Tiancheng Li

Department of Computer Science
305 N. University Street
Purdue University
West Lafayette, IN 47907

Email: li83@cs.purdue.edu
Phone: (765)586-7289
Fax: (765)494-0739
Homepage: http://www.cs.purdue.edu/homes/li83/

Research Interests

Databases and data mining, with a focus on security and privacy in data publishing, data mining, and databases outsourcing. I am also interested in security and privacy issues in other fields such as healthcare, genome computation, cloud computing, location services, and social networks.

Education

Purdue University, West Lafayette, IN — August 2005 – Present
Ph.D., Computer Science — May 2010 (expected)

- PhD Thesis: Privacy Preservation in Data Publishing and Sharing
- Advisor: Ninghui Li

M.S., Computer Science — December 2008

Zhejiang University, Hangzhou, China — September 2001 – July 2005
B.S., Computer Science — July 2005

Research Experience

Purdue University — West Lafayette, IN
Research Assistant — August 2005 – Present

Privacy preserving data publishing.

- ◇ Designed three new generalization schemes and enumeration and pruning techniques.
- ◇ Proposed the *t-closeness* privacy measure for data publishing.
- ◇ Developed several prevention methods for privacy protection in dynamic data publication.
- ◇ Proposed the *Injector* framework for modeling and integrating background knowledge.
- ◇ Proposed an integrated framework for evaluating privacy utility tradeoff.
- ◇ Proposed the *slicing* technique for anonymizing sparse high-dimensional transaction data.

Role mining and role engineering.

- ◇ Designed role mining algorithms for incorporating user attribute semantics.
- ◇ Developed a methodology for evaluating a wide variety of role mining algorithms.

Security and practicality of outsourcing association rule mining.

- ◇ Analyzed both the security and the cost associated with outsourcing association rule mining.

AT&T Labs - Research — Florham Park, NJ
Research Intern — May 2009 – August 2009

Algorithm-based privacy attacks.

- ◇ Analyzed and developed techniques for defending against method-based attacks.

IBM Almaden Research Center — San Jose, CA
Research Intern — May 2008 – August 2008

Data de-duplication for bandwidth savings.

- ◇ Designed and implemented an effective de-duplication algorithm using data mining techniques.
- ◇ Designed a system model incorporating de-duplication into online client-server services.

IBM Almaden Research Center — San Jose, CA
Research Intern — May 2007 – August 2007

Trust preservation and verification for regulatory compliance.

- ◇ Developed and implemented WORMSEAL for trust preservation and verification.
- ◇ Designed a new authenticated data structure called *Homomorphic Hash Tree (HHT)*.

Teaching Experience

Teaching Assistant, Purdue University — August 2006 – May 2007

CS 426: Computer Security — Spring 2007

◇ Developed course materials, including homeworks, exams, and course projects. Conducted weekly lab sessions and led discussions. Taught students to use software packages and libraries for course purposes. Graded homeworks and programming projects.

CS 483: Computational Complexity — Fall 2006

◇ Led course reviews. Designed and graded assignments.

CS 381: Algorithm Design — Fall 2006

◇ Designed and delivered lectures for weekly recitation sessions. Assisted in the development and maintenance of the course website. Designed, developed, and graded problem sets and exams.

Training Program in Teaching, Purdue University — January 2009 – May 2009

GRAD 590: Preparing Future Faculty — Spring 2009

Workshops: Excellence in Teaching — Spring 2009 & Fall 2009

◇ Participated in a training program offered by Purdue Center for Instructional Excellence (CIE) and the Graduate School. Took formal training in teaching theories and methodologies for college-level students. Enrolled in a series of workshops on improving teaching effectiveness.

Publications

Conferences/Workshop Papers:

17. Ian Molloy, Ninghui Li, and Tiancheng Li. On the (In)Security and (Im)Practicality of Outsourcing Precise Association Rule Mining. To appear in the *IEEE International Conference on Data Mining (ICDM)*, 2009.
16. Tiancheng Li, Ninghui Li, Jian Zhang, and Ian Molloy. Slicing: A New Approach to Privacy Preserving Data Publishing. In the *Preprints arXiv:0909.2290v1 [cs.DB]*, 2009.
15. Tiancheng Li, Xiaonan Ma, and Ninghui Li. WORM-SEAL: Trustworthy Data Retention and Verification for Regulatory Compliance. In the *European Symposium on Research in Computer Security (ESORICS)*, pp. 472-488, 2009.
14. Tiancheng Li and Ninghui Li. On the Tradeoff Between Privacy and Utility in Data Publishing. In the *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 517-526, 2009.
13. Ian Molloy, Ninghui Li, and Tiancheng Li, Evaluating Role Mining Algorithms, In the *ACM Symposium on Access Control Models and Technologies (SACMAT)*, pp. 95-104, 2009.
12. Tiancheng Li, Ninghui Li, and Jian Zhang. Modeling and Integrating Background Knowledge in Data Anonymization. In the *IEEE International Conference on Data Engineering (ICDE)*, pp. 6-17, 2009.
11. Cornel Constantinescu, Jan Pieper, and Tiancheng Li. Block Size Optimization in Deduplication Systems. In the *Data Compression Conference (DCC)*, pp. 442, 2009.
10. Ian Molloy, Hong Chen, Tiancheng Li, Qihua Wang, Ninghui Li, Elisa Bertino, Seraphin Calo, and Jorge Lobo. Mining Roles with Semantic Meanings. In the *ACM Symposium on Access Control Models and Technologies (SACMAT)*, pp. 21-30, 2008.
9. Tiancheng Li and Ninghui Li. Injector: Mining Background Knowledge for Data Anonymization. In the *IEEE International Conference on Data Engineering (ICDE)*, pp. 446-455, 2008.
8. Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-Closeness: Privacy Beyond k-Anonymity and ℓ-Diversity. In the *IEEE International Conference on Data Engineering (ICDE)*, pp. 106-115, 2007.
7. Tiancheng Li and Ninghui Li. Optimal k-Anonymity with Flexible Generalization Schemes through Bottom-up Searching. In the *IEEE International Workshop on Privacy Aspects of Data Mining (PADM)*, in conjunction with ICDM, pp. 518-523, 2006.

Journal Articles:

6. Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. Closeness: A New Privacy Measure for Data Publishing. To appear in the *IEEE Transaction on Knowledge and Data Engineering (TKDE)*, 2009.
5. Ji-Won Byun, Tiancheng Li, Elisa Bertino, Ninghui Li, and Yonglak Sohn. Privacy Preserving Incremental Data Dissemination. In the *Journal of Computer Security (JCS)*, 17(1): 43-68, 2009.
4. Tiancheng Li and Ninghui Li. Towards Optimal k-Anonymization. In the *Data & Knowledge Engineering Journal (DKE)*, 65(1): 22-39, 2008.

Manuscripts Under Review:

3. Tiancheng Li, Ninghui Li, and Jian Zhang. Background Knowledge in Privacy Preserving Data Publishing. Submitted to the *ACM Transactions on Database Systems (TODS)*, 2009.
2. Paper title and author names are removed due to the double-blind review process. Submitted to the *ACM International Conference on Management of Data (SIGMOD)*, 2010.
1. Paper title and author names are removed due to the double-blind review process. Submitted to the *ACM International Conference on Management of Data (SIGMOD)*, 2010.

Patents

2. Tiancheng Li and Cornel Constantinescu. Data De-Duplication for Bandwidth Savings. US Patent Pending, Filed by IBM Almaden Research Center, 2008.
1. Tiancheng Li and Xiaonan Ma. System and Method for Efficient Trust Preservation in Data Stores. US Patent Pending, IBM Docket No. ARC920080022US1, Filed by IBM Almaden Research Center, 2007.

Presentations and Talks

9. Topic: WORM-SEAL: Trustworthy Data Retention and Verification for Regulatory Compliance
   Location: European Symposium on Research in Computer Security (ESORICS), Saint Malo, France.
   Date: September 2009.
8. Topic: On the Tradeoff between Privacy and Utility in Data Publishing.
   Location: AT&T research Labs, Florham Park, New Jersey.
   Date: July 2009.
7. Topic: On the Tradeoff between Privacy and Utility in Data Publishing.
   Location: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD), Paris, France.
   Date: June 2009.
6. Topic: Modeling and Integrating Background Knowledge in Data Anonymization.
   Location: IEEE International Conference on Data Engineering (ICDE), Shanghai, China.
   Date: April 2009.
5. Topic: Data De-duplication Algorithms and Applications to Online Client-Server Services.
   Location: IBM Almaden Research Center, San Jose, California.
   Date: August 2008.
4. Topic: Injector: Mining Background Knowledge for Data Anonymization.
   Location: IEEE International Conference on Data Engineering (ICDE), Cancun, Mexico.
   Date: April 2008.
3. Topic: Effective Trust Preservation and Verification in Regulatory Compliance.
   Location: IBM Almaden Research Center, San Jose, California.
   Date: August 2007.
2. Topic: t-Closeness: Privacy Beyond k-Anonymity and l-Diversity .
   Location: IEEE International Conference on Data Engineering (ICDE), Istanbul, Turkey.
   Date: April 2007.

1. Topic: Optimal k-Anonymity with Flexible Generalization Schemes through Bottom-Up Searching.
   Location: IEEE International Conference on Data Mining (ICDM), Hong Kong, China.
   Date: December 2006.

Professional Services

Program Committee Member

ACM GIS Workshop on Security and Privacy in GIS and LBS (SPRINGL 2009).

External Reviewer

Conferences: SIGMOD (2008, 2007), VLDB (2009, 2008), ICDE 2008, WWW 2009.
Conferences: S&P (2008, 2007), CCS 2006.
Journals: TODS 2008, VLDB Journal 2009, TKDE (2009, 2008), TKDD 2009.

Honors and Awards

Purdue Research Foundation (PRF) Scholarship, 2008.
Honored Undergraduate of Zhejiang University, 2005.
Government Scholarship for Excellent Students, China, 2004.
National Scholarship for Excellent Students, China, 2003.
First Prize in National Contest in Mathematical Modeling, China, 2003.
Undergraduate Student Fellowship, Zhejiang University, 2001, 2002, 2003, 2004.
First Prize in National Contest in Mathematics for high school students, China, 1999, 2000.

References

Prof. Ninghui Li
Associate Professor
Department of Computer Science
Purdue University
Phone: (765)496-6756
ninghui@cs.purdue.edu

Dr. Graham Cormode
Research Scientist
Database Research Department
AT&T Labs - Research
Phone: (973)360-8957
graham@research.att.com

Prof. Elisa Bertino
Professor
Department of Computer Science
Purdue University
Phone: (765)496-2399
bertino@cs.purdue.edu

Dr. Divesh Srivastava
Department Head
Database Research Department
AT&T Labs - Research
Phone: (973)360-8776
divesh@research.att.com

Prof. Chris Clifton
Associate Professor
Department of Computer Science
Purdue University
Phone: (765)494-6005
clifton@cs.purdue.edu