



Department of Computer Science

Tsinghua University, Beijing 100084, China

Guoliang Li

Associate Professor

Department of Computer Science, Tsinghua University, Beijing, China

<http://dbgroun.cs.tsinghua.edu.cn/ligl>

Address: Room 10-204, East Main Building, Department of Computer Science, Tsinghua University, Beijing 100084, China

Tel: +86-10-6278-9150 **Fax:** +86-10-6277-1138

Email: liguoliang@tsinghua.edu.cn



Education:

- 2004.9 — 2009.7 **Doctor of Science, Computer Science**
Department of Computer Science, **Tsinghua University**, Beijing, China
- 2000.9 — 2004.7 **Bachelor Degree, Computer Science**
Department of Computer Science, **Harbin Institute of Technology**, Harbin, China

Experience:

- 2012.12—Now **Associate Professor**, Dept. of Computer Science, Tsinghua University
- 2009.7—2012.11 **Assistant Professor**, Dept. of Computer Science, Tsinghua University
- 2008.1--2008.6 Visitor, University of California, Irvine (Prof. Chen Li)
- 2006.4--2006.11 Visitor, National University of Singapore (Prof. Beng Chin Ooi)

Research Interests:

- Database Usability, Data Cleaning and Data Integration
- Hybrid Human-Machine Data Management

Awards:

- 2013 New Century Excellent Talents in University (Ministry of Education, China)
- 2012 Young Researcher Award (Science News & Elsevier China)
- 2010 Best PhD Thesis Award in Beijing

Professional Services:

- PC Member:
 - SIGMOD 2014
 - VLDB: 2012, 2013, 2014
 - ICDE: 2012, 2013, 2014
 - KDD: 2012, 2013, 2014
 - IJICA:2013
 - CIKM: 2010, 2012,2014
- PC Co-Chair: WAIM 2014, WebDB2014
- SIGMOD2014 Publicity Co-Chair
- CIKM2014 Publication Co-Chair
- WAIM2013 Distinguished Young Lecturer



Publications:

➤ Journal Papers:

1. Dongzhe Ma, Jianhua Feng, **Guoliang Li**. A Survey of Address Translation Technologies for Flash Memories. **ACM Computing Surveys**, 2014. (To appear)
2. Weiwei Hu, **Guoliang Li**, Jiakai Ni, Dalie Sun, Kian-lee Tan. Bp-tree: A Predictive B+-tree for Reducing Writes on Phase Change Memory. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2014. (To appear)
3. Jiannan Wang, **Guoliang Li**, Jianhua Feng. Extending String Similarity Join to Tolerant Fuzzy Token Matching. **ACM Transactions on Database Systems (TODS)**, 2014. (To appear)
4. Sitong Liu, **Guoliang Li**, Jianhua Feng. A Prefix-Filter based Method for Spatio-Textual Similarity Join. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2014 (To appear.)
5. Jiakai Ni, **Guoliang Li**, Lijun Wang, Jianhua Feng. Adaptive Database Schema Design for Multi-Tenant Data Management. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 2014 (To appear)
6. **Guoliang Li**, Dong Deng, Jianhua Feng. Pass-Join+: A Partition-based Method for String Similarity Joins with Edit-Distance Constraints. **ACM Transactions on Database Systems (ACM TODS)**, 38(2):9-40,2013.
7. **Guoliang Li**, Jianhua Feng, Chen Li. Supporting Search-As-You-Type Using SQL in Databases. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 25(2): 461-475 (2013).
8. Yang Cao, Ju Fan, **Guoliang Li**: A User-Friendly Patent Search Paradigm **IEEE Transactions on Knowledge and Data Engineering. (TKDE)** 25(6):1439-1443 (2013)
9. Jianhua Feng, Jiannan Wang, **Guoliang Li**. Trie-join: a trie-based method for efficient string similarity joins. **The VLDB Journal**, 21(4):437-461, 2012
10. Jianhua Feng, **Guoliang Li**: Fuzzy Type-Ahead Search in XML Data. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**, 24(5):882-895, 2012.
11. **Guoliang Li**, Jianhua Feng, Jianyong Wang, Lizhu Zhou: KEMB: A Keyword-Based XML Message Broker. **IEEE Transactions on Knowledge and Data Engineering (TKDE)** 23(7):1035-1049 (2011)
12. **Guoliang Li**, Shengyue Ji, Chen Li, Jianhua Feng. Efficient Fuzzy Full-Text Type-Ahead Search. **The VLDB Journal**, 20(4): 617-640, 2011.
13. **Guoliang Li**, Jianhua Feng, Xiaofang Zhou, Jianyong Wang. Providing Built-in Keyword Search Capabilities in RDBMS. **The VLDB Journal**, 20(1):1-19, 2011.
14. Jianhua Feng, **Guoliang Li**, Jianyong Wang: Finding Top-k Answers in Keyword Search over Relational Databases Using Tuple Units. **IEEE Transactions on Knowledge and Data Engineering (TKDE)**23(12):1781-1794 (2011)
15. **Guoliang Li**, Jianhua Feng, Beng Chin Ooi, Jianyong Wang, and Lizhu Zhou. An effective 3-in-1 keyword search method over heterogenous data sources. **Information Systems**, Volume 36, Issue 2, Pages 248-266, 2011.
16. Jiannan Wang, Inci Cetindil, Shengyue Ji, Chen Li, Xiaohui Xie, **Guoliang Li**, and Jianhua Feng. Interactive and fuzzy search: a dynamic way to explore MEDLINE. **Bioinformatics (2010) 26(18)**: 2313-2320.



17. Jianhua Feng, **Guoliang Li**, Lizhu Zhou: Finding and Ranking Compact Connected Trees for Effective Keyword Proximity Search over XML Documents. **Information Systems**, Volume 35, Issue 2, Pages 186-203, 2010.
18. **Guoliang Li**, Jianhua Feng, Jianyong Wang, and Lizhu Zhou. Incremental sequence-based frequent query pattern mining from xml queries. **Data Mining and Knowledge Discovery**, Volume 18, Number 3, pages 472-516, 2009.
19. **Guoliang Li**, Jianhua Feng, Chen Li, Lizhu Zhou: SAIL: Structure-aware Indexing for Effective Keyword Search over XML Documents. **Information Sciences**, Volume 179, Issue 21, Pages 3745-3762, 2009.

➤ **Conference Papers (Full Research Papers)**

1. **Guoliang Li**, Jun Hu, Kian-lee Tan, Zhifeng Bao, Jianhua Feng. Effective Location Identification from Microblogs. **ICDE 2014**. (To appear)
2. Dong Deng, **Guoliang Li**, Shuang Hao, Jiannan Wang, Jianhua Feng, Wen-syan Li. MassJoin: A MapReduce-based Algorithm for String Similarity Joins. **ICDE 2014**. (To appear)
3. Yong Zeng, Zhifeng Bao, H V Jagadish, Tok Wang Ling, **Guoliang Li**. Breaking out of the MisMatch Trap. **ICDE 2014**. (To appear)
4. Jiannan Wang, **Guoliang Li**, Tim Kraska, Michael J. Franklin, Jianhua Feng. Leveraging Transitive Relations for Crowdsourced Joins. **SIGMOD 2013**:229-240..
5. **Guoliang Li**, Ting Wang, Jianhua Feng. Location-Aware Publish/Subscribe. **SIGKDD 2013**:802-810
6. Dong Deng, Yu Jiang, **Guoliang Li**, Jian Li, Cong Yu. Scalable Column Concept Determination for Web Tables Using Large Knowledge Bases. **VLDB 2013**: 1606-1617.
7. Dong Deng, **Guoliang Li**, Jianhua Feng: Top-k String Similarity Search with Edit-Distance Constraints. **ICDE 2013**:925-936.
8. Ruicheng Zhong, **Guoliang Li**, Kian-Lee Tan, Lizhu Zhou: G-tree: an efficient index for KNN search on road networks. **CIKM 2013**: 39-48
9. Jiannan Wang, **Guoliang Li**, Jianhua Feng. Can We Beat The Prefix Filtering? An Adaptive Framework for Similarity Join and Search. **SIGMOD 2012**:85-96
10. Ju Fang, **Guoliang Li**, Lizhu Zhou. SEAL: Spatio-Textual Similarity Search. **VLDB 2012**:253-264
11. **Guoliang Li**, Dong Deng, Jiannan Wang, Jianhua Feng: PASS-JOIN: A Partition-based Method for Similarity Joins. **VLDB 5(3)**:253-264 (2012)
12. **Guoliang Li**, Jiannan Wang, Chen Li, Jianhua Feng: Supporting Efficient Top-k Queries in Fuzzy Type-Ahead Search. **SIGIR**, 2012: 355-364.
13. **Guoliang Li**, Jing Xu, Jianhua Feng: Desks: Direction-Aware Spatial Keyword Search. **ICDE**, 2012:474-485
14. **Dong Deng**, Guoliang Li, Jianhua Feng: An Efficient Trie-based Method for Approximate Entity Extraction with Edit-Distance Constraints. **ICDE**, 2012:762-773
15. Weihuang Huang, **Guoliang Li**, Kian-lee Tan, Jianhua Feng: Efficient Safe-Region Construction for Moving Top-K Spatial Keyword Queries. **CIKM 2012**:932-941
16. Ruicheng Zhong, Ju Fan, **Guoliang Li**, Kian-lee Tan, Lizhu Zhou: Location-Aware Instant Search. **CIKM**, 2012:385-394



17. Jiannan Wang, **Guoliang Li**, Jeffrey Xu Yu, Jianhua Feng: Entity Matching: How Similar Is Similar. **VLDB** 4(10):622-633 (2011)
18. **Guoliang Li**, Dong Deng, Jianhua Feng: Faerie: Efficient Filtering Algorithms for Approximate Dictionary-based Entity Extraction. **SIGMOD**, 2011:529-540.
19. Dongzhe Ma, Jianhua Feng, **Guoliang Li**: LazyFTL: A Page-level Flash Translation Layer Optimized for NAND Flash Memory. **SIGMOD**, 2011:1-12.
20. **Guoliang Li**, Ju Fan, Hao Wu, Jiannan Wang, Jianhua Feng. DBease: Making Databases User-Friendly and Easily Accessible. **CIDR**, 2011: 45-56
21. Ju Fan, **Guoliang Li**, Lizhu Zhou. Interactive SQL Query Suggestion: Making Databases User-Friendly. **ICDE**, 2011:351-362
22. Jiannan Wang, **Guoliang Li**, Jianhua Feng. Extending String Similarity Join to Allow Fuzzy Token Matching. **ICDE**, 2011:458-469
23. Jiannan Wang, **Guoliang Li**, Jianhua Feng, Trie-Join: Efficient Trie-based String Similarity Joins with Edit-Distance Constraint. **VLDB**, 2010:1219-1230
24. **Guoliang Li**, Shengyue Ji, Chen Li, Jianhua Feng. Efficient Type-Ahead Search on Relational Data: a TASTIER Approach. **SIGMOD 2009**:695-706.
25. Shengyue Ji, **Guoliang Li**, Chen Li, Jianhua Feng. Effective Fuzzy Interactive Search. **WWW 2009**: 371-380.
26. Mei Hui, **Guoliang Li**, Yuan Zhou, Dawei Jiang. Supporting Database Applications as a Service. **ICDE**, 2009:832-843.
27. **Guoliang Li**, Beng Chin Ooi, Jianhua Feng, Jianyong Wang, Lizhu Zhou: EASE: an effective 3-in-1 keyword search method for unstructured, semi-structured and structured data. **SIGMOD 2008**: 903-914.
28. **Guoliang Li**, Jianhua Feng, Jianyong Wang, Lizhu Zhou: Effective keyword search for valuable lcas over xml documents. **ACM CIKM 2007**: 31-40
29. Beng Chin Ooi, Bei Yu, **Guoliang Li**: One table stores all: Enabling painless free-and-easy data publishing and sharing. **CIDR 2007**: 142-153.
30. Bei Yu, **Guoliang Li**, Karen R. Sollins, Anthony K. H. Tung: Effective keyword-based selection of relational databases. **SIGMOD 2007**: 139-150.
31. **Guoliang Li**, Jianhua Feng, Jianyong Wang, Yong Zhang, Lizhu Zhou: Incremental Mining of Frequent Query Patterns from XML Queries for Caching. **ICDM 2006**: 350-361.

➤ **Conference Demo and Short Papers**

32. **Guoliang Li**, Nan Zhang, Ruicheng Zhong, Kian-lee Tan, Lizhu Zhou, Jianhua Feng. TsingNUS: A Location-Based Service System Towards Live City. **SIGMOD** 2013. (Demo)
33. Yong Zeng, Zhifeng Bao, Tok Wang Ling, **Guoliang Li**: Removing the mismatch headache in XML keyword search. **SIGIR** 2013: 1109-1110 (Demo)
34. Hao Wu, **Guoliang Li**, Chen Li, Lizhu Zhou: Seaform: Search-As-You-Type in Forms. **VLDB**, 2010:1565-1568 (Demo)
35. **Guoliang Li**, Shengyue Ji, Jiannan Wang, Chen Li, Jianhua Feng. Efficient Fuzzy Type-Ahead Search in TASTIER. **ICDE 2010**:1105-1108 (Demo)
36. **Guoliang Li**, Jianhua Feng, Jianyong Wang, Lizhu Zhou. An Effective and Versatile



Keyword Search Engine on Heterogeneous Data Sources. **VLDB**, 2008:1452-1455.

37. **Guoliang Li**, Xiaofang Zhou, Jianhua Feng, Jianyong Wang. Progressive Top-k Keyword Search in Relational Database. **ICDE**, 2009:1183-1186 (Short).

Projects

1. Data Integration on Location-based Social Network. National Science Foundation of China (**NSFC**). 2014-2017. 800K RMB. **PI**
2. Location-Aware Data Extraction and Cleaning. National Science Foundation of China (**NSFC**). 2013-2016. 840K RMB. **Co-PI**
3. New User-Friendly Search Paradigm in Databases. National Science Foundation of China (**NSFC**). 2011-2013. 210K RMB. **PI**
4. In-memory Computing. **SAP**. 2013-2014. 300K USD. **PI**
5. Direction-Aware Search. **Samsung**. 2013-2014. 400K RMB. **PI**
6. Location-Aware Search. **Tencent**. 2012-2014. 500K RMB. **Co-PO**
7. Location-Aware Integration. **Microsoft**. 2013-2014. 10K USD. **PI**
8. Massive Data Cleaning and Integration. **Huawei**. 2012-2013. 120K RMB. **PI**.
9. Data Extraction in Hidden Databases. **Sogou**. 2012-2013. 300K RMB. **Co-PI**.
10. Outsourcing databases. **Siemens**. 2013-2014. 20K USD. **Co-PI**
11. Trading Consistency for Performance and Availability in Large-Scale Distributed Databases. **IBM SUR Project**. 10K USD, 2012. **PI**
12. Massive Data Management. Research Foundation of Tsinghua. 2011-2013, 390K RMB. **PI**
13. Next-Generation Search Engine and Browser. Key Projects of Core-High-Tech-Basic Program. 2011-2012. 1M RMB. **Co-PI**.
14. Indexing, Query Processing and Optimizations on Flash Databases. National Science Foundation of China (**NSFC**). 2010-2012. 300K RMB. **Co-PI**
15. Keyword Search in Large Scale Data. 863 Project of China. 2009-2011. 1M RMB. **Co-PI**

Affiliated Students

1. Ju Fan. Ph.D Student. Now Postdoc at National University of Singapore
2. Jiannan Wang. Ph.D Student. Now Postdoc at University of California, Berkeley
3. Hao Wu. Ph.D Student. Now Researcher at SAP, Beijing
4. Nan Zhang. Master Student. Now Researcher at Microsoft Seattle
5. Ting Wang. Master Student. Now Researcher at Google Palo Alto
6. Jing Xu. Master Student. Now Researcher at Google Shanghai