

厦门大学计算机科学系研究生辅导课

论文是怎样炼成的 (2011新版)

林子雨

厦门大学计算机科学系

E-mail: ziyulin@xmu.edu.cn ▶▶





大纲

□ 专题一：谈谈做研究

□ 专题二：学术规范

□ 专题三：基本常识

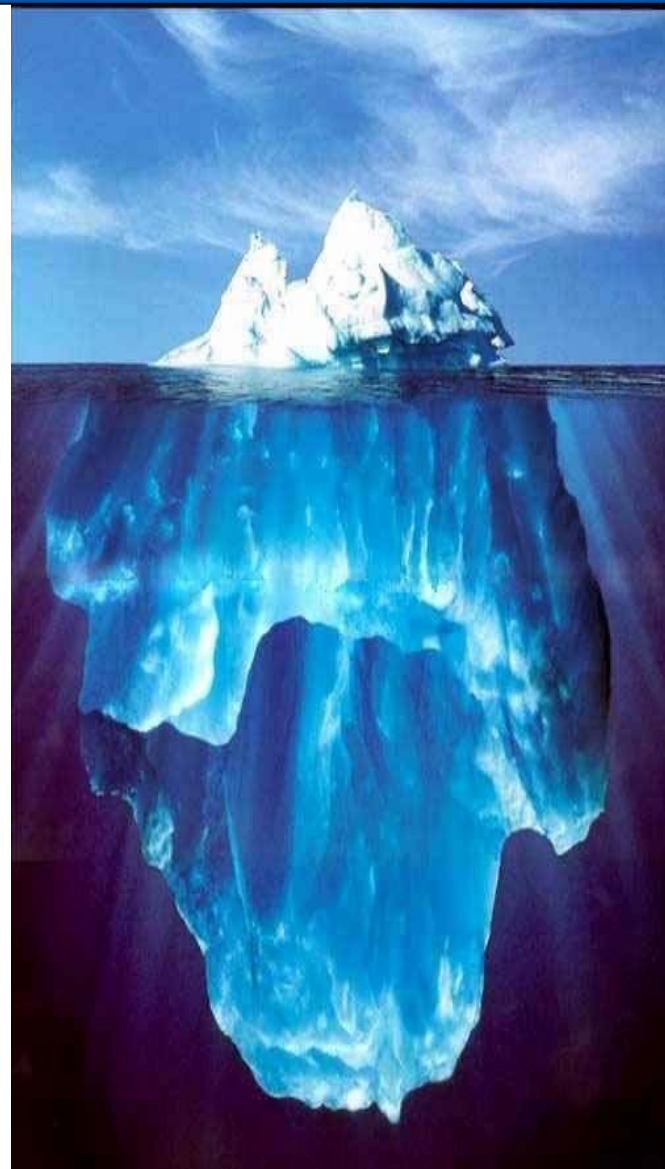
□ 专题四：论文发表

□ 专题五：论文写作

□ 专题六：硕士论文

□ 附件

说明：本PPT部分内容来自周志华老师的PPT
《如何做研究，如何写论文》





专题一：谈谈做研究

- 1.1 为什么要做研究
- 1.2 为什么要写论文
- 1.3 论文好写吗
- 1.4 如何做研究

说明：本部分内容来自周志华老师的PPT《如何做研究，如何写论文》





1.1为什么要做研究

- 研究 \neq 研发
 - 研究：目的是发现新知识、发明新技术
 - 研发：基于已有的知识和技术进行研制、开发
- 科学研究扩展人类的知识
- 没有科学研究就没有技术进步





1.2为什么要写论文

- 把你的工作告诉同行
 - 经过同行评审（**peer-review**），成为科学文献
- 基础研究的主要成果
 - 基础研究通常离实际应用有较大距离
 - 只有很少的研究工作能很快进入实际应用

前沿研究-> 实验室成熟技术/工业界新技术-> 工业界成熟技术
- 评比和考核的重要依据





1.3论文好写吗

- 很容易!
 - 只要有了好的研究工作，写论文不过就是用文字把你的工作描述出来
- 很难!
 - 如果没有研究工作支撑的话
 - 论文是“做”出来的，而不是“写”出来的，“写”的时间其实最多只占10%





1.4如何做研究

- 研究活动的大致过程-TPIC:
Topic -> Problem -> Idea -> Concrete work
(theoretical analysis, experiments, etc.)
-> Paper writing -> Submit





1.4如何做研究(Topic)

Topic

- 从Topic开始，计算机科学发展到今天，已经是一个非常广袤的学科
- 先要进入一个具体的分支学科和领域，并获得必要的了解
- 即使在一个分支学科和领域中，也有太多的话题





1.4如何做研究(Topic)

How ?

- 通常情况：导师给你一个topic
 - 导师往往是该领域的资深学者，对topic可能有较好的把握能力
 - 研究领域不存在“好”、“坏”之分，只要做得足够深入，都能做出好的工作
 - 在特定的时期，某些领域可能更活跃，相对来说杰出成果出得更多





1.4如何做研究(Topic)

How ? (续)

- 然后，阅读关于该话题的重要文献，了解该话题的研究历程、研究现状
 - 请导师或该领域资深学者推荐读物
 - 基于导师推荐的读物，顺藤摸瓜（例如从参考文献）找到尽可能多的重要文献
 - 读不懂的先跳过去，多读几遍
- 每个**topic**发展的历程都是无数聪明人智慧和汗水铺就的，这里面会有一条前后衔接的线索
- 某一天当你突然把头脑中杂乱的东西串成一线，你会觉得顿时豁然开朗：“啊，原来是这样！”，这时你就开始尝到研究带来的乐趣了





1.4如何做研究(Topic)

How ? (续)

如果因为种种原因，你需要自己去找topic、自己找东西读，那该怎么办？

办法1：找人请教（自己要加以判断）

- 师兄、师姐
- 该领域的著名学者
 - 一定要有礼貌
 - 没有回音也不要难过
 - 名人可能每天会收到几十封类似你这样的信，不可能都回复
 - 名人自己可能有一群学生嗷嗷待哺
- 网络： **BBS, mailing list, etc.**





1.4如何做研究(Topic)

How ? (续)

办法2：自力更生

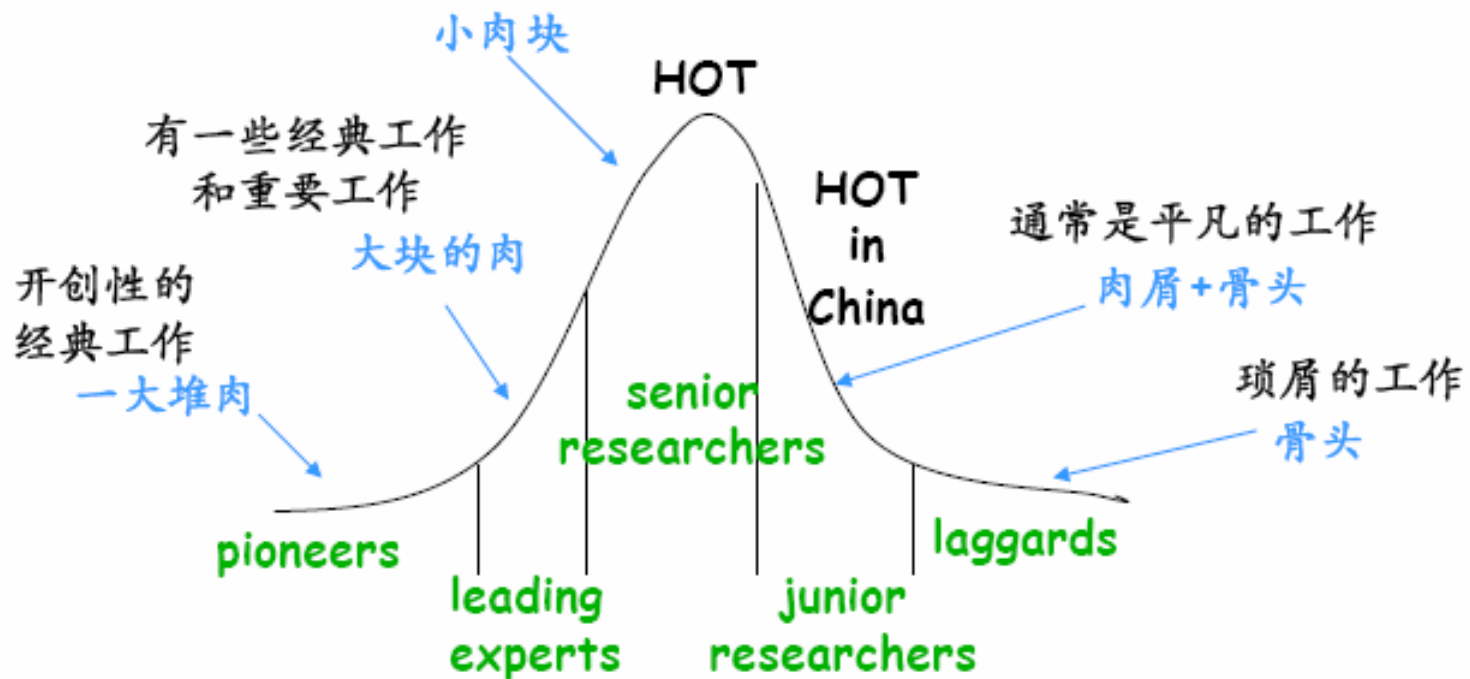
- ①搞清楚自己的领域里最重要的刊物、会议
 - 问人、网上搜索、**BBS, etc.**
- ②找来那上面最近几年的文章
 - **NJU**数字图书馆、系图书馆、**Internet, etc.**
- ③读!
 - 目的是为了大概知道有哪些**topic**，读摘要就可以了
 - 可能要花很多时间





1.4如何做研究(Topic)

How ? (续) 还有一些重要工作



注意:

- 这只是目前**CS**中的一般情况，国内往往滞后一段时间
- 科学发展往往是螺旋式上升的：“冷”了很久的一個**topic**，可能由于新的经典工作而重新“热”起来





1.4如何做研究(Topic)

How ? (续)

选择最适合你的**topic**:

- ①自己的兴趣
 - 这是最重要的!
- ②自己的知识结构
 - 没有必要的知识积累, 一切从头开始的话,
- ③能否获得必要的资源
 - 例如数据





1.4如何做研究(Problem)

Problem

- “问题”是科学研究的心脏！
 - 任何有价值的研究，都是为了解决某个问题
 - 提出一个好的问题，已经成功了一半
- “问题”其实才是研究的真正开始
 - 这可能是**CS**研究中最困难的部分
 - 会找问题，是具有独立研究能力的标志





1.4如何做研究(Problem)

How ?

- 导师给你一个问题
 - 你很幸运!
 - 赶紧做, 说不定导师忍不住自己去做了
- 自己产生问题, 导师帮你判断
 - 这是通常的情况
 - 导师的研究经验和见解, 会使得你少做无用功
- 有了问题以后:
 - 搞清楚该问题上已经有过的所有工作
 - 导师给你推荐重要文献, 然后自己顺藤摸瓜
 - 优点是什么? 缺点是什么? 为什么没有彻底解决问题?
 - 悟出已有工作的发展线索





1.4如何做研究(Problem)

How ?

- 完全自力更生
 - 读发表在重要刊物和会议上的有关你的topic的文献
 - 2、30篇读下来，你大概能知道有哪些问题是没解决的了
 - 关注这个topic上活跃的leading expert的工作、他们的文章中可能会指出一些需要解决的重要问题
 - 兴趣
 - 有价值的问题
 - 知识结构
 - 资源
 - 宜“小题大做”，忌“大题小做”

对研究能力极大的锻炼

你可能会走很多弯路，有可能一无所成，但也有可能练就一身硬功夫





1.4如何做研究(Idea)

Idea

- 这是考验你聪明才智的时候了！
- 不要指望导师给你**idea**！
 - 如果导师连**idea**都给你，你的运气实在是好到家了
- 你也许会有很多**idea**，导师可以帮助你判断、改善有了好的**idea**，问题就解决了一大半
- 如果一下想不到好的**idea**，不要着急





1.4如何做研究(Idea)

How ?

没人帮你判断idea怎么办？

- 自力更生：
 - 是新的idea吗？
 - 是不是有道理？
 - 不是瞎蒙出来的
 - 是否可行？
 - 把这个idea先放到一边，过一周再考虑一下：
 - 是否有更好的办法？
 - 如果答案都是“yes”，那就赶紧动手吧！





1.4如何做研究(Concrete work)

Concrete work

- **Idea**需要得到支持
 - **CS**里面通常是理论分析和实验验证
- 理论分析往往需要较好的数学功底
 - 没有怎么办: **1.学! 2.找人合作**
- 实验验证需要较好的实验设计能力
 - 没有怎么办: **1.学! 2.找人合作**
- 两者都不容易





1.4如何做研究(Concrete work)

How ?

- 理论分析
 - 周全、不要有漏洞
 - 尽可能简单的工具
- 实验验证
 - 实验方案周全仔细
 - 基准测试
 - 其他学者也能使用的数据
 - 不可缺少的比较
 - 实验是可重复的
- 必要的分析和解释
 - Strength/weakness
 - How/When strong?
 - Why strong?
 - How/When weak?
 - Why weak?





1.4如何做研究(小结)

- 小结一下
 - Topic -> Problem -> Idea -> Concrete work
 - 对一项具体的研究工作来说:
 - 找到好的问题是非常重要的
 - 有了好的idea, 问题就解决了一大半
 - 具体的工作是必不可少的





大纲

□ 谈谈做研究

□ 学术规范

□ 基本常识

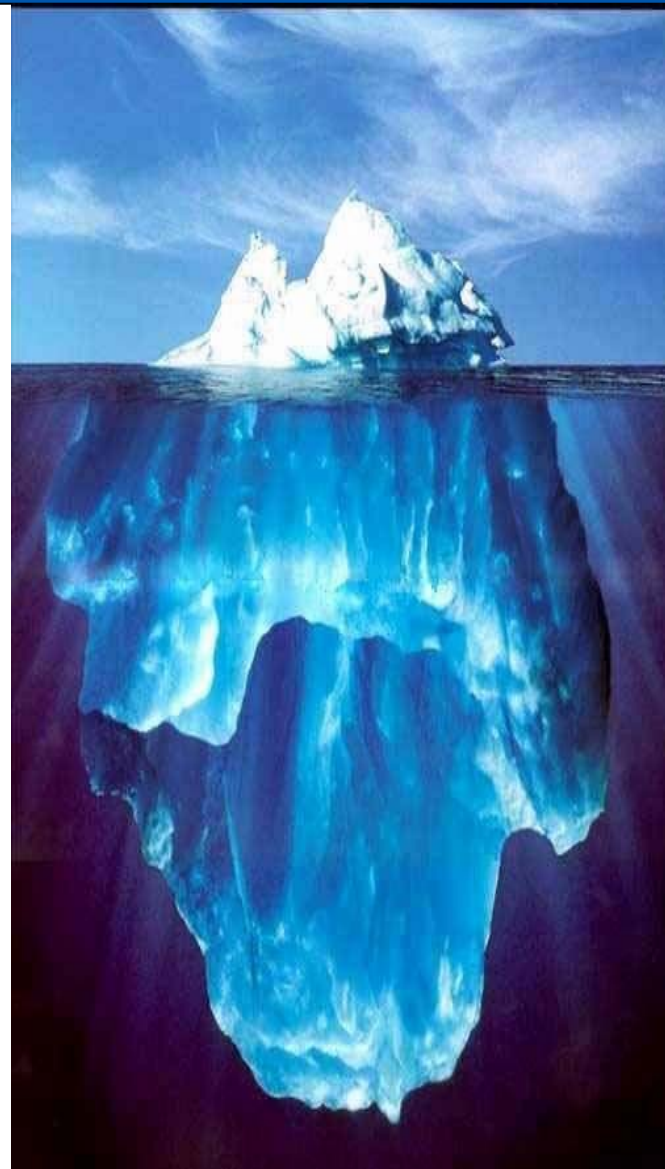
□ 论文发表

□ 论文写作

□ 硕士论文

□ 附件

说明：本PPT部分内容来自周志华老师的PPT
《如何做研究，如何写论文》





专题二 学术规范

2.1 惨痛教训

2.2 学术论文借鉴与剽窃的区别

2.3 学术论文剽窃的主要表现形式

2.4 学术论文剽窃的责任承担

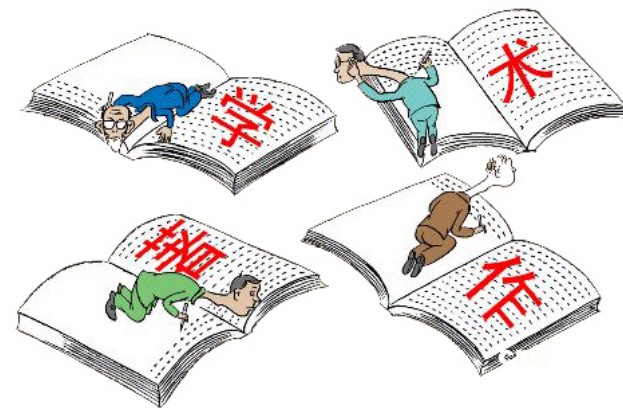
本专题主要参考文献[MaL05]





2.1 惨痛教训

- 2009年2月，浙江大学副教授贺海波因为学术造假而被浙江大学开除
- 2009年3月，因毕业论文涉嫌抄袭，被原作者和校方发现后，武汉某高校研究生不堪压力跳湖自杀
- 2009年4月，研究生论文抄袭惹祸，上海大学陈湛匀教授承认自己把关不严
- 2009年6月，辽宁大学副校长陆杰荣与其学生杨伦共同署名发表的一篇论文被指抄袭
- 2009年6月，广州中医药大学校长被指博士学位论文抄袭
- 2009年8月，武汉理工大学校长、中国科学院二〇〇九年院士候选人周祖德及其学生的论文收录某学术会议论文集，被怀疑指出抄袭后删除该文推出“正式版”论文集
- 2009年10月18日，复旦新闻学院副教授许燕涉嫌抄袭被举报
- 2010年9月，北京科技大学信息工程学院院长尹怡欣教授署名的论文被指剽窃
- 2011年9月，中国农业大学教授李季伦等6人实名举报：中国农业大学校长石元春在任期内利用职权窃取他人治理黄淮海盐碱土的成果，当选中国科学院、中国工程院和第三世界科学院院士。





2.2 学术论文借鉴与剽窃的区别

- 在科学研究中借鉴或合理使用他人的学术成果,是科学发展、学术进步的必由之路,是著作权法所允许的行为
- 所谓借鉴,就是以介绍、评论或者说明为目的在作品中直接或者间接引用他人已经发表的作品,并注明作品来源、出处等,不侵犯著作权人合法权益的合理使用
- 如何做到科学的、合理的使用,而不是将他人成果据为己有,就成为判断学术论文借鉴与剽窃的基本标准
- 依照《著作权法》第二十二条第一款第(二)项的规定,“为介绍、评论某一作品或者说明某一问题,在作品中适当引用他人已发表的作品”,这属于法定的“**合理使用**”,不能算抄袭。
- 关键是对他人已经发表的作品只能适当引用,如果自己的作品基本或大部分是从他人的作品中拿来的,就算抄袭。





2.3 学术论文剽窃的主要表现形式

侵犯学术论文著作权的行为方式有多种,包括

- 擅自发表他人论文
- 擅自修改他人论文
- 擅自进行网络传播
- 歪曲篡改论文内容
- 擅自增删署名

剽窃是侵犯学术论文著作权的重要形式

抄袭又是剽窃的主要形式





2.3 学术论文剽窃的主要表现形式

- “抄袭”是把别人的作品抄来当作自己的作品的一种行为,导致其作品内容的具体表达形式与被抄作品相同。其实质表现在:作品是作者自己独创的,还是照抄照搬他人的。
- 如果使用了他人的作品,并且注明出处,就属于创作上的借鉴、引用。这种“名抄”是合法的。
- 如果使用他们的作品,不注明,就是“暗抄”,等于将别人的学术成果变为自己的,属于偷、窃的范畴,因此称为剽窃。





2.3 学术论文剽窃的主要表现形式

具体来看,常见的剽窃形式包括哪些呢?

- 直接照搬原文的语句、段落,甚至整篇文章
- 将原文删节、顺序调整,重新组合成文
- 将多篇文章组合成一篇
- 仅对原文做叙述方式上的改变,仍然沿袭原文的写作思路、结构
- 仿照一篇论文的写作方式,虽然改变了所要论述的对象,但是使用原文的论证方法、图表





2.3 学术论文剽窃的主要表现形式

认定学术论文剽窃的困难

1. 剽窃的隐蔽性强,有些剽窃从字面上难以看出。
2. 判定论文的抄袭与否需要对相关的学术领域具有较深入的了解,才能分辨出二者之间的内在联系
3. 抄袭手段的多种多样,一篇论文中可能综合运用了多种抄袭手段,比如有的内容直接翻译自外文资料,有的内容是节选自某学术论著,有的是“巧妙”的将多篇论文嫁接、组合
4. 抄袭的欺骗性。抄袭论文往往也列出该文的引文出处,将被抄袭论文列入其中。乍一看是合理使用,但是仔细比对就会发现实际引用的文字数量远远大于标注的引文数量
5. 以抄袭文为标准或是以被抄袭文为基准,可能会得出截然相反的结论





2.4 学术论文剽窃的责任承担

(1) 剽窃者承担的侵权责任

最普遍的形式是,权利人主张权利,提出侵权损害赔偿之诉。要求剽窃者承担停止侵害、消除影响、赔礼道歉、赔偿损失等民事责任。

除了需追究剽窃者的侵权赔偿责任,还应当对剽窃者因剽窃所得的不当得利进行纠正。如果剽窃者并未发表,而是凭借因该论文获得了有效的考试成绩,或者通过了学位论文答辩,该论文对于权利人的侵权诉讼请求,只要求剽窃给予赔偿、赔礼道歉等的话,那么剽窃者基于剽窃行为的不当得利仍然存在,就使得被侵权人的精神损失并未得到有效抚慰,不能有效杜绝违法行为再次发生。所以,还应当取消学生因剽窃论文所获得的资格或荣誉以及其他的不当得利。





2.4 学术论文剽窃的责任承担

(2) 刊登剽窃论文的刊物承担的责任

- 从现有的案例来看,发表剽窃论文的期刊机构有时并没有受到起诉
- 有些期刊以声明的方式免责
- 有些期刊与作者签订保证原创性的协议,来避免陷入侵权纠纷





2.4 学术论文剽窃的责任承担

(3) 教育机构以及相关人员的责任

由于有些剽窃论文并不是为了公开发表,而是为了应付学校的作业、毕业论文等,权利人往往无法知道自己的论文被剽窃。但是学校在教学、考核等过程中,一经发现学生剽窃论文的行为,查证属实后,会给予纪律处分。由于未造成重大社会影响,学校一般不会将剽窃事实告知权利人。但是由于性质恶劣,剽窃者往往被学校处以较重的处罚。从法律上来讲,学校没有告知权利人的义务,也就不存在以行政处罚规避法律制裁的问题。

对于这类剽窃,校方没有直接的故意(如果存在直接故意,可以作为共同侵权人追究责任,自不待言),但如果是教师指导过、机构审查过的论文,就另当别论了。指导教师有义务对论文进行指导、修改,相关的论文评审机构成员也有能力对论文进行审查,作为该领域的专家,能够判断学生是否剽窃,而出于疏忽大意没有判断。由于这是指导教师或评审人员的职务行为,相应的教育机构应当承担连带责任。教育机构还可以依职权进一步追究直接责任人的责任。比如对学术论文指导失误的老师,在其招生资格上给予限制。

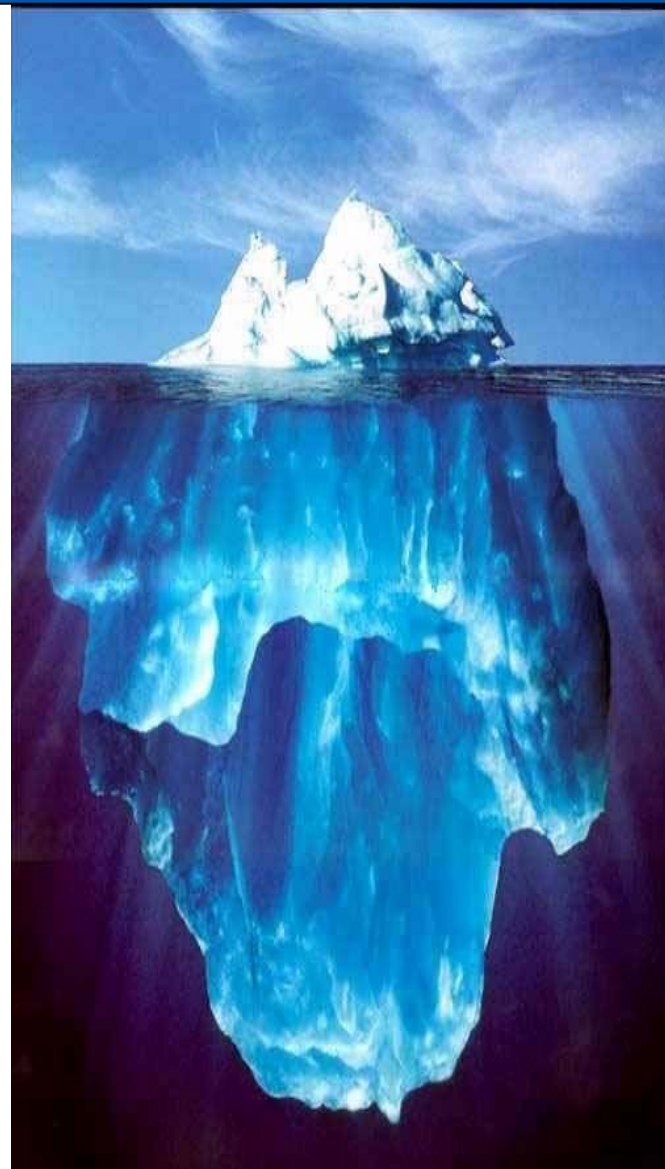




大纲

- 谈谈做研究
- 学术规范
- 基本常识
- 论文发表
- 论文写作
- 硕士论文
- 附件

说明：本PPT部分内容来自周志华老师的PPT
《如何做研究，如何写论文》





专题三 基本常识

3.1 常用的数据库有哪些？

3.2 有哪些实用的论文相关网站？

3.3 什么是核心期刊？

3.4 什么是SCI、EI、ISTP？

3.5 什么是期刊的影响因子？





3.1 常用的数据库有哪些

- 中文数据库
 - [万方数据资源系统](#)
 - [中国期刊全文数据库 \(cnki\)](#)
- 外文数据库
 - [ACM Digital Library](#)
 - [IEEE/IEE IEL](#)
 - [Springer](#)





3.1 常用的数据库有哪些

万方数据资源系统



万方数据库资源系统包含中国学位论文全文数据库、中国数字化期刊群、科技信息子系统以及商务信息子系统。

中国学位论文全文数据库由国家法定学位论文收藏机构——中国科技信息研究所提供，并委托万方数据加工建库，收录了自1980年以来我国自然科学领域博士、博士后及硕士研究生论文，其中全文60余万篇，每年稳定新增15余万篇，是我国收录数量最多的学位论文全文库。





3.1 常用的数据库有哪些

中国期刊全文数据库 (cnki)



<http://dlib.edu.cnki.net>

清华大学主办，收录1979年至今约7200种期刊全文,至2005年8月止，累积期刊全文文献1500多万篇。是目前世界上最大的连续动态更新的中国期刊全文数据库。期刊文献出版平均不迟于纸质期刊出版之后2个月。

内容覆盖理工A（数理科学）理工B（化学化工能源与材料）、理工C（工业技术）、农业、医药卫生、文史哲、经济政治与法律、教育与社会科学、电子技术与信息科学，分九大专辑，126个数据库。更新频率 日更新、月更新。





3.1 常用的数据库有哪些

ACM（美国计算机学会） Digital Library



ACM数据库镜像站点: <http://acm.lib.tsinghua.edu.cn>

ACM 创立于1947年，是全球历史最悠久和最大的计算机教育和科研机构。目前提供的服务遍及100余国家，会员达85,000多位专业人士，涵盖工商业，学术界及政府单位。它致力于发展信息技术教育、科研和应用，出版最具权威和前瞻性的出版物，如专业期刊、会议录和新闻报道；并于1999年开始提供电子数据库服务—ACM Digital Library全文数据库。

在过去的几年里，ACM全文数据库增加了1950年代至今的所有出版物的全文内容，以及Special Interest Group的出版文献，包括快报和会议录。同时ACM还整合了第三方出版社的内容，全面集成“在线计算机文献指南”（The Guide to Computing Literature），这是一个书目资料和文摘数据库，集合了ACM和其他3,000多家出版社的出版物，旨在为专业和非专业人士提供了解计算机和信息技术领域资源的窗口。





3.1 常用的数据库有哪些

IEEE/IEE IEL



<http://ieeexplore.ieee.org/>

提供1988年以来，美国电气电子工程师学会和英国电气工程师学会出版的120多种期刊、600多种会议录、近900种标准的全文信息。





3.1 常用的数据库有哪些

Springer电子期刊



清华镜像: <http://springer.lib.tsinghua.edu.cn>
主站点: <http://www.springerlink.com>

德国施普林格(Springer-Verlag)是世界上著名的科技出版集团, 所提供的全文电子期刊按学科分为以下11个“在线图书馆”: 生命科学、医学、数学、化学、计算机科学、经济、法律、工程学、环境科学、地球科学、物理学与天文学, 是科研人员的重要信息源。





3.2 有哪些实用的论文相关网站



<http://www.informatik.uni-trier.de/~ley/db/>

C-DBLP

<http://www.cdblp.cn/>

CiteSeer^x beta

<http://citeseerx.ist.psu.edu/>





3.3 什么是核心期刊

■“**核心期刊**”是20世纪30年代由英国文献学家布拉德福(B.C.Bradford)提出的。他按照载文密度将专业期刊划分为对该专业最有贡献的核心区域(nucleus)和论文数量与之相等的几个相继区域,将密度最大的核心区域的期刊称为“核心期刊”。60年代,美国文献计量学家加菲尔德(E.Garfield)用类似方法从引文角度证实了核心期刊的存在。后来,众多学者对以上经验定律提出各种修正方案,从多方面发展了相关的文献计量理论与模型,形成了诸多公认的定律和数学模型。我们的定义是: **某学科(或某领域)的核心期刊,是指那些发表该学科(或该领域)论文较多、使用率(含被引率、摘转率和流通率)较高、学术影响较大的期刊。**

■核心期刊包含了学科的主要期刊,具有数量少,信息量大的特点,占整个期刊量的5%,信息量占整个期刊的50%,浏览核心期刊,可以在较短的时间内掌握较多的信息,了解最新的科研动态等。

■我国科技界从上世纪60—70年代开始引进核心期刊的理论与方法,到90年代,推广到人文社会科学界。较早的研究成果当属北京大学图书馆编制的《中文核心期刊要目总览》,它为方便读者查找专业论文和优化馆藏提供了很好的参考依据。





3.3 什么是核心期刊

- **核心期刊**只是一种相对统计的概念，没有完全绝对的量化指标，其研究成果可以作为评价学术研究成果的参考工具，可以作为读者投稿的参考工具，但如果作为衡量论文水平的绝对标准来用，就可能发生谬误。
- 从个体角度看，发表在核心期刊上的论文未必每篇学术水平都高，而发表在非核心期刊上的文章未必每篇学术水平都低，应该根据评价项目具体情况请学科专家来评价论文本身学术价值。
- 使用这种期刊表时应避免因片面夸大其作用而带来的不良影响。





3.3 什么是核心期刊

目前国内有7大核心期刊（或来源期刊）遴选体系：北京大学图书馆“中文核心期刊”、南京大学“中文社会科学引文索引（CSSCI）来源期刊”、中国科学技术信息研究所“中国科技论文统计源期刊”（又称“中国科技核心期刊”）、中国社会科学院文献信息中心“中国人文社会科学核心期刊”、中国科学院文献情报中心“中国科学引文数据库（CSCD）来源期刊”、中国人文社会科学学报学会“中国人文社科学报核心期刊”以及万方数据股份有限公司正在建设中的“中国核心期刊遴选数据库”。

国家级艺术类核心期刊,面向国内外公开发行。通常所说的中文核心期刊,是指被北大每年出版的《中国核心期刊要目总览》中被列出的期刊。





3.3 什么是核心期刊

关于公布厦门大学研究生申请学位发表学术论文核心刊物目录的通知

各学院、研究院：

根据《厦门大学博士、硕士研究生申请学位发表学术论文的规定》（厦大研〔2006〕9号）文件精神，经校学位评定委员会审议认定，我校博士、硕士研究生申请学位发表学术论文的核心刊物目录除继续采用《[中文核心期刊要目总览](#)》（北京大学出版社2004年版）外，同时启用《[厦门大学核心学术刊物目录](#)》（2005年版）。其中，2004年版《中文核心期刊要目总览》（北京大学出版社）于2005年9月2日启用（厦大研〔2005〕19号），2000年版《中文核心期刊要目总览》（北京大学出版社）在新版启用后的一年内有效，即有效期至2006年9月1日。特此公布，并于公布之日起执行。

请各单位及时通知导师与学生，遵照执行。

研究生院

2006-3-27





3.3 什么是核心期刊

理工科一类核心学术刊物

被SCI、EI、ISTP和SSCI、A&HCI、ISSHP收录的学术论文，均为理工科一类核心学术刊物论文

理工科二类核心学术刊物

未被收进一类核心学术刊物的CSCD（《中国科学引文数据库》）中文核心库期刊均为理工科二类核心学术刊物（共606种）





3.4 什么是SCI、EI、ISTP

2.4.1 什么是SCI?

2.4.2 什么是EI?

2.4.3 什么是ISTP?

2.4.4 SCI收录号查询方法

2.4.5 EI收录号查询方法

2.4.6 ISTP收录号查询方法

2.4.7 什么是ISSN/ISBN/IDS number/DOI?

2.4.8 什么是JCR分区?





3.4.1 什么是SCI?

《科学引文索引》(Science Citation Index, SCI)是由美国科学信息研究所(ISI)1961年创办出版的引文数据库,其覆盖生命科学、临床医学、物理化学、农业、生物、兽医学、工程技术等方面的综合性检索刊物,尤其能反映自然科学研究的学术水平,是目前国际上三大检索系统中最著名的一种,其中以生命科学及医学、化学、物理所占比例最大,收录范围是当年国际上的重要期刊,尤其是它的引文索引表现出独特的科学参考价值,在学术界占有重要地位。许多国家和地区均以被SCI收录及引证的论文情况来作为评价学术水平的一个重要指标。从SCI的严格的选刊原则及严格的专家评审制度来看,它具有一定的客观性,较真实地反映了论文的水平和质量。根据SCI收录及被引证情况,可以从一个侧面反映学术水平的发展情况。特别是每年一次的SCI论文排名成了判断一个学校科研水平的一个十分重要的标准。





3.4.1 什么是SCI?

版本	出版周期	收录期刊数
印刷版(SCI)	双月刊	3,500种
联机版(SciSearch)	周更新	5,600种
光盘版(带文摘)(SCICDE)	月更新	3,500种
网络版(SCIExpanded)	周更新	5,600种





3.4.1 什么是SCI?

什么是SCI分区?

《JCR 期刊影响因子及分区情况》，由中国科学院文献情报中心提供。按照该期刊的影响因子，投稿难度和影响力，分为1~4区：一区一般是各领域的top期刊，二区是高水平期刊，三区次之，四区则更普通。





3.4.1 什么是EI?

《工程索引》(**EngineeringIndex, EI**), 1884年创刊, 由美国工程信息公司出版, 报道工程技术各学科的期刊、会议论文、科技报告等文献。

版本	出版周期	收录文献源
光盘版(EI Compendex)	双月刊	2600种
网络版(EI Compendex Web)	季度更新	5,600
种光盘版(带文摘)(SCICDE)	周更新	5000种





3.4.1 什么是ISTP?

《科技会议录索引》（**Index to Scientific & Technical Proceedings**, 简称**ISTP**）创刊于1978年，由美国科学情报研究所编辑出版。该索引收录生命科学、物理与化学科学、农业、生物和环境科学、工程技术和应用科学等学科的会议文献，包括一般性会议、座谈会、研究会、讨论会、发表会等。其中工程技术与应用科学类文献约占**35%**，其他涉及学科基本与**SCI**相同。

ISTP收录论文的多少与科技人员参加的重要国际学术会议多少或提交、发表论文的多少有关。我国科技人员在国外举办的国际会议上发表的论文占被收录论文总数的**64.44%**。

版本	出版周期收录	文献源
印刷版(ISTP)	月刊	每年报导4,700多种会议录
光盘版(ISTP)	季度更新	年报导10,000多种会议录
网络版(WOSP—S/T)	周更新	同光盘版





3.4.4 SCI收录号查询方法

ISI Web of KnowledgeSM

(1) 进入<http://isiknowledge.com/>；选择web of science

(2) Select a search option选“ADVANCED SEARCH”

按照右边Field Tags和 Booleans提示，在输入方框中输入检索信息，比如AU=fang AND TI=capacitor，然后点击“SEARCH”。

(3) 点击“RESULTS”里的检索结果数字链接，如“3”，打开自己发表论文题目的链接,此时在Full record（包括Title、Author、Source... IDS number等）中有IDS Number: 226QR，这个IDS number 并不是SCI的收录号。

（提醒：注意TI里的关键词一定要与你发表论文的词语匹配，包括单复数。）

(4) 点击右边“Output This Record”下面的“SAVE”按钮，“Select a data format for the saved file.”有四个选项，选择HTML格式。

(5) 选择HTML格式后，点击下面按钮“CONTINUE”,出现文件下载对话框（默认文件名为

“savedrecs.html”），点“保存为”，保存到自选的目录文件夹里。

(6) 打开保存的HTML格式网页，或者下载完毕后直接打开保存的文件，这时得到的就是包含有形如**UT ISI: 000250604100008**收录号的记录格式。获取ISTP的方法与此完全相同。





3.4.5 EI收录号查询方法



(1) 进入图书馆网站，电子资源里点击“EI village2”

<http://www.engineeringvillage2.org.cn/>

(2) 在“SEARCH FOR”中输入主题词，并在“SEARCH IN”中限制。如果是搜索姓名，由于国内外期刊在发表时姓名写法不太一样，比如张三丰，有 San-Feng Zhang, Zhang Sanfeng, S.-F zhang等，如果检索自己文章收录结果和自己发表的文章数目不一样，注意姓名检索的拼法

(3) 在“Search Results”里点击“Detailed”，“**Accession number**”就是文章收录号





3.4.6 ISTP收录号查询方法

ISI基于Web-of-Science-的检索平台，因为SCI和ISTP是由同一家数据商经营，设在同一个检索平台上，所以具有完全相同的检索界面，所以获取ISTP的方法与此完全相同。





3.4.7 什么是ISSN/ISBN/IDS number/DOI

ISSN--国际标准连续出版物号（International Standard Serial Number）的简称，是ISDS(国际连续出版物数据系统)国际中心为在该系统登记的连续出版物分配的号码。采用ISSN编码系统的出版物有：期刊、会议录等。ISSN由8位数字组成，分两个部分：序号和检验码。在联机书目中，ISSN可以作为一个检索字段，从而为用户增加了一种检索途径。

ISBN又称国际标准书号，它是英文International Standard Book Number的缩写。ISBN由四段共十位数字组成，第一段为组号，是国家、地区、语言或其他组织集团的代号，由国际书号中心负责分配，中国的组号为数字“7”；第二段为出版社号，由国家标准书号中心负责分配，其位数视申请出版社图书出版量多少而异，山东人民出版社的出版社号为“209”；第三段为书序号，由出版社负责管理分配；第四段为校验码，其数值由前九位数字依次以10~1加权之和并以11为模计算得到。在联机书目中，ISBN可以作为一个检索字段，从而为用户增加了一种检索途径。

IDS number：SCI的“收录检索号”有些人填文献记录中的IDS Number，但每种期刊每一期上的文献IDS umber都相同，IDS Number并不是SCI的收录号。在ISI Web of Science中，IDS number的解释为：Document Solution number. This number uniquely identifies the journal and issue.





3.4.7 什么是ISSN/ISBN/IDS number/DOI

DOI 是“Digital Object Identifier”的简写，用来标识在数字环境中的内容对象。DOI可以用来揭示有关该数字对象的一些信息 包括从INTERNET哪里可以找到它等。随着时间推移数字对象的某些有关信息可能会有变化 包括从哪里可以找到它，但是DOI不会改变。

■ DOI是一个永久性的标识号 由International DOI Foundation管理。通过doi查找原始文献的方法很简单，只要将下面的url地址加在已知的doi前面就得到该文献在doi系统中的url. <http://dx.doi.org/> 例如已知某文献的doi:

10.1103/PhysRevLett.95.253601 则该文献在doi系统中的url就是
<http://dx.doi.org/10.1103/PhysRevLett.95.253601>

■ 事实上 如果已知doi, 也可通过以下地址由doi系统自动帮助查找url.
<http://dx.doi.org/> 在"Resolve A DOI"的提示框内输入已知doi, 点击"Go"按钮 doi系统就会自动链接到该文献的url,并显示相应的页面。另外 CrossRef网站也提供了通过doi查找url的功能 <http://www.crossref.org> 在"DOI Resolver"输入提示框中输入已知doi "submit"即可。

其他DOI工具网站

英文: <http://sherman.library.nova.edu/doi/>

中文: <http://www.doi.org.cn/>





3.4.8 什么是JCR分区

- ISI每年出版JCR（《期刊引用报告》，全称Journal Citation Reports）。JCR对包括SCI收录的3800种核心期刊(光盘版)在内的8000多种期刊(网络版)之间的引用和被引用数据进行统计、运算，并针对每种期刊定义了影响因子（Impact Factor）等指数加以报道。
- JCR分区表是指把其SCI数据库中收录的不同期刊，按不同专业分类排列的。每个学科都有自己的分区，一般分为1，2，3，4区。1区均是本领域顶级刊物。
- 附件：[2011年最新JCR分区表.xls](#)





3.5 什么是期刊的影响因子

影响因子(Impact Factor)是一个国际上通行的期刊评价指标，即某期刊前两年发表的论文在统计当年的被引用总次数除以该期刊在前两年内发表的论文总数。该指标是相对统计值，可克服大小期刊由于载文量不同所带来的偏差。一般来说，影响因子越大，其学术影响力也越大。

查询外文期刊影响因子，可使用外文数据库Web of Science中的JCR（Journal of Reports），其中JCR Science Edition 用于查询自然科学类期刊，JCR Social Sciences Edition用于查询人文社会科学类期刊。（[2011年SCI影响因子](#)）。

查询中文期刊的影响因子，可使用中国学术期刊（光盘版）电子杂志社和中国科学文献计量评价中心联合推出的《中国学术期刊综合引证报告》（万锦堃主编，科学出版社）。

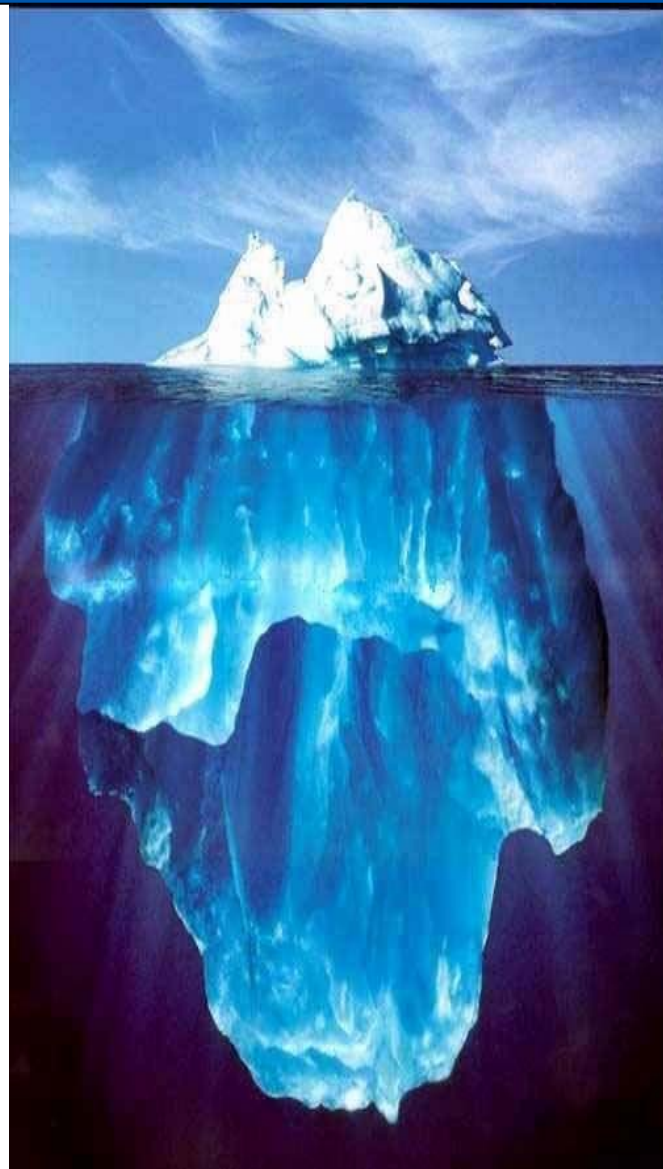




大纲

- 谈谈做研究
- 学术规范
- 基本常识
- 论文发表
- 论文写作
- 硕士论文
- 附件

说明：本PPT部分内容来自周志华老师的PPT
《如何做研究，如何写论文》





专题四：论文发表

- 4.1确定作者署名
- 4.2确定投稿目标
- 4.3发表源的类型
- 4.4发表源的档次
- 4.5文章类型
- 4.6论文发表注意事项
- 4.7审稿流程
- 4.8Reviewer如何审稿
- 4.9审稿表





4.1 确定作者署名

确定论文署名：

- 一般情况，按贡献大小排序
 - 如果是普通合作者，则考虑P, I, C
 - 如果是导师给问题，则主要考虑I, C
 - 导师是在培养学生
 - 综合考虑
 - 提出idea的人经常成为第一作者
- 通常由第一作者执笔
- 通讯作者通常是整个研究工作的负责人
- 作者不要太多，小的贡献可以放在致谢里





4.2确定投稿目标

确定投稿目标：

- 根据工作的水准，挑选合适的发表源
 - 除非有特殊的考虑，否则：
 - 低投—遗憾
 - 高投—延误发表
 - 针对不同的发表源可能有不同的写法
- 通常情况：导师帮你确定
- 如果需要自力更生，那么.....





4.3发表源的类型

- 发表源的类型
- Journal: 例如AIJ, IEEE Trans. PAMI
- Conference: 例如SIGMOD, SIGKDD, VLDB
- Magazine: 例如AI Mag, IEEE Intelligent System
- 计算机科学界很重视会议论文，高档次会议并不亚于高档次刊物
 - 这是个例外，其他学科通常是不把会议当回事的





4.4发表源的档次

■权威刊物

《[中国科学](#)》 《[科学通报](#)》 《[自然科学进展](#)》

《[计算机学报](#)》 《[软件学报](#)》 《[自动化学报](#)》 《[电子学报](#)》

《[计算机研究与发展](#)》

《[Journal of Computer Science and Technology](#)》 (JCST)

■一般核心刊物

《[计算机科学](#)》 《[计算机工程与应用](#)》

《[厦门大学学报](#)》

■学术会议（[数据库类：CCF推荐会议类别](#)）

一级：VLDB、SIGMOD、ICDE、SIGKDD

二级：WWW、PODS、EDBT、SIGIR、ICDM、ICDT、CIKM、ER

三级：DASFAA、PAKDD、DOLAP

四级：WAIM、APWEB、WISE

五级：NDBC





4.4发表源档次



《**计算机学报**》是中国计算机领域权威性学术刊物。其宗旨是报道中国计算机科学技术领域最高水平的科研成果。它由中国计算机学会与中国科学院计算技术研究所主办、科学出版社出版，以中文编辑形式与读者见面，同时以英文摘要形式向国际各大检索系统提供基本内容介绍。

《计算机学报》始创于**1978**年,刊期为月刊，每期**160**面。

《计算机学报》由资深专家组成编辑委员会。其中包括两院院士十几名。主编由人工智能专家高文教授担任。副主编由中国科学院院士张钹教授、中国工程院院士何新贵教授、网络专家林闯教授、信息安全专家冯登国研究员、系统结构专家孟丹研究员、唐志敏研究员担任。

《计算机学报》一直在计算机领域保持着领先水平,刊登的文章被国际多种著名检索刊物所收录。包括《中国学术期刊文摘》,美国《EI》，英国《SA》,美国《数学评论》,日本《科技文献速报》,俄罗斯《文摘杂志》等。

《计算机学报》刊登的内容覆盖计算机领域的各个学科，以论文、技术报告、短文、研究简报、综论等形式报道以下方面的科研成果：计算机科学理论、计算机硬件体系结构、计算机软件、人工智能、数据库、计算机网络与多媒体、计算机辅助设计与图形学以及新技术应用等。

《计算机学报》拥有众多读者，本刊的内容常被研究人员、大学教授及研究生们作为选题的依据。《计算机学报》的引用率在计算机相关领域的刊物中最高,发行量名列全国学术性刊物的榜首。

《计算机学报》愿意和国内外同仁建立友好往来，相互促进了解，加强合作，共同为计算机科学事业的发展而努力。





4.4发表源档次



《**软件学报**》是一本刊登计算机软件各领域原创性研究成果的期刊,所刊登的论文均经过严格的同行专家评议.《软件学报》主要面向全球华人计算机软件学者,致力于创办与世界计算机科学和软件技术发展同步的以中文为主的"中文国际软件学术期刊",为全球华人同行提供学术交流平台.

《软件学报》创刊于1990年,由中国科学院软件研究所和中国计算机学会联合主办.CN11-2560/TP, ISSN1000-9825, CODEN RUXUEW.月刊,每期176面,每月6号出版.

《软件学报》注重刊登反映计算机科学和计算机软件新理论、新方法和新技术以及学科发展趋势的文章,主要涉及理论计算机科学、算法设计与分析、系统软件与软件工程、模式识别与人工智能、数据库技术、计算机网络、信息安全、计算机图形学与计算机辅助设计、多媒体技术及其他相关的内容.





4.4发表源档次



中国科学 F 辑: 信息科学

【主办单位】中国科学院、国家自然科学基金委员会

【刊载内容】

主要报道计算机科学与技术、控制科学与控制工程、通信与信息系统、电子科学与技术、生物信息学等方面具有重要意义的研究成果，是信息科学领域的综合性学术期刊。

【文章收录】

《中国科学论文与引文数据库》、《中国科学引文数据库》、《中国期刊全文数据库》、《中国数字化期刊群》等。

注：该期刊只能LATEX投稿，LATEX模板很难用





4.4发表源档次



地位及水平：多年来，《计算机研究与发展》一直被评为我国计算机类核心期刊，是国务院学位办指定的评估学位与研究生教育的“中文重要期刊”；已被多个国际著名的检索系统收录，如美国《工程索引》（Ei）、日本《科学文献速报》、俄罗斯《文摘杂志》等；其影响因子和总被引频次在同类学术期刊中名列前茅；连年被评为“百种中国杰出学术期刊”。

刊登内容：计算机科学技术领域高水平的学术论文、最新科研成果和重大应用成果。刊登内容：述评、计算机基础理论、软件技术、信息安全、计算机网络、图形图象、体系结构、人工智能、计算机应用、数据库技术、存储技术及计算机相关领域。





4.4发表源档次



JCST, the only English Language Journal of computer fields published in China, is an international forum for scientists and engineers in all aspects of computer science and technology to publish high quality and refereed papers. The papers of original research and innovatory applications from all parts of the world are welcome. The papers for publication in JCST are selected through rigorous peer reviews to ensure originality, timeliness, relevance, and readability. While it emphasizes publication of previously unpublished materials, selected conference papers with exceptional merit that require wider exposure are, at the discretion of the editors, also published provided they meet the journal's peer review standard. JCST also seeks clearly written survey and review articles from experts in the field to promote insightful understanding of the state-of-the-art and the technology trends. JCST is covered in many authoritative indexing and abstracting systems, such as SCIE, Web of Science, Research Alert, CompuMath Citation Index, Engineering Information, Computing Reviews, INSPEC, CBST, Abstracts Magazine, Computer Abstracts, Mathematica Review, Mathematics Abstracts, and DBLP.





4.4发表源档次



《**计算机科学**》由国家科技部主管，西南信息中心主办，系“中文科技核心期刊”、“中国科技论文统计与分析用期刊”、“中国科学引文数据库来源期刊”、“中国期刊方阵双效期刊”。主要报导国内外计算机科学与技术的发展动态，涉及面广的方法论与技术，和反映新苗头、能起承先启后作用的研究成果。内容涉及程序理论、计算机软件、计算机网络与信息、数据库、人工智能、人机界面、国际会议、应用等。





4.4发表源档次



《**计算机工程与应用**》杂志是由中华人民共和国工业和信息化部华北计算技术研究所主办的、面向中高级计算机专业工作者的学术刊物，系中国计算机学会会刊、中国电子学会一级会刊，计算机工程与应用学会学报、计算机类中文核心期刊。被认定为“中国科学引文数据库来源期刊”“中国学术期刊综合评价数据库来源期刊”，并收录在《中国期刊网》、《中国学术期刊（光盘版）》。

《计算机工程与应用》是一本面向计算机全行业的综合性学术刊物，覆盖面宽、信息量大、报道及时是本刊的服务宗旨。多年来，本刊坚持走学术与实践相结合的道路，在内容上既注重理论的先进性又兼顾实用技术的广泛性，在促进学术交流的同时，推进了科技成果的转化。





4.4发表源档次



《厦门大学学报（自然科学版）》于1931年创刊，由教育部主管，厦门大学主办，是国内外公开发行的综合性学术期刊（双月刊），本刊以印刷版、光盘版、网络版同时出版。

本刊坚持党的基本路线，提倡“双百”方针，以促进科学技术发展，传播科技成果，发现和培养科技人才提供了一个总结经验、相互交流的园地。

本刊主要刊载数学，计算机科学，物理学，技术科学，化学，化工，海洋学，环境科学，生命科学等学科的最新研究成果。所刊载的论文分三大类型：（1）“快讯”：报道具有原创性的最新研究成果或阶段性研究成果。

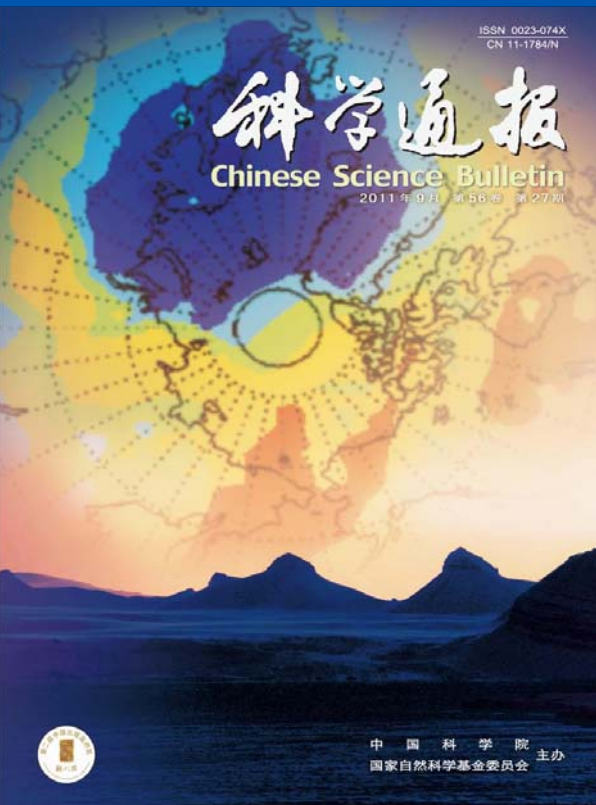
（2）各学科“研究论文”：刊载理工科基础理论与实验研究学术论文。

（3）“研究简报”：刊载内容新颖，实用的阶段性成果。





4.4发表源档次



主管单位：中国科学院
主办单位：中国科学院
主编：夏建白
刊号：ISSN0023-074X/CN11-1784/N
创刊时间：1950年
出版周期：旬刊
开本：大16开
页数：不固定（80~160）
网址：<http://www.scichina.com> csb.scichina.com

期刊宗旨及涵盖范围：

自然科学综合性学术刊物。力求及时报道自然科学各学科基础理论和应用研究方面具有创新性和高水平的、有重要意义的最新研究成果，要求文章的可读性强，能在一个比较宽泛的学术领域产生深刻的影响。读者对象为国内外科技工作者。





4.4发表源档次



- 《电子学报》是中国电子学会主办的高级学术刊物,刊登电子与信息科学及相邻领域的原始 (original) 科研成果。1962年创刊,每月25日出版。
- 《电子学报》为中国自然科学核心期刊之一;科技部科技论文统计源期刊;中国科学引文数据库来源期刊。多次获得全国优秀科技期刊奖、中国科协优秀学术期刊奖、国家期刊奖,并得到2000年、2004年及2010年国家自然科学基金重点学术期刊专项基金支持。

《电子学报》现被世界主要检索系统收录:(1)美国工程索引 (Engineering Index, 1993~); (2)美国化学文摘CA; (3)英国科学文摘INSPEC; (4)日本科技文献快报。国内重要的检索工具如《中国无线电电子学文摘》,《中国电子科学文摘》,《中国科学文摘》,《中国机械工程文摘》,《中国学术期刊文摘》等均收录《电子学报》。





4.4发表源档次



自动化学报 ACTA AUTOMATICA SINICA

[首页](#) | [English](#)

- 《自动化学报》（ACTA AUTOMATICA SINICA）是由中国自动化学会、中国科学院自动化研究所共同主办的高级学术期刊。本刊于1963年创刊，1966年停刊，1979年复刊，现为大16开本，月刊，每期112页。科学出版社与Elsevier合作出版，国内外公开发行。
- 《自动化学报》的读者对象为从事自动化科学技术的科研、教学、开发人员和研究生、大学生等。《自动化学报》被国外著名检索刊物，如美国工程索引（EI）、英国科学文摘（SA）、日本科学技术文献速报（JICST）、俄罗斯文摘杂志（AJ）等多种检索刊物和数据库收录。





4.4发表源档次



- 《自然科学进展》[1]期刊是国家自然科学基金委员会和中国科学院主办的综合性学术月刊,刊登自然科学各学科领域的基础研究和应用基础研究方面的高水平、有创造性和重要意义的最新研究成果论文,以促进国内外学术交流.中文版由各地邮局公开发行,英文版由英国 Taylor & Francis Ltd总代理,在世界各地发行.
- 《自然科学进展》期刊中文版为《中国科技期刊引证报告》的源期刊,并被《中文核心期刊要目总览》、“生物学文摘”等数据库和检索系统收录;英文版(Progress in Natural Science)被SCI Expanded, Chemical Abstracts(CA), Engineering Index (EI), 俄罗斯《文摘杂志》, 美国《数学评论》和日本《科技文献速报》等多种国际检索系统收录.





4.5文章类型

- Journal
 - Regular paper
 - Short paper/Concise paper
 - Letter/Correspondence
- Conference
 - Research
 - Industrial
 - Demo
 - Poster
- 注意：
 - 1.不要自行与AE/Reviewer联系
 - 2.超过原定的审稿期后，可以写一封有礼貌的信给主编





4.6 论文发表注意事项

论文发表注意事项

- 同一工作不能在刊物上发表两次
- 在会议上发表的工作，经过扩充后可以再投往刊物
 - 应该有至少**30%**的新东西（例如更详尽的理论分析、更多的实验结果等）
 - 投稿时根据刊物的要求明确说明
- **CS**刊物上文章篇幅通常比较长，会议文章篇幅较短
- **CS**刊物发表周期通常比较长，一些刊物可能要**2-3**年
- 绝不要一稿多投！
 - 现在有一些比较好的会议已经开始允许同时向刊物投稿





4.7 审稿流程

稿件处理流程

- Journal

- 编辑部/主编(Editor-in-Chief)收到稿件
- 转给合适的Associate Editor 处理
- AE找reviewer审稿
- Reviewer的意见到AE手里汇总
- AE做出处理意见: accept, revision, reject
- Editor-in-Chief审定





4.7 审稿流程

处理意见的大致类型:

- **Accept without revision**
 - 罕见的情况
- **Minor revision**
 - 比较少见的情况，已经录用了
- **Major revision / Reject but resubmission encouraged**
 - 多数情况；修改后再进行一轮审稿
- **Reject**
 - 大多数文章，尤其是对很好的刊物来说
 - 很多刊物一期只有5、6篇文章
 - 即使不能被录用，好刊物的审稿意见会对你有很大帮助





4.7 审稿流程

Conference

- 一般来说：
 - PC member对论文进行投标
 - PC Chair根据投标情况，把论文分配给适当的PC member
- 有的会议设置了Area Chair，协助PC Chair处理某一领域的论文；有的会议没有投标过程
- PC member 忙不过来时，往往会找external reviewer帮忙审稿
- 会议的处理意见通常是：录用或拒绝





4.8Reviewer如何审稿

- 一般来说，主要考察以下方面：
 - Relevance: 是否相关
 - Originality: 是否有创新
 - Significance: 是否重要
 - Soundness: 是否有足够强的支持
 - Presentation: 是否表达得好





4.8Reviewer如何审稿

In summary, reviewers will check:

- Does the paper introduce a new problem or provide a new solution to an existing one?
- What is the main result of this paper?
- Is the result significant?
- Is the paper technically sound?
- Does the paper provide an assessment of the strength and weakness of the results?
- Is the paper clearly written so as to accessible by most researchers in this area?
- Does the paper refer appropriate related works?
-
- **Don't attempt to fool the reviewers !**





4.9审稿表

Organization

(Poor) ☐1 ☐2 ☒3 ☐4 ☐5 (Excellent)

Clarity

(Low) ☐1 ☐2 ☒3 ☐4 ☐5 (High)

Length

(Too Long) ☐1 ☒2 ☐3 ☐4 ☐5 (Too Short)

References

(Incomplete) ☐1 ☒2 ☐3 ☐4 ☐5 (Adequate)

Correctness

(Incorrect) ☐1 ☐2 ☒3 ☐4 ☐5 (Correct)

Significance

(Low) ☐1 ☐2 ☒3 ☐4 ☐5 (High)

Originality

(Low) ☐1 ☐2 ☒3 ☐4 ☐5 (High)

Attachments

(Unnecessary) ☐1 ☐2 ☐3 ☒4 ☐5 (Helpful)

If Survey Coverage

(Shallow) ☐1 ☐2 ☐3 ☐4 ☐5 (Broad)

Contribution

(No New Results) ☐1 ☐2 ☒3 ☐4 ☐5 (Significant)

*What are the contributions of the paper?

*What are the additional ways in which the paper could be improved?





专题五 论文写作

5.1 论文结构剖析

5.2 论文典型元素

5.3 论文写作方法

5.4 实例讲解

- 论文写作工具
- 论文写作时间安排
- 如何写读书笔记
- 如何做PPT





5.1 论文结构剖析

- 题目
- 作者
- 摘要
- 简介
- 研究方法
- 实验数据
- 相关工作
- 结束语
- 致谢
- 参考文献
- 附件





5.2 论文典型元素

- ✓ 术语和符号表
- ✓ 问题描述
- ✓ 本文贡献
- ✓ 本文结构
- ✓ 定义
- ✓ 定理
- ✓ 算法
- ✓ 实例
- ✓ 公式
- ✓ 图
- ✓ 表





5.2 论文典型元素

✓ 图

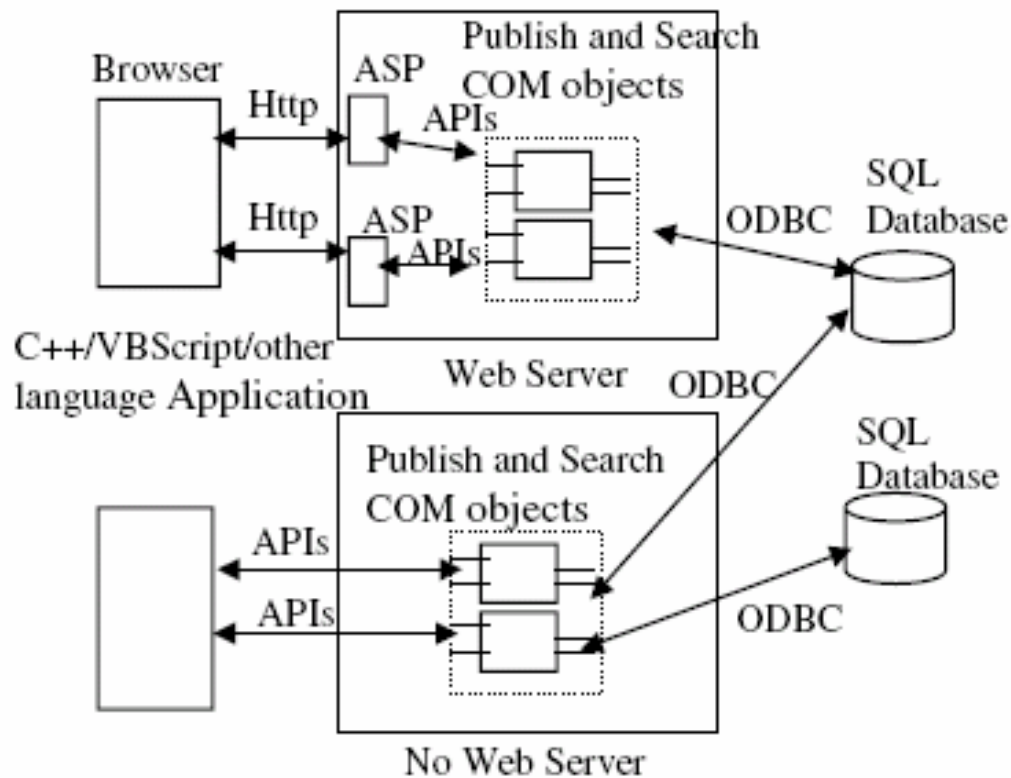


Figure 2. Architecture of DBXplorer





5.2 论文典型元素

✓表

HashVal	ColId
v ₁	c ₁
v ₂	x
v ₃	x
v ₄	x
v ₅	c ₂

Table 1. Uncompressed hash table

NewColId	ColId
x	c ₁
x	c ₂

Table 3. ColumnsMap table

HashVal	ColId
v ₁	c ₁
v ₂	c ₁
v ₃	c ₁
v ₄	c ₁
v ₂	c ₂
v ₃	c ₂
v ₄	c ₂
v ₅	c ₂

**Table 2.
Compressed hash table**





5.2 论文典型元素

✓ 算法

Algorithm Computing a Pub-Col Symbol Table

Inputs: A database

Outputs: A symbol table S and ColumnsMap table

//Compute hash table S :

Set S to empty

Scan database, and for each keyword K in column c

Insert ($\text{hash}(K)$, c) into S if it does not already occur

//Compress S using Algorithm CP-Comp:

Set table ColumnsMap to empty

Compute $\{\text{ColId}_1, \dots, \text{ColId}_p\}$ and $\{\text{HV}_1, \dots, \text{HV}_p\}$

For $i = 1$ to p

If $|\text{HV}_i| * |\text{ColId}_i| > |\text{HV}_i| + |\text{ColId}_i|$

Remove from S all entries involving HV_i

Create artificial column x_i

For each v in HV_i , insert (v , x_i) into S

For each c in ColId_i , insert (c , x_i) into

ColumnsMap

Output S and ColumnsMap

Figure 3. Constructing Pub-Col symbol table





5.2 论文典型元素

✓ 实例

For example, consider the schema graph G in Figure 4 over five tables. Let the query keywords be $\{K_1, K_2, K_3\}$. The black nodes represent the MatchedTables set, while the rest are white. Assume that K_1, K_2 and K_3 all occur in T_2 (in different columns), K_2 occurs in T_4 , and K_3 occurs in T_5 . The four possible join trees are shown on the right (including the singleton node, T_2). In contrast, for example the tree induced by $\{T_4, T_3, T_5\}$ cannot be a join tree, as these tables do not contain all keywords.

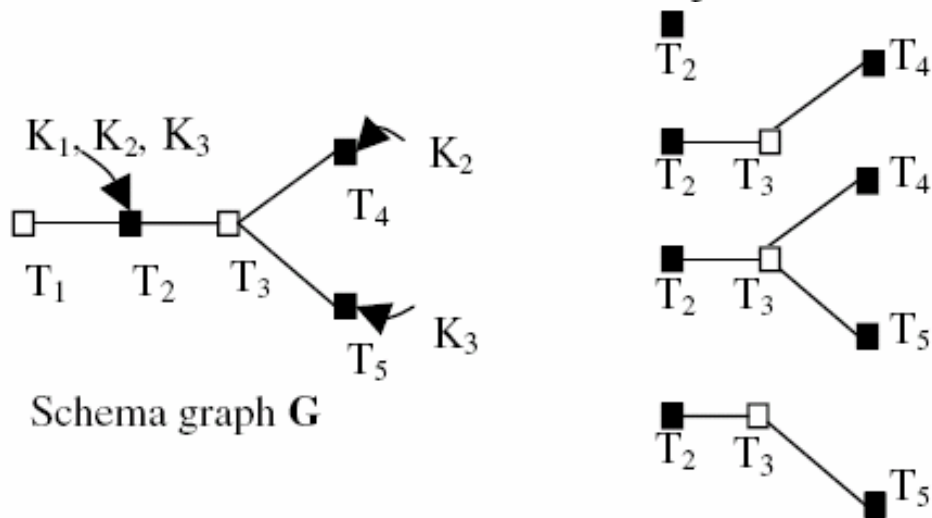


Figure 4. Join trees





5.2 论文典型元素

✓ 本文结构

The rest of this paper is organized as follows. In Section 3, we present an overview of DBXplorer. Sections 4 and 5 describe the preprocessing component responsible for creating the symbol table. Section 6 describes the search component that answers keyword searches once the symbol table has been built. Section 7 discusses extensions needed for generalized keyword matches described above. Section 8 presents experiments that demonstrate the effectiveness of our solution. An appendix containing screen shots of user interactions with DBXplorer is included to provide a better feel for the front-end.





5.2 论文典型元素

✓ 本文贡献

The contributions of this paper are as follows:

1. We outline a framework for keyword querying of relational databases. Our framework makes joins implicit and transparent, and incorporates notions of proximity and prestige when ranking answers.

There has been a fair amount of earlier work on keyword querying of databases, including [6, 7, 12, 13]. We describe the connections of the BANKS model to related work on keyword search in Section 6.

2. We present novel, efficient heuristic algorithms for executing keyword queries.
3. We describe key features of the BANKS system.





5.2 论文典型元素

✓ 定义

Definition 1 (Joining network of tuples) *A joining network of tuples j is a tree of tuples where for each pair of adjacent tuples $t_i, t_j \in j$, where $t_i \in R_i, t_j \in R_j$, there is an edge (R_i, R_j) in G_u and $(t^i \bowtie t^j) \in (R_i \bowtie R_j)$.*

Definition 3 (Joining Network of Tuple Sets) *A joining network of tuple sets J is a tree of tuple sets where for each pair of adjacent tuple sets R_i^K, R_j^M in J there is an edge (R_i, R_j) in G_u .*





5.2 论文典型元素

✓ 公式

The tuple sets are obtained from the basic tuple sets using the following formula.

$$R_i^K = \bigcap_{k \in K} \bar{R}_i^k - \bigcup_{k \in \{k_1, \dots, k_m\} - K} \bar{R}_i^k \quad (2)$$





5.2 论文典型元素

✓ 定理

Theorem 4 T_{max} is unbounded if and only if G has one of the following properties:

- There is a node of G that has at least two incoming edges.
- G has a directed cycle.





5.2 论文典型元素

✓ 问题描述

问题描述：假设当前的实视图集合是 M ，新到达的查询是 q ，用户的访问具有局部性和可导航性，实视图的动态选择问题就是，采用一定的视图替换策略，用与查询 q 相关的视图来替换 M 中一部分与 q 最不相关的视图。

Problem Statement. Given a set of frequent closed patterns $CP = X_1, X_2, \dots, X_m$ that are mined from database $DB = t_1, t_2, \dots, t_n$, pattern summarization is to group the m closed frequent patterns into k pattern profiles, each of which is represented with a probability model.





5.2 论文典型元素

✓ 术语和符号表

Acronym	Term
GVW	Global View Window
PVW	Public View Window
UVW	User View Window
CV	Completely-aggregated View
PV	Partly-aggregated View
CVD	Completely-aggregated View District
PVD	Partly-aggregated View District
CVS	Completely-aggregated View Set
PVS	Partly-aggregated View Set

Table 1. The term acronym list





5.3 论文结构剖析

5.3.1 题目

5.3.2 作者

5.3.3 摘要

5.3.4 简介

5.3.5 研究方法

5.3.6 实验数据

5.3.7 相关工作

5.3.8 结束语

5.3.9 致谢

5.3.10 参考文献





5.1.1 题目

基本要求：简明、达意

字数要求：《软件学报》规定题目不能超过20个中文字符

实例：

《应用于移动互联网的Peer-to-Peer关键技术》

. 《基于分离逻辑的程序验证技术》

《从不确定图中挖掘频繁子图模式》

《基于蚁群优化解决传感器网络中的能量洞问题》

《Building a Text Classifier by a Keyword and Wikipedia Knowledge》

《XTree: A New XML Keyword Retrieval Model》

《Constructing PEKS schemes secure against keyword guessing attacks is possible?》





5.1.2 作者

作者信息：姓名、单位、城市、邮编、E-mail

中文姓名：林子雨 方正

英文姓名：LIN Zi-Yu Ziyu Lin Fang Zheng Zheng Fang

相关问题：

- 共同作者
- 通讯作者和第一作者
- 作者排名





5.1.3 摘要

- **定义**：对文章内容准确概括而不加诠释或评论的简短陈述。
- **重要性**：结构清晰、内容完整的摘要，易于(1)被二次文献数据库收录(EI等数据库)；(2)引起读者的兴趣。
- **内容**：包括研究目的、方法、结果和结论等，即信息性摘要。
- **如何做**：按下面(1)(2)(3)的要求进行撰写，然后按顺序把它们合成在一起，就形成了最终的摘要内容。别忘记在合成时注意保持层次之间语句的衔接和连贯性。
 - (1) 目的：主要说明作者写此文的目的，或本文主要解决的问题。
 - (2) 过程及方法：主要说明作者主要工作过程及所用的技术或方法，也包括众多的边界条件、使用的主要设备和仪器。
 - (3) 结果或结论：作者在此工作过程最后得出的结果和结论，如有可能，尽量提一句作者结论和结果的应用范围和应用情况。

摘要范例：【范例论文】





5.1.4 简介

- 引出研究话题
- 指出当前主要研究方法和不足
- 提出自己的研究方法和优点
- 给出文章的结构

论文范例：[基于形状特征k-d树的多维时间序列相似搜索](#)





5.1.5 研究方法

论文的核心内容，阐述自己的研究思路和具体方法，包括理论框架、定理、算法等内容

写作套路

- ✓ 由框架性描述到细节论述
- ✓ 去繁就简，去粗取精
- ✓ 问题采用形式化定义
- ✓ 定理有证明
- ✓ 算法有分析
- ✓ 难懂内容有实例说明
- ✓ 格式要规范（定义、算法、定理）





5.1.6 实验数据

通过实验数据，证明论文方法的优越性

- 实验环境
- 测试数据集
- 对照方法
- 实验方法
- 实验数据展示
- 实验结果说明





5.1.6 实验数据

#keyw	JNTS_K	JNTS_L	CNs	neTS's
2	25	5.355	4.485	2.96
3	55.22	13.86	9.27	4.35
4	85.69	33.88	24.03	5.91
5	101	37.3	26	7.12

(a) Fix maximum candidate networks' size to 3

MaxCNsize	JNTS_K	JNTS_L	CNs	neTS's
1	0.95	0.95	0.95	2.96
2	3.72	2.36	2.12	2.96
3	29.22	4.74	3.7	2.96
4	422.88	10.36	6.4	2.96
5	6941	24.75	11.45	2.96

(b) Fix number of keywords to 2

MaxCNsize	JNTS_K	JNTS_L	CNs	neTS's
1	0.59	0.59	0.59	4.35
2	5.01	3.91	3.35	4.35
3	55.22	13.86	9.27	4.35
4	639.61	50.49	29.51	4.35
5	7532	223	103.66	4.35

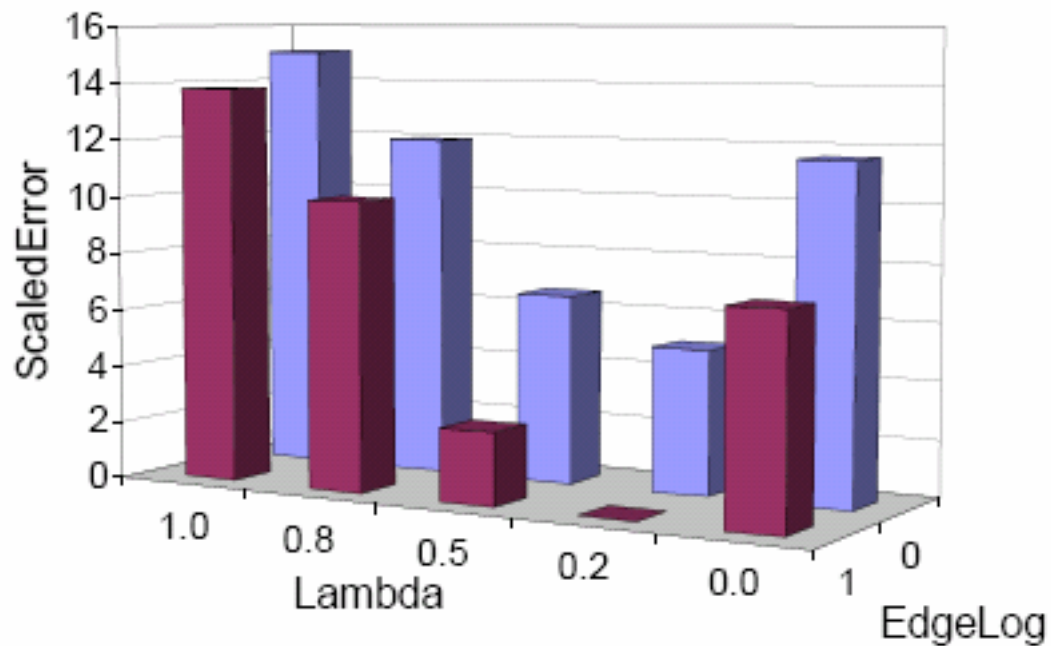
(c) Fix number of keywords to 3

Figure 9: Evaluation of the candidate network generator





5.1.6 实验数据





5.1.6 实验数据

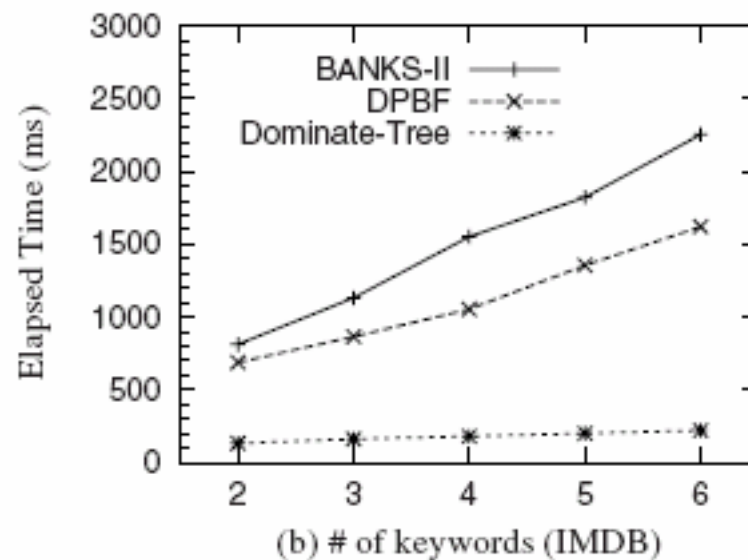
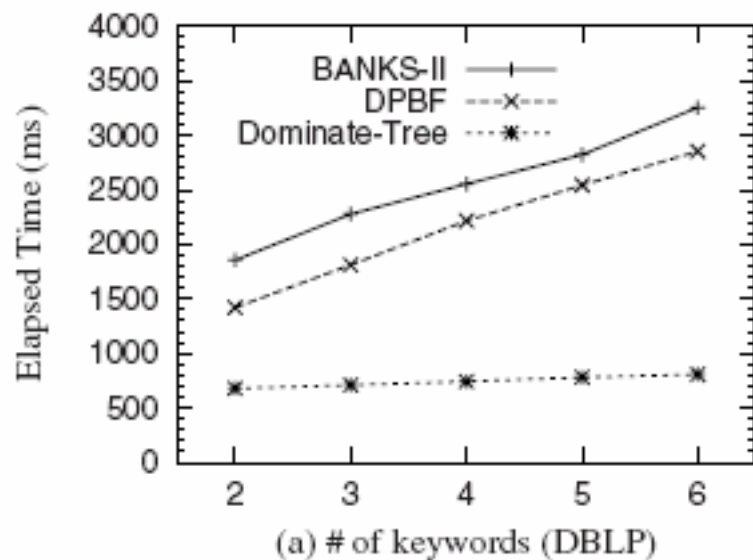
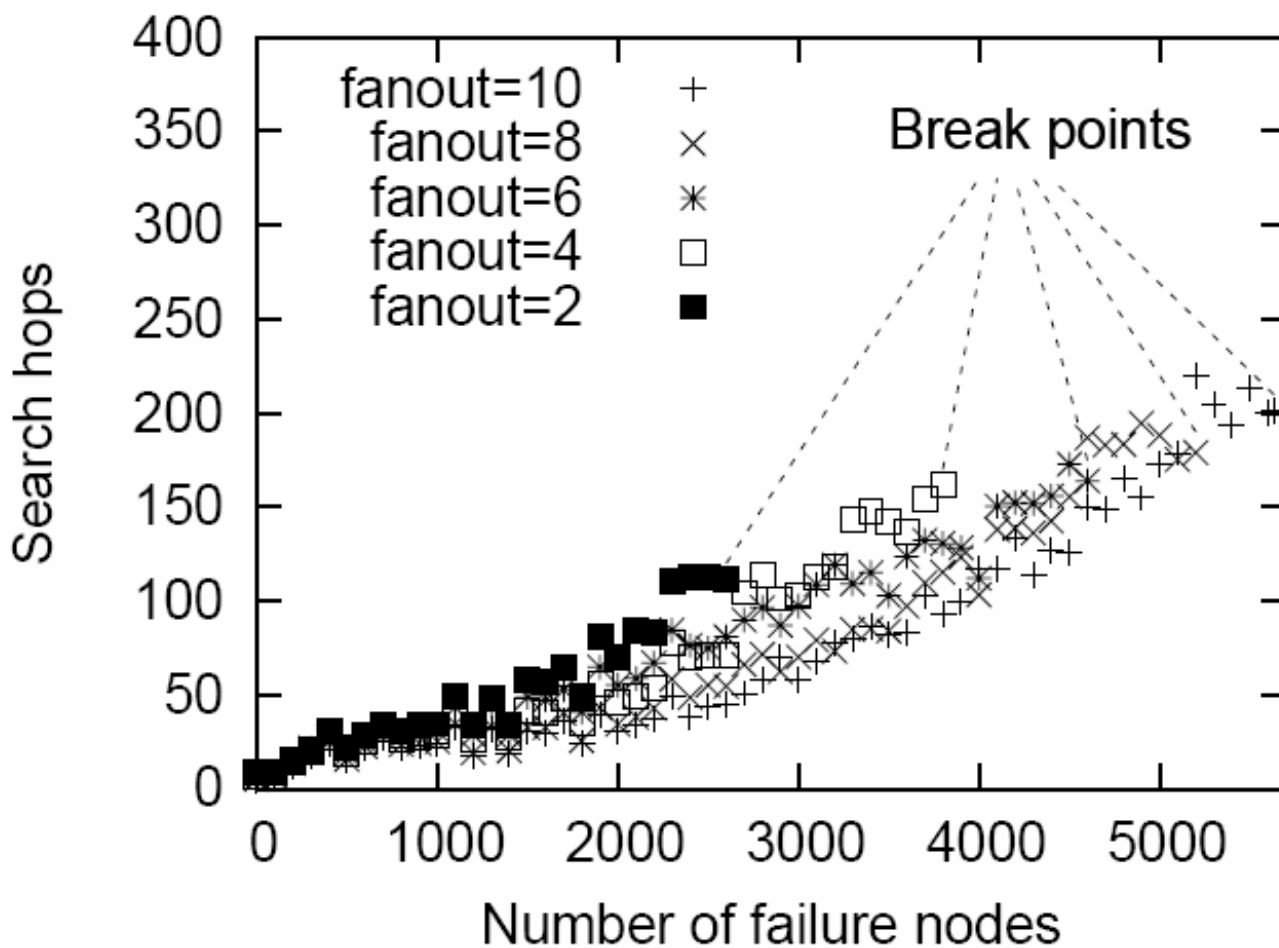


Fig. 2. Elapsed Time of Identifying Top-100 Results

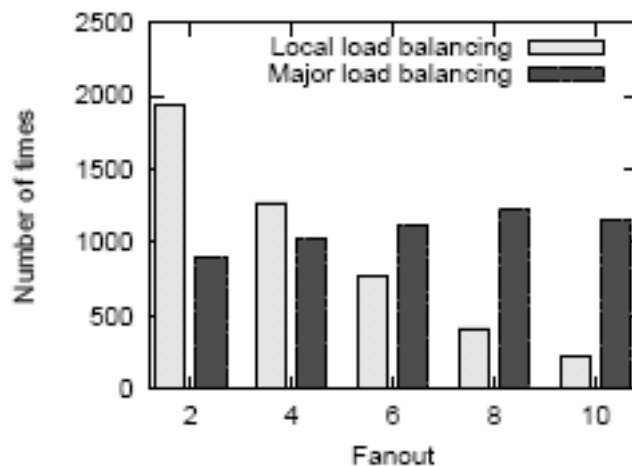


5.1.6 实验数据

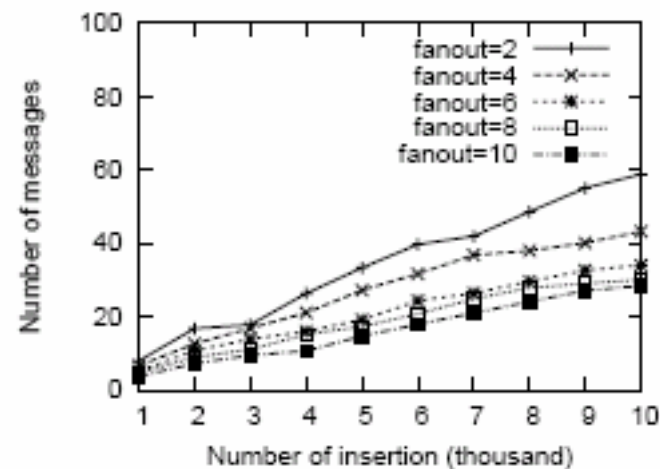




5.1.6 实验数据



(a) Number of times each load balancing scheme is invoked



(b) Average additional messages required for doing load balancing





5.1.6 实验数据

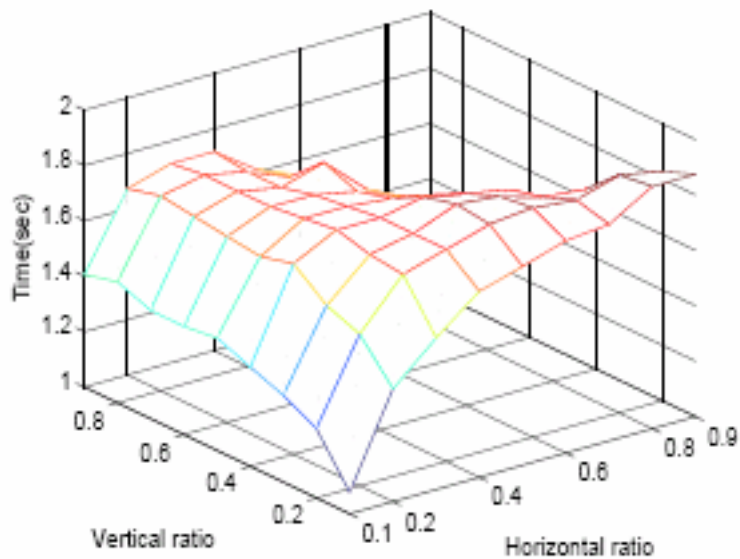


图 4.29 同时变化水平和垂直比例





5.1.6 实验数据

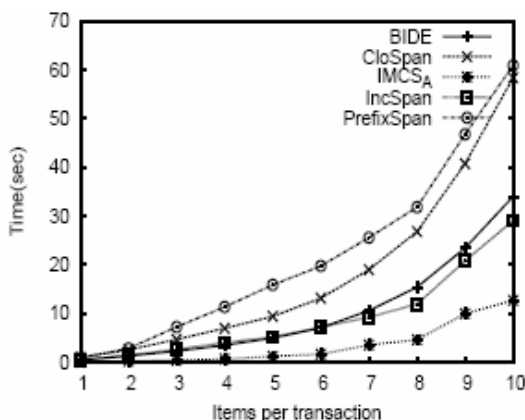


图 4.13 每个事务的项数变化时算法运行情况

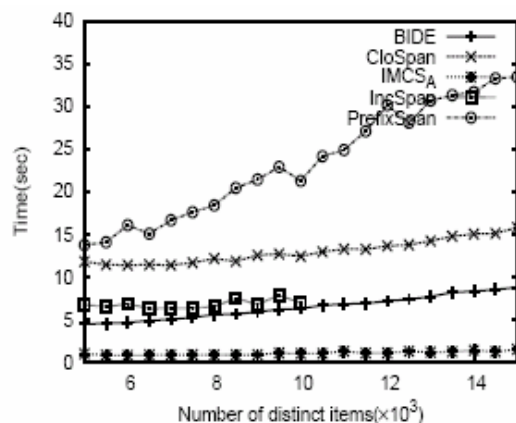


图 4.14 数据集中不同项的个数改变时算法运行情况

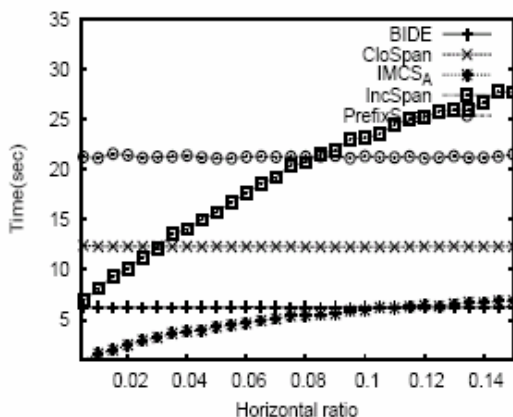


图 4.15 水平比例变化时算法运行情况

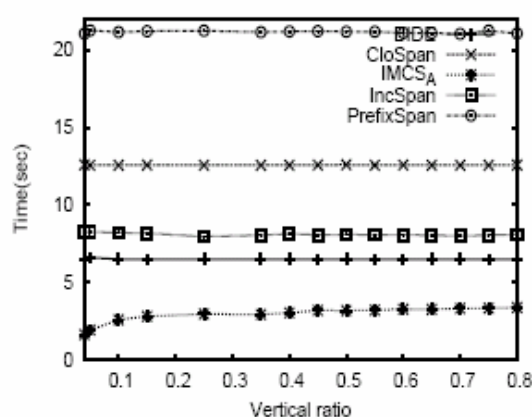


图 4.16 垂直比例变化时算法运行情况





5.1.6 实验数据

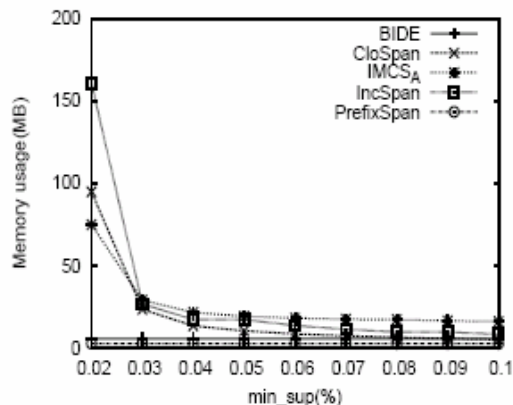


图 4.19 内存使用情况

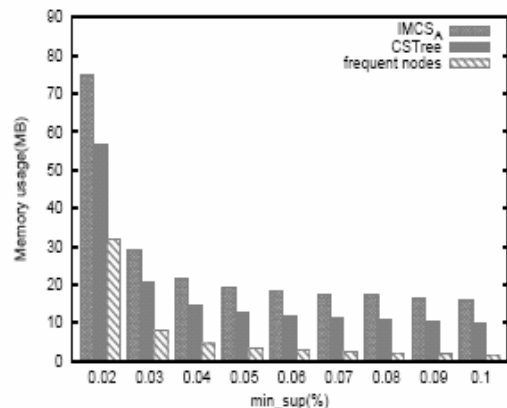


图 4.20 CSTree内存使用

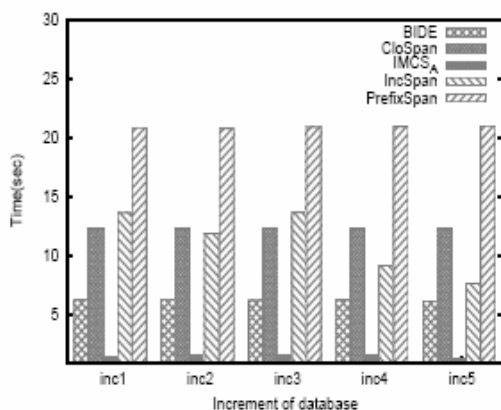


图 4.21 累积更新情况

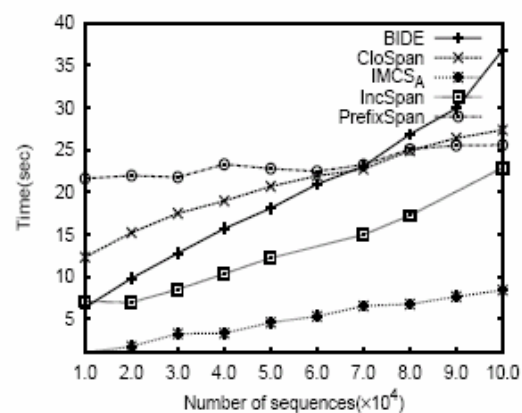


图 4.22 序列数量变化时算法运行情况





5.1.7 相关工作

介绍本研究领域相关的研究成果

- 全面性
- 新颖性
- 分主次
- 有对比
- 有归纳





5.1.8 结束语

- 总结本文研究内容
- 指出未来研究方向

实例[HuangSZ06]

相似搜索是多维时间序列应用的一个重要方面.现有的一些方法大多都采用欧几里德距离作为计算相似度的依据.但欧几里德距离不能顾及序列本身局部的起伏变化,极大地降低了它在比较相似度方面的准确性,从而会导致返回错误的结果.针对这一问题,本文提出一种抽取序列(或子序列)的形状特征的方法,并将形状特征向量与检索结构k-d树有机地结合,使得在搜索k-d树的同时实现了序列(子序列)的局部变化匹配,以弥补欧几里德距离的不足,从而极大地提高了查询效率和正确率.我们下一步的工作是:针对高维数据,考虑使用数据降维方法;在相似度比较时,考虑不同数据维所占的权重.





5.1.9 致谢

- 提供课题基金支持信息
- 向对论文做出贡献的相关人员表示感谢

原则：简洁

实例1

This research was supported by the National Natural Science Foundation of China under Grant No. 60603045, No. 608703063, and the National High-tech Research and Development Project of China under Grant No. 2007AA01Z153.

实例2[YangSWC09]

We express our grateful thanks to Prof. Jian Pei from Simon Fraser University for his valuable suggestions on this work.





5.1.10 参考文献

列出本文引用的相关参考文献

规范性要求

①会议论文	引用号	②期刊论文	引用号	③书籍	引用号
作者姓名		作者姓名		作者	
标题		标题		书名	
论文集名称		期刊名称		版次	
会议地点		年份		出版地	
出版商		期号		出版者	
年份		卷号		年份	
页数		页数		页数	

① 期刊类

如: Sayah JY, Kime CR. Test scheduling in high performance VLSI system implementations. IEEE Trans. on Computers, 1992,41(1):52-67.

作者 论文题目 刊名

② 书籍类

如: Eissen HN. Immunology: An Introduction to Molecular and Cellular Principles of the Immune Responses. 5th ed., New York: Harper and Row, 1974. 40-56.

作者 书名 版次 出版地 出版者

③ 会议录、论文集、论文汇编: 多编者用eds; 文集名中出现的缩写词要加注全称。

如: Hunninghaks GW. _____ In: Harris CC, ed. _____ New York: Academic Press, 1980. 54-56.

作者 题目 论文集编者 文集名 出版地 出版者





专题六：硕士论文

6.1 硕士论文工作程序

6.2 硕士论文选题

6.3 硕士论文评价





6.1 硕士学位论文工作程序

- (1) 文献阅读和问题研究——找题
- (2) 选题与定题——与导师讨论
- (3) 文献研究
- (4) 初纲——与导师讨论
- (5) 细纲（可以初纲合并）
- (6) 初稿：不是草稿
- (7) 导师审阅——2-3个工作周
- (8) 导师反馈——约见学生讨论
- (9) 修改初稿——根据导师意见修改
- (10) 导师二审——.....
- (11) 定稿和打印——提交论文、等候答辩
- (12) 论文送审——专家审议
- (13) 准备答辩——制作Presentation，了解注意问题
- (14) 正式答辩——介绍、陈述、提问、回答、表决、结束

摘自《厦门大学吴世农副校长给厦大2009年新聘教师的讲座》





6.2 硕士论文选题

论文选题应注意的原则

- 1、确定论文的基本类型：学术型或应用型
- 2、学术型：作好文献收集、阅读和评述；
应用型：理论联系实际，联系自己的体会或工作经验；
- 3、研究的问题不要太大、太宏观、太复杂；
“可以通过研究解决的问题”；
- 4、避免重复选题，即使选题重复，研究的角度或切入点应有所不同；
- 5、选题前后要对该领域或“相关研究”有所了解——难点、热点、关键点；
- 6、把握和控制选题：可以“多选”，与导师和同学充分讨论，
基本把握研究要点、研究过程和预期结果。

摘自《厦门大学吴世农副校长给厦大2009年新聘教师的讲座》





6.3 硕士论文评价

序号	评价指标	评 价 要 素	评 判 分 数				综合得分 (权重1件判分型)
			(选择以下优秀、良好、合格或不合格四类之一, 然后打分)				
			优 秀 (90分以上)	良 好 (80-89分)	合 格 (70-79分)	不 合 格 (70分以下)	
1	选题 (权重: 10%)	论文的选题具有较大的理论意义或应用价值;研究方向和论题明确。					
2	文献评述 (权重: 10%)	文献收集和整理全面、准确、系统、即时, 综合反映该学科及相关领域的发展和最新成果, 归纳和评述正确。					
3	理论基础和专业知识 (权重: 10%)	论文反映了学生系统地掌握了与本学科专业相关理论合专业知识, 并能够应用专业理论和知识开展科学研究或解决实际问题。					
4	研究设计和研究方法 (权重: 10%)	论文的研究设计合理;研究方法科学、规范, 达到解决问题的目标。					
5	分析、研究和论证 (权重: 15%)	论文的分析框架合理、分析逻辑清晰、论证严谨。					
6	资料和论据 (权重: 10%)	研究的资料、数据或论据翔实、有效、充分。					
7	研究成果和结论 (权重: 10%)	论文的研究成果或结论可靠, 具有一定的理论价值或应用价值。					
8	改进和创新 (权重: 15%)	论文在基础理论、专业理论或研究方法等方面具有明显的改进或一定的创新。					
9	写作能力 (权重: 10%)	论文的结构合理、层次分明、概念清晰、格式(图表、公式、参考文献等)规范、语言表述准确、文笔流畅。					

摘自《厦门大学吴世农副校长给厦大2009年新聘教师的讲座》





附件

【参考文献】

[MaL05]马治国,李晓鸣.学术论文剽窃的认定及法律责任.重庆交通大学学报(社会科学版),2005:25(4), 68-73.

[YangSWC09]Bishan Yang, Jian-Tao Sun, Tengjiao Wang, Zheng Chen: Effective multi-label active learning for text classification. KDD 2009: 917-926.

[HuangSZ06]黄 河, 史忠植, 郑 征.基于形状特征k-d树的多维时间序列相似搜索.软件学报, 2006:17(10), 2048-2056.

[Zhou05]周志华.如何做研究, 如何写论文.

The background is a solid blue color. It features several faint, light-blue silhouettes of people. In the top left, a group of people is holding hands in a circle. In the top right, another group of people is standing together. On the right side, a large silhouette of a person is shown from the side, looking towards the center. In the bottom left, there are silhouettes of two people, one of whom appears to be holding a phone to their ear.

Thank You!