CURRICULUM VITAE of Feifei Li

ACM Fellow   CCF Fellow   IEEE Fellow

Research Interests

Database/big data systems and large-scale data management, systems, and analytics. Security issues in data management and systems. Machine learning/AI methods for systems and data analytics.

Education

September 2002 - August 2007 Ph.D. in Computer Science, Computer Science Department, Boston University, PhD Thesis: "Query and Data Security Issues in the Data Outsourcing Model".

September 1998 - Jan 2002 B.S. in Computer Engineering, School of Computer Engineering, Nanyang Technological University, Singapore (transferred from Tsinghua University, China).

September 1997 - January 1998 Study in the Department of Electrical Engineering, Tsinghua University (transferred to NTU), Beijing, China.

Honors and Awards

- ACM SIGMOD Best Paper Award, 2023.
- EDBT Test of Time Award, EDBT 2022.
- CCF Fellow, 2022.
- ACM Fellow, 2021.
- IEEE Fellow, 2021.
- IEEE ICDCS Best Paper Award, 2020.
- ACM SoCC Best Paper Award (runner-up), 2019.
- The Global Leading Technology Achievement Award at the World Internet Conference, 2019.
- ACM Distinguished Scientist, 2018.
- Visa Research Faculty Award, 2017.
- Presidential Scholar Award, University of Utah, 2017.
- ACM SIGMOD Research Highlight Award, 2017.
- ACM SIGMOD Best Paper Award, 2016.
- Teaching Dean's List, Top Instructors in the College of Engineering, CS 6530 and CS 6950, Fall 2017.
- SIGMOD Best System Demonstration Award, ACM SIGMOD 2015.
- Google Faculty Award, 2015.
- IEEE ICDE 2014 10+ Years Most Influential Paper Award.
- Teaching Dean's List, Top 15% Class in the College of Engineering, CS 6530, Fall 2013.
- SIGMOD Best Undergraduate Research Poster Award (Advisor), SIGMOD 2013.
- Google App Engine Education Award, 2012.
- HP Labs Innovation Research Award (61 awards selected from more than 500 worldwide submissions), HP, 2012.
- HP Labs Innovation Research Award (62 awards selected from 626 worldwide submissions), HP, 2011.
- COFRS (Committee on Faculty Research Support) Award, Florida State University, 2011.
- NSF Career Award, 2011.
- First Year Assistant Professor Award, Florida State University, 2008.
- Best Presenter Award, in research summer interns Seminar@Luch Series, IBM T.J. Watson Research Center, August, 2006.
- Best Paper Award, in 20th IEEE International Conference on Data Engineering (ICDE), 2004.
- First Class Honor (Accelerating Honor), School of Computer Engineering, Nanyang Technological University, Singapore, 2002.
- Defense Science & Technology Agency Gold Medal, Singapore, 2001.
- Merit Award in Singapore Advanced e-Business Applications Competition, 2000.
- National Science Class, China, 1997.

Professional Experience

- **Alibaba Group**, Sunnyvale, CA, USA & Hangzhou, CN

  Vice President, Chief Database Scientist, President of the Database Products Business Unit; Director of the Database and Storage Research Lab **2018 - present**

- **MariaDB Corporation**, Redwood City, CA, USA

  Board Memeber **2019 - 2022**

- **SelectDB Corporation**, Beijing, CN,

  Board Memeber **2023 - Present**

- **School of Computing, University of Utah**, Salt Lake City, UT, USA

  Professor (on leave) **2018 - 2020**

- **School of Computing, University of Utah**, Salt Lake City, UT, USA

  Associate Professor **July, 2013 - April, 2018**

- **Department of Computer Science and Engineering and ZhiYuan College, Shanghai Jiao Tong University**, Shanghai, China,

  Visiting Professor **July, 2014 - July, 2015**

- **School of Computing, University of Utah**, Salt Lake City, UT, USA

  Assistant Professor **Aug, 2011 - June, 2013**

- **Computer Science Department, Florida State University**, Tallahassee, FL, USA

  Assistant Professor & Director for the SAIT Laboratory **Aug, 2007 - Aug, 2011**

- **Database Research Group, Microsoft Research**, Redmond, WA, USA

  Research Summer Intern **May, 2007 - August, 2007**

- **Database Management Research Department, AT&T Shannon Research Lab**, Florham Park, NJ, USA

  Research Consultant **2007 - 2009**

- **Database Research Group, IBM T.J.Watson Research Center**, Hawthorne, NY, USA

  Research Summer Intern **June, 2006 - September, 2006**

- **SQL Server Group, Microsoft**, Redmond, WA, USA

  Summer Intern **June, 2005 - September, 2005**

Grants

1. " Neighborhood Looking Glass: 360 Degree Automated Characterization of the Built Environment for Neighborhood Effects Research", *co-PI*, NIH R01, PI: Quynh Nguyen, co-PI Tolga Tasdizen, Kim Brunisholz, 09/01/18-08/31/22, $1,340,980 (Utah's portion: $550,710)

2. "III: Small: Persistent Data Summaries: Temporal Analytics on Big Data Histories", *PI*, NSF IIS, co-PI Jeff Phillips, 08/01/18-08/31/21, $499,934.

3. "SaTC: CORE: Medium: Large-Scale Data Driven Anomaly Detection and Diagnosis from System Logs", *PI*, NSF SaTC, co-PI Rob Ricci, Vivek Srikumar, 07/16/18-08/31/22, $1,100,000.

4. "Data driven transaction analytics", *Single PI*, Visa Research, Gift from the Visa Research Faculty Award program 2017, $35,000.

5. "SaTC: CORE: Small: Efficient Hardware-Aware and Hardware-Enabled Algorithms for Secure In-Memory Databases", *PI*, NSF SaTC, PI Rajeev Balasubramonian, co-PI, 09/01/17-08/31/20, $516,000.

6. "Presidential Scholar Award", *Single PI*, University of Utah, 2017. $10,000.

7. "Analytics and Integration of Large Scale Heterogeneous Spatio-Temporal Data", *single PI*, NSFC (NSF China) Oversea Collaboration Grant, 01/01/2018-12/31/2022, Chinese RMB 1,800,000.

8. "Scalability and Performance Improvement for Apache Carbon Data", *PI*, Gift from Huawei, Single PI, 03/01/17-02/28/18, $80,000.

9. "Towards a Database Engine for Interactive and Online Sampling and Analytics", *PI*, NSF IIS, Single PI, 09/01/16-08/31/19, $500,000.

10. "NSF DSSP Workshop: Data Science for Secure and Privacy-Aware Big Data Management and Mining", *sole PI*, NSF, 09/07/16- 09/06/17, $98,000.

11. Research Experience for Undergraduate (for the NSF SEAL project), *PI*, NSF, 09/16-08/17, $16,000.

12. Research Experience for Undergraduate (for the NSF STORM project), *PI*, NSF, 09/16-08/17, $16,000.

13. "TWC: Medium: Collaborative: Seal: Secure Engine for AnaLytics - From Secure Similarity Search to Secure Data Analytics", *PI*, NSF SaTC, co-PI Jeff Phillips, PI at UCSB Rachel Lin, 07/15/15-06/30/19, $942,056: $600,007 at Utah and $342,049 at UCSB.

14. "Spatio-Temporal Online Analytics With Concept Enriched Text", *PI*, Google Research Awards program, 02/2015-02/2016, $57,640.

15. "Automated Query Engine for Large Heterogeneous Data", *PI*, NSFC (NSF China) Oversea Collaboration Grant, co-PI Bin Yao at Shanghai Jiao Tong University, 01/01/2015-12/31/2016, Chinese RMB 200,000.

16. "CIF21 DIBBs: STORM: Spatio-Temporal Online Reasoning and Management of Large Data", *PI*, NSF DIBBs program, co-PI John Horel, Jeff Phillips, Paul Rosen, 11/01/2014-10/31/2017, $1,157,975.

17. Research Experience for Undergraduate (REU Supplement for the NSF BIGDATA grant), *PI*, NSF, 04/14-04/15, $16,006.

18. "BIGDATA: Small: DCM: DA: Building a Mergeable and Interactive Distributed Data Layer for Big Data Summarization Systems", *PI*, NSF BIGDATA program, co-PI Jeff Phillips, 09/15/2013-08/31/2016, $685,380.

19. "TWC: Medium: TCloud: A Self-Defending, Self-Evolving and Self-Accounting Trustworthy Cloud Platform", *co-PI*, NSF SaTC, PI Jacobus Van Der Merwe, co-PIs Robert Ricci, Eric Eide, 09/01/2013-08/31/2017, $999,991.

20. "CSR: Medium: Energy-Efficient Architectures for Emerging Big-Data Workloads", *co-PI*, NSF CNS, PI Rajeev Balasubramonian, co-PIs Al Davis, Mary Hall, 07/01/2013-06/30/2017, $873,286.

21. Research Experience for Undergraduate (REU Supplement for the CAREER grant), *sole PI*, NSF, 07/12-07/13, $16,000.

22. "Time-Sensitive and Keyword-Aware Recommendation in Large Data", *sole-PI*, HP, 09/01/2011-08/31/2013, $125,000 (supported by two HP IRP awards at $45,000 in 2011-2012 with an additional $5,000 for collaboration support, and $70,000 in 2012-2013 with an additional $5,000 for collaboration support).

23. "Non-Conventional Search and Identification of Delinquent Parents", *co-PI (PI at the University of Utah)*, PI: Sudhir Aggarwal (FSU), co-PI: Piyush Kumar (FSU), Florida Department of Revenue, 2/16/2011-9/29/2013, $632,465.

24. "Efficient Aggregate Similarity Search", *sole PI*, FSU COFRS (Committee on Faculty Research Support) Award, 2011, $14,000.

25. "CAREER: Novel Query Processing Techniques for Distributed Probabilistic Data", *sole PI*, NSF IIS Program, 02/11-01/16, $498,138.

26. Industry Gift, *sole PI*, HP Labs, 08/10, $3,000.

27. "Scholarship for Service at Florida State University", *co-PI*, PI: Mike Burmester (FSU), co-PIs: Sudhir Aggarwal (FSU), Xiuwen Liu (FSU), NSF Division of Undergraduate Education, 09/10-08/14, $1,853,894.

28. "III:Small: Efficient Ranking and Aggregate Query Processing for Probabilistic Data", *sole PI*, NSF IIS Program, 09/09-09/12, $328,831.

29. Research Experience for Undergraduate (REU Supplement for the CT-ISG grant), *sole PI*, NSF, 06/09-09/11, $12,000.

30. "CT-ISG: Collaborative Research: Towards Trustworthy Database Systems", *sole PI at FSU*, NSF Cyber Trust Program, 10/08-09/11, $150,620.

31. DoD 2009 IASP (Information Assurance Scholarship Program) Grant, *co-PI*, PI: Mike Burmester (FSU), 08/09-08/10, $46,977.

32. "Building Trustworthy Database Systems", *sole PI*, FSU CRC Planning Grant, 12/01/07-11/30/08, $12,000.

33. "Query Verification for Distributed Databases", *sole PI*, FSU CRC First Year Assistant Professor Grant, 05/08-08/08, $16,000.

34. "Ranking and Aggregate Query Processing for Large Scientific Data with Fuzzy Information", Source: FSU Undergraduate Research Award, *sole PI*, Student: Justin DeBrabant, $4,000, 2009.

Publications

#: alphabetic ordering of authors.
Book Chapters

1. *Privacy Preservation on Time Series*, by S. Papadimitriou, **F. Li**, G. Kollios, P. S. Yu, *Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques*, ISBN: 978-1-439-803653, *Chapman & Hall/CRC Press*, December, 2010. Link: http://www.crcpress.com/ product/isbn/9781439803653

2. *Authenticated Index Structures for Outsourced Databases*, by **F. Li**, M. Hadjieleftheriou, G. Kollios, L. Reyzin, *Handbook of Database Security*, ISBN: 978-0-387-48532-4, *Springer*, November, 2007. Link: http://www.springer.com/computer/security+and+cryptology/book/978-0-387-48532-4

3. *Trip Planning Queries in Road Network Databases*, by **F. Li**, M. Hadjieleftheriou, G. Kollios, D. Cheng, S. H. Teng, *Encyclopedia of GIS*, ISBN: 978-0-387-30858-6 (Print), 978-0-387-35973-1 (Online), *Springer*, February, 2008. Link: http://www.springerlink.com/content/h01539/

Refereed Journals

4. *ROVEC: Runtime Optimization of Vectorized Expression Evaluation for Column Store*, by M. Li, Z. Miao, D. Wu, **F. Li**, S. Wang, W. Cao, Z. Qiao, B. Ruan, K. Liang, J. Yang, H. Dai, G. Chen. 35(3):3045-3058, 2023, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

5. *Learning-based query optimization for multi-probe approximate nearest neighbor search*, by P. Zhang, B. Yao, C. Gao, B. Wu, X. He, **F. Li**, Y. Lu, C. Zhan, F. Tang. 32(3):623-645, 2023, **The International Journal on Very Large Data Bases (VLDBJ)**.

6. *Efficient Oblivious Query Processing for Range and kNN Queries*, by Z. Chang, D. Xie, J. Phillips, **F. Li**, R. Balasubramonian. 34(12):5741-5754, 2022, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

7. *Semantic Embedding for Regions of Interest*, by D. Paul, D. Xie, **F. Li**, J. Phillips. 30(3):311-331, 2020, **The International Journal on Very Large Data Bases (VLDBJ)**.

8. *SolarDB: Towards a Shared-Everything Database on Distributed Log-Structured Storage*, by T. Zhu, Z. Zhao, **F. Li**, W. Qian, A. Zhou, D. Xie, R. Stutsman, H. Li, H. Hu. Vol. 25, No. 2, 11:1-11:26, 2019, **ACM Transactions on Storage (ACM TOS)**.

9. # *Wander Join and XDB: Online Aggregation via Random Walks*, by **F. Li**, B. Wu, K. Yi, Z. Zhao. Vol. 44, No. 1, 2:1-2:41, 2019, **ACM Transactions on Database Systems (ACM TODS)**.

10. *Spell: Online Streaming Parsing of Large Unstructured System Logs*, by M. Du, **F. Li**. 31(11):2213-2227, 2019, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

11. *Twitter-derived neighborhood characteristics associated with obesity and diabetes*, by Q. Nguyen, K. D. Brunisholz, W. Yu, M. McCullough, H. A. Hanson, M. L. Litchman, **F. Li**, Y. Wan, J. A. VanDerslice, M. Wen, K. R. Smith. **Scientific Reports by Nature**, Vol. 7, Article number: 16425, 2017.

12. *Geotagged US Tweets As Predictors Of County-Level Health Outcomes*, by Q. Nguyen, M. McCullough, H. Meng, D. Paul, D. Li, S. Kath, G. Loomis, E. Nsoesie, M. Wen, K. Smith, **F. Li**. 2017, **American Journal of Public Health (AJPH)**.

13. *# Wander Join and XDB: Online Aggregation via Random Walks*, by **F. Li**, B. Wu, K. Yi, Z. Zhao. Vol. 46, No. 1, pages 33-40, 2017. **SIGMOD Record**.

14. *ATOM: Efficient Tracking, Monitoring, and Orchestration of Cloud Resources*, by M. Du, **F. Li**. Vol. 28, No. 8, pages 2172-2189, 2017. **IEEE Transactions on Parallel and Distributed Systems (IEEE TPDS)**.

15. *Exact and Approximate Flexible Aggregate Similarity Search*, by **F. Li**, K. Yi, Y. Tao, B. Yao, Y. Li, D. Xie, M. Wang. Vol. 25, No. 4, pages 317-338, 2016, **The International Journal on Very Large Data Bases (VLDBJ)**.

16. *Building a National Neighborhood Dataset From Geotagged Twitter Data for Indicators of Happiness, Diet, and Physical Activity*, by Q. Nguyen, D. Li, H. W. Meng, S. Kaith, E. Nsoesie, **F. Li**, M. Wen. Vol. 2, No. 2, 2016, **JMIR Public Health & Surveillance**.

17. *Leveraging Geotagged Twitter Data to Examine Neighborhood Happiness, Diet, and Physical Activity*, by Q. Nguyen, S. Kath, H. W. Meng, D. Li, K. R. Smith, J. A. VanDerslice, M. Wen, **F. Li**. Vol. 73, pages 77-88, 2016, **Applied Geography**.

18. *Comparing Different Implementations of Near Data Computing with In-Memory MapReduce Workloads*, by S. Pugsley, J. Jestes, R. Balasubramonian, V. Srinivasan, A. Buyuktosunoglu, A. Davis, **F. Li**, IEEE Micro Special Issue on Big Data (**IEEE Micro**), Vol. 34, No. 4, pages 44-52, 2014.

19. *Scalable Keyword Search on Large RDF Data*, by W. Le, **F. Li**, A. Kementsietsidis, S. Duan. Vol. 26, No. 11, pages 2774–2788, 2014. **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

20. *Dynamic Monitoring of Optimal Locations in Road Network Databases*, by B. Yao, X. Xiao, **F. Li**, Y. Wu, W. Le. Vol. 23, No. 5, pages 697–720, 2014, **The International Journal on Very Large Data Bases (VLDBJ)**.

21. *Spatial Approximate String Search*, by **F. Li**, B. Yao, M. Tang, M. Hadjieleftheriou. Vol. 25, No. 6, pages 1394-1409, 2013, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

22. *Query Access Assurance in Outsourced Databases*, by W. Le, **F. Li**. Vol. 5, No. 2, pages 178-191, 2012, **IEEE Transactions on Services Computing (IEEE TSC)**.

23. *Semantics of Ranking Queries on Probabilistic Data*, by J. Jestes, G. Cormode, **F. Li**, K. Yi. Vol. 23, No. 12, pages 1903-1917, 2011, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

24. *Group Enclosing Queries*, by **F. Li**, B. Yao, P. Kumar. Vol. 23, No. 10, pages 1526-1540, 2011, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

25. *Top-k Queries on Temporal Data*, by **F. Li**, K. Yi, W. Le. Vol. 19, No. 5, pages 715-733, 2010, **The International Journal on Very Large Data Bases (VLDBJ)**.

26. *Authenticated Index Structures for Aggregation Queries*, by **F. Li**, M. Hadjieleftheriou, G. Kollios, L. Reyzin. Vol. 13, No. 4, pages 32:1-32:35, 2010, **ACM Transactions on Information and System Security (ACM TISSEC)**.

27. *The World in a Nutshell: Concise Range Queries*, by K. Yi, X. Liang, **F. Li**, L. Chen. Vol. 23, No. 1, pages 139-154, 2011, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

28. *Small Synopses for Group-By Query Verification on Outsourced Data Streams*, by K. Yi, **F. Li**, G. Cormode, M. Hadjieleftheriou, G. Kollios, D. Srivastava. Vol. 34, No. 3, pages 1–42, 2009, **ACM Transactions on Database Systems (ACM TODS)**.

29. *Robust Approximate Aggregation in Sensor Data Management Systems*, by J. Considine, M. Hadjieleftheriou, **F. Li**, J. Byers, G. Kollios. Vol. 34, No. 1, pages 1–35, 2009, **ACM Transactions on Database Systems (ACM TODS)**.

30. *Efficient Processing of Top-k Queries in Uncertain Databases with x-Relations*, by K. Yi, **F. Li**, G. Kollios, D. Srivastava. Vol. 20, No. 12, pages 1669-1682, December, 2008, **IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)**.

31. *Robust Aggregation in Sensor Networks, by G. Kollios, J. Byers, J. Considine*, M. Hadjieleftheriou, and **F. Li**. **IEEE Data Engineering Bulletin**, Vol. 28, No. 1, March 2005.

32. *Towards Building Logical Views of Websites*, by Z. Liu, W. Ng, E. Lim, **F. Li**. **Data & Knowledge Engineering (DKE)**, Vol. 49, No 2, pages 197-222, May 2004, Elsevier. (Invited by WIDM 2002).

Refereed Conferences

33. *Detecting Logic Bugs Of Join Optimizations In DBMS*, by X. Tang, S. Wu, X. Zhang, **F. Li**, G. Chen. In Proceedings of 42nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2023**), **Best Paper Award**, pages 1(1): 55:1-55:26, Seattle, USA, June, 2023.

34. *A Unified And Efficient Coordinating Framework For Autonomous DBMS Tuning*, by X. Zhang, C. Zhang, H. Wu, J. Tan, **F. Li**, B. Cui. In Proceedings of 42nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2023**), pages 1(2): 186:1-186:26, Seattle, USA, June, 2023.

35. *PolarDB-IMCI: A Cloud-Native HTAP Database System at Alibaba*, by J. Wang, T. Li, H. Song, X. Yang, W. Zhou, **F. Li**, B. Yan, Q. Wu, Y. Liang, C. Ying, Y. Wang, B. Chen, C. Cai, Y. Ruan, X. Weng, S. Chen, L. Yin, C. Yang, X. Cai, H. Xing, N. Yu, X. Chen, D. Huang, J. Sun:. In Proceedings of 42nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2023**), pages 1(2): 199:1-199:25, Seattle, USA, June, 2023.

36. *CatSQL: Towards Real World Natural Language to SQL Applications*, by H. Fu, C. Liu, B. Wu, **F. Li**, J. Tan, J. Sun. In Proceedings of 48th International Conference on Very Large Data Bases (**VLDB 2023**), pages 16(6): 1534-1547, Vancouver, Canada, August, 2023.

37. *OneShotSTL: One-Shot Seasonal-Trend Decomposition For Online Time Series Anomaly Detection And Forecasting*, by X. He, Y. Li, J. Tan, B. Wu, **F. Li**. In Proceedings of 48th International Conference on Very Large Data Bases (**VLDB 2023**), pages 16(6): 1399-1412, Vancouver, Canada, August, 2023.

38. *Encrypted Databases Made Secure Yet Maintainable*, by M. Li, X. Zhao, L. Chen, C. Tan, H. Li, S. Wang, Z. Mi, Y. Xia, **F. Li**, H. Chen. In Proceedings of 17th USENIX Symposium on Operating Systems Design and Implementation (**OSDI 2023**), pages 117-133, Boston, USA, June, 2023.

39. *Persistent Memory Disaggregation for Cloud-Native Relational Databases*, by C. Ruan, Y. Zhang, C. Bi, X. Ma, H. Chen, **F. Li**, X. Yang, C. Li, A. Aboulnaga, Y. Xu. In Proceedings of 28th conference on ACM Architectural Support for Programming Languages and Operating Systems (**ACM ASPLOS 2023**), pages 498-512, Vancouver, Canada, March 2023.

40. *Knock Out 2PC with Practicality Intact: a High-performance and General Distributed Transaction Protocol*, by Z. Lai, H. Fan, W. Zhou, Z. Ma, X. Peng, **F. Li**, E. Lo. In Proceedings of 38th IEEE International Conference on Data Engineering, (**IEEE ICDE 2023**), pages 2317-2331, California USA, April, 2023.

41. *Active Sampling for Sparse Table by Bayesian Optimization with Adaptive Resolution*, by X. He, J. Tan, B. Wu, **F. Li**, X. Zhang, G. Liang, J. Xu. In Proceedings of 38th IEEE International Conference on Data Engineering, (**IEEE ICDE 2023**), pages 816-828, California USA, April, 2023.

42. *EulerFD: An Efficient Double-Cycle Approximation of Functional Dependencies*, by Q. Lin, Y. Gu, J. Sai, J. Liu, K. Ren, L. Xiong, T. Wang, Y. Pang, S. Wang, **F. Li** . In Proceedings of 38th IEEE International Conference on Data Engineering, (**IEEE ICDE 2023**), pages 2878-2891, California USA, April, 2023.

43. *Database Workload Characterization with Query Plan Encoders*, by D. Paul, J. Cao, **F. Li**, V. Srikumar. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(4) pages 923-935, Sydney, Australia, August, 2022.

44. *Facilitating Database Tuning with Hyper-Parameter Optimization: A Comprehensive Experimental Evaluation*, by X. Zhang, Z. Chang, Y. Li, H. Wu, J. Tan, **F. Li**, B. Cui. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(9) pages 1808-1821, Sydney, Australia, August, 2022.

45. *AB-tree: Index for Concurrent Random Sampling and Updates*, by Z. Zhao, D. Xie, **F. Li**. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(9) pages 1835-1847, Sydney, Australia, August, 2022.

46. *SA-LSM : Optimize Data Layout for LSM-tree Based Storage using Survival Analysis*, by T. Zhang, J. Tan, X. Cai, J. Wang, **F. Li**, J. Sun. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(10) pages 2161-2174, Sydney, Australia, August, 2022.

47. *Operon: An Encrypted Database for Ownership-Preserving Data Management*, by S. Wang, Y. Li, H. Li, **F. Li**, C. Tian, L. Su, Y. Zhang, Y. Ma, L. Yan, Y. Sun, X. Cheng, X. Xie, Y. Zou. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(12) pages 3332-3345, Sydney, Australia, August, 2022.

48. *HEDA: Multi-Attribute Unbounded Aggregation over Homomorphically Encrypted Database*, by X. Ren, L. Su, Z. Gu, S. Wang, **F. Li**, Y. Xie, S. Bian, C. Li, F. Zhang. In Proceedings of 48th International Conference on Very Large Data Bases (**VLDB 2022**), 16(4) pages 601-614, Sydney, Australia, August, 2022.

49. *CloudJump: Optimizing Cloud Databases for Cloud Storages*, by Z. Chen, X. Yang, **F. Li**, X. Cheng, Q. Hu, Z. Miao, R. Xie, X. Wu, K. Wang, Z. Song, H. Sun, Z. Zhuang, Y. Yang, J. Xu, L. Yin, W. Zhou, S. Wang. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(12) pages 3432-3444, Sydney, Australia, August, 2022.

50. *Tair-PMem: a Fully Durable Non-Volatile Memory Database*, by C. Gong, C. Tian, Z. Wang, S. Wang, X. Wang, Q. Fu, W. Qin, Q. Long, R. Chen, J. Qi, R. Wang, G. Zhu, C. Yang, W. Zhang, **F. Li**. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(12) pages 3346-3358, Sydney, Australia, August, 2022.

51. *VRE: A Versatile, Robust, and Economical Trajectory Data System*, by H. Lan, J. Xie, Z. Bao, **F. Li**, W. Tian, F. Wang, S. Wang, A. Zhang. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(12) pages 3398-3410, Sydney, Australia, August, 2022.

52. *Ganos: A Multidimensional, Dynamic, and Scene-Oriented Cloud-Native Spatial Database Engine*, by J. Xie, Z. Chen, J. Liu, F. Wang, **F. Li**, Z. Chen, Y. Liu, S. Cai, Z. Fan, F. Xiao, Y. Chen. In Proceedings of 47th International Conference on Very Large Data Bases (**VLDB 2022**), 15(12) pages 3483-3495, Sydney, Australia, August, 2022.

53. *Releasing Private Data for Numerical Queries*, by Y. Qiu, W. Dong, K. Yi, B. Wu, **F. Li**. In Proceedings of 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**ACM SIGKDD 2022**,), pages 1410-1419, Washington DC, US, Aug, 2022.

54. *Remus: Efficient Live Migration for Distributed Databases with Snapshot Isolation*, by J. Kang, L. Cai, **F. Li**, X. Zhou, W. Cao, S. Cai, D. Shao. In Proceedings of 41st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2022**), pages 2232-2245, Philadelphia, USA, June, 2022.

55. *ESDB: Processing Extremely Skewed Workloads in Real-time*, by J. Zhang, S. Cheng, Z. Xue, J. Deng, C. Fu, W. Zhou, S. Wang, C. Chen, **F. Li**. In Proceedings of 41st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2022**), pages 2286-2298, Philadelphia, USA, June, 2022.

56. *Towards Practical Oblivious Join*, by Z. Chang, D. Xie, S. Wang, **F. Li**. In Proceedings of 41st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2022**), pages 803-817, Philadelphia, USA, June, 2022.

57. *PreQR: Pre-training Representation for SQL Understanding*, by X. Tang, S. Wu, M. Song, S. Yin, **F. Li**, G. Chen. **F. Li**. In Proceedings of 41st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2022**), pages 204-216, Philadelphia, USA, June, 2022.

58. *Towards Dynamic and Safe Configuration Tuning for Cloud Databases*, by X. Zhang, H. Wu, Y. Li, J. Tan, **F. Li**, B. Cui. In Proceedings of 41st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2022**), pages 631-645, Philadelphia, USA, June, 2022.

59. *PolarDB-X: An Elastic Distributed Relational Database for Cloud-Native Applications*, by W. Cao, **F. Li**, G. Huang, J. Lou, J. Zhao, D. He, M. Sun, Y. Zhang, S. Wang, X. Wu, H. Liao, Z. Chen, X. Fang, M. Chen, C. Liang, Y. Luo, H. Wang, S. Wang, Z. Ma, X. Yang, X. Peng, Y. Ruan, Y. Wang, J. Zhou, J. Wang, Q. Hu, J. Kang. In Proceedings of 37th IEEE International Conference on Data Engineering, (**IEEE ICDE 2022**), pages 2859-2872, Kuala Lumpur, Malaysia, May, 2022.

60. *PinSQL: Pinpoint Root Causes to Resolve SQL Performance Issues in Cloud Databases*, by X. Liu, Z. Yin, C. Zhao, C. Ge, L. Chen, Y. Gao, D. Li, Z. Wang, G. Liang, J. Tan, **F. Li** . In Proceedings of 37th IEEE International Conference on Data Engineering, (**IEEE ICDE 2022**), pages 2549-2561, Kuala Lumpur, Malaysia, May, 2022.

61. *A Comparative Study of in-Database Inference Approaches*, by Q. Lin, S. Wu, J. Zhao, J. Dai, **F. Li**, G. Chen. In Proceedings of 37th IEEE International Conference on Data Engineering, (**IEEE ICDE 2022**), pages 1794-1807, Kuala Lumpur, Malaysia, May, 2022.

62. *Ubiquitous Verification in Centralized Ledger Database*, by X. Yang, S. Wang, **F. Li**, Y. Zhang, W. Yan, F. Gai, B. Yu, L. Feng, Q. Gao, Y. Li. In Proceedings of 37th IEEE International Conference on Data Engineering, (**IEEE ICDE 2022**), pages 1808-1821, Kuala Lumpur, Malaysia, May, 2022.

63. *A Sampling-based Learning Framework for Big Databases*, by J. Zhang, S. Wu, J. Zhao, Z. Xie, **F. Li**, Y. Gao, G. Chen. In Proceedings of 18th The WEB Conference, (**WWW 2022**), pages 1871-1881, Lyon, France, April, 2022.

64. *Constrained Non-Affine Alignment of Embedding*, by Y. Wang, Y. Zheng, Y. Peng, M. Yeh, Z. Zhuang, D. Mahashweta, B. Mangesh, **F. Li**, W. Zhang, and J. Phillips In Proceedings of 21st IEEE International Conference on Data Mining (**IEEE ICDM 2021**), pages 1403-1408, Auckland, New Zealand, December, 2021.

65. *Towards Cost-Effective and Elastic Cloud Database Deployment via Memory Disaggregation*, by Y. Zhang, C. Ruan, C. Li, X. Yang, W. Cao, **F. Li**, B. Wang, J. Fang, Y. Wang, J. Huo, C. Bi. In Proceedings of 46th International Conference on Very Large Data Bases (**VLDB 2021**), pages 1900-1912, Copenhagen, Denmark, August, 2021.

66. *Revisiting the Design of LSM-tree Based OLTP Storage Engine with Persistent Memory*, by B. Yan, X. Cheng, B. Jiang, S. Chen, C. Shang, J., G. Huang, X. Yang, W. Cao, **F. Li**. In Proceedings of 46th International Conference on Very Large Data Bases (**VLDB 2021**), pages 1872-1885, Copenhagen, Denmark, August, 2021.

67. *Building Enclave-Native Storage Engines for Practical Encrypted Databases*, by Y. Sun, S. Wang, H. Li, **F. Li**. In Proceedings of 46th International Conference on Very Large Data Bases (**VLDB 2021**), pages 1019-1032, Copenhagen, Denmark, August, 2021.

68. *Cquirrel: Continuous Query Processing over Acyclic Relational Schemas*, Q. Wang, C. Zhang, D. Alsayed, K. Yi, B. Wu, **F. Li**, C. Zhan. In Proceedings of 46th International Conference on Very Large Data Bases (**VLDB 2021**), Demo paper, pages 2667-2670, Copenhagen, Denmark, August, 2021.

69. *PolarDB Serverless: A Cloud Native Database for Disaggregated Data Centers*, by W. Cao, Y. Zhang, X. Yang, **F. Li**, S. Wang, Q. Hu, X. Cheng, Z. Chen, Z. Liu, J. Fang, B. Wang, Y. Wang, H. Sun, Z. Yang, Z. Cheng, S. Chen, J. Wu, W. Hu, J. Zhao, Y. Gao, S. Cai, Y. Zhang, J. Tong. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 2447-2489, Xi'An, China, June, 2021.

70. *LogStore: A Cloud Native and Multi-Tenant Log Database*, by W. Cao, X. Feng, B. Liang, T. Zhang, Y. Gao, Y Zhang, **F. Li**. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 2464-2476, Xi'An, China, June, 2021.

71. *VeriDB: An SGX-based Verifiable Database*, by W. Zhou, Y. Cai, Y. Peng, S. Wang, K. Ma, **F. Li**. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 2182-2194, Xi'An, China, June, 2021.

72. *ResTune: Resource Oriented Tuning Boosted by Meta-Learning for Cloud Databases*, by X. Zhang, H. Wu, Z. Chang, S. Jin, J. Tan, **F. Li**, T. Zhang, B. Cui. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 2102-2114, Xi'An, China, June, 2021.

73. *At-the-time and Back-in-time Persistent Sketches*, by B. Shi, Y. Peng, Z. Zhao, J. Phillips, **F. Li**. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 1623-1636, Xi'An, China, June, 2021.

74. *Weighted Distinct Sampling: Cardinality Estimation for SPJ Queries*, by Y. Qiu, Y. Wang, K. Yi, **F. Li**, B. Wu, C. Zhan. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 1465-1477, Xi'An, China, June, 2021.

75. *Spatial Independent Range Sampling*, by D. Xie, J. Phillips, M. Matheny, **F. Li**. In Proceedings of 40th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2021**), pages 2023-2035, Xi'An, China, June, 2021.

76. *SlimStore: A Cloud-based Deduplication System for Multi-version Backups*, by Z. Zhang, H. Hu, Z. Xue, C. Chen, Y. Yu, C. Fu, X. Zhou, **F. Li**. In Proceedings of 36th IEEE International Conference on Data Engineering, (**IEEE ICDE 2021**), pages 1841-1846, Crete, Greece, April, 2021.

77. *HybrIDX: New Hybrid Index for Volume-hiding Range Queries in Data Outsourcing Services*, by K. Ren, Y. Guo, J. Li, X. Jia, C. Wang, Y. Zhou, S. Wang, N. Cao, **F. Li**. In Proceedings of 40th IEEE International Conference on Distributed Computing Systems, (**IEEE ICDCS 2020**), pages 23-33, **Best Paper Award**, Singapore, November, 2020.

78. *Efficient Join Synopsis Maintenance for Data Warehouse*, by Z. Zhao, **F. Li**, Y. Liu. In Proceedings of 39th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2020**), pages 2027–2042, Portland, USA, June, 2020.

79. *FalconDB: Blockchain-based Collaborative Database*, by Y. Peng, M. Du, **F. Li**, R. Cheng, D. Song. In Proceedings of 39th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2020**), pages 637–652, Portland, USA, June, 2020.

80. *Timon: A Timestamped Event Database for Efficient Telemetry Data Processing and Analytics*, by W. Cao, Y.Gao, **F. Li**, S. Wang, B. Lin, K. Xu, X. Feng, Y. Wang, Z. Liu, G. Zhang. In Proceedings of 39th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2020**), pages 739–753, Portland, USA, June, 2020.

81. *Leaper: A Learned Prefetcher for Cache Invalidation in LSM-tree based Storage Engines*, by L. Yang, H. Wu, T. Zhang, X. Cheng, **F. Li**, L. Zou, Y. Wang, R. Chen, J. Wang, G. Huang. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2020**), pages 13(11): 1976-1989, Tokyo, August, 2020.

82. *AnalyticDB-V: A Hybrid Analytical Engine Towards Query Fusion for Structured and Unstructured Data* , by C. Wei, B. Wu, S. Wang, R. Lou, C. Zhan, **F. Li**, Y. Cai. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2020**), pages 13(12): 3152-3165, Tokyo, August, 2020.

83. *LedgerDB: A Centralized Ledger Database for Universal Audit and Verification*, by X. Yang, Y. Zhang, S. Wang, B. Yu, **F. Li**, Y. Li, W. Yan. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2020**), pages 13(12): 3138-3151, Tokyo, August, 2020.

84. *Diagnosing Root Causes of Intermittent Slow Queries in Large-Scale Cloud Databases*, by M. Ma, Z. Yin, S. Zhang, S. Wang, C. Zheng, X. Jiang, H. Hu, C. Luo, Y. Li, N. Qiu, **F. Li**, C. Chen, D. Pei. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2020**), pages 13(8): 1176–1189, Tokyo, August, 2020.

85. *Two-Level Data Compression using Machine Learning in Time Series Database*, by X. Yu, Y. Peng, **F. Li**, S. Wang, X. Shen, H. Mai, Y. Xie, In Proceedings of 35th IEEE International Conference on Data Engineering, (**IEEE ICDE 2020**), pages 1370-1381, Dallas, USA, April, 2020.

86. *HotRing: A Hotspot-Aware In-Memory Key-Value Store*, by J. Chen, L. Chen, S. Wang, G. Zhu, Y. Sun, H. Liu, **F. Li**. In Proceedings of 18th USENIX Conference on File and Storage Technologies (**USENIX FAST 2020**), pages 239–252, Santa Clara, USA, February, 2020.

87. *FPGA-Accelerated Compactions for LSM-based Key Value Store*, by T. Zhang, J. Wang, X. Cheng, H. Xu, N. Yu, G. Huang, T. Zhang, D. He, **F. Li**, W. Cao, Z. Huang, J. Sun. In Proceedings of 18th USENIX Conference on File and Storage Technologies (**USENIX FAST 2020**), pages 225–237, Santa Clara, USA, February, 2020.

88. *Narrowing the Gap Between Serverless and its State with Storage Functions*, by T. Zhang, D. Xie, **F. Li**, R. Stutsman. In Proceedings of 10th ACM Symposium of Cloud Computing (**ACM SoCC 2019**), pages 1–12, Amsterdam, Netherland, June, 2019. **Socc Best Paper Award** (runner-up)

89. *Role of the TPC in the Cloud Age*, by A. Crolotte, **F. Li**, M. Poess, P. Boncz, R. Nambiar. In Proceedings of Technology Conference on Performance Evaluation and Benchmarking (**joint with VLDB 2019**), pages 133-138, Los Angeles, California, August, 2019.

90. *Cloud Native Database Systems at Alibaba: Opportunies and Challenges*, by **F. Li**, In Proceedings of 45th International Conference on Very Large Data Bases (**PVLDB 2019**), 12(12) pages 2263-2272, Los Angeles, California, August, 2019.

91. *iBTune: Individualized Buffer Tuning for Large-scale Cloud Databases*, by J. Tan, T. Zhang, **F. Li**, J. Chen, Q. Zheng, P. Zhang, H. Qiao, Y. Shi, W. Cao, R. Zhang. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2019**), 12(10) pages 1221-1234, Los Angeles, California, August, 2019.

92. *AnalyticDB: Real-time OLAP Database System at Alibaba Cloud*, by C. Zhan, M. Su, C. Wei, X. Peng, L. Lin, S. Wang, Z. Chen, **F. Li**, Y. Pan, F. Zheng, C. Chai. In Proceedings of 45th International Conference on Very Large Data Bases (**VLDB 2019**), pages 12(12) 2059-2070, Los Angeles, California, August, 2019.

93. *X-Engine: An Optimized Storage Engine for Large-scale E-Commerce Transaction Processing*, by G. Huang, X. Cheng, J. Wang, Y. Wang, D. He, T. Zhang, **F. Li**, S. Wang, W. Cao, Q. Li. In Proceedings of 38th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2019**), pages 651-665, Amsterdam, Netherland, June, 2019.

94. *Bursty Event Detection Throughout Histories*, by D. Paul, Y. Peng, **F. Li**, In Proceedings of 34th IEEE International Conference on Data Engineering, (**IEEE ICDE 2019**), pages 1370-1381, Macau, China, April, 2019.

95. *AI Pro: Data Processing Framework for AI Models*, by R. Frost, D. Paul, **F. Li**, In Proceedings of 34th IEEE International Conference on Data Engineering, (**IEEE ICDE 2019**), pages 1980-1983, Macau, China, April, 2019.

96. *Pcard: Personalized Restaurants Recommendation from Card Payment Transaction Records*, by M. Du, R. Christensen, W. Zhang, **F. Li**, In Proceedings of 15th The WEB Conference, (**WWW 2019**), pages 2687-2693, San Francisco, May, 2019.

97. *OpenTag: Open Attribute Value Extraction from Product Profiles*, by G. Zheng, S. Mukherjee, X. Dong, **F. Li**. In Proceedings of 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**ACM SIGKDD 2018**, Applied Data Science Track, Oral Presentation), pages 1049-1058, London, UK.

98. *Solar: Towards a Shared-Everything Database on Distributed Log-Structured Storage*, by T. Zhu, Z. Zhao, **F. Li**, W. Qian, A. Zhou, D. Xie, R. Stutsman, H. Li, H. Hu. In Proceedings of the 2018 USENIX Annual Technical Conference (**USENIX ATC 2018**), pages 795-807, Boston, USA, July, 2018.

99. *Persistent Bloom Filter: Membership Testing for the Entire History*, by Y. Peng, J. Guo, **F. Li**, W. Qian, A. Zhou. In Proceedings of 37th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2018**), pages 1037-1052, Houston, USA, June, 2018.

100. *Random Sampling Over Joins Revisited*, by Z. Zhao, R. Christensen, **F. Li**, X. Hu, K. Yi. In Proceedings of 37th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2018**), pages 1525-1539, Houston, USA, June, 2018.

101. *Secure DIMM: Moving ORAM Primitives Closer to Memory*, by A. Shafiee, R. Balasubramonian, M. Tiwari, **F. Li**, In Proceedings of 24th IEEE International Symposium on High-Performance Computer Architecture (**IEEE HPCA 2018**), pages 428–440, Vienna, Austria, February, 2018.

102. *DeepLog: Anomaly Detection And Diagnosis From System Logs Through Deep Learning*, by M. Du, **F. Li**, G. Zheng, V. Srikumar. In Proceedings of 24th ACM Conference on Computer and Communications Security (**CCS 2017**), pages 1285–1298, Dallas, 2017.

103. *Distributed Trajectory Similarity Search*, by D. Xie, **F. Li**, J. Phillips. In Proceedings of 43rd International Conference on Very Large Data Bases (**VLDB 2017**), PVLDB 10(11), pages 1478-1489, Munich, 2017.

104. *Compass: Spatio Temporal Sentiment Analysis of US Election*, by D. Paul, **F. Li**, K. Teja, X. Yu, R. Frost. In Proceedings of 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**ACM SIGKDD 2017**, Applied Data Science Track, Oral Presentation), pages 1585-1594, Halifax, Canada.

105. *Spell: Streaming Parsing of System Event Logs*, by M. Du, **F. Li**. In Proceedings of 16th IEEE International Conference on Data Mining (**IEEE ICDM 2016**), pages 859-864, Barcelona, Spain, December, 2016.

106. *Fast And Concurrent RDF Queries With RDMA-Based Distributed Graph Exploration*, by J. Shi, Y. Yao, R. Chen, H. Chen, **F. Li**. In Proceedings of 12th USENIX Symposium on Operating Systems Design and Implebbmentation (**OSDI 2016**), pages 317-332, Savannah, USA, November, 2016.

107. *Simba: Efficient In-Memory Spatial Analytics*, by D. Xie, **F. Li**, B. Yao, G. Li, L. Zhou, M. Guo In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 1071-1085, San Francisco, USA, June, 2016.

108. #*Wander Join: Online Aggregation via Random Walks*, by **F. Li**, B. Wu, K. Yi, Z. Zhao In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 615-629, San Francisco, USA, June, 2016. **SIGMOD Best Paper Award**

109. *Graph Analytics Through Fine-Grained Parallelism*, by Z. Shang, **F. Li**, J. X. Yu, Z. Zhang, H. Cheng. In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 463-478, San Francisco, USA, June, 2016.

110. *Matrix Sketching Over Sliding Windows*, by Z. Wei, X. Liu, **F. Li**, S. Shang, X. Du, J. Wen. In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 1465-1480, San Francisco, USA, June, 2016.

111. *Privacy Preserving Subgraph Matching on Large Graphs in Cloud*, by Z. Chang, L. Zou, **F. Li**, In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 199-213, San Francisco, USA, June, 2016.

112. # *Wander Join: Online Aggregation for Joins*, by **F. Li**, B. Wu, K. Yi, Z. Zhao Demo Ppaer, In Proceedings of 35th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2016**), pages 2121-2124, San Francisco, USA, June, 2016.

113. *Oblivious RAM: A Dissection and Experimental Evaluation*, by Z. Chang, D. Xie, **F. Li**, In Proceedings of 42nd International Conference on Very Large Data Bases (**VLDB 2016**), PVLDB 9(12), pages 1113-1124, New Delhi India, 2016.

114. *Spatial Online Sampling and Aggregation*, by L. Wang, R. Christensen, **F. Li**, K. Yi, In Proceedings of 42nd International Conference on Very Large Data Bases (**VLDB 2016**), PVLDB 9(3), pages 84-95, New Delhi India, 2016.

115. *Simba: Spatial In-Memory Big Data Analytics*, by D. Xie, **F. Li**, B. Yao, G. Li, L. Zhou, Z. Chen, M. Guo. Demo Paper. In Proceedings of 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (**ACM SIGSPATIAL 2016**), pages 86:1-86:4, San Francisco, USA, 2016.

116. *ATOM: Automated Tracking, Orchestration, and Monitoring of Resource Usage in Infrastructure as a Service Systems*, by M. Du, **F. Li**, In Proceedings of IEEE International Conference on Big Data (**IEEE BigData 2015**), pages 271-278, Santa Clara CA, November, 2015.

117. *Fixed-Function Hardware Sorting Accelerators for Near Data MapReduce Execution*, by S. Pugsley, A. Deb, R. Balasubramonian, **F. Li**, In Proceedings of 33rd IEEE International Conference on Computer Design (**IEEE ICCD-33**,), pages 439-442, New York, October, 2015.

118. *Distributed Online Tracking*, by M. Tang, **F. Li**, Y. Tao, In Proceedings of 34th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2015**), pages 2047-2061, Melbourne, Australia, June 2015.

119. *STORM: Spatio-Temporal Online Reasoning and Management of Large Spatio-Temporal Data*, by R. Christensen, L. Wang, **F. Li**, K. Yi, J. Tang, N. Villa, In Proceedings of 34th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2015**), system demo, pages 1111-1116, Melbourne, Australia, June 2015 (**SIGMOD Best Demonstration Award**).

120. *Scalable Histograms on Large Probabilistic Data*, by M. Tang, **F. Li**, In Proceedings of 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**ACM SIGKDD 2014**), pages 631–640, NYC, NY, August 2014.

121. *Continuous Matrix Approximation on Distributed Data*, by M. Ghashami, J. Phillips, **F. Li**, In Proceedings of 40th International Conference on Very Large Databases (**VLDB 2014**), pages 809–820, Hangzhou, China, September 2014.

122. *NDC: Analyzing the Impact of 3D-Stacked Memory+Logic Devices on MapReduce Workloads*, by S. Pugsley, J. Jestes, H. Zhang, R. Balasubramonian, V. Srinivasan, A. Buyuktosunoglu, A. Davis, **F. Li**, In Proceedings of 2014 IEEE International Symposium on Performance Analysis of Systems and Software (**IEEE ISPASS 2014**), pages 190–200, Monterey, CA, March 2014.

123. *Optimal Splitters for Temporal and Multi-version Databases*, by W. Le, **F. Li**, Y. Tao, R. Christensen, In Proceedings of 32nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2013**), pages 109-120, NYC, NY, June 2013.

124. *Quality and Efficiency for Kernel Density Estimates in Large Data*, by Y. Zheng, J. Jestes, J. Phillips, **F. Li**, In Proceedings of 32nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2013**), pages 433-444, NYC, NY, June 2013.

125. *Adaptive Log Compression for Massive Log Data*, by R. Christensen, **F. Li**, In Proceedings of 32nd ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2013**, Undergraduate Research Poster), **SIGMOD Best Undergraduate Research Poster Award**, pages 1283-1284, NYC, NY, June 2013.

126. *Secure Nearest Neighbor Revisited*, by B. Yao, **F. Li**, X. Xiao, In Proceedings of 29th IEEE International Conference on Data Engineering, (**IEEE ICDE 2013**), pages 733-744, Brisbane, Australia, April, 2013.

127. *LogKV: Exploiting Key-Value Stores for Log Processing*, by Z. Cao, S. Chen, **F. Li**, M. Wang, X. Sean Wang, In Proceedings of 6th Biennial Conference on Innovative Data Systems Research (**CIDR 2013**), pages NA, Asilomar, California, January 2012.

128. *Ranking Large Temporal Data*, by J. Jestes, J. Phillips, **F. Li**, M. Tang, In Proceedings of 38th International Conference on Very Large Databases (**VLDB 2012**), PVLDB 5(11) pages 1412-1223, Istanbul, Turkey, August 2012.

129. *Building Wavelet Histograms on Large Data in MapReduce*, by J. Jestes, K. Yi, **F. Li**, In Proceedings of 38th International Conference on Very Large Databases (**VLDB 2012**), PVLDB 5(2): 109-120, Istanbul, Turkey, August 2012.

130. *Towards Fair Sharing of Block Storage in a Multi-tenant Cloud*, by X. Ling, Y. Mao, **F. Li**, R. Ricci, In Proceedings of 4th USENIX Workshop on Hot Topics in Cloud Computing (**USENIX HotCloud 2012**, Boston, MA, June 2012.

131. *ColumbuScout: Towards Building Local Search Engines over Large Databases*, by C. Hansen, **F. Li**, In Proceedings of 31st ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2012**, system demo), pages 617-620, Scottsdale, Arizona, May 2012.

132. *Efficient Parallel kNN Joins for Large Data in MapReduce*, by C. Zhang, **F. Li**, J. Jestes, In Proceedings of 15th International Conference on Extending Database Technology (**EDBT 2012**), pages 38-49, Berlin, Germany, March 2012.

133. *Efficient Threshold Monitoring for Distributed Probabilistic Data*, by M. Tang, **F. Li**, J. Phillips, J. Jestes, In Proceedings of 28th IEEE International Conference on Data Engineering, (**IEEE ICDE 2012**), pages 1120-1131, Washington DC, April, 2012.

134. *Scalable Multi-Query Optimization for SPARQL*, by W. Le, A. Kementsietsidis, S. Duan, **F. Li**, In Proceedings of 28th IEEE International Conference on Data Engineering, (**IEEE ICDE 2012**), pages 666-677, Washington DC, April, 2012.

135. *Multi-Approximate-Keyword Routing in GIS Data*, by B. Yao, M. Tang, **F. Li**, In Proceedings of 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (**ACM SIGSPATIAL GIS 2011**), pages 201-210, Chicago, Illinois, November 2011.

136. *Flexible Aggregate Similarity Search*, by Y. Li, **F. Li**, K. Yi, B. Yao, M. Wang, In Proceedings of 30th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2011**), pages 1009-1020, Athens, Greece, June 2011.

137. *Rewriting Queries on SPARQL Views*, by W. Le, S. Duan, A. Kementsietsidis, **F. Li**, M. Wang, In Proceedings of 20th International World Wide Web Conference, (**WWW 2011**), pages 655–664, Hyderabad, India, March, 2011.

138. *Optimal Location Queries in Road Network Databases*, by X. Xiao, B. Yao, **F. Li**, In Proceedings of 27th IEEE International Conference on Data Engineering, (**IEEE ICDE 2011**), pages 804-815, Hannover, Germany, April, 2011.

139. *#Probabilistic String Similarity Joins*, by J. Jestes, **F. Li**, Z. Yan, K. Yi, In Proceedings of 29th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2010**), pages 327-338, Indianapolis, June 2010.

140. *Logging Every Footstep: Quantile Summaries for the Entire History*, by Y. Tao, K. Yi, C. Sheng, J. Pei, **F. Li**. In Proceedings of 29th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2010**), pages 639-650, Indianapolis, June 2010.

141. *Approximate String Search in Spatial Databases*, by B. Yao, **F. Li**, M. Hadjieleftheriou, K. Hou, (Full Paper) In Proceedings of 26th IEEE International Conference on Data Engineering, (**IEEE ICDE 2010**), pages 545-556, Long Beach, California, March 2010.

142. *K Nearest Neighbor Queries and kNN-Joins in Large Relational Databases (Almost) for Free*, by B. Yao, **F. Li**, P. Kumar, (Full Paper) In Proceedings of 26th IEEE International Conference on Data Engineering, (**IEEE ICDE 2010**), pages 4-15, Long Beach, California, March 2010.

143. *Ranking Distributed Probabilistic Data*, by **F. Li**, K. Yi, J. Jestes, In Proceedings of 28th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2009**), pages 361-374, Providence, June 2009.

144. *#Semantics of Ranking Queries for Probabilistic Data and Expected Ranks*, by G. Cormode, **F. Li**, K. Yi, (Full Paper) In Proceedings of 25th IEEE International Conference on Data Engineering (**IEEE ICDE2009**), pages 305-316, Shanghai, China, April 2009.

145. *Reverse Furthest Neighbors in Spatial Databases*, by B. Yao, **F. Li**, P. Kumar, (Full Paper) In Proceedings of 25th IEEE International Conference on Data Engineering, (**IEEE ICDE 2009**), pages 664-675, Shanghai, China, April 2009.

146. *Improving Transaction-time DBMS Performance and Functionality*, by D. Lomet, **F. Li**, (Full Paper) In Proceedings of 25th IEEE International Conference on Data Engineering, (**IEEE ICDE 2009**), pages 581-591, Shanghai, China, April 2009.

147. *A Concise Representation of Range Queries*, by K. Yi, X. Lian, **F. Li**, L. Chen, (Short Paper) In Proceedings of 25th IEEE International Conference on Data Engineering, (**IEEE ICDE 2009**), pages 1179-1182, Shanghai, China, April 2009.

148. *Finding Frequent Items in Probabilistic Data*, by Q. Zhang, **F. Li**, K. Yi, In Proceedings of 27th ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2008**), pages 819-832, Vancouver, Canada, June 2008.

149. *Randomized Synopses for Query Assurance on Data Streams*, by K. Yi, **F. Li**, M. Hadjieleftheriou, D. Srivastava, G. Kollios, (Full Paper, Full Presentation) In Proceedings of 24th International Conference on Data Engineering (**IEEE ICDE 2008**), pages 416-425, Cancun, Mexico, April 2008.

150. *Efficient Processing of Top-k Queries in Uncertain Databases*, by K. Yi, **F. Li**, D. Srivastava, G. Kollios, (Poster Paper) In Proceedings of 24th IEEE International Conference on Data Engineering (**IEEE ICDE 2008**), pages 1406-1408, Cancun, Mexico, April 2008.

151. *Proof-Infused Streams: Enabling Authentication of Sliding Window Queries On Streams*, by **F. Li**, K. Yi, M. Hadjieleftheriou, G. Kollios, In Proceedings of the 33rd International Conference on Very Large Databases(**VLDB 2007**), pages 147-158, Vienna, Austria, September 2007.

152. *Time Series Compressibility and Privacy*, by S. Papadimitriou, **F. Li**, G. Kollios, P. S. Yu, In Proceedings of the 33rd International Conference on Very Large Databases (**VLDB 2007**), pages 459-470, Vienna, Austria, September 2007.

153. *Hiding in the Crowd: Privacy Preservation on Evolving Streams through Correlation Tracking*, by **F. Li**, J. Sun, S. Papadimitriou, G. Mihaila, and I. Stanoi. In Proceedings of the 23rd IEEE International Conference on Data Engineering (**IEEE ICDE 2007**), pages 686-695, Istanbul, Turkey, April 2007.

154. *Dynamic Authenticated Index Structures for Outsourced Databases*, by **F. Li**, M. Hadjieleftheriou, G. Kollios, and L. Reyzin. In Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data (**ACM SIGMOD 2006**), pages 121-132, Chicago, Illinois, USA, June 2006.

155. *Characterizing and Exploiting Reference Locality in Data Stream Applications*, by **F. Li**, C. Chang, G. Kollios, and A. Bestavros. In Proceedings of the 22nd IEEE International Conference on Data Engineering (**IEEE ICDE 2006**), pages 81-92, Atlanta, Georgia, USA, April 2006.

156. *On Trip Planning Queries in Spatial Databases*, by **F. Li**, D. Cheng, M. Hadjieleftheriou, G. Kollios, and S.-H. Teng. In Proceedings of the 9th International Symposium on Spatial and Temporal Databases (**SSTD 2005**), pages 273-290, Angra dos Reis, Brazil, August, 2005.

157. *Approximate Aggregation Techniques for Sensor Databases*, by J. Considine, **F. Li**, G. Kollios, and J. Byers. In Proceedings of the 20th IEEE International Conference on Data Engineering (**IEEE ICDE 2004**), pages 449-460, Boston, MA, March 30 - April 2, 2004. **Best Paper Award**

158. *Spatio-Temporal Aggregation Using Sketches*, by Y. Tao, G. Kollios, J. Considine, **F. Li**, and D. Papadias. In Proceedings of the 20th IEEE International Conference on Data Engineering (**IEEE ICDE 2004**), pages 214-226, Boston, MA, March 30 - April 2, 2004.

159. *WICCAP Data Model: Mapping Physical Websites to Logical Views*, by Z. Liu, **F. Li**, and W. Ng. In Proceedings of the 21st International Conference on Conceptual Modeling (**ER 2002**), pages 120-134, Tampere, Finland, October 8-10, 2002. Springer.

160. *Web Information Collection, Collaging and Programming*, by **F. Li**, Z. Liu, Y. Huang, and W. Ng. In Proceedings of the 3rd IEEE International Conference on Information, Communications & Signal Processing (**IEEE ICICS 2001**), Singapore, October 15-18, 2001.

Refereed Workshops

1. *A Visual Tool for Building Logical Data Models of Websites*, by Z. Liu, W. Ng, **F. Li**, and E. Lim. In Proceedings of the Fourth International Workshop on Web Information and Data Management (**WIDM02**), in conjunction with the Eleventh International Conference on Information and Knowledge Management (CIKM2002), pages 92-95, McLean, Virginia, November 4-9, 2002.

2. *An Information Concierge for the Web*, by **F. Li**, Z. Liu, Y. Huang, and W. Ng. In Proceedings of the First International Workshop on Internet Bots: Systems and Applications (**INBOSA 2001**), in conjunction with the 12th International Conference on Database and Expert System Applications (**DEXA 2001**), pages 672-676, Munich, Germany, September 3-8, 2001.

Technical Reports

1. *Authenticated Index Structures for Aggregation Queries in Outsourced Databases*, by **F. Li**, M. Hadjieleftheriou, G. Kollios, and L. Reyzin, BUCS-TR-2006-011, May, 2006.

2. *GreedyDual-Join: Locality-Aware Buffer Management for Approximate Join Processing Over Data Streams*, by **F. Li**, C. Cheng, A. Bestavros , and G. Kollios, BUCS-TR-2004-028, April, 2004.

Professional Service

- Services

1. Associate Editor, IEEE Transactions on Cloud Computing, Since 2019.
2. Member of the Steering Committee, IEEE Transactions on Big Data (IEEE, TBD), since June 2018.
3. Associate Editor, ACM Transactions on Database Systems (ACM TODS), since January, 2017.
4. Associate Editor, IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), since November, 2013.
5. Editor in Chief, Survey Track, Distributed and Parallel Databases (DAPD), since August 2017.
6. Member of the SIGMOD Jim Gray PhD Dissertation Award Committee, (Starting 2017).
7. Associte Managing Director, Data Science and Engineering (DSE) Journal by Springer, since 2016.
8. Senior Area Chair, SIGKDD'22, 28TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining.
9. PC Area Chair, ACM SIGMOD 2024, 42nd ACM SIGMOD/PODS International Conference on Management of Data and Principle of Database Systems.
10. Senior Area Chair, SIGKDD'23, 29TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining.
11. PC co-Chair, Industry Track, VLDB'2022, 48th International Conference on Very Large Databases.
12. PC co-Chair, IEEE ICDE'2022, 38th IEEE International Conference on Data Engineering.
13. PC co-Chair, Industry Track, VLDB'2021, 47th International Conference on Very Large Databases.
14. PC co-Chair, Industry Track, IEEE ICDE'2021, 37th IEEE International Conference on Data Engineering.
15. Senior Area Chair, SIGKDD'21, 27TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining.
16. Senior Area Chair, SIGKDD'19, 25TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining.
17. PC co-Chair, PhD Symposium, VLDB'2019, 45th International Conference on Very Large Databases.
18. PC co-Chair, PhD Symposium, ICDE'2019, 35th IEEE International Conference on Data Engineering.
19. Senior PC member, ACM SIGMOD 2019, 38th ACM SIGMOD/PODS International Conference on Management of Data and Principle of Database Systems.
20. PC co-Chair, IEEE DCS 2019, 4th IEEE International Conference on Data Science in Cyberspace.
21. Demo PC co-Chair, ACM SIGMOD 2018, 37th ACM SIGMOD/PODS International Conference on Management of Data and Principle of Database Systems.
22. Member of the Best Paper Award selection committee, ACM CIKM 2017.
23. PC co-Chair, NDBC 2017, National Database Conference of China.
24. PC co-Chair, NSF 2016 workshop on Data Science for Secure and Privacy-aware (DSSP) Large Data Management and Mining.
25. PC co-Chair, APWeb 2016, The 18th Asia Pacific Web Conference.
26. PC Associate Chair, ACM SIGMOD 2015, 34th ACM SIGMOD/PODS International Conference on Management of Data and Principle of Database Systems.
27. PC co-Chair, WAIM'14, 15th International Conference on Web Age Information Management.
28. General co-Chair, SIGMOD'14, 33rd ACM SIGMOD/PODS International Conference on Management of Data and Principle of Database Systems.
29. PC Area Vice Chair, ICDE'14, 30th IEEE International Conference on Data Engineering.
30. Demo PC Chair, VLDB'14, 40th International Conference on Very Large Databases.
31. Publicity co-Chair, DASFAA'14, 19th International Conference on Database Systems for Advanced Applications.
32. PC co-Chair, CloudDB'13, 5th International Conference on Cloud Data Management.
33. co-PC Chair, MobiDE'13, 12th International ACM Workshop on Data Engineering for Wireless and Mobile Access (in conjunction with ACM SIGMOD/PODS 2013).
34. Proceedings co-Chair, ACM GIS'13, 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2013).

35. SIGMOD 2012 Mentorship Chair (in charge of the SIGMOD travel awards, the SIGMOD recruiting event, and the SIGMOD vis-a-vis meeting).1
36. Member of Editorial Advisory Board, Information Systems, Elsevier.
37. Guest Editor, Distributed and Parallel Databases, Special Issue of Scalable Data Summarization on Big Data.
38. Associate Editor, Journal of Computer Science and Technology (Database and Data Management area).
39. Publicity co-Chair, WISE'10, 11th IEEE International Conference on Web Information System Engineering.
40. Proceedings co-Chair, ICDE'10, IEEE 26th International Conference on Data Engineering.
41. Publicity Chair, MobiDE'09, 8th International ACM Workshop on Data Engineering for Wireless and Mobile Access (in conjunction with ACM SIGMOD/PODS 2009).

- Invited Reviewers/Panelists for Grant Proposals:

  1. NSF: USA National Science Foundation (08, 09, 10, 11, 12, 13, 14, 15, 16, 17, 18).
  2. Hongkong RGC: Hongkong Research Grants Council (08, 09, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19).
  3. Qatar National Research Fund (QNRF), 12, 14, 15, 16.
  4. Vienna Science and Technology Fund, 12.
  5. Romanian National Council for Research and Development, 12.
  6. Kazakhstan, National Centre of Science and Technology, 14.

- Program Committees:

  1. SIGMOD'24, 42nd ACM SIGMOD International Conference on Management of Data.
  2. VLDB'23, 48th International Conference on Very Large Databases.
  3. SIGKDD'22, 28TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining, Senior Area Chair.
  4. SIGMOD'22, 40th ACM SIGMOD International Conference on Management of Data.
  5. VLDB'22, 47th International Conference on Very Large Databases, PC co-Chair, Industry Track.
  6. IEEE ICDE'22, IEEE 39th International Conference on Data Engineering, PC co-Chair.
  7. SIGMOD'21, 39th ACM SIGMOD International Conference on Management of Data.
  8. VLDB'21, 46th International Conference on Very Large Databases, PC co-Chair, Industry Track.
  9. IEEE ICDE'21, IEEE 38th International Conference on Data Engineering, PC co-Chair, Indutry Track.
  10. SIGKDD'19, 25TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining, Senior Area Chair.
  11. SIGMOD'20, 38th ACM SIGMOD International Conference on Management of Data (senior PC).
  12. VLDB'20, 45th International Conference on Very Large Databases.
  13. SoCC'20, 11th ACM Symposium on Cloud Computing.
  14. SIGKDD'19, 25TH ACM SIGKDD Conference on Knolwedge Discovery and Data Mining.
  15. SoCC'19, 10th ACM Symposium on Cloud Computing.
  16. SIGMOD'19, 37th ACM SIGMOD International Conference on Management of Data (senior PC).
  17. VLDB'19, 44th International Conference on Very Large Databases.
  18. SIGMOD'18, 36th ACM SIGMOD International Conference on Management of Data.
  19. VLDB'18, 43rd International Conference on Very Large Databases.
  20. SIGMOD'17, 35th ACM SIGMOD International Conference on Management of Data.
  21. VLDB'17, 42nd International Conference on Very Large Databases.
  22. IEEE ICDE'17, IEEE 34th International Conference on Data Engineering.
  23. EDBT'16, 19th International Conference on Extending Database Technology.
  24. SIGMOD'16, 34th ACM SIGMOD International Conference on Management of Data.
  25. VLDB'16, 41st International Conference on Very Large Databases.
  26. IEEE ICDE'16, IEEE 33rd International Conference on Data Engineering.
  27. SIGMOD'15, 33rd ACM SIGMOD International Conference on Management of Data.
  28. VLDB'15, 40th International Conference on Very Large Databases.
  29. IEEE ICDE'15, IEEE 32nd International Conference on Data Engineering.
  30. IEEE ICDE'14, IEEE 31st International Conference on Data Engineering.

31. GeoRich 2014, 1st International ACM Workshop on Managing and Mining Semantically Rich Geo-Spatial Data.
32. VLDB'14, 39th International Conference on Very Large Databases.
33. SOCC'13, 4th ACM Symposium on Cloud Computing.
34. ACM CIKM'13, ACM 22nd Conference on Information and Knowledge Management (Demo and Poster Tracks).
35. IEEE BigData'13, 2013 IEEE International Conference on Big Data.
36. WebDB'13, 16th International Workshop on the Web and Databases.
37. SSTD'13, 13th International Symposium on Spatial and Temporal Databases.
38. DASFAA'13, 18th International Conference on Database Systems for Advanced Applications (Demo Track).
39. PAKDD'13, 17th Pacific-Asia Conference on Knowledge Discovery and Data Mining.
40. SIGMOD'13, 32nd ACM SIGMOD International Conference on Management of Data.
41. VLDB'13, 38th International Conference on Very Large Databases.
42. IEEE ICDE'13, IEEE 29th International Conference on Data Engineering.
43. IEEE ICDE'13 PhD Symposium, PhD Symposium of IEEE 29th International Conference on Data Engineering.
44. ICDCS'13, 33rd International Conference on Distributed Computing Systems.
45. WAIM'13, 14th International Conference on Web-Age Information Management.
46. QUeST'12, 3rd SIGSPATIAL ACM GIS 2011 International Workshop on Querying and Mining Uncertain Spatio-Temporal Data.
47. DBRank'12, 6th International Workshop on Ranking in Databases (in conjunction with VLDB'12).
48. ACM CIKM'12, ACM 21th Conference on Information and Knowledge Management.
49. ACM GIS'12, 20th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.
50. WAIM'12, 13th International Conference on Web-Age Information Management.
51. IEEE ICDM'12, IEEE 12th International Conference on Data Mining.
52. PAKDD'12, 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining.
53. VLDB'12, 38th International Conference on Very Large Data Bases, Demo Committee.
54. MAMIVE'11, 1st ACM SIGSPATIAL International Workshop on Managing and Mining Virtual Environments.
55. CANS'11, 10th International Conference on Cryptography and Network Security.
56. QUeST'11, 2nd SIGSPATIAL ACM GIS 2011 International Workshop on Querying and Mining Uncertain Spatio-Temporal Data.
57. ACM GIS'11, 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.
58. IEEE ICDE'12, IEEE 28th International Conference on Data Engineering.
59. IEEE ICDE'12, IEEE 28th International Conference on Data Engineering, Demo Committee.
60. ECML-PKDD'11, 2011 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases.
61. ACM CIKM'11, ACM 20th Conference on Information and Knowledge Management.
62. DBPL'11, The 13th International Symposium on Database Programming Languages.
63. IEEE ICDM'11, IEEE 11th International Conference on Data Mining.
64. WAIM'11, 12th International Conference on Web-Age Information Management.
65. SSDBM'11, 23rd International Conference on Scientific and Statistical Database Management.
66. ACM MobiDE'11, 10th International ACM Workshop on Data Engineering for Wireless and Mobile Access.
67. SSTD'11, 12th International Symposium on Spatial and Temporal Databases, PC member and Demo Committee.
68. PETSE'11, 3rd International Workshop on Privacy Enhanced Technology and Security Engineering.
69. ACM SIGMOD'11, 30th ACM SIGMOD International Conference on Management of Data, Best Demo Committee.
70. ACM SIGMOD'11, 30th ACM SIGMOD International Conference on Management of Data, Demo

Committee.

71. ICS'10, 13th Information Security Conference.
72. KMIS'10, International Conference on Knowledge Management and Information Sharing.
73. WCMT'10, 2nd International Workshop on Web-based Contents Management Technologies.
74. ACM SIGMOD'10, 29th ACM SIGMOD International Conference on Management of Data, Best Demo Committee.
75. PDCAT-10, 11th International Conference on Parallel and Distributed Computing, Applications and Technologies.
76. IEEE HPCC'10, IEEE 12th International Conference on High Performance Computing and Communications.
77. IEEE ICDE'11, IEEE 27th International Conference on Data Engineering.
78. IEEE ICDM'10, IEEE 10th International Conference on Data Mining.
79. WAIM'10, 11th International Conference on Web-Age Information Management.
80. ACM CIKM'10, ACM 19th Conference on Information and Knowledge Management.
81. ACM MobiDE'10, 9th International ACM Workshop on Data Engineering for Wireless and Mobile Access.
82. UMMM'10, 1st International Workshop on Uncertain Mobile Data Management and Mining.
83. DASFAA'10, 15th International Conference on Database Systems for Advanced Applications (Demon Track).
84. MDM'10, 11th International Conference on Mobile Data Management.
85. QUeST'09, 1st SIGSPATIAL ACM GIS 2009 International Workshop on Querying and Mining Uncertain Spatio-Temporal Data.
86. MOUND'10, 2nd International Workshop on Management and mining Of UNcertain Data (in conjunction with ICDE 2010).
87. ACM WIDM'09, ACM 11th International Workshop on Web Information and Data Management, co-located with CIKM'09.
88. IEEE ICDM'09, IEEE 9th International Conference on Data Mining.
89. KMIS'09, International Conference on Knowledge Management and Information Sharing.
90. ACM CIKM'09, ACM 18th Conference on Information and Knowledge Management.
91. ACM MobiDE'09, 8th International ACM Workshop on Data Engineering for Wireless and Mobile Access.
92. SDM'09, 9th SIAM International Conference on Data Mining.
93. PAKDD'09, 13th Pacific-Asia Conference on Knowledge Discovery and Data Engineering.
94. IEEE HPCC'09, IEEE 11th International Conference on High Performance Computing and Communications.
95. IEEE ICDE'09, IEEE 25th International Conference on Data Engineering.
96. IEEE ICDM'08, IEEE 8th International Conference on Data Mining.
97. ACM WIDM'08, ACM 10th International Workshop on Web Information and Data Management, co-located with CIKM'08.
98. ACM CIKM'08, ACM 17th Conference on Information and Knowledge Management.
99. SSDBM'08, 20th International Conference on Scientific and Statistical Database Management.
100. IEEE NetDB'08, Networking Meets Databases, co-located with IEEE ICDE 2008.

- Invited Reviewers for the following journals:
  1. ACM Transactions on Database Systems (ACM TODS).
  2. Data & Knowledge Engineering (DKE).
  3. IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE).
  4. IEEE Transactions on Parallel and Distributed Systems (IEEE TPDS).
  5. IEEE Transactions on Vehicular Technology (IEEE TVT).
  6. IEEE Transactions on Knowledge Discovery from Data (IEEE TKDD).
  7. IEEE Transactions on Dependable and Secure Computing (IEEE TDSC).
  8. IEEE/ACM Transactions on Networking.
  9. Distributed and Parallel Databases (DAPD by Springer).
  10. International Journal of Information Security.
  11. International Journal of Approximate Reasoning.

12. Journal of Medical System.
13. Journal of Computer Science and Technology.
14. Journal of STEM Education.
15. Very Large Database Journal (VLDBJ).
16. Knowledge and Information Systems.
17. Information Systems.
18. Statistical Analysis and Data Mining.
19. Electronics and Telecommunications Research Institute Journal (ETRI Journal).
20. Algorithms (one of MDPI's open access journals).
21. Geographical Information Science.
22. International Journal of Geographical Information Science.

Teaching Experience

- Track Director, School of Computing, University of Utah.
  Director for the Data Track Graduate Program. Fall 2011 to Fall 2013.

- Instructor, ZhiYuan College, Shanghai Jiao Tong University.
  ACM Class, Database Systems and Big Data Management.
  Fall 2014 and Spring 2015.

- Instructor, School of Computing, University of Utah.
  Graduate Course: CS6530, Advanced Database Systems.
  Fall 2012, Fall 2013, Fall 2015, Fall 2016, Fall 2017.

- Instructor, School of Computing, University of Utah.
  Undergraduate Course: CS5530 Database Systems.
  Spring 2012, Spring 2013, Spring 2014, Spring 2016, Spring 2017

- Instructor, School of Computing, University of Utah.
  Graduate Course: CS7941 Data Reading Group.
  Spring 2012, Spring 2013, Fall 2013, Fall 2015.

- Instructor, School of Computing, University of Utah.
  Graduate Course: CS6931 Database Seminar.
  Spring 2012.

- Instructor, School of Computing, University of Utah.
  Graduate Course: CS6960 Database Kernels and Large Data Management.
  Fall 2011.

- Instructor, Computer Science Department, Florida State University.
  Undergraduate Course: COP4710 Theory and Structure of Databases.
  Fall 2010, Spring 2010, Fall 2009, Fall 2008, Spring 2008.

- Instructor, Computer Science Department, Florida State University.
  Graduate Course: COP5725 Database Systems.
  Spring 2011, Spring 2010, Spring 2009, Fall 2007.

- Instructor, Computer Science Department, Florida State University.
  Graduate Course: CIS5930 Advanced Topics in Data Management.
  Fall 2008.

Faculty Mentor

- Quynh Nguyen, Faculty Mentor, NIH K01 Early Career Award, University of Utah, Since Fall 2014.
- Tucker Hermans, Faculty Mentor, School of Computing, University of Utah, Since Fall 2015.

Research Advising

Doctoral Advising

- Bin Yao, September 2007 to July 2011, "Spatial Approximate String Search" (Associate Professor, Computer Science Department, Shanghai Jiao Tong University).
- Jeffrey Jestes, September 2008 to July 2013, "Efficient Summarization Techniques for Massive Data" (First employment: Cerner).
- Wangchao Le, September 2008 to July 2013, "Supporting Scalable Data Analytics on Large Linked Data", (First employment: Microsoft).
- Chi Zhang, September 2010 to Decemeber 2013, "Parallel Similarity Join" (First employment: Walmart Labs).
- Mingwang Tang, "New Problems in Exploring Distributed Data", May 2010 to November 2014 (Uber).

- Robert Christensen, Since September 2013 (First employment: Visa Research).
- Min Du, "Large scale system log monitoring and analysis: a data-driven approach", January 2014 to July 2018 (First employment: UC Berkeley).
- Dong Xie, "Large scale spatial data management and analysis", September 2015 to May 2020 (First employment: Assistant Professor@CSE, Penn State University), MSR PhD Fellows (Class 2018-2019).
- Debjyoti Paul, "Data-driven spatio-temporal analysis for multimode data", September 2015 to March 2020 (First empoyment: Facebook Research).
- Zhao Chang, Since September 2014.
- Guineng Zheng, Since September 2015.
- Zhuoyue Zhao, Since September 2016, Google PhD Fellows (Class 2019).
- Yanqing Peng, Since September 2016, Google PhD Fellows (Class 2020).
- Yuwei Wang, Since January 2019.

Master's Project, Major Advisor

- Richie Frost, August 2017 to August 2019.
- Neeka Embrahimi, August 2017 to August 2019.
- Ya Gao, August 2017 to August 2019.
- Jason Vuong, August 2016 to December 2017.
- Yuwei Wang, August 2017 to August 2019.
- Kilari Murali Krishna Teja, August 2016 to August 2017 (Finicity).
- Rui Dai, August 2015 to August 2017 (Airbnb).
- Jiyuan Li, August 2015 to August 2017 (Amazon).
- Jun Tang, August 2014 to August 2016 (Google).
- Jonathan Dunstan, August 2014 to August 2016 (Maverik).
- Yupeng Zhang, August 2015 to December 2015 (Amazon).
- Mengyang Wang, September 2012 to July 2015 (Microsoft).
- Klemen Simonic, September 2013 to May 2015 (Facebook).
- Natalee Ann Villa, September 2013 to May 2015 (Adobe).
- Fei Luo, September 2012 to May 2014 (Fuhu).
- Quinton Williams, August 2012 to May 2013.
- Cody Hansen, January 2012 to December 2013 (Disney Interactive).
- Limou Wang, January 2012 to May 2013 (Turn, Yahoo).
- Chenxu Ding, January 2012 to May 2013 (Epic).
- Namita Mahtta, August 2012 to May 2013 (Goldman Sachs).
- Daniel Gomez, January 2010 to December 2010 (Raytheon).
- Kun Hou, January 2008 to December 2010 (Teradata).
- Vanessa Cedeno, September 2009 to May 2010 (First employment: professor at ESPOL, Ecuador).
- Diming Lu, January 2009 to August 2009.
- Yashas Shankar, August 2008 to August 2009.

Undergraduate Student Advising, Research Advisor

- Richie Frost, August 2016 to August 2017.
- Dyllon Gagnier, August 2016 to June 2017.
- Elijah Grubb, August 2016 to May 2017.
- Neeka Embrahimi, August 2016 to August 2017.
- Oscar Marshall, May 2012 to September 2015.
- Jason Vuong, May 2012 to August 2016.
- Yi Li, May 2012 to May 2014.
- Jun Tang, May 2012 to May 2014.
- Wendy Lynne Chidester, September 2012 to December 2012.
- Robert Christensen, June 2012 to Aug 2013.
- Cody Hansen, January 2011 to December 2011, received a BrightFutures scholarship.
- Steven Burgart, January 2010 to August 2010, received a scholarship from the DoD's Information Assurance Scholarship Program (First employment: DoD).
- Justin DeBrabant, January 2009 - May 2010, received a 2009 Undergraduate Research and Creative Endeavors Award from the Office of National Fellowships (First employment: PhD student at Brown).
- Jeffrey Jestes, January 2009 - May 2010, received a GAANN (Graduate Assistance in Areas of National Need) fellowship.

Master Student Graduated, Thesis or Project Based

- Neeka Embrahimi, "The GradTracker System", 2019.
- Ya Gao, "Optimizing the performance of PGXL", 2019.
- Rhichie Frost, "AI Pro: A Dynamic Data Pipeline Generator", 2018.
- Kilari Murali Krishna Teja, "Ternion Networks for Large Scale Paraphrase detection", 2017.
- Robert Christensen, "Sampling index for large spatio-temporal data", 2017.
- Jiyuan Li, "Academic research keywords and fields detection", 2017.
- Jonathan Dunstan, "WiFi Data Extraction Using Raspberry Pi 2", August, 2016.
- Jun Tang, "ETL For Large Spatial Temporal Data", August, 2016.
- Yupeng Zhang, "Profile Description for Tweet Users", December, 2015.
- Klemen Simonic, "Concept Aware Co-occurrence and Its Applications", May, 2015.
- Natalee Villa, "Spatio-temporal Analysis and Visualization of Social Media Behavior", May, 2015.
- Kun Hou, "Approximate String Search in Spatial Data", December, 2010.
- Daniel Gomez, "iRideBus - FSU Bus Route App", December, 2010.
- Vanessa Cedeno, "Comparing HadoopDB: A hybrid of DBMS and MapReduce Technologies with the DBMS PostgreSQL", April, 2010.
- Diming Lu, "Study of OpenMoko Power Management for Linux Mobile Operating System" (Project-based), August, 2009.
- Yashas Shankar, "Implementing authenticated indexing structures inside the database engines", (student chose to graduate with the course-only option), August, 2009.

Master's Thesis and Project, Examining Committee Member

- Aniraj Kesavan, "NIC Aware Transfer of Voluminous Data for In-Memory Databases", 2017. .
- Kaiqing Wang, "Stability of the Null Distribution for Spatial Scan Statistics", 2016.
- Liang Zhang, "Bounds on Estimated Spatial Scan Statistic", 2015.
- Raghvendra Singh, "Scalable Spatial Scan Statistics", 2015.
- Amey Desai, "Streaming Algorithms for Matrix Approximation", 2014.
- Ya Li, "An Online Real Estate Management System", 2010.
- Vineet Sahu, "Java Applet based Database Management Interface", 2010.
- Midhula Devabhaktuni, "Design and Automation of the Florida State University Medical School Database", 2010.
- Bharathi Tenneti, "Graduate Student Tracking", 2009.
- Jacob Rettig, "SAMOS/NGI Data Management System", 2009.
- Cory Hekimian-Williams, "A Study on Localization of RFID Tags Using Phase-Difference", 2009.
- Bharathi Tenneti, "Graduate Student Tracking", 2009.
- Nicholas Wallen, "Graduate Student Database Project", 2009.
- Michael Connor, "A Simple, Thread-Safe, Approximate Nearest Neighbor Algorithm", 2008.
- Evan Hollander, "e-Voting Software Review Methodologies", 2008.

PhD's Thesis Defense Committee Member

- Michael Matheny, October, 2019, "Approximate Statistical Discrepancy".
- Haibo Ding, August, 2018, "Affective Polarity Recognition and Human Needs Categorization for Affective Events".
- Hyun-wook Baek, May, 2018, "Inter-party Visibility in a Cloud Computing Platform".
- Yan Zheng, April 2017, " Algorithms and Coresets for Large Scale Kernel Smoothing".
- Wathsala Widanagamaachchi, May 2017, "Interactive Visualization and Exploration of Feature Evolution in Dynamic Data",
- Ashequl Qadir, May 2016, "Acquiring Knowledge for Affective Sate Recognition in Social Media".
- Sidharth Kumar, November, 2015. "A Scalable and Tunable Adaptive Resolution Parallel I/O Framework".
- Amirali Abdullah, August, 2015. "Bounds for Nearest Neighbor Algorithms and Embeddings".
- Xing Lin, July, 2015. "Using Similarity in Content and Access Patterns to Improve Space Efficiency and Performance in Storage Systems".
- Qian Chen, July, 2015, "Query Authentication in Data Outsourcing and Integration Services".
- Zechao Shang, July, 2015, "Efficient Graph Computing".
- Xiaocheng Hu, July, 2015, "On the I/O Complexities of Several Fundamental Problems with Applicaitons in Databases".
- Parasaran Raman, October, 2013, "Exploring the Landscape of Clusterings ".
- Mingwang Tang, *Chair*, October, 2013. "New Problems in Exploring Distributed Data".
- Chi Zhang, *Chair*, October, 2013. "Parallel Similarity Join".
- Parasaran Raman, October, 2013. "Exploring the Landscape of Clusterings".
- Wangchao Le, *Chair*, "Supporting Scalable Data Analytics on Large Linked Data", July, 2013.
- Jeffrey Jestes, *Chair*, "Efficient summarization Techniques for Massive Data", July, 2013.
- Bin Yao, *Chair*, "Spatial Approximate String Search", June, 2011.
- Michael Connor, "Algorithms for Solving Near Point Problems", April, 2011.
- Yiming Wu, "Efficient Learning of Optimal Linear Representation for Object Classification", March, 2010.
- Yuhua Zhu, "Optimal Linear Features for Content Based Image Retrieval and Application", April, 2010.

PhD's Thesis, Qualifying, Area Examining Committee Member

- Meysam Taassori, Qualifying Exam, 2017.
- Baek Hyun-wook, Qualifying Exam, 2016.
- Haibo Ding, Qualifying Exam, 2016.
- Wathsala Widanagamaachchi, Proposal Defense, "Interactive Visualization and Exploration of Tracking Graphs for Understanding Evolution of Features", 2015.
- Yan Zheng, Proposal Defense, "Large Scale Kernel Smoothing", 2015.
- Ashequl Qadir, Proposal Defense, "Acquiring Knowledge for Affective State Recognition in Tweets", 2015.
- Lalindra de Silva, Proposal Defense, "Extracting Precursor Activities for Societal Events", 2015.
- Wathsala Widanagamaachchi, Qualifying Exam, 2014.
- Ashequl Qadir, Qualifying Exam, 2014.
- Lalindra de Silva, Qualifying Exam, 2014.
- Sidharth Kumar, Qualifying Exam, 2013.
- Seth Pugsley, Proposal Defense, "Architectures to Exploit Data Locality In Big Data Workloads", 2013.
- Seth Pugsley, Qualifying Exam, 2013.
- Xing Lin, Qualifying exam, 2013.
- Parasaran Raman, Proposal Defense, "Exploring the Landscape of Clusterings: Towards Better Integration of Data, Clusterings and the User", 2011.
- Chi Zhang, Prospectus Defense, "Parallel Similarity Join", advisor, 2011.
- Jeffrey Jestes, Area Survey Defense, "Summarizing Massive Datasets", advisor, 2011.
- Chi Zhang, Area Survey Defense, "Survey on Parallel Similarity Join", advisor, 2011.
- Wangchao Le, Area Survey Defense, "Query Processing for Large Scale Semantic Web and Graph Data", advisor, 2011.
- Bin Yao, Prospectus Defense, "Spatial Approximate String Search", advisor, 2011.
- Jiangbo Yuan, Area Survey Defense, "Survey on Approximate Nearest Neighbor Search Algorithms in Dimensional Space", 2011.
- Michael Connor, Prospectus Defense, "Proposal to Study Near Point Problems", 2011.
- Bin Yao, Area Survey Defense, "Spatial Keyword Search", advisor, 2010.
- Michael Connor, Area Survey Defense, "A survey of nearest neighbor problems", 2010.
- Yuhua Zhu, Area Survey Defense, "Ideas, Algorithms, Systems in Content Based Image Retrieval", 2009.
- Yiming Wu, Area Survey Defense, "Linear Representations for Object Recognition", 2009.
- Yuhua Zhu, Prospectus Defense, "Optimal Linear Features for Content Based Image Retrieval and Application", 2009.
- Yiming Wu, Prospectus Defense, "Efficient and Discriminative Component Analysis", 2009.

External Presentations

1. 08/2023, Modernization of Databases in the Cloud Era: Building Databases that Run Like Legos, Invited Keynote Talk, VLDB 2023.
2. 08/2021, Benchmarking Cloud-Native Database Systems in the Cloud Era, Invited Keynote Talk, TPCTC 2021 associated with VLDB 2021.
3. 10/2020, Cloud native database systems for enterprise applications, Conference Keynote Talk, ACM SoCC 2020.
4. 09/2019, Cloud native database systems: challenges and opportunities, Industry Keynote Talk, VLDB 2019.
5. 08/2018, Spatio-temporal data analytics, FFD, invited talk.

6. 07/2018, Data Science Research and Education, Beijing Institute of Big Data Research at PKU, invited talk.

7. 03/2018, Sampling over Join, UIUC, invited talk.

8. 12/2017, Interactive Data Analytics, HKU, invited talk.

9. 11/2017, Interactive Data Analytics, UIUC, invited talk.

10. 06/2017, Online aggregation and analytics, Beijing Institute of Big Data Research at PKU, invited talk.

11. 01/2017, NSF DIBBs Workshop, Panelist.

12. 12/2016, In-memory spatial databases and analytics. International Conference on Data Science.

13. 12/2016, Simba: In-memory Spatial Big Data Analytics, HKUST.

14. 08/2016, Interactive Data Management and Analytics, Ali Cloud, Alibaba.

15. 07/2016, Towards Interactive, Online, and Secure Big (Spatial) Data Analytics, VLDB Summer School.

16. 06/2016, Online Aggregation over Joins, SIGMOD.

17. 10/2015, Interactive Online Data Exploration and Analytics, Symposium on Visualization in Data Science, VDS at IEEE VIS.

18. 04/2015, Interactive Online Data Analytics, Database Technology Conference China.

19. 12/2014, STORM: Towards Interactive Data Analytics, Harbin Institute of Technology, Harbin, China.

20. 11/2014, Interactive Data Analytics on Big Data, Hong Kong University of Science and Technology, Hong Kong.

21. 11/2014, Interactive Data Analytics on Big Data, AnHui University, China.

22. 11/2014, Interactive Data Analytics on Big Data, Beijing University of Posts and Telecommunications, China.

23. 10/2014, Interactive Data Analytics on Big Data, Nanyang Technological University, Singapore.

24. 10/2014, Interactive Data Analytics on Big Data, XiDian University, China.

25. 10/2014, Interactive Data Analytics on Big Data, Suchow University, China.

26. 10/2014, Interactive Data Analytics on Big Data, Wuhan University, China.

27. 09/2014, Secure Data Analytics in the Cloud, Invited Seminar, National Database Conference of China (NDBC) 2014.

28. 09/2014, Interactive Data Analytics on Big Data, Distinguished Lecture Series, ZhiYuan College, Shanghai Jiao Tong University, China.

29. 09/2014, Interactive Data Analytics on Big Data, Invited Seminar, ApWeb 2014.

30. 08/2014, Secure Data Analytics in the Cloud, Jiangsu University, China.

31. 08/2014, Interactive Data Analytics on Big Data, Invited Talk, Forum of Future Data (FFD) Workshop 2014.

32. 12/2013, Summarizing and interacting with large data, Big Data Workshop, East China Normal University, China.

33. 08/2013, Summarizing large data and towards building local search engines, Database summer school, Fudan University, China.

34. 06/2013, Towards Building an Automatic Query Engine for Large Heterogeneous Data, Microsoft Research Asia Database Summer School, in Conjunction with WAIM'13.

35. 06/2013, Query Processing Techniques for Probabilistic Data, Hong Kong University of Science and Technology, Hong Kong.

36. 09/2012, Flexible and Scalable Similarity Matching, Database Colloquium, Cornell University.

37. 07/2012, Parallel kNN Join for Large Data in MapReduce, HP Labs China, Beijing, China.

38. 01/2012, Efficient Threshold Monitoring for Distributed Probabilistic Data, International Shonan meeting/seminar on "Large-scale Distributed Computation" organized by the National Institute of Informatics of Japan, the Shonan Village, Japan. *Invited, but did not make it due to the schedule conflict and visa issue.*

39. 01/2012, Building Wavelet Histograms for Large Data in MapReduce, Shanghai Jiao Tong University, Shanghai, China.

40. 12/2011, Building Wavelet Histograms for Large Data in MapReduce, HP Labs China, Beijing, China.

41. 12/2011, Building Wavelet Histograms for Large Data in MapReduce, Tsinghua University, Beijing, China.

42. 12/2011, Building Wavelet Histograms for Large Data in MapReduce, Renming University, Beijing, China.

43. 11/2011, Building Wavelet Histograms for Large Data in MapReduce, Indiana University-Purdue University Indianapolis, Indianapolis.

44. 11/2011, Multi-Approximate-Keyword-Routing in GIS Data, ACM SIGSPATIAL GIS 2011, Chicago, IL.

45. 06/2011, Flexible Aggregate Similarity Search, SIGMOD 2011, Athens, Greece.

46. 05/2011, International Collaborations on Database Research, Invited Group Discussion, NSF workshop on US-China research collaborations, Arlington, VA.

47. 10/2010, Security Issues in Databases, Invited Group Discussion, NSF Workshop on Trustworthy Computing, Arlington, VA.

48. 07/2010, Query Authentication for Outsourced Databases, HP Lab China, Beijing, China.

49. 07/2010, Query Authentication for Outsourced Databases, Tsinghua University, Beijing, China.

50. 07/2010, Modeling, Using, and Querying Probabilistic Strings, China's PhD Summer School of Computer Science, Beijing, China.

51. 03/2010, K Nearest Neighbor Queries and kNN-Joins in Large Relational Databases (Almost) for Free, ICDE 2010, Long Beach, CA.

52. 03/2010, Approximate String Search in Spatial Databases, ICDE 2010, Long Beach, CA.

53. 03/2010, Probabilistic String Similarity Joins, Yahoo! Research, Santa Clara, CA.

54. 12/2009, Query and Data Security in the Database-as-a-Service Model, Computer Science Department, Shanghai Jiao Tong University, Shanghai, China.

55. 08/2008, Approximate Aggregation in Sensor Networks, China 973 Wireless Sensor Network Lab, Shanghai Jiao Tong University, Shanghai, China.

56. 06/2008, Finding Frequent Item in Probabilistic Data, School of Computing Science, Simon Fraser University, Vancouver, Canada.

57. 04/2008, PIRS: Query Verification on Data Streams, ICDE 2008, Cancun, Mexico.

58. 04/2008, Efficient Query Processing of Top-k Queries on Uncertain Data (poster), ICDE 2008, Cancun, Mexico.

59. 03/2008, Data and Query Security for the Data Publishing Model, AT&T Labs - Research, Florham Park, NJ.

60. 09/2007, Authenticating Sliding Window Queries over Data Streams, VLDB 2007, Vienna, Austria.

61. 07/2007, Optimized Transaction Time Versioning Inside a Database Engine, Microsoft Research, Redmond, WA.

62. 04/2007, Privacy Preservation on Evolving Streams through Correlation Tracking, ICDE 2007, Istanbul, Turkey.

63. 03/2007, Data and Query Security for the Data Publishing Model, IBM T. J. Watson Research Center, Hawthorne, NY.

64. 08/2006, Privacy Preservation on Evolving Streams through Correlation Tracking, IBM T. J. Watson Research Center, Hawthorne, NY.

65. 06/2006, Dynamic Authenticated Index Structures for Outsourced Databases , SIGMOD 2006, Chicago, IL.

66. 04/2006, Characterizing and Exploiting Reference Locality in Data Stream Applications, ICDE 2006, Atlanta, GA.

67. 04/2004, Spatio-Temporal Aggregation Using Sketches, ICDE 2004, Boston, MA.

Patents Issued

- Co-Inventor, Scalable Multi-Query Optimization for SPARQL, US patent number 9,542,444.
- Co-Inventor, Scalable Multi-Query Optimization for SPARQL, US patent number 9,280,583.
- Co-Inventor, Scalable summarization of data graphs, US Patent number 8,984,019.
- Co-Inventor, Scalable summarization of data graphs, US Patent number 8,977,650.
- Co-Inventor, Methods and Apparatus for Ranking Uncertain Data in a Probabilistic Database, US Patent number 8,825,640.
- Co-Inventor, Verification of Outsourced Data Streams, US patent number 8.112.802.
- Co-Inventor, Preserving Privacy of One-Dimensional Data Streams Using Dynamic Correlations, US patent number 7.853.545.
- Co-Inventor, Preserving Privacy of One-Dimensional Data Streams Using by Perturbing Data With Noise and Using Dynamic Autocorrelation, US patent number 7.840.516.

Patents Under Application

- Co-Inventor, Wander Join Methods for Online Aggregation over Joins, US patent application filed by the University of Utah.

Software Issued

- Simba: Spatial In-Memory Big Data Analytics:
  `https://github.com/InitialDLab/Simba`
- XDB: approXimate DataBase for Interactive and Online Analytics:
  `https://github.com/InitialDLab/XDB`
- Compass: Spatiotemporal Sentiment Analysis for US Election:
  `htt://www.estorm.org`
- Sampling index for spatio-temporal data:
  `https://github.com/InitialDLab/SONAR-SamplingIndex`.
- HashTagHealth: A Social Media Big Data Resource for Neighborhood Effects Research:
  `https://hashtaghealth.github.io/`
- Seal-ORAM: An Oblivious RAM library
  `https://github.com/InitialDLab/SEAL-ORAM`
- ATOM for automated tracking and monitoring in CloudWatch:
  `https://github.com/ricosfeifei/CloudWatch-ATOM`
- Library for kernal density estimate:
  `http://www.cs.utah.edu/~lifeifei/kde/`
- Library for partitioning large multi-versioned databases:
  `http://www.cs.utah.edu/~lifeifei/split/`
- Library for building wavelet histograms for big data in MapReduce:
  `http://www.cs.utah.edu/~lifeifei/histogram/`.
- Authenticated index structure library:
  `http://www.cs.utah.edu/~lifeifei/aisl/`.
- Library for ranking probabilistic data:
  `http://www.cs.utah.edu/~lifeifei/rankaggprob/`.
- Library for ranking distributed probabilistic data:
  `http://www.cs.utah.edu/~lifeifei/pdtopk/`.
- Library for ranking large temporal data:
  `http://www.cs.utah.edu/~lifeifei/rankt/`.
- Library for kNN-joins for large data in MapReduce:
  `http://www.cs.utah.edu/~lifeifei/knnj/`.

- Library for probabilistic string similarity joins:
  `http://www.cs.utah.edu/~lifeifei/probstring/`.
- Library for finding frequent items in probabilistic data:
  `http://www.cs.utah.edu/~lifeifei/phitter/`.
- Library for multi-keyword routing in large spatial road networks:
  `http://www.cs.utah.edu/~lifeifei/makr/`.
- Library for approximate string search in large spatial data:
  `http://www.cs.utah.edu/~lifeifei/sas/`.
- Library for optimal location queries:
  `http://www.cs.utah.edu/~lifeifei/olq/`.
- Library for randomized synopses for query assurance on data streams:
  `http://www.cs.utah.edu/~lifeifei/pirs/`.

Major Media Exposure

- SIGMOD Research Highlight Award Technical Perspective summary,
  `https://sigmodrecord.org/publications/sigmodRecord/1703/pdfs/09_Wander_TP_Naughton.pdf`
- News release for sentiment analysis over geo-tagged tweets for analyzing and predicting 2016 election,
  `http://www.sltrib.com/news/4536497-155/university-of-utah-software-gauges-popular`
- News release for using social media data to build healthy/unhealthy geo-indicators by Time.com,
  `http://time.com/4533176/coffee-pizza-tweet-twitter-social-media/`
- News release for Google Faculty Award, 2015,
  `http://googleresearch.blogspot.com/2015/02/google-faculty-research-awards-winter.html`
- News release for the Data-Driven Science by NSF, 2014,
  `http://www.nsf.gov/news/news_summ.jsp?cntn_id=132880&org=NSF&from=news`
- News release for the HP Labs Innovation Research Award Program, 2012,
  `http://www.hpl.hp.com/open_innovation/irp/2012_results.html`
- News release for the HP Labs Innovation Research Award Program, 2011,
  `http://www.hp.com/hpinfo/newsroom/press/2011/110907xb.html`

Departmental and University Service

- Project Lead, The GradTrack System, School of Computing and College of Engineering, since Summer 2016.
- Member of the College of Engineering TFR Committee (Tenured Faculty Review Policy Cmmittee), since Summer 2017.
- Chair of the Data Management Faculty search committee, Spring 2017.
- Director of Graduate Study, since November 2015.
- Chair of the academic misconduct and appeals committee, College of Engineering, since Spring 2015.
- Member of the 7-Year GradReview Preparation Committee, School of Computing, 2016.
- Member of RPT Review Committee, (Chair of a subcommittee), 2017.
- Member of RPT Review Committee, (Chair of a subcommittee), 2016.
- Member of the graduate admission committee, Spring 2017.
- Member of the graduate admission committee, Spring 2016.
- Member of the graduate admission committee, Spring 2015.
- Member of the faculty search committee, Spring 2016.
- Chair of the big data faculty search committe, Spring 2014.
- Member of the graduate admission committee, Spring 2014.
- Member of Informal RPT Review Committee, Spring 2014.
- Member of the College of Engineering Scholarship Review/Selection Committee, 03/2014.
- Co-chair of the SoC graduate poster competition, 02/2014.
- Member of RPT Review Committee, (Chair of a subcommittee), 2013.
- Member of the College of Engineering Scholarship Review/Selection Committee, 02/2013.
- Member of the Due Progress Committee, School of Computing, University of Utah, 11/2012.
- Panel member in "Grad School Application Pointers", 10/2012.

- Member of a team to improve the access experience to the Unite database, University of Utah.
- Chair of the data track, School of Computing, University of Utah, September 2011 to December 2013.
- Member of the graduate admission committee, School of Computing, University of Utah, 2011-2012.
- Judge for the undergraduate senior projects, School of Computing, University of Utah, April 2012.
- Judge for the School of Computing poster competition, University of Utah, February 2012.
- Member of the graduate studies committee (responsible for the due progress evaluation), School of Computing, University of Utah, Fall, 2011.
- Member, the hosting team for the Fulbright Scholars from Iraq to the Florida State University, Florida State University, May-July 2011.
- Member of the graduate admission committee, Department of Computer Science, Florida State University, 2008, 2009, 2010, 2011.
- Member of the faculty search committee, Department of Computer Science, Florida State University, 2008, 2011.
- Member of the PhD portfolio committee, Department of Computer Science, Florida State University, 2008, 2009, 2010, 2011.
- Chair of the database and data management PhD qualification exam committee, Florida State University, 2008, 2009, 2010, 2011.
- Member of the faculty executive committee, Department of Computer Science, Florida State University, 2010-2011.
- Member of the annual faculty evaluation committee, Department of Computer Science, Florida State University, 2009-2010.
- Director for the SAIT (Security and Assurance in IT) lab, Department of Computer Science, Florida State University, 2008-2011.
- Founding member, Center of Academic Excellence in Information Assurance Education (CAE/IAE) and Center of Academic Excellence in Research (CAE-R), designations by DoD and NSA, Florida State University, 2009-2011.
- Founding member, the Scholarship for Service program at the Florida State University, Florida State University, 2009-2011.
- Judge and faculty representatives for the annual ACM programming contest, Department of Computer Science, Florida State University, 2008, 2009, 2010.