How to Write and Publish
Research Papers for the Premier Forums
in Knowledge & Data Engineering

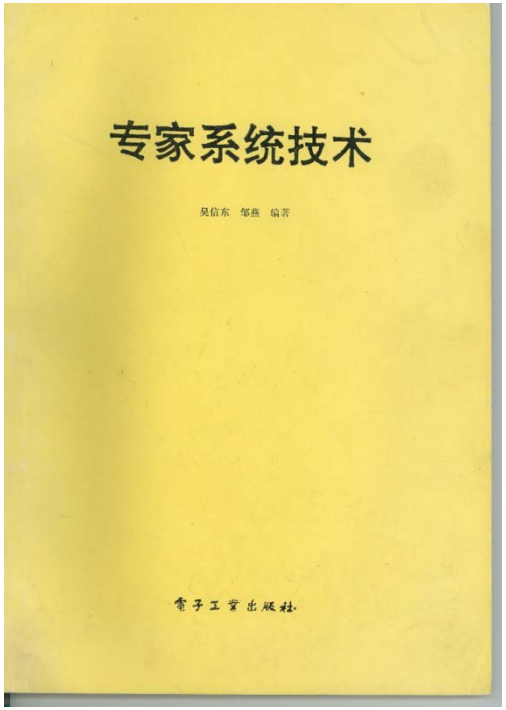
Xindong Wu (吴信东)

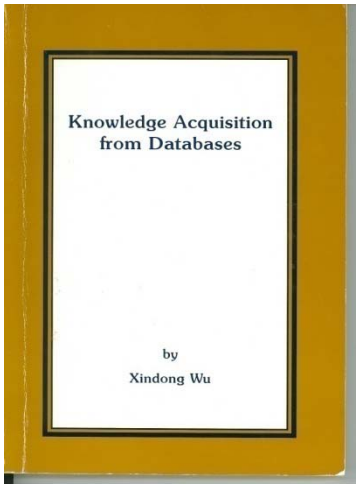
[1] Department of Computer Science,
University of Vermont, USA;

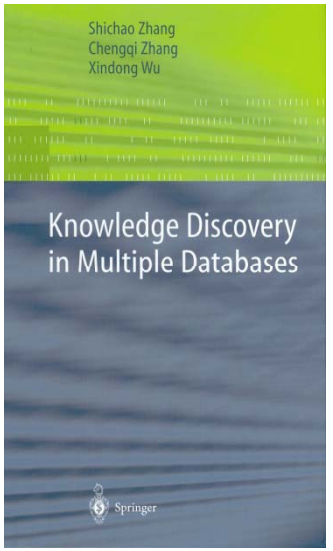
[2] 中国合肥工业大学计算机与信息学院

xwu@uvm.edu
www.cs.uvm.edu/~xwu

Deduction ⇨ Induction: My Research Background







…

1988 Expert Systems

1990 Expert Systems

1995 数据挖掘

2004 数据挖掘

…

Contents

- **Some TKDE and ICDM statistics**
- Scientific writing and paper structure
- What to know and how to write a top-quality paper
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- How to publish at ICDM and TKDE
- Paper reviewing and its feedback
- Summary of take-home messages

Focused Areas in Knowledge & Data Engineering

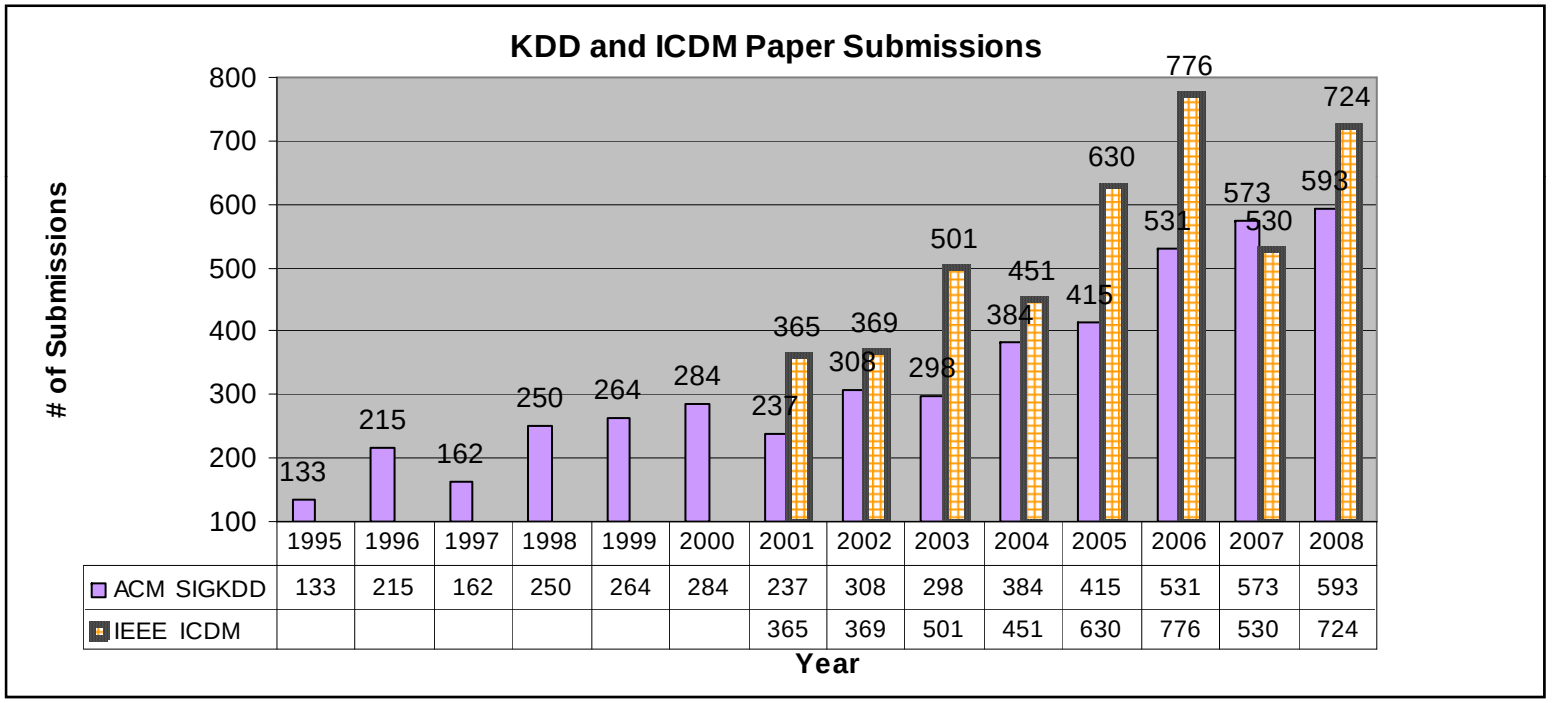
IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING

- Data Mining
  - Knowledge Discovery in Databases (KDD)
  - Intelligent Data Analysis
- Database Systems
  - Data Management
  - Data Engineering
- Knowledge Engineering
  - Semantic Web
  - Knowledge-Based Systems
  - Soft Computing

Major Forums in Data Mining

- **Conferences (conference publications are extremely important in Computer Science):**
  - **The birth of data mining/KDD: 1989 IJCAI Workshop on Knowledge Discovery in Databases**
    - **1991-1994 Workshops on Knowledge Discovery in Databases**
  - **1995 – date: ACM International Conferences on Knowledge Discovery in Databases and Data Mining (KDD)**
  - **2001 – date: IEEE International Conference on Data Mining (ICDM) and SIAM-DM (SDM)**
  - **Several regional conferences, incl. PAKDD (since 1997) & PKDD (since 1997)**
- **Journals (top journals vs high-impact journals):**
  - **Data Mining and Knowledge Discovery (DMKD, since 1997)**
  - **Knowledge and Information Systems (KAIS, since 1999, Impact Factor: 1.733)**
  - **IEEE Transactions on Knowledge and Data Engineering (TKDE, Impact Factor: 2.236)**
  - **ACM Trans. on Knowledge Discovery from Data (TKDD, since 2007, not SCI indexed)**
  - **Many others, incl. TPAMI, ML, IDA, …**

ACM KDD vs. IEEE ICDM



TKDE Submission Numbers and Acceptance Rates

2001	294	25.50%
2002	233	24.00%
2003	355	26.40%
2004	347	21.00%
2005	480	30.00%
2006	588	23.00%
2007	625	22.00%
2008	680	being accpt'd, 0.06% @ 1/23/09
Year	**New Submission #**	**(Current) Accpt Rate**

Contents

- Some TKDE and ICDM statistics
- **Scientific writing and paper structure**
- What to know and how to write a top-quality paper
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- How to publish at ICDM and TKDE
- Paper reviewing and its feedback
- Summary of take-home messages

Why Write a Scientific Paper

- Advance knowledge in your research field with evidence
- Explain your ideas and make them accessible to others
- Two key components in a research paper:
  - An explicit claim on your contribution on a research problem
  - Evidence to support your claim
- Your contribution can possibly be a refutation of a hypothesis on the research problem
- [Take-Home Message #1] **It is NOT enough to design yet another technique or system without convincing evaluation.**

What to Claim for a Scientific Paper

- Your technique solves a problem for the first time
- Your technique performs better, in one or more of the following dimensions [Alan Bundy, How-To Guides, homepages.inf.ed.ac.uk/bundy/howtos/writingGuide.html], than its rivals:
  - Behaviour: X has a higher success rate then Y or produces better quality outputs, e.g. shorter,easier to understand, more similar to human outputs, etc.
  - Coverage: X is applicable to a wider range of examples than Y
  - Efficiency: X is faster or uses less space than Y
  - Useability: Users find X easier to use than its rivals
- [Take-Home Message #2] **You should avoid claiming too many dimensions, but one or two with in-depth evidence.**

Typical Structure of a Research Paper (1)

- Title: Catchy and indicative of your research contribution
  - ICDM Data Mining on ICDM Paper Submissions: The shorter a paper title, the better its acceptance chance (less possibility for being incremental work)
- Abstract: A summary of the research problem, your claim, and the evidence
- Introduction: **Motivation**, a re-statement of the abstract information, **significance**, an outline of the rest of the paper
- Related work:
  a. **A critical review** on the rival approaches that supports the motivation
  b. How to differentiate existing work with your own **creative contributions**.

Research Paper Structure (2)

- Problem statement and algorithm design:
  - Explain your ideas in detail, with examples
  - Highlight your contributions
  - Do ***NOT*** simply put your algorithms in pseudo code!
- Evaluation: Evidence to support the claim of your research contribution
  - Unless you can provide proofs for a theoretical paper on theorems, experimental results are always expected
- Conclusion: A summary of the research contribution, a discussion on its significance, and a mention of future work
- References: List and ***cite*** related work.

Contents

- Some TKDE and ICDM statistics
- Scientific writing and paper structure
- **What to know and how to write a top-quality paper**
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- How to publish at ICDM and TKDE
- Paper reviewing and its feedback
- Summary of take-home messages

What to Know Before You Write

- Assess the audience: To whom are you writing? Why will they be reading your writing?
- Assess the purpose: What should the reader take away?
- Read other people's writing from the forums that you are targeting
  - Language skills and the writing style are always important
  - A paper published in one top journal can easily get rejected by another top journal – community difference or cultural difference
- [Take-Home Message #3] **Know your enemy: Check who are on the program committee or editorial board, and cite their relevant work with due credit**
- Follow the rules – length limits, formatting standards etc.

How to Write a Top-Quality Paper

- [Take-Home Message #4] **Choose a promising topic**
  - 10 Challenging Problems in Data Mining Research (presented by Qiang Yang & Xindong Wu at ICDM '05) http://www.cs.uvm.edu/~icdm/
    - A topic of your interest
    - Your background for the topic
    - Advice from your advisor and senior researchers
- Present a convincing case
- Provide in-depth analysis of empirical results
- Spend more time on the introduction.

How to Present a Convincing Case

- What exactly is the problem being solved?
- How are your ideas **significant** (to justify a paper)?
  - Some ideas are so simple that have been used many times w/o being published
- Is all related work referenced and reviewed?
- Are the comparative studies with previous work convincing?
- Has your system been implemented and used, and if so what did it demonstrate from the real world (for you and the reader to learn)?

In-Depth Analysis of Empirical Results

- Enough details for (a) your experiment settings (so that other researchers can verify and improve your results), and (b) your experimental objectives
- What were the alternatives considered at various points of your experiments? Why and how have you made the choices for your experiments?
- [Take-Home Message #5] **Are the experimental results consistent and conclusive?**
- Can you fine-tune some key parameters to get better or worse results? If so, use figures and tables to show their impacts on your system performances
- How do the experimental results correspond to the motivation of the paper?
- What have you found surprising and tried to avoid in these experiments? How generally applicable are these lessons?

The Most Important Part of Your Paper: the Introduction

- The 1/3 – 2/3 Rule from a reviewer's perspective:
  - 1/3 time to read your introduction and make a decision
  - Remaining 2/3 time to find evidence for the decision
- [Take-Home Message #6] **A good introduction with a good motivation is half of your success!**
- What to cover in the introduction
  - The research problem
  - The motivation of your research on the research problem
  - The claim of your contribution
  - A summary of your evidence to support your claim
  - The significance of your contribution
  - An outline of the rest of the paper.

Contents

- Some TKDE and ICDM statistics
- Scientific writing and paper structure
- What to know and how to write a top-quality paper
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- **How to publish at ICDM and TKDE**
- Paper reviewing and its feedback
- Summary of take-home messages

How to Publish at ICDM and TKDE (1)

- ICDM and TKDE both look for significant technological contributions
- ICDM and TKDE are both *very* tough, expecting *best* results in their respective research field
- [Take-Home Message #7] **Reading and citing relevant papers from the premier forums (incl. ICDM/KDD and TKDE) is a must**
- A possible way to publish in both ICDM/KDD and TKDE:
  - Submit to ICDM/KDD to get (quick) feedback
  - Expand & submit to TKDE if positive feedback from ICDM/KDD, with
    a. at least 30% new material, and
    b. a title footnote to state the conference acceptance/publication.

How to Publish at ICDM and TKDE (2)

- How about application papers?
  - Application papers are always invited, but innovations are necessary. A case of an innovative application must be presented, for the ICDM/TKDE audience.
- How about data analysis w/o large volumes of data?
  - Experiments on large databases are not always required, but generally expected
  - Reasons on why not large data sets should be explained.
- Most important of all: the uniqueness of your research in the field!
  - You work has to be (1) technically sound, (2) relevant, (3) original, (4) significant, and (5) well clarified.

Contents

- Some TKDE and ICDM statistics
- Scientific writing and paper structure
- What to know and how to write a top-quality paper
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- How to publish at ICDM and TKDE
- **Paper reviewing and its feedback**
- Summary of take-home messages

The Review Process

- TKDE
  - EiC: Assign papers to AEs, and handle inconsistency between the AE and reviewers
  - AE: Solicit reviewers, and coordinate the review process
  - Reviewers: Read and provide reviews
- ICDM
  - PC Chairs: Assign papers to Vice Chairs and PC members
  - Vice Chairs: Resolve conflicting reviews and make paper acceptance recommendations
  - PC members: Reviewers.

How to Deal with Feedback (1)

- How to deal with Handling Editors
  - Be polite, but to the point
  - Ask for a change, if a clearly biased or unfair case.
- How to deal with conflicting review reports
  - For journal submissions
    - Try every effort to address every concern
    - [Take-Home Message #8] **Provide a point-by-point statement of changes**
    - Use other reviewers' comments to disagree with the negative ones
  - For conference submissions
    - Rebut if you think you have a reasonable chance to win – Nothing to lose
    - Get senior authors involved in the rebuttal.

How to Deal with Feedback (2)

- How to deal with “arrogant” and “ignorant” reviewers
  - If there is no chance to win them over, provide a gentle statement for the “unreasonable” criticisms that you are not addressing
  - You should still try and resolve some of their comments
  - Your attitude towards the reviewers’ comments is important – all reviewers will read your statement of changes, and an accommodating approach is useful.
- Critical reviews are always expected from first-rate journals and conferences – Don’t get emotional with negative comments
- [Take-Home Message #9] **Be accommodating and persistent in journal submissions** & good luck! !

Contents

- Some TKDE and ICDM statistics
- Scientific writing and paper structure
- What to know and how to write a top-quality paper
  - A promising topic
  - A convincing case
  - In-depth analysis of empirical results
  - The most important part: the introduction
- How to publish at ICDM and TKDE
- Paper reviewing and its feedback
- **Summary of take-home messages**

Summary of Take-Home Messages

1. It is NOT enough to design yet another technique or system without convincing evaluation
2. You should avoid claiming too many dimensions, but one or two with in-depth evidence
3. Know your enemy: Check who are on the program committee or editorial board, and cite their relevant work with due credit
4. Choose a promising topic
5. Are the experimental results consistent and conclusive?
6. A good introduction with a good motivation is half of your success!
7. Reading and citing relevant papers from the premier forums is a must
8. Provide a point-by-point statement of changes (when dealing with journal feedback)
9. Be accommodating and persistent in journal submissions.