

# 图灵奖得主CCF论文发表情况分析\*

韩京洋 陆嘉恒 杜小勇  
中国人民大学

关键词：ACM 图灵奖 CCF 论文 研究领域

中国计算机学会 (CCF) 在 2010 年公布了“CCF 推荐国际学术会议和期刊目录”。此目录包含了计算机领域重要的学术会议和期刊，具有一定的代表性，一经公布，就引起了巨大反响。特别是被教育部第三次学科评估采用之后，逐渐为国内高校所接受，在国际上也产生了较大影响力。这对我国计算机学科的发展起到了积极的推动作用。本文对 ACM 图灵奖得主在 CCF 推荐的国际学术会议和期刊目录上发表论文的情况进行了统计分析，希望能为我国计算机学科的评价（包括评奖和优秀人才选拔）提供参考。

CCF 推荐的国际学术会议和期刊目录涵盖了计算机领域内的 10 个研究方向：计算机系统与高性能计算、计算机网络、网络与信息安全、软件工程 / 系统软件 / 程序设计语言、数据库 / 数据挖掘 / 内容检索、计算机科学理论、计算机图形学与多媒体、人工智能与模式识别、人机交互与普适计算、前沿 / 交叉 / 综合。刊物和会议分为 A, B, C 三类。A 类指国际上极少数的顶级期刊和会议；B 类指国际上著名的期刊和会议；C 类指国际学术界所认可的重要期刊和会议。下文将在 CCF 推荐的 A, B, C 三类期刊和会议上发表的论文统称为 CCF 论文。

## 图灵奖得主CCF论文发表情况

ACM 图灵奖从 1966 年设立，截至 2014 年共

49 届，有 62 名科学家获此殊荣。我们对这些图灵奖得主发表的 CCF 论文总数以及 CCF A 类论文总数进行了统计分析，详见附录。统计分析结果显示：图灵奖得主平均每人发表 81 篇论文，24 篇 CCF A 类论文，48 篇 CCF 论文。CCF A 类论文约占发表论文总数的 30%，占 CCF 论文的 50% 左右。这充分说明图灵奖得主的论文质量较高。

表1 t-检验结果中的双尾P值<sup>1</sup>

	CCF论文数>0	CCF论文数>10	CCF论文数>20	CCF论文数>30	CCF论文数>40
双尾P值	5.181E-151	2.16E-12	0.089	0.068	5.34572E-06

对图灵奖得主发表 CCF 论文情况进行分析，可以发现，不同图灵奖得主发表的 CCF 论文数目相差较大，这与“普遍认为图灵奖得主发表的论文一定很多”的想法不吻合。我们采用 t-检验双样本等方差假设的方法分析图灵奖得主与其发表的 CCF 论文数目是否具有显著的相关性。设定检验水准  $\alpha=0.05$ ，“样本 1”为 62 位图灵奖得主发表的 CCF 论文数；“样本 2”为 5 组数据，按照 CCF 论文数大于 0、大于 10、大于 20、大于 30、大于 40 的分类，随机抽取 1500 位非图灵奖得主发表的 CCF 论文数。当  $P<0.05$  时，证明样本 1 和样本 2 有显著差异；当  $P>=0.05$  时，样本 1 和样本 2 无显著差异。t-

\* 这里的 CCF 论文指的是，在 CCF 推荐的国际学术会议和期刊上发表的论文。

<sup>1</sup> P 值(P value)是指，当原假设为真时所得到的样本观察结果或更极端结果出现的概率。

检验的结果见表1。

根据t-检验结果(图1),当样本2中CCF论文数小于20时, $P < 0.05$ ,证明样本1和样本2有显著差异。

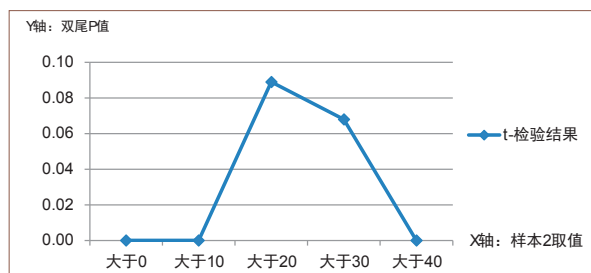


图1 t-检验结果

著差异,即图灵奖得主与其发表的CCF论文数具有显著的相关性;当样本2中CCF论文数大于20并且小于40时, $P > 0.05$ ,证明样本1和样本2无显著差异,即图灵奖得主与其发表的CCF论文数无显著的相关性。有趣的是,当样本2中CCF论文数大于40时, $P$ 又一次明显小于0.05,证明样本1和样本2又有了显著差异。通过以上结论我们发现,学者的论文数太多或者太少都与图灵奖得主发表的论文数有显著差异,图灵奖得主的CCF论文数实际上既不是很少,也不是特别多。这个实验结果说明了,不能单凭CCF论文数目来决定一位计算机学者的成果质量和影响力。

我们还统计了全世界CCF论文总数排名前100位的学者,发现图灵奖得主仅占6位。其中CCF论文总数最多的图灵奖得主是排在第32位的罗伯特·塔扬(Robert Tarjan),共发表了199篇CCF论文。

所有的图灵奖得主都为计算机科学作出了巨大贡献,但他们发表论文的情况却大不相同。我们对CCF论文总数以及A类论文总数前三位和后三位的图灵奖得主进行了深入的分析。表2~5是有关数据。

图灵奖得主发表的CCF论文数目差异较大。我们通过列举几位图灵奖得主的研究工作来具体分析造成这种差异的原因。

### 1. 发表CCF论文数目较多的图灵奖得主

罗伯特·塔扬以在数据结构和图论上的开创性

表2 CCF论文总数前三位图灵奖得主

图灵奖得主	获奖年份	CCF论文总数
Robert Tarjan	1986	199
Amir Pnueli	1996	174
Edmund M. Clarke	2007	169

表3 CCF论文总数后三位图灵奖得主

图灵奖得主	获奖年份	CCF论文总数
Richard Hamming	1968	3
Fernando J. Corbató	1990	3
Kristen Nygaard	1977	3

表4 CCF A类论文总数前三位图灵奖得主

图灵奖得主	获奖年份	A类论文总数
Robert Tarjan	1986	112
Michael Stonebraker	2014	112
Adi Shamir	2002	83

表5 CCF A类论文总数后三位图灵奖得主

图灵奖得主	获奖年份	A类论文总数
Douglas Engelbart	1997	0
Ole-Johan Dahl	2001	0
Robert E. Kahn	2004	0

工作而闻名,他是许多图论算法的发明者。他将数学理论通过计算机科学的方式来实现,提出了许多广泛有效的算法和数据结构,发表了数篇高水平的理论文章。爱德蒙·克拉克(Edmund M. Clarke)的主要贡献是首次提出了模型检查的方法,并应用在自动机并发系统的验证研究上,成为形式逻辑研究方面模型检查的开创者,对模型检查技术应用于硬件和软件工业中的算法验证技术作出了奠基性的贡献。阿米尔·伯努利(Amir Pnueli)的主要贡献是在计算机科学中引入时序逻辑所进行的开创性的研究工作,以及在编程语言和系统验证方面取得的进展,发表了多篇理论研究文章。

以上3位科学家均为数学专业出身,后转入计算机领域进行研究工作,发表了很多高质量的论文。

### 2. 发表CCF论文数目较少的图灵奖得主

道格拉斯·恩格尔巴特 (Douglas Engelbart) 是计算机鼠标的发明者, 他开发了超文本系统、网络计算机以及图形用户界面, 对人机交互领域作出了重要的贡献。费尔南多·科尔巴 (Fernando J. Corbató) 领导开发了通用大规模分时系统、资源共享计算机系统——CTSS 和 Multics。CTSS 系统的开发对 UNIX 系统的诞生具有重要意义。奥利·约翰·达尔 (Ole-Johan Dahl) 和克利斯登·奈加特 (Kristen Nygaard) 设计了编程语言 SIMULA 1 和 SIMULA 67。SIMULA 67 被认为是最早的面向对象程序设计语言, 它引入了所有后来面向对象程序设计语言遵循的基础概念。

以上 4 位科学家发表的论文数目虽然不多, 但是论文质量很高, 而且他们开发的系统或者编程语言成果显著、贡献巨大。

由此得出结论, 研究领域为科学理论的图灵奖得主发表的 CCF 论文数目相对较多, 而研究领域为系统开发和编程语言的图灵奖得主发表的 CCF 论文数目相对较少。但毫无疑问, 他们均为计算机领域作出了杰出的贡献。

## 图灵奖得主的研究领域介绍

表6 图灵奖得主研究领域分布情况

研究领域	人数
Software engineering, Systems software and Programming language	21
Computer theory	21
Artificial intelligence and pattern recognition	7
Database, Data mining and information retrieval	4
Computer network	2
Interdisciplinary	2
Network and information security	2
Computer graphics and multimedia	1
Human-computer interaction and pervasive computing	1
High-performance computing	1
总计	62

我们根据学者已发表的 CCF 论文所属的研究领域, 分析出每一位学者的研究领域。如果某位学者的论文发表在多个领域, 则仅统计发表论文最多的一个领域。62 位图灵奖得主的研究领域分类统计情况见表 6。

从表 6 中可以看出, 软件工程 / 系统软件 / 编程语言、计算机理论研究领域人数最多, 排在前两位; 人工智能与模式识别排在第三位; 数据库、计算机网络、信息安全等领域排在之后。这说明图灵奖得主的研究领域侧重于软件与计算机科学理论, 而计算机系统和应用技术类所占的比例很小。

## 图灵奖得主取得巨大成果时的年龄统计

毫无疑问, 图灵奖得主的获奖原因也就是他们取得的最大成就。所以, 我们对 62 位图灵奖得主作出巨大贡献时的年龄进行了统计 (见表 7)。

表7 图灵奖得主作出巨大贡献时的年龄统计

年龄段	20~29岁	30~39岁	40~49岁	50岁以上	合计
人数	9	41	12	0	62
百分比	15%	66%	19%	0	100%

从表 7 可以看出, 图灵奖得主作出巨大贡献时的年龄段集中在 30~39 岁, 这说明一个人成就事业的黄金阶段是中青年时期。在图灵奖得主中, 有超过 81% 的人在 40 岁以前到达事业的顶峰, 30~39 岁是科学创造的最佳年龄段。图灵奖得主的大部分研究成果是在著名大学和研究机构工作时取得的, 这也反映出好的研究条件和体制环境对科技创新具有积极影响。

## CCF十大研究领域的从业人员统计分析

我们对全世界 144 万名计算机学者的研究领域进行了分析。

表8 CCF十大研究领域研究人员比例

研究领域	比例
Artificial intelligence and pattern recognition	23.05%
High-performance computing	17.12%
Interdisciplinary	13.44%
Computer graphics and multimedia	9.85%
Database, Data mining and information retrieval	8.63%
Computer network	8.26%
Software engineering, Systems software and programming language	6.98%
Computer theory	5.23%
Human-computer interaction and pervasive computing	4.52%
Network and information security	2.92%
合计	100%

由表8可以看出,当今研究人员数目最多的领域是人工智能与模式识别,其次是高性能计算与前沿/交叉/综合。此统计结果和图灵奖得主的领域分布不同,这说明,未来很有可能在这两个领域产生更多的图灵奖得主。

## 总结

本文通过对62位图灵奖得主的CCF论文进行分类统计,得到以下结论:

第一,发表论文的数目不能简单作为评价一个学者是否杰出的评价指标。大多数图灵奖得主的CCF论文数目,既不是很多也不是很少。学者论文数太多或者太少都与图灵奖得主的论文数有统计上的显著差异。论文只是评价体系中的一个度量指标,有些成就和贡献并不能通过论文数目度量,比如鼠标的发明等。

第二,不同的研究领域,图灵奖得主的CCF论文数目不同。计算机理论领域发表的论文数目较多,计算机应用技术领域发表的论文数目较少。不同的计算机研究领域评定人才的标准应具有差异性。

第三,过去40多年,图灵奖得主多为在计算机科学理论和软件语言方面作出贡献的科学家,但这

并不意味着今后的图灵奖得主也来自这些研究领域。人工智能与模式识别、高性能计算、前沿/交叉/综合已经成为当前计算机科学领域研究的热点。相信不久的将来,这些领域会诞生更多的图灵奖得主。

回顾62位图灵奖得主获得的成就,就是对计算机科学技术发展史的总结。通过对这些图灵奖得主发表的CCF论文的研究分析,对我国评奖和优秀人才选拔具有参考价值。■

## 致谢:

以上分析数据,均来自中国人民大学自主研发的Panda Search学术搜索平台。该搜索平台目前共收录论文2658485篇,全世界计算机学者1439786位,并且为每位计算机学者建立了CCF论文主页,统计论文详细信息。用户可以根据需求快速查找学者发表的CCF推荐A,B,C三个类别的论文信息。在线访问的网址为<http://pandasearch.ruc.edu.cn/>。本文的研究获国家863高科技计划资助(课题编号:2012AA011001),特此致谢!



韩京洋

中国人民大学硕士研究生。主要研究方向为元数据管理与集成技术。  
jingyang\_han@163.com



陆嘉恒

CCF高级会员。中国人民大学教授。主要研究方向为云数据管理、XML数据管理和空间数据管理。  
jiahenglu@ruc.edu.cn



杜小勇

CCF常务理事、会士、CCF专委工委主任、本刊编委。中国人民大学教授。主要研究方向为数据管理、智能信息检索。  
duyong@ruc.edu.cn

附录:

图灵奖得主	获奖年份	CCF A类论文	CCF 论文总数
Michael Stonebraker	2014	112	123
Leslie Lamport	2013	32	73
Shafi Goldwasser	2012	77	98
Silvio Micali	2012	83	107
Judea Pearl	2011	72	137
Leslie G. Valiant	2010	48	86
Charles P. Thacker	2009	5	5
Barbara Liskov	2008	47	82
Joseph Sifakis	2007	14	73
E. Allen Emerson	2007	28	76
Edmund M. Clarke	2007	37	169
Frances E. Allen	2006	2	6
Peter Naur	2005	1	4
Robert E. Kahn	2004	0	15
Vinton G. Cerf	2004	1	5
Alan Kay	2003	3	5
Leonard M. Adleman	2002	31	36
Adi Shamir	2002	83	140
Ronald L. Rivest,	2002	62	112
Kristen Nygaard	2001	1	3
Ole-Johan Dahl	2001	0	9
姚期智(Andrew Chi-Chih Yao)	2000	74	109
Frederick P. Brooks	1999	15	61
James Gray	1998	38	48
Douglas Engelbart	1997	0	16
Amir Pnueli	1996	56	174
Manuel Blum	1995	35	51
Raj Reddy	1994	13	15
Edward Feigenbaum	1994	10	11
Richard E. Stearns	1993	20	66
Juris Hartmanis	1993	22	60
Butler W. Lampson	1992	19	29

图灵奖得主	获奖年份	CCF A类论文	CCF 论文总数
Robin Milner	1991	16	50
Fernando J. Corbató	1990	3	3
William Kahan	1989	4	4
Ivan Sutherland	1988	3	6
John Cocke	1987	2	6
Robert Tarjan	1986	112	199
John Hopcroft	1986	30	59
Richard M. Karp	1985	60	138
Niklaus Wirth	1984	7	20
Dennis Ritchie	1983	1	6
Ken Thompson	1983	2	12
Stephen A. Cook	1982	51	82
E. F. Codd	1981	3	23
C. Antony R. Hoare	1980	7	61
Kenneth E. Iverson	1979	1	6
Robert W. Floyd	1978	10	15
John Backus	1977	2	10
Dana S. Scott	1976	5	18
Michael O. Rabin	1976	20	39
Herbert A. Simon	1975	22	40
Allen Newell	1975	25	34
Donald E. Knuth	1974	9	37
Charles W. Bachman	1973	2	18
Edsger W. Dijkstra	1972	1	19
John McCarthy	1971	17	26
James H. Wilkinson	1970	3	4
Marvin Minsky	1969	9	12
Richard Hamming	1968	3	3
Maurice Wilkes	1967	8	13
Alan J. Perlis	1966	4	6
平均		24	48