

深圳工作坊  
2015年8月28日

## How To Do High Quality Research And Write Acceptable Papers?

### “内在美”与“外在美”的追求

Michael Rung-Tsong Lyu

**吕荣聪**

Computer Science & Engineering Department  
The Chinese University of Hong Kong  
香港中文大学计算机科学暨工程系

1

## Agenda 大纲

- How to do high quality research?  
■ **如何做高品质的研究?**
- How to write acceptable papers?  
■ **如何写会被接受的文章?**

2

## Motivation: Why Should I Do Research? 为什么我要作科研?

- External drive 外在驱策
  - Award, degree and diploma 奖项、文凭、学位
  - Parents, teachers, friends 父母、老师、朋友
  - Peer pressure 同学压力
- Internal drive 内在动机
  - Sense of honor and responsibility 名誉与责任感
  - Strong ambition (self-expectation) 自我期许
  - Research interest (sense of achievement/fulfillment, curiosity) 兴趣, 成就感, 好奇心

3

## Two Major Research Elements and How to Approach Them 研究要素

- Define a problem 定义问题
  - To find it 找到问题
  - To survey it 查阅问题
  - To express it 表达问题
  - To nurture it 深入问题
- Find the solution 找到答案
  - To derive it 推论答案
  - To validate it 证实答案
  - To complete it 完成答案
  - To elaborate it 推展答案

4

## Research Problem Selection 选择问题

- Good research largely depends on the selected problem **好的研究多数取决于所选择的问题**
  - 90% of a research job is done when you find a good problem.
  - A good problem is difficult to find **好的问题不好找**
    - Not too easy or too difficult **不能太难也不能太容易**
- How to select a problem? **如何找到问题**
  - Is it an old problem or a new problem? **是新的还是旧的**
    - Usually, new problems have more opportunities **新问题较有机会**
  - Is it a significant problem? **有意义的问题吗？**
    - Practically important yet technically challenging **实际重要且有科研上的挑战**

5

## The Key Ingredients of Research: Contribution (Where is the beef?) 贡献

- One major contribution is better than many small ones **一个重要贡献胜过许多小的贡献**
- What is the contribution type?
  - Knowledge discovery **发现**
  - Knowledge invention **发明**
  - Knowledge integration **整合**
  - Knowledge application **应用**
- Idea! Idea! Idea! **创意！创意！创意！**
- Identify, describe, and demonstrate the big idea **识别，描述，展示你的创意**
- Asking the right problem, then asking the problem right **用对的方式问对的问题**
- Innovation! Innovation! Innovation! **创新，创新，创新**

6

## The Right Problem and the Right Way to Ask Problem **用对的方式问对的问题**

- Asking the problem right can lead to asking the right problem **用对的方式问会引导出问对的问题**
- Can we predict reliability? **预测可靠性？**
- Can we predict reliability with models? **用模式预测？**
- Can we predict reliability of software? **软体可靠性可测？**
- Can we predict reliability with user experience? **以用户体验来预测？**

## Innovation! Innovation! Innovation! **创新，创新，创新**

How to Find Research Problems?

**如何找到研究问题？**

## 1. New Solution to Old Problems 老问题，新答案

- New solution to a reduced problem **老问题，新答案**
  - Fermat's last theorem **费玛最后定理**
  - For all positive integers a, b, c and n, there is no solutions to  $a^n + b^n = c^n$  when  $n \geq 3$
  - Java and C alias/pointer analysis example (PLDI2013)
  - Title: DBSCAN Revisited: Mis-Claim, Un-Fixability, and Approximation (Gan and Tao), SIGMOD 2015

9

## 1. New Solution to Old Problems 老问题，新答案

- New solution from the same area **同一领域，新答案**
  - Coolstreaming (INFOCOM 2015 Test of Time Paper Award)
- New solution from other areas **其他领域，新答案**
  - Recommendation applied to Software Reliability Engineering (ICSE2010)

10

## 2. New Problems (Usually with a Twist) 新问题，加点更新

- Extension 扩展
  - $M^4$  (Maxi-Min Margin Machine) work
- Summarization 总结
  - my thesis – Confucius
- Variation 变化
  - linear combination model

11

## 2. New Problems (Usually with a Twist) 新问题，加点更新

- Refinement 更新
  - Refine existing problem: Concept of social recommendation
  - Refine existing approach: Missing data prediction with CF

12

## Reliability Prediction of Web Services

- Approach 1: Neighborhood-based approach – to consider [users](#)
- Approach 2: Model-based approach – to consider [data sparsity](#)
- Approach 3: Time-aware approach – to consider [temporal](#) factor
- Approach 4: Network coordinate based approach – to consider [spatial](#) factor
- Approach 5: Ranking-based approach – to consider [ranking](#)
- Approach 6: Reputation-aware approach – to consider [reputation](#)

## Reliability Prediction of Web Services

- Approach 1: Neighborhood-based approach – to consider [users](#) [ICSE'10, ACM TOSEM]
- Approach 2: Model-based approach – to consider [data sparsity](#) [IEEE TSC'13]
- Approach 3: Time-aware approach – to consider [temporal](#) factor [ISSRE'11]
- Approach 4: Network coordinate based approach – to consider [spatial](#) factor [ICWS'12]
- Approach 5: Ranking-based approach – to consider [ranking](#) [TPDS'13]
- Approach 6: Reputation-aware approach – to consider [reputation](#) [SCC'13]

## 2. New Problems (Usually with a Twist) 新问题，加点更新

- Generalization 通用
  - nonparametric work
- Specialization 专门
  - online algorithms

15

## 3. New Areas 新的领域

- Exploration 探索
  - Data Analytics research
- Cross Disciplinary Research 跨越领域
  - Machine Learning in Software Engineering and Distributed Systems
- Disruptive Research 颠覆性研究
  - Quantum computing (*new theory*)
  - Internet and WWW (*new user experience*)
  - Apple's revolution (*new business model*)

16

## Good versus Bad Research Problems

### 好的研究问题及坏的研究问题

- Good research problem **好问题**
    - After the research, more people will be doing it – you opened the door **人会继续做，因你开了一扇门**
    - After the research, nobody can improve it – you closed the door **没人可以再改进，因你关了一扇门**
  - Bad research problem **坏的问题**
    - Nobody will follow the research **没人会跟进**
    - Nobody really cares the research **没人在乎**
  - Research subject hotspot index:  $df(t)/dF(t)$   
**研究课题热点指标**
- 
- $f(t)$ =No. of top papers and  $F(t)$ =No. of total papers

## More about Ambition **雄心**

- Principle of “aim high, accept low”
  - Use problem selection as example
    - Aim high **期望高超**
      - Do not patch a small hole left by leading researchers
      - Find a more fundamental problem which may have a long impact
    - Accept low **接纳妥协**
      - If it is difficult to find a fundamental problem, then we need a compromise
      - Advice from professor/colleague is important
-

## Literature Survey 文献考察

- Research lifecycle: imitating, remembering, analyzing and innovating 模仿, 记忆, 分析, 创新
- Use tools 使用工具
  - Trace backward 回顾
    - Tutorial paper and reference list
  - Trace forward 前瞻
    - Use Google scholar to find papers that cite the current work
- Proactive vs. passive reading 积极及消极阅读
  - Reading with a critical attitude 评断/独立思考心态
  - Reading according to your own agenda 有自己的时间表
  - Reading between lines (not only what was said but what was not said) 读出字里行间的信息
- Form a study group 成立学习小组

19

## Nurturing Good Taste 培养好的品味

- There are many mediocre papers published 平庸文章
  - Do not ruin your taste by poor-quality papers 不要破坏品味
- Read selectively 精选地读
  - Highly cited papers and papers from first-tier journals and top top-ranked conferences 被引用的文章, 高水平的期刊及会议
- Classification of papers 文章分类
  - Type A: 80% understanding (main idea, solution method and main results) 理解思路, 方法及结果
  - Type B: 50% understanding (idea & results) 理解思路及结果
  - Type C: 20% understanding (only introduction) 理解简介
- Learn to appreciate good papers and criticize poor papers 学会欣赏好的文章, 评断不好的文章

20

## Monitoring Activities of Leading Research Group in Your Field **关注你的专业领域里面的主要研究团队**

- Identify leading research groups in your field  
**找出你专业的主要研究团队**
- Find out their recent research focus  
**跟进他们最近的研究课题**

21

## Two Major Research Elements and How to Approach Them **研究要素**

- Define a problem **定义问题**
  - To find it **找到问题**
  - To survey it **查阅问题**
  - To express it **表达问题**
  - To nurture it **深入问题**
- Find the solution **找到答案**
  - To derive it **推论答案**
  - To validate it **证实答案**
  - To complete it **完成答案**
  - To elaborate it **推展答案**

22

## How to Find the Solution?

### 如何找到解答

- Five typical steps in scientific research
  - Observation/Exploration      观察/探索
  - Assumption/Proposition      假设/提议
  - Theory/Methodology      理论/方法
  - Verification/Experimentation      验证/实验
  - Conclusion/Elaboration      结论/推演
  
- Thinking outside the box      超越框框限制的思考

23

## Research Environment 研究环境

- Large group can be a blessing      大组可以祝福你
  - More resourceful in terms of interaction (now) and networking (future)      目前多有互动，将来多有联结
- Senior students can be very helpful to junior students      学长对学弟妹很有帮助
  - Experience sharing & encouragements      分享和鼓励
  - More tolerant to mistakes      容忍错误
  - More accessible      容易交谈
- Good versus bad environments      好的环境和差的环境
- Each group has its own culture      各有文化
  - Each colleague is a teacher; learn from them      每个同事都是老师
  - Every big project comes from a small team; find the ones who suit you      大的课题来自小组，找到适合你的课题

24

## Guidance and Feedback 指导与回应

- Role of Advisor 导师的角色
  - Joint decision on problem selection 一起决定研究问题
  - Set up the research standard 设定研究水平
  - Help when students get stuck 卡住时的帮助
    - Find out why 找出为什么
    - Re-directing 重订方向
- Feedback on research results 研究成果的回应
  - Positive and negative feedback 正面及负面
- Help in oral presentation and written reports 口头报告及书面报告的帮助

25

## Agenda 大纲

- How to do high quality research?
- 如何做高品质的研究?
  
- How to write acceptable papers?
- 如何写会被接受的文章?

26

## Writing 写文章

- Critical to the sale of your ideas/results  
**很重要！明确地指出文章的卖点（想法/结果）**
- Paper organization **架构：文字，图表**
  - Proper arrangement of texts, figures and tables
- Multi-pass writing style
  - 1<sup>st</sup> pass: Detailed outline **详细大纲**
  - 2<sup>nd</sup> pass: Rapid writing **快速写成**
  - 3<sup>rd</sup> pass: Fine Fine-tuning **详细修改**
  - 4<sup>th</sup> pass: cross-reading **交叉阅读**

27

## Writing Procedure 写作过程

- Carefully determine the paper title **仔细决定题目**
- Proper use of names and notations **正确引用**
  - **Refer to Irwin's website**
- Tell them what you are going to do, tell them what you are doing, tell them what you have done. **你想作什么，正在作什么，已经作了什么**
- Motivation! Motivation! Motivation! **动机！动机！动机！**

28

## Title 文章标题

- A title needs to be catchy but precise
  - Like Likes Like
  - “All Models Are Wrong; Some Are Worse Than Others”
  - STELLAR: Spatial-TEmporaL LAtent Ranking for Successive Point-of-Interest Recommendation.
- Use acronyms
  - KEEP: The Knowledge and Education Exchange Platform

29

## Motivation! Motivation! Motivation! 动机! 动机! 动机!

- The introduction is by far the most important section in the entire paper, especially for conferences. 简介最重要
- Reviewers are always very busy. 评审很忙
- If a reviewer can reject your paper without reading it all, it saves time! 可以不读就知道要拒绝, 那很省时间
- The introduction is the first section they read, so make sure your paper does not get killed in Section 1. 不要在简介一被读完就出局
- “5 years ago I used to write the introduction last. Now it is always the first section I write.” “以前我是最后写简介, 现在我总是最先写”

30

## Strong Statements Are Dangerous

### 强烈声明的危险

- Be very careful when you make strong statements about some research issue: there are people that think otherwise.
- Be especially careful when taking position on some hotly debated topics in the community, like: **尤其是会引起争论**
  - Supervised learning vs. non supervised learning
  - Parametric vs. non parametric
  - Statistical vs. analytical
  - Partitioned vs global multiprocessor scheduling
  - Hard real-time wireless
  - Testing vs static analysis
  - Etc. etc. etc.
- Instead of saying “X is black”, say “X is usually black, but in some cases that are not considered in this paper it is white”. **X通常是黑色，不在本篇考虑中的情况下也会是白色**

31

## ... But If You Are Confident, Go For It!

### 但是，你若有信心。 。 。 。 。 。 。

- However, high impact papers are those that successfully challenge existing preconceptions. **有影响力的文章是成功地挑战现存概念**
- So do not be shy when you state the main contribution of your paper! **大胆声明本篇文章的主要贡献**
  - If it is somehow controversial, you might have some troubles getting the paper accepted at first, but it is well worth in term of impact. **争论性的文章不一定被接受，但是引起反响**
  - If it is not, you should still stress your contribution so the reviewer gets more interested in the paper. **强调你的贡献让评审有兴趣**
- Just be sure to prove your point well enough; the keyword here is “**successfully challenge.**” **成功地挑战**

32

## Criticize Your Writing in the Reviewer's View 从评审的眼光来读你的文章

- A main factor of your success is to know how others think and feel **知道别人怎么看怎么想**
- Reviewers' mindset: "You are assumed guilty until proven innocent" **你若不能证明你是清白的，他们将对你充满怀疑**
- Remind instead of assume, but don't humiliate their intelligence **提醒他们，不是想当然耳；但别把他们看扁了**
- Proper use of citation **正确引用**
- Clearly and articulately indicate your contributions **贡献**
- Criticize yourself first, and leave reviewers no room for further criticism **自己先评断，使得他人没什么好批评**
- Remember, our reviewers are hostile ... **假设评审是你的敌人**
- But they should not be malicious ... **但他们不应心怀恶意**

33

## You Cannot Make Everybody Happy 不能使得每个人开心

- Different people are looking for different things. **不同人不同想法**
- Also they are often biased. **他们常有偏见**
- You must accept that it is simply impossible to make everybody perfectly happy; you are forced to make trade-offs. **不可能个个开心，你要取舍**
  - For the same reason, take all people's reaction with a grain of salt.
- The key: two half glasses of water are better than one full and one empty glass here. **两瓶半杯水比较好**
  - Just one negative review is enough to kill a conference paper.
- The lesson: **bad results can turn out good, so don't loss your heart. 失败可以成为祝福，不要怀忧丧志**

34

## Plagiarism 作弊/抄袭

- Academic honesty is everything; we can't emphasize it too much **学术诚实是一切**
- Plagiarism is a severe problem **抄袭/作弊非常严重**
- If you are not caught early, you will be caught eventually, only with a higher price to pay **迟早被抓，越晚越糟**
- Intentionally and un-intentionally **故意和不小心**
  - Properly cite and paraphrase **适当地引用或重述**
  - Always be critical **总是要小心**

35

## Conclusions 结论

- **对计算机科研而言，表达和内容一样的重要。所以外在和内在美都要追求**
- **好的科研成果要靠长期累积的功夫，不是一蹴可及**
- **创新的想法有时来自灵感，但要经过具体实现和验证**

***"Remember, Red, hope is a good thing, maybe the best of things, and no good thing ever dies."***

*Andy Dufresne, in "The Shawshank Redemption"*

**继续努力，不要轻易放弃！**

36