媒体计算研究概览：核心问题、研究领域及应用

本文将引述其中的基本观点，并加入最近的研究成果与观点，以之抛砖引玉，供中国计算机学会的同行评鉴。

什么是社会多媒体计算？

除包含不同媒体内容在时间与空间上的关联与交互之外，社会多媒体还包括了如下两类社会交互：

以多媒体为手段的社会交互 多媒体是人与人之间交往与信息传递的媒介。视频博客和用户协作标注（collaborative tagging）的图片是两种典型的社会多媒体；而纯文本的博客和短信则属于社会媒体，不是社会多媒体。

社会交互的多媒体影像 指现实世界的社会活动以及人际关系在照片、视频等媒体中的真实记录或“影像”。如事件或活动现场的照片、监控视频等。这是传统的社会媒体所不具有的功能。

因此，我们可以将社会多媒体定义为“通过社会交互进行传播或用来记录社会交互的多媒体资源和应用”。从这一定义出发，我们可以从如下两个视角来认识和理解社会多媒体计算：

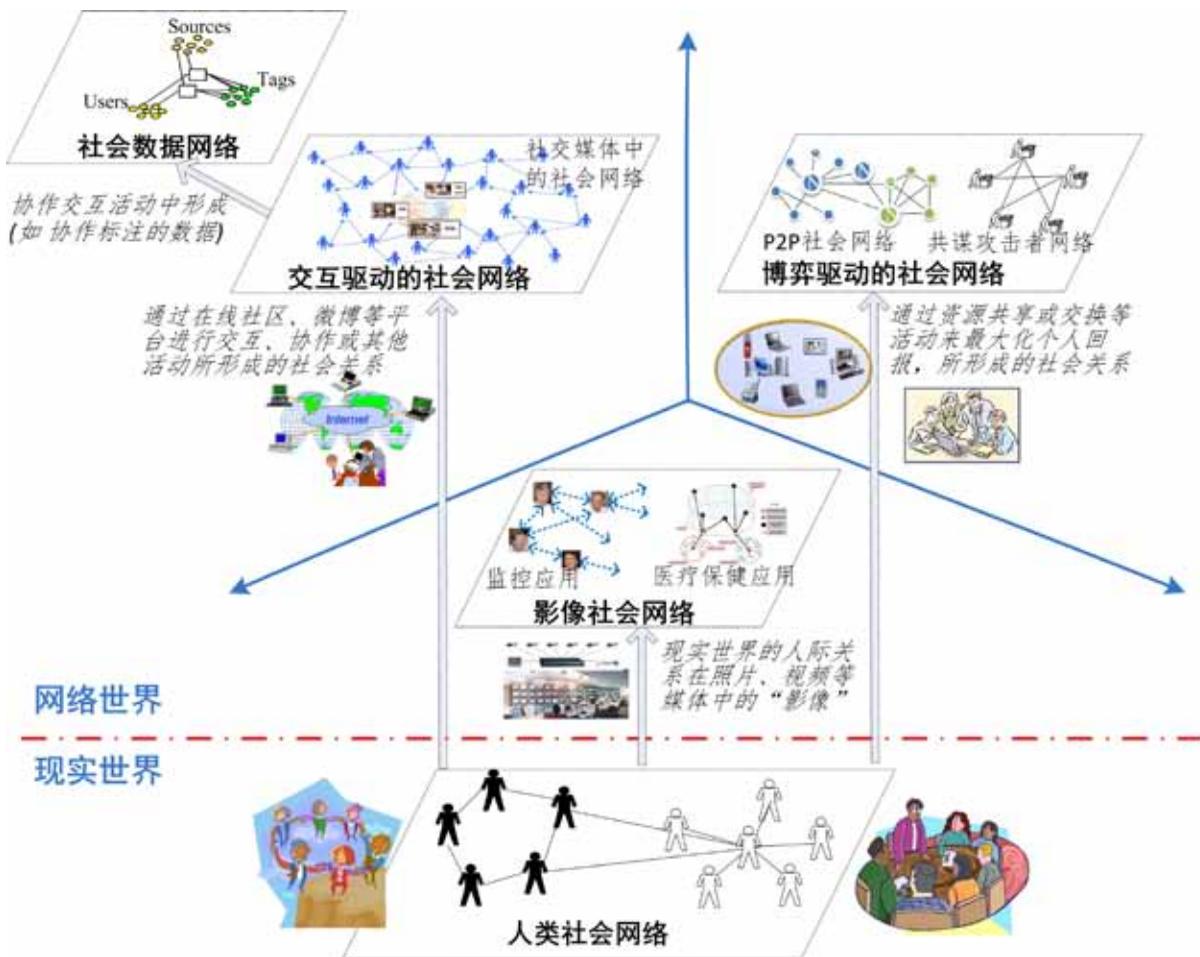


图3 三种多媒体社会网络的示例

以社会学为中心、基于多媒体的社会计算利用多媒体技术和工具来支持更强大的社会交互，来获取大量的用户交互数据和行为数据，来分析社会多媒体中用户行为和群体动力学，以及开展大规模社会学实验。显然，这不仅扩展了社会计算的研究范围，也为计算社会学提供了新的技术手段和工具。如图1的例子，除可以基于协作标注数据来进行社会网络分析外，还可以基于图像分析来建立多媒体社会网络，即通过对合影照片中个体的同现关系进行统计建模并推理他们之间是否存在某种社会关系。

社会学助力的多媒体计算 利用社会学理论来设计或改进现有的多媒体应用，将社会网络分析和社交数据用于媒体内容理解，以及利用群体间的社会动力学模型来提高多媒体通信和内容保护系统的

性能等。如图1所示，我们可以使用协作标注数据来推断照片的语义（如宴会、派对或会谈），甚至利用社会关系网络来辅助人脸识别。

因此，社会多媒体计算是一个多媒体技术和社会学深度融合后产生的跨学科研究与应用领域。图2概括了社会多媒体计算中涉及的核心问题、研究领域及潜在应用等，下面分别一一简单讨论。

多媒体社会网络：从交互驱动到影像刻画

多媒体社会网络（multimedia social networks）是社会多媒体计算的核心。直观地讲，多媒体社会网络是用户在共享和交换多媒体内容和资源的过程

中所形成的社会网络^[5]。进一步分析发现至少有如下3种多媒体社会网络(如图3所示):

交互驱动的社会网络(interaction-driven social networks) 用户间交互、协作或参与在线社区活动所形成的社会网络,是社会多媒体中最常见的社会网络形式。同时,在线社会活动(如协作标注)也会产生社会数据网络(social data networks),它虽然不是用户间的社会网络,但从它出发很容易分析得到一个用户社会网络。

博弈驱动的社会网络(gaming-driven social networks) 在这种社会网络中,用户通过交换和共享资源来期望获得最大化的回报,为此他们需要观察和学习别人的游戏规则并相应地调整自己的策略,因此社会关系变化通常可以建模为一个博弈过程。典型的例子包括P2P社会网络^[4]和共谋攻击者网络(colluder social networks)^[5]。

影像社会网络(imagery social networks) 现实世界中用户间的社会关系与社会活动在照片、监控视频或无线传感数据等中的真实记录或“影像”。影像社会网络提供了一种“虚实结合”的方式来分析现实世界中的人类社会网络。

上述三类社会多媒体网络对社会多媒体计算提出了一些挑战性问题:

多媒体社会动力学建模 建模和分析用户行为及其社会动力学是多媒体社会网络分析中的基础性理论问题。在多媒体社会网络中,用户行为的建模和分析与媒体内容紧密相关,在某些情况下用户行为甚至具有高度动态性(如P2P社会网络中用户行为可能随网络带宽、位置等随时变化),从而使多媒体社会网络分析有别于传统的社会网络分析。

关系的发现和预测 多媒体社会网络中的关系发现和预测必须考虑以多媒体为手段的社会交互和记录在多媒体内容中的社会交互,从而具有很高的复杂性和挑战性。例如,在影像社会网络分析中,理论上我们可以通过提取具有足够描述性和判别性的特征(如人脸、动作或活动线索)来分析和理解人们在某个活动或场景下的社会关系,然而无论是基于网上照片还是监控视频,要利用现有的人脸与动作识别技

术来建模这种关系网络将非常困难。

内容、网络和上下文的融合分析 通常三种社会数据和知识可直接用于社会多媒体计算:(1)用户数据,如用户行为、偏好或兴趣;(2)社会上下文信息,如社会网络、结构和关系;(3)以社会化方式收集的数据,如协作标注数据等。要使用这三类数据,一个可能的技术挑战是传统统计学习方法中的数据独立同分布(independent and identically distributed, I.I.D)假设不再成立,而其正是大多数基于统计学习的多媒体内容分析系统之基础。另一个相关问题则与社会数据的质量有关。以协作标注数据为例,由于协作标注过程与用户的知识、文化背景等紧密相关,因此标注的数据往往存在不一致、二义性等问题。因此,如何提高社会数据的质量是社会多媒体内容理解系统中的重要问题。

创新应用

除在线内容分享和社交服务、社会多媒体搜索外,社会多媒体计算还能在如下领域中广泛应用:

多媒体通信 在P2P网络等多媒体通信系统中,可以对网络中用户的行为用多媒体社会网络来建模,在此基础上使用奖惩、积分等机制奖励优质节点、惩罚违规行为,使节点间更趋向于合作而非竞争关系,从而能有效地提高网络通信系统的效率。最近一项对Skype用户的调查表明,在P2P系统中应用社会网络能够减少投机取巧和NAT问题的出现^[4]。

互动媒体服务和娱乐 互动媒体服务和娱乐是未来几年社会多媒体计算具有广泛应用前景的两个领域,它们都强调运用用户行为模型和社会网络分析来设计交互式多媒体系统。协同智能游戏(collective intelligence gaming)、用户兴趣驱动的自动视频广告投放(user-targeted video advertising)是两个典型的例子^[6]。在协同智能游戏中,用户通过玩游戏的方式来帮助系统实现对多媒体内容的分析与理解;而用户兴趣驱动的自动视频广告投放则通过学习用户的兴趣与偏好,并进行多媒体社会网络的建模与分析,从而为不同用户投放适合其兴趣和购

买意愿的广告节目。这种新广告模式不仅对用户的干扰程度较低，在实际应用中往往能取得更好的广告效果。

医疗保健 医疗健康和家庭保健是社会多媒体计算的一项极具发展前景的创新应用。通过大量的无线传感器和摄像机来监控个人的健康状态，并通过与其社会网络相结合，能及时发现疾病发生并做出相应的应急预警。例如，利用老年人或残疾人携带的心电传感器来捕捉到任何不正常的心率信息，并将该数据与环境上下文信息（如位置信息、现场图片等）发送到包括携带者家庭成员、医生、朋友、急救中心等的社会网络上，有利于让医生或其亲友在关键时刻做出正确的决定^[7]。

安全应用 近年来社会群体性事件的“网上网下联动”发生趋势表明，只有将网络空间的社会媒体信息与物理世界获取的监控视频数据融合起来分析，才能对可疑行为进行较准确的分析与预警。因此，通过研究监控视频数据和社会媒体信息在社会行为和群体交互行为表达上的关联性，建立个体和人际关系的跨数据域映射关系，可以在此基础上构建影像社会网络，为可疑行为和异常事件的事前预警和事后分析提供支持。

未来研究热点

对于这个新兴的跨学科研究领域，我们选择性地讨论一些社会多媒体计算研究未来可能的热点。

上下文情景化媒体理解

以近年来引起广泛关注的社会多媒体协同标注为代表，社会多媒体计算为多媒体内容的分析与理解研究引入了一个全新的视角^[8]，称之为上下文情景化媒体理解（contextualized multimedia understanding）。它包含三层涵义。第一，直接利用社会网络关系和社会数据（如协同标注的标签）来进行多媒体内容的分析与理解；第二，媒体分析与理解应融入用户的社会人文因素（如文化、地域等在图像语义理解上的差异）和上下文情景信息（例如

GPS等位置信息）；第三，可以将多个局部语义理解引擎集成到一个媒体语义理解网络中，其中每个节点能从每个本地用户的操作和可用的局部上下文情景信息中获得知识，同时不同节点间能够相互学习以提升对媒体内容的理解水平。这种网络协同式的多媒体信息检索方法在今后几年有可能会改变多媒体理解和检索的计算模型。

多媒体交互动力学建模

研究开发新的多媒体社会交互工具需要紧密结合社会学、多媒体技术和通信技术，而社会学中关于大规模社会交互动力学的研究则可能为新型社会交互方法和工具的研究注入新的动力。从这个意义上讲，社会学家应在其中发挥更积极的作用。未来的自然用户界面将允许用户与远程虚拟环境进行无缝通信与交互，使得我们真正能够实现“天涯若比邻”。实际上，这一愿景可望通过融合新型传感器（如触觉、嗅觉和动作传感器）、高逼真显示设备和3D技术来实现。可以预见，这种新型的交互模式将会极大地改变我们的生活与交流方式。

盗版追踪及版权保护

多媒体版权保护不是一个新问题，但是随着社会多媒体内容在互联网上的爆炸式增长，它又重新受到广泛关注。在多媒体社会网络中，盗版和共谋攻击者间的动力学分析将有助于研究开发面向社会多媒体的版权保护技术。正是从这一思想出发，文献^[5]提出一个基于多媒体社会网络的多媒体取证框架，其中利用数字指纹来追踪非法拷贝的元凶。然而，由于互联网上大部分社会多媒体内容很少携带水印或者数字指纹，因此这一框架的应用范围仍然受限。因此，需要进一步研究多媒体社会网络中的用户行为建模，并与基于媒体指纹（mediaprinting）的副本检测等数字版权管理技术相结合^[9]，探索适合社会多媒体的版权保护技术。

社会媒体社区分析

社会多媒体为群体事件分析等问题的解决提供

了一种新的思路。面对大规模的社会多媒体数据，研究的重点将集中在社会学理论的大规模验证，以及分析这类数据的新理论和推理方法等方面。如果考虑视频博客或微博中内容、社会和时间三个维度间的相互作用关系，社会行为和群体动力学分析问题（如结构和演进模型、信息扩散模型等）就会更为复杂^[6]。无论在联网设备方面从固定PC发展到可移动的手机，还是在媒体形式方面从纯文字发展为文字、图像、视频等多种媒体混合，客观上都加快了信息传播的速度，从而有效缩短网络中人际关系的距离，因此“六度分离”等社会学原则在微博社区是否仍然成立还需要大规模的实验验证。

隐私保护

无论是对工业界还是学术界，妥善处理好隐私保护问题都很有必要。然而，与一般社会媒体相比，社会多媒体数据的隐私保护问题更加复杂。例如，很多用户通常热衷于到Facebook等社交网站分享照片，但是这些照片可以轻易地被用于人脸识别和图像检索之类的额外用途，特别是在照片中标注了诸如姓名、电子邮箱和实际地址等信息的情况下。这也是近年来“人肉搜索”问题令普通公众、学术界和政府部门一直很头疼的部分原因。因此，研究社会多媒体中的隐私保护问题以及相应技术手段具有很强的科学与现实意义。

结语

虽然目前社会多媒体计算的研究仅仅处于起步阶段，但是其前景光明。相信通过社会学家、计算科学家及多媒体领域研究人员的共同努力，这一领域的研究将很快就能取得丰硕的成果，同时也将为社会学和多媒体领域的研究注入新的活力。■



田永鸿

北京大学信息科学技术学院副教授。主要研究方向为机器学习、图像和视频信息处理、社会多媒体计算。
yhtian@pku.edu.cn



黄铁军

北京大学信息科学技术学院教授，数字视频编解码国家工程实验室副主任。主要研究方向为图像分析与理解、视频编码和媒体识别方法。
tjhuang@pku.edu.cn



高文

CCF会士，2010年CCF王选奖获得者。中国工程院院士。北京大学信息科学技术学院教授，数字视频编解码国家工程实验室副主任。主要研究方向为计算机视觉、视频编解码、数字媒体技术等。
wgao@pku.edu.cn

参考文献

- [1] F.-Y. Wang. Social Computing: Concepts, Contents, and Methods, *International Journal of Intelligent Control and Systems*, 9(2), 2004, 25~35
- [2] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy and M. Van Alstyne. *Computational Social Science*, *Science*, vol. 323, 6 Feb 2009, 721~723
- [3] Y. H. Tian, J. Srivastava, T.J. Huang, and N. Contractor. *Social Multimedia Computing*, *Computer*, 43(8), Aug. 2010, 27~36
- [4] J. Altmann, Z. B. Bedane. A P2P File Sharing Network Topology Formation Algorithm Based on Social Network Information, *Proc. IEEE Int' l NetSciCom Workshop*, 2009
- [5] H. V. Zhao, W. Lin and K. J. R. Liu. Behavior Modeling and Forensics for Multimedia Social Networks: A Case Study in Multimedia Fingerprinting, *IEEE Signal Processing Magazine*, 26(1), Jan. 2009, 118~139

更多参考文献：www.ccf.org.cn/cccf