

# Keke Chen

Phone: +1 (937) 212-5919  
E-Mail: [kekechen@cc.gatech.edu](mailto:kekechen@cc.gatech.edu) ,  
[gtkeke@yahoo.com](mailto:gtkeke@yahoo.com)  
WWW: [www.cc.gatech.edu/~kekechen](http://www.cc.gatech.edu/~kekechen)  
Citizenship: P.R. China

## RESEARCH INTERESTS

---

- Database Systems, data mining, and information visualization;
- Privacy&Trust in information sharing and service computing; secure data outsourcing;
- Web science: web search ranking, social computing, anti-spamming, and cloud computing.

## EDUCATION

---

**Georgia Institute of Technology, Atlanta GA** (9/2001 – 8/2006)

Ph.D. degree in Computer Science, received in 08/2006

Advisor: Professor Ling Liu

PhD dissertation: *Geometric Methods for Mining Large, Private, and Distributed Datasets*

Area: Databases and Data Mining

Minor: Industrial Systems and Engineering

**State University of New York at Buffalo, Buffalo NY** (8/2000 – 5/2001)

Ph.D. Student, Computer Science

Area: Algorithms and Information Visualization

**Zhejiang University, China** (9/1996 – 4/1999)

M.Sc. Degree in Computer Science

Advisor: Professor Chun Chen

Area: Databases, Distributed Computing and Multimedia

**Tongji University, China** (9/1992 – 7/1996)

B.Sc. Degree in Computer Science

## PROFESSIONAL EXPERIENCE

---

**Assistant Professor, Department of Computer Science and Engineering, Wright State University, Dayton OH** (8/2008 – present)

**Senior Research Scientist, Yahoo! Labs, Santa Clara CA** (8/2006 – 8/2008)

*Mining user preference in web search; Anti link-spamming; Ranking function adaptation; Cloud-based learning and data analysis; Cross-Domain Ranking*

**Graduate Research Assistant, Georgia Institute of Technology, Atlanta GA** (5/2002 – 8/2006)

Advisor: Professor Ling Liu

- *Random Geometric Perturbation and Collaborative Multiparty Privacy-Preserving Data Mining*, 2004-2007
- *Finding the Best K for Clustering Large Categorical Datasets and Data Streams*, 2003-2005
- *iVIBRATE Visual Framework for Cluster Analysis of Large Numerical Datasets*, 2002-2005

**Research Intern, Yahoo!Search, Sunnyvale CA** (Summers of 2004 and 2005)

*Web Search Relevance Analysis*; prototypes: *Online Topic Clustering on Search Results* and *Web Search Organizer*

**Research Scientist/Lecturer, Zhejiang University, China** (5/1999 – 8/2000)

**Research Assistant, Zhejiang University, China** (9/1997 – 4/1999)

**Project Lead, Huaguang Software, China** (5/1999 – 8/2000)

---

#### TEACHING EXPERIENCE

---

**Lecturer, Zhejiang University** (9/1999 – 7/2000)

**Graduate Teaching Assistant, State University of New York at Buffalo** (8/2000 – 5/2001)

**Graduate Teaching Assistant, Georgia Institute of Technology** (8/2001 – 5/2002)

---

#### PUBLICATIONS

---

##### Book Chapter:

1. **Keke Chen** and Ling Liu, *A Survey of Multiplicative Data Perturbation for Privacy Preserving Data Mining*. *Privacy-Preserving Data Mining: Models and Algorithms*. Edited by Charu C. Aggarwal and Philip S. Yu, February, 2008, Springer

##### Journal Papers:

2. **Keke Chen** and Ling Liu, *VISTA: Validating and Refining Clusters via Visualization*. *Information Visualization*, Vol.3, No.4, Sept. 2004.
3. **Keke Chen** and Ling Liu, *iVIBRATE: Interactive Visualization Based Framework for Clustering Large Datasets*, *ACM Transactions on Information Systems (TOIS)*, Vol. 24, Issue 2, April, 2006.
4. Yongjing Lin, Wenyuan Li, **Keke Chen** and Ying Liu, *Document Clustering and Ranking System for Exploring MEDLINE Citations*, *Journal of the American Medical Informatics Association (JAMIA)*, Volume 14, Issue 5, 2007,
5. **Keke Chen** and Ling Liu, *Best K: the Critical Clustering Structure in Categorical Dataset*, to appear in *Journal of Knowledge and Information Systems (KAIS)*, 2008

##### Conference Papers:

6. **Keke Chen**, Rongqing Lu, CK Wong, Gordon Sun, Larry Heck and Belle Tseng: *Trada: Tree-Based Ranking Function Adaptation*, *ACM Conference on Information and Knowledge Management (CIKM)*, Oct. 2008
7. **Keke Chen**, Ya Zhang, Zhaohui Zheng, Hongyuan Zha, and Gordon Sun, *Adapting Ranking Functions to User Preferences*, *DBRank workshop at IEEE conference on Data Engineering (ICDE)*, 2008
8. **Keke Chen** and Ling Liu, *Towards Attack-Resilient Geometric Data Perturbation*, in *Proceedings of the SIAM Data Mining Conference (SDM07)*, Minneapolis, MN, April 2007.
9. **Keke Chen** and Ling Liu, *Space Adaptation: Privacy-preserving Multiparty Collaborative Mining with Geometric Perturbation*, in *Proceedings of ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2007)* (Brief Announcement)

10. Zhaohui Zheng, Hongyuan Zha, **Keke Chen**, and Gordon Sun, *A Regression Framework for Learning Ranking Functions Using Relative Preferences*, in Proceedings of ACM SIGIR Conference, Amsterdam, Netherlands, July 2007
11. Zhaohui Zheng, Hongyuan Zha, Tong Zhang, Olivier Chapelle, **Keke Chen**, Gordon Sun *A General Boosting Method and its Application to Learning Ranking Functions for Web Search*, in Proceedings of Neural Information Processing Systems (NIPS), Canada, 2007
12. Hua Yan, **Keke Chen** and Ling Liu, *Efficiently Clustering Transactional Data with Weighted Coverage Density*, in Proceedings of ACM Conference on Information and Knowledge Management (CIKM06), Chicago, November 2006.
13. **Keke Chen** and Ling Liu, *Detecting the Change of Clustering Structure in Categorical Data Streams*, in Proceedings of SIAM Data Mining Conference (SDM06), Washington DC, 2006.
14. **Keke Chen** and Ling Liu, *Privacy-preserving Data Classification with Rotation Perturbation*, in Proceedings of the 5<sup>th</sup> IEEE International Conference on Data Mining (ICDM05), Huston, TX, 2005
15. **Keke Chen** and Ling Liu, *The 'Best K' for Entropy-based Categorical Data Clustering*, in Proceedings of the 17<sup>th</sup> Scientific and Statistical Database Management Conference (SSDBM05), Santa Barbara, CA, June 2005.
16. **Keke Chen** and Ling Liu, *ClusterMap: Labeling Clusters in Large Dataset via Visualization*, in Proceedings of ACM 13<sup>th</sup> Conference on Information and Knowledge Management (CIKM04). Washington DC, November 2004.
17. **Keke Chen** and Ling Liu, *Validating and Refining Clusters via Visual Rendering*, in Proceedings of the 3<sup>rd</sup> IEEE International Conference on Data Mining (ICDM03). Melbourne, FL, November 2003.
18. **Keke Chen** and Ling Liu, *A Visual Framework Invites Human into the Clustering Process*, in Proceedings of the 15<sup>th</sup> Scientific and Statistical Database Management Conference (SSDBM03). Cambridge, Boston, July 2003.
19. **Keke Chen** and Ling Liu, *Cluster Rendering of Skewed Datasets via Visualization*, in Proceedings of ACM Symposium on Applied Computing (ACM SAC2003). Melbourne, FL, March 2003.

#### **FILED PATENTS**

---

- [PT.1] **Keke Chen**, Ya Zhang, Zhaohui Zheng, Hongyuan Zha, and Gordon Sun: Systems and Methods for Learning a Balanced Relevance Function for a Content Item from Expert and User Judgments
- [PT.2] Zhaohui Zheng, Hongyuan Zha, **Keke Chen**, and Gordon Sun: A Regression Framework for Learning Ranking Functions Using Relative Preferences
- [PT.3] Keke Chen, Rongqing Lu, CK Wong, Gordon Sun, Larry Heck, and Belle Tseng: Methods and Apparatuses for Adapting a Ranking Function of a Search Engine for Use with a Specific Domain
- [PT.4] Keke Chen, Jing Bai, Srihari Reddy, Belle Tseng: Systems and Methods for Learning Ranking Functions for a Group of Domains

## **PROFESSIONAL ACTIVITIES**

---

### *Reviewer for Professional Conferences*

International Conference on Web Information System Engineering (WISE), ACM Symposium on Applied Computing (SAC), International Joint Conference on AI (IJCAI), International Conference on Information and Knowledge Management (CIKM), International Conference on Extended Database Technology (EDBT), ACM Conference on Management of Data (SIGMOD), ACM Conference on Knowledge Discovery and Data Mining (SIGKDD)

### *Reviewer for Professional Journals*

SIGMOD Record, Knowledge and Information Systems (KAIS), Journal of Data Mining and Knowledge Discovery (DMKD), IEEE Transactions on Knowledge and Data Engineering (TKDE), IEEE Internet Computing, Pattern Recognition

*Publicity Chair:* DOA08, ODBASE08, CoopIS08, GADA08

*PC Member:* CoopIS08, ICDM08, FIS08

*Member of the Editorial Board:* The Open Information Systems Journal, Bentham Science Publishers