# Chen Li

Curriculum Vitae, As of January, 2017
Professor
Department of Computer Science, University of California, Irvine, CA 92697-3435
chenli @ ics.uci.edu, http://www.ics.uci.edu/~chenli

| | | | |
|---|---|---|---|
| **Education** | 1996 – 2001 | Stanford University | CA |

**Ph.D. Computer Science**
Thesis: *Query Processing and Optimization in Information-Integration Systems.*
Advisor: Professor Jeffrey D. Ullman

| | | |
|---|---|---|
| 1994 – 1996 | Tsinghua University | Beijing, China |

**M.S. Computer Science**
Dissertation: *A Communication Interface based on SCI.*
Advisor: Professor Meiming Shen.

| | | |
|---|---|---|
| 1989 – 1994 | Tsinghua University | Beijing, China |

**B.S., Computer Science. Minor: Enterprise Management**

**Research Interests**

Databases, information systems, search, data-intensive computing, data integration, data warehousing, data quality

**Experience**

| | | |
|---|---|---|
| 11/2012–current | UC Irvine, Department of Computer Science | UCI, Departmen |

**Professor**

| | |
|---|---|
| 7/2007 – 11/2012 | UC Irvine, Department of Computer Science |

**Associate Professor**

| | |
|---|---|
| 2001 – 7/2007 | UC Irvine, Department of Computer Science |

**Assistant Professor**

8/2008 – 12/2014
**Founder/CTO**
Started a company specialized in information access for enterprises.
Took a leave from UCI to work at SRCH2 from 7/2013 to 12/2014.

| | | |
|---|---|---|
| 9/2005–12/2007 | Google | Santa Monica/Irvine, CA |

**Visiting Research Scientist**
Full time before 12/31/2005, part time after 1/1/2006.

| | | |
|---|---|---|
| 1996 – 2001 | Stanford University, Computer Science | CA |

**Graduate Student Researcher**
Focused on database research, including data integration, web databases, data warehouses, and multimedia databases.

| | | |
|---|---|---|
| 2000 | IBM Almaden Research Center | CA |

**Supplemental Research Associate**
Researched on a web-based database system

2000           Common Object Inc.           San Mateo, CA
**Consultant**
Participated in the design and implementation of a corporate database system using data-integration technologies.

1998           Hewlett Packard Labs           Palo Alto, CA
**Research Intern**
Ported a well-known main-memory database system to a shared-memory distributed system.

1995           Chinese University of Hong Kong           Hong Kong
**Student Visitor**
Studied how to reduce communication cost using a simplified transfer protocol in parallel processing.

1993 – 1996           Tsinghua University           Beijing, China
**Research Assistant**
Researched on parallel and distributed processing and high-performance networking.

**Publications**

*Journal Articles:*

J1. Answering Queries with Useful Bindings. Chen Li and Edward Chang. ACM Transactions on Database Systems (TODS), Volume 26, Issue 3, September 2001.

J2. Clustering for Approximate Similarity Search in High-Dimensional Spaces. Chen Li, Edward Chang, Hector Garcia-Molina, and Gio Wiederhold. IEEE Transaction on Knowledge and Data Engineering, Volume 14, Number 4, pp.792-808, July/August 2002.

J3. Computing Complete Answers to Queries in the Presence of Limited Access Patterns. Chen Li. VLDB J. 12(3): 211-227 (2003).

J4. Answering Queries Using Materialized Views with Minimum Size, Rada Chirkova, Chen Li, and Jia Li, VLDB Journal (2006), Volume 15, Number 3, 191-210.

J5. Achieving Communication Efficiency through Push-Pull Partitioning of Semantic Spaces in Client-Server Architectures, Amitabha Bagchi, Amitabh Chaudhary, Michael T. Goodrich, Chen Li, and Michal Shmueli-Scheuer. TKDE, October 2006 (Vol. 18, No. 10).

J6. Supporting Efficient Record Linkage for Large Data Sets Using Mapping Techniques, Liang Jin, Chen Li, and Sharad Mehrotra, World Wide Web Journal, Volume 9, Issue 4, 557-584, 2006.

J7. Rewriting Queries Using Views in the Presence of Arithmetic Comparisons, Foto Afrati, Chen Li, and Prasenjit Mitra, Theoretical Computer Science, Volume 368, Numbers 1-2, pages 88-123, 2006.

J8. Using Views to Generate Efficient Evaluation Plans for Queries, Foto Afrati, Chen Li, and Jeff Ullman, Journal of Computer and System Sciences, 73(5), 703-724, 2007.

J9. SEPIA: Estimating Selectivities of Approximate String Predicates in Large Databases, Liang Jin, Chen Li, and Rares Vernica, VLDB Journal, Volume 17, Issue 5, pages 1213-1229, August 2008.

J10. Scott Christley, Yiming Lu, Chen Li, Xiaohui Xie: Human genomes as email attachments. Journal of Bioinformatics 25(2): 274-275 (2009)

J11. Guoliang Li, Chen Li, Jianhua Feng, Lizhu Zhou: SAIL: Structure-

aware indexing for effective and progressive top-k keyword search over XML documents. Inf. Sci. 179(21): 3745-3762 (2009)

J12. Interactive and Fuzzy Search: A Dynamic Way to Explore MEDLINE, Jiannan Wang, Inci Cetindil, ShengyueJi, Chen Li, Xiaohui Xie, Guoliang Li, Jianhua Feng, Journal of Bioinformatics, 2010.

J13. Supporting BioMedical Information Retrieval: The BioTracer Approach, Heri Ramampiaro and Chen Li, In Transactions on Large-Scale Data- and Knowledge-Centered Systems (TLDKS), 2011, No.4. Vol. 6990, Springer. pp. 73–94.

J14. ASTERIX: towards a scalable, semistructured data platform for evolving-world models. Alexander Behm, Vinayak R. Borkar, Michael J. Carey, Raman Grover, Chen Li, Nicola Onose, Rares Vernica, Alin Deutsch, Yannis Papakonstantinou, Vassilis J. Tsotras, Distributed and Parallel Databases, 2011, 29(3), 185-216.

J15. Efficient fuzzy full-text type-ahead search, Guoliang Li, Shengyue Ji, Chen Li, Jianhua Feng:. VLDB J. 20(4): 617-640 (2011).

J16. Ahmadi, A., Behm, A., Honnalli, N., Li, C., Weng, L., Xie, X. (2011). Hobbes: optimized gram-based methods for efficient read alignment. Nucleic Acids Research.

J17. SKIF-P: a point-based indexing and ranking of web documents for spatial-keyword search, Ali Khodaei, Cyrus Shahabi, Chen Li, GeoInformatica 16(3): 563-596 (2012)

J18. Speeding Up Chemical Searches Using the Inverted Index: The Convergence of Chemoinformatics and Text Search Methods, Ramzi Nasr, Rares Vernica, Chen Li, Pierre Baldi. Journal of Chemical Information and Modeling 52(4): 891-900 (2012)

J19. Supporting Search-As-You-Type Using SQL in Databases, Guoliang Li, Jianhua Feng, Chen Li, IEEE Trans. Knowl. Data Eng. 25(2): 461-475 (2013)

J20. Improving read mapping using additional prefix grams, Jongik Kim, Chen Li, Xiaohui Xie. BMC Bioinformatics 15: 42 (2014)

J21. Negative Factor: Improving Regular-Expression Matching in Strings, Xiaochun Yang, Tao Qiu, Bin Wang, Baihua Zheng, Yaoshu Wang, Chen Li, ACM TODS, 2015.

***Peer-reviewed Conferences Full Papers:***

C1. Performance Analysis of the Communication Mechanism for POE Workstation Cluster. Weiqiang Zhuang, Chen Li, Meiming Shen. Microcomputer & Micro-system, Jan, 1995.

C2. 2D BubbleUp: Managing Parallel Disks for Media Servers. Edward Chang, Hector Garcia-Molina, and Chen Li. Proc. of International Conference of Foundations of Data Organization (FODO'98), pages 221-230, Kobe, Japan, 1998.

C3. RIME: A Replicated Image Detector for the World-Wide Web. Edward Chang, James Ze Wang, Chen Li, and Gio Wiederhold. Proceedings of SPIE Symposium of Voice, Video, and Data Communications, pages 58—67, Boston, MA, November 1998.

C4. Searching Near-Replicas of Images via Clustering. Edward Chang, Chen Li, James Wang, Peter Mork, and Gio Wiederhold. Proc. of SPIE Symposium of Voice, Video, and Data Communications, Multimedia Storage and Archiving Systems VI, pages 281-292, Boston, MA, September, 1999.

C5. Optimizing Large Join Queries in Mediation Systems. Ramana Yerneni, Chen Li, Jeffrey Ullman, Hector Garcia-Molina. International

Conference on Database Theory (ICDT), Jerusalem, Israel, January, 1999. (29% accepted)

C6. Computing Capabilities of Mediators. Ramana Yerneni, Chen Li, Hector Garcia-Molina, Jeffrey Ullman. SIGMOD'99, Philadelphia, PA, May 1999. (20% accepted)

C7. Query Planning with Limited Source Capabilities. Chen Li and Edward Chang. Proc. of International Conference on Database Engineering (ICDE), pages 401-412, San Diego, CA, February, 2000. (14% accepted)

C8. On Answering Queries in the Presence of Limited Access Patterns. Chen Li and Edward Chang. Proc. of International Conference on Database Theory (ICDT), pages 219-233, London, UK, January, 2001. (35% accepted)

C9. Minimizing View Sets without Losing Query-Answering Power. Chen Li, Mayank Bawa, and Jeff Ullman. Proc. of International Conference on Database Theory (ICDT), pages 99-113, London, UK, January, 2001. (35% accepted)

C10. Generating Efficient Plans for Queries Using Views. Foto Afrati, Chen Li, and Jeff Ullman. Proc. of ACM SIGMOD, pages 319 - 330, Santa Barbara, CA, May, 2001. (15% accepted)

C11. Answering Queries Using Views with Arithmetic Comparisons. Foto Afrati, Chen Li, and Prasenjit Mitra. In Proc. of ACM PODS, pages 209-220, June 2002, Madison, Wisconsin. (22% accepted)

C12. Executing SQL over Encrypted Data in the Database-Service-Provider Model. Hakan Hacigumus, Bala Iyer, Chen Li, and Sharad Mehrotra. In Prof. of ACM SIGMOD, pages 216-227, June 2002, Madison, Wisconsin. (18% accepted). **Received SIGMOD 2012 10-year Test-of-Time Award.**

C13. Efficient Record Linkage in Large Data Sets, Liang Jin, Chen Li, and Sharad Mehrotra, Proc. of the 8th International Conference on Database Systems for Advanced Applications (DASFAA), pages 137-146, March, 2003, Kyoto, Japan. (33% accepted). **Received DASFAA 2013 10-year Best Paper Award.**

C14. Materializing Views with Minimal Size to Answer Queries. Rada Chirkova and Chen Li. Proc. of ACM PODS, pages 38-48, San Diego, CA, June 2003. (20% accepted).

C15. On Containment of Conjunctive Queries with Arithmetic Comparisons. Foto Afrati, Chen Li, Prasenjit Mitra. Proc. of EDBT'04, pages 459-476, Crete, Greece, March 2004 (14% accepted).

C16. NNH: Improving Performance of Nearest-Neighbor Searches Using Histograms. Liang Jin, Nick Koudas, Chen Li. Proc. of EDBT'04, pages 385-402, Crete, Greece, March 2004. (14% accepted).

C17. Secure XML Publishing without Information Leakage in the Presence of Data Inference. Xiaochun Yang and Chen Li, Proc. of VLDB'04, pages 96-107, August 29 -- September 3, Toronto, Canada. (16% accepted).

C18. Indexing Mixed Types for Approximate Retrieval, Liang Jin, Nick Koudas, Chen Li, Anthony K.H. Tung. VLDB 2005. (16% accepted).

C19. Selectivity Estimation for Fuzzy String Predicates in Large Data Sets, Liang Jin and Chen Li. VLDB 2005. (16% accepted).

C20. Relaxing Join and Selection Queries. Nick Koudas, Chen Li, Anthony Tung, and Rares Vernica. VLDB 2006, Seoul, Korea, 2006. (13.2% accepted)

C21. Supporting Approximate Similarity Queries with Quality Guarantees in P2P Systems, Qi Zhong, Iosif Lazaridis, Mayur Deshpande, Chen Li, Sharad Mehrotra, Hal Stern, COMAD 2006, December 14-16, 2006, Delhi, India. (26% accepted)

C22. Protecting Individual Information Against Inference Attacks in Data Publishing, Chen Li, Houtan Shirani-Mehr, and Xiaochun Yang. DASFAA 2007. (18.7% accepted)

C23. VGRAM: Improving Performance of Approximate Queries on String Collections Using Variable-Length Grams, Chen Li, Bin Wang, and Xiaochun Yang. VLDB 2007.

C24. Processing Spatial-Keyword (SK) Queries in Geographic Information Retrieval (GIR) Systems, Ramaswamy Hariharan, Bijit Hore, Chen Li, Sharad Mehrotra, SSDBM, 2007.

C25. Data Exchange with Arithmetic Comparisons, Foto Afrati, Chen Li, and Vassia Pavlaki. EDBT 2008.

C26. Efficient Merging and Filtering Algorithms for Approximate String Searches, Chen Li, Jiaheng Lu, Yiming Lu, ICDE 2008, 257-266.

C27. Supporting Keyword Queries on Structured Databases with Limited Search Interfaces Nurcan Yuruk, Xiaowei Xu, Chen Li, Jeffrey Xu Yu, DASFAA 2008.

C28. Cost-Based Variable-Length-Gram Selection for String Collections to Support Approximate Queries Efficiently, Xiaochun Yang, Bin Wang, and Chen Li, ACM SIGMOD 2008.

C29. Space-Constrained Gram-Based Indexing for Efficient Approximate String Search, Alexander Behm, Shengyue Ji, Chen Li, and Jiaheng Lu, ICDE 2009.

C30. Best-Effort Top-k Query Processing Under Budgetary Constraints, Michal Shmueli-Scheuer, Chen Li, Yosi Mass, Haggai Roitman, Ralf Schenkel, and Gerhard Weikum, ICDE 2009.

C31. Guoliang Li, Shengyue Ji, Chen Li, Jianhua Feng: Efficient type-ahead search on relational data: a TASTIER approach, SIGMOD Conference, 2009.

C32. Shengyue Ji, Guoliang Li, Chen Li, Jianhua Feng: Efficient interactive fuzzy keyword search, International World Wide Web Conference, 2009.

C33. Efficient Approximate Search on String Collections (Tutorial), Marios Hadjieleftheriou, Chen Li, PVLDB 2009.

C34. Efficient parallel set-similarity joins using MapReduce, Rares Vernica, Michael J. Carey, Chen Li. SIGMOD 2010.

C35. Hybrid Indexing and Seamless Ranking of Spatial and Textual Features of Web Documents, Ali Khodaei, Cyrus Shahabi, Chen Li, DEXA 2010.

C36. Supporting Location-Based Approximate-Keyword Queries. Sattam Alsubaiee, Alexander Behm, Chen Li, ACM GIS 2010.

C37. Answering Approximate String Queries on Large Data Sets Using External Memory, Alexander Behm, Chen Li, and Michael Carey, ICDE 2011.

C38. CHIME: An Efficient Error-Tolerant Chinese Pinyin Input Method, Yabin Zheng, Chen Li, Maosong Sun, IJCAI 2011, 2551-2556.

C39. Location-Based Instant Search, Shengyue Ji, Chen Li, SSDBM 2011: 17-36.

C40. Inside "Big Data Management": Ogres, Onions, or Parfaits? Vinayak Borkar, Michael J. Carey, and Chen Li, EDBT 2012

C41. Supporting Efficient Top-k Queries in Type-Ahead Search, Guoliang Li, Jiannan Wang, Chen Li, Jianhua Feng, SIGIR 2012.

C42. Efficient direct search on compressed genomic data, Xiaochun Yang, Bin Wang, Chen Li, Jiaying Wang, Xiaohui Xie, ICDE 2013: 961-972

C43. Improving regular-expression matching on strings using negative factors, Xiaochun Yang, Bin Wang, Tao Qiu, Yaoshu Wang, Chen Li, SIGMOD Conference 2013: 361-372

C44. String similarity measures and joins with synonyms, Jiaheng Lu,

Chunbin Lin, Wei Wang, Chen Li, Haiyong Wang, SIGMOD Conference 2013: 373-384

C45. Storage Management in AsterixDB, Sattam Alsubaiee, Alexander Behm, Vinayak R. Borkar, Zachary Heilbron, Young-Seok Kim, Michael J. Carey, Markus Dreseler, Chen Li, PVLDB 7(10): 841-852 (2014)

C46. AsterixDB: A Scalable, Open Source BDMS, Sattam Alsubaiee, Yasser Altowim, Hotham Altwaijry, Alexander Behm, Vinayak R. Borkar, Yingyi Bu, Michael J. Carey, Inci Cetindil, Madhusudan Cheelangi, Khurram Faraaz, Eugenia Gabrielova, Raman Grover, Zachary Heilbron, Young-Seok Kim, Chen Li, Guangqiang Li, Ji Mahn Ok, Nicola Onose, Pouria Pirzadeh, Vassilis J. Tsotras, Rares Vernica, Jian Wen, Till Westmann, PVLDB 7(14): 1905-1916 (2014)

C47. Efficient instant-fuzzy search with proximity ranking, Inci Cetindil, Jamshid Esmaelnezhad, Taewoo Kim, Chen Li. ICDE 2014: 328-339.

C48. LSM Based Storage and Indexing: An Old Idea with Timely Benefits, Alsubaiee, S., Carey, M. J., Li, C. GeoRich (2015).

C49. Hobbes3: Dynamic Generation of Variable-Length Signatures for Efficient Approximate Subsequence Mappings, Jongik Kim, Chen Li, Xiaohui Xie, ICDE 2016.


### *Peer-reviewed Workshop Publications and Demos (appeared in Proceedings):*

W1.  HiComm -- A New Technique for Improving Communication Performance in Workstation Cluster. Chen Li, Weiqiang Zhuang, Meiming Shen, Dingxing Wang, Weimin Zheng, APPT'95, October, 1995, Beijing, China.

W2.  Capability Based Mediation in TSIMMIS. Chen Li, Ramana Yerneni, Vasilis Vassalos, Hector Garcia-Molina, Yannis Papakonstantinou, Jeffrey Ullman, Murty Valiveti. Proc. of ACM SIGMOD, Demo track, pages 564 – 566, Seattle, WA, June, 1998.

W3.  Towards Perception-Based Image Retrieval. Edward Chang, Beitao Li, and Chen Li. Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries, pages 401-412, South Carolina, June, 2000.

W4.  Answering Queries with Database Restrictions (Research Summary). Chen Li. Symposium on Abstraction, Reformulation and Approximation (SARA), pages 328 – 329, July 2000, Horseshoe Bay (Lake LBJ), Texas.

W5.  Describing and Utilizing Constraints to Answer Queries in Data-Integration Systems. Chen Li. Online Proc. of IJCAI 2003 workshop on Information Integration on the Web, pages 163-168, August 2003, Acapulco, Mexico.

W6.  Schema-Guided Wrapper Maintenance for Web-Data Extraction. Xiaofeng Meng, Dongdong Hu, Chen Li. Proc. of International Workshop on Web Information and Data Management (WIDM), pages 1-8, New Orleans, Louisiana, Nov. 2003. (28% accepted)

W7.  A Supervised Visual Wrapper Generator for Web-Data Extraction. Xiaofeng Meng, Haiyan Wang, Dongdong Hu, Chen Li: Proc. of the 28th Annual International Computer Software and Applications Conference (COMPSAC), pages 657 – 662, Dallas, TX, Nov. 2003.

W8.  RACCOON: A Peer-Based System for Data Integration and Sharing. Chen Li, Jia Li, Qi Zhong. Proc. of ICDE'2004, page 852, Boston, MA, March 2004, demo track. (58% accepted)

W9.  XGuard: A System for Publishing XML Documents without Information Leakage in the Presence of Data Inference. Proc. of

ICDE'2005, Demo track, Tokyo Japan, March 2005.

W10. Answering Aggregation Queries on Hierarchical Web Sites Using Adaptive Sampling. Foto Afrati, Paraskevas Lekeas, and Chen Li. Technical Report, UCI ICS, August 2005. A short version appears in CIKM'2005, 31st October - 5th November, 2005 Bremen, Germany.

W11. Quality-Driven Approximate Methods for GIS Data Integration. Ramaswamy Hariharan, Michal Schmueli-Scheuer, Chen Li, and Sharad Mehrotra. ACM GIS 2005, November 4-5th, 2005 Bremen, Germany.

W12. Communication-Efficient Query Answering with Quality Guarantees in Client-Server Applications. Michal Shmueli-Scheuer, Amitabh Chaudhary, Avigdor Gal, Chen Li. WebDB 2007

W13. Quality-Aware Retrieval of Data Objects from Autonomous Sources for Web-Based Repositories, Houtan Shirani-Mehr, Chen Li, Gang Liang, Michal Shmueli-Scheuer, ICDE 2008.

W14. Rares Vernica, Chen Li: Efficient top-k algorithms for fuzzy search in string collections. KEYS 2009: 9-14

W15. Automatic URL Completion and Prediction Using Fuzzy Type-Ahead Search, Jiannan Wang, Guoliang Li, Jianhua Feng, Chen Li, SIGIR 2009 (Poster)

W16. Efficient fuzzy type-ahead search in TASTIER, Guoliang Li, Shengyue Ji, Chen Li, Jiannan Wang, Jianhua Feng, ICDE 2010, Demo.

W17. Fuzzy Keyword Search on Spatial Data, Sattam Alsubaiee, Chen Li, DASFAA 2010. (Excellent Demo Award)

W18. Seaform: Search-As-You-Type in Forms, Hao Wu, Guoliang Li, Chen Li, Lizhu Zhou, VLDB 2010 (Demo).

W19. The Flamingo Software Package on Approximate String Queries. Chen Li, DASFAA Workshops 2011, 477.

W20. Big data platforms: what's next? Vinayak R. Borkar, Michael J. Carey, Chen Li, ACM Crossroads 19(1): 44-49 (2012)

W21. ASTERIX: An Open Source System for "Big Data" Management and Analysis, Sattam Alsubaiee, Yasser Altowim, Hotham Altwaijry, Alexander Behm, Vinayak R. Borkar, Yingyi Bu, Michael J. Carey, Raman Grover, Zachary Heilbron, Young-Seok Kim, Chen Li, Nicola Onose, Pouria Pirzadeh, Rares Vernica, Jian Wen. PVLDB 5(12): 1898-1901 (2012)

W22. ASTERIX: Scalable Warehouse-Style Web Data Integration, Sattam Alsubaiee, Alexander Behm, Raman Grover, Rares Vernica, Vinayak Borkar, Michael J. Carey, Chen Li. IIWeb 2012 (co-located with SIGMOD 2012).

W23. Towards Interactive Analytics and Visualization on One Billion Tweets, Jianfeng Jia, Chen Li, Xi Zhang, Chen Li, Michael J. Carey, Simon Su, ACM GIS 2016

W24. A Demonstration of TextDB: Declarative and Scalable Text Analytics on Large Data Sets, Zuozhi Wang, Flavio Bayer, Seungjin Lee, Kishore Narendran, Xuxi Pan, Qing Tang, Jianwen Wang and Chen Li, ICDE 2017.

**Non-refereed Publications and Book Chapters:**

O1. Managing Parallel Disks for Continuous Media Data. Edward Chang, Chen Li, and Hector Garcia-Molina. A Book Chapter in Information Organization & Databases, p.107-120, Kluwer Publisher, 2000

O2. Using Constraints to Describe Source Contents in Data Integration Systems. Chen Li, IEEE Intelligent Systems 18(5): 49-53 (2003).

O3. Recent Progress on Selected Topics on Database Research -- A Report from Nine Young Chinese Researchers Working in the United States.

The Journal of Computer Science and Technology, Zhiyuan Chen, Chen Li, Jian Pei, Yufei Tao, Haixun Wang, Wei Wang, Jiong Yang, Jun Yang, and Donghui Zhang. Journal of Computer Science and Technology, Volume 18, Issue 5, Pages 538 – 552, September 2003.

O4. Report on the First International VLDB Workshop on Clean Databases (CleanDB 2006), Dongwon Lee and Chen Li, ACM SIGMOD Record, 36(1), 54 – 56, 2007.

O5. Chen Li: Rewriting Queries using Views. Encyclopedia of Database Systems 2009: 2438-2441

O6. Search-As-You-Type: Opportunities and Challenges, Chen Li, Guoliang Li, IEEE Data Eng. Bull. 2010.

O7. Record Linkage: A 10-Year Retrospective. Chen Li, Sharad Mehrotra, Liang Jin, DASFAA (1) 2013: 3-12

**Patents**

- Li, C., Wang, B., Yang, X., Patent, "IMPROVING PERFORMANCE OF APPROXIMATE STRING QUERIES USING VARIABLE-LENGTH HIGH-QUALITY GRAMS", US Patent 7996369, http://www.google.com/patents/US7996369
- Li, C., Li, G., Ji, S., Wang, J., Feng, J., Patent, "SUPPORTING INTERACTIVE, FUZZY SEARCH ON STRUCTURED DATA EFFICIENTLY", US Patent 8631035, http://www.google.com/patents/US8631035

**Professional activities**

o July 2016 – now: ICDE 2017 Area Chair
o July 2016 – now: SIGMOD 2017 committee member
o Oct. 2015 – now: PVLDB advisory committee member.
o Oct. 2015 – now: Guest editor of VLDBJ Special Issue for VLDB 2015.
o Co-PC Chair of VLDB2015.
o **Recent Program Member:** Journal of Distributed and Parallel Databases (Editor Board member), PVLDB 2011, SIGMOD 2010 (Group Leader), VLDB 2010, ICDE 2010 (also Local Arrangements Chair), PVLDB 2009, ICDE 2009 Demo Track, ADC 2009, CIKM 2008, VLDB 2008, WWW 2008, ICDE 2008, KDD 2007, SIGMOD 2007, ICDT 2007, WebDB 2007, KDD 2006, CIKM 2006, VLDB 2005
o **Program Member and Proceedings Chair**, ACM PODS 2005.
o **Program Member of Demonstrations Track**, ACM SIGMOD 2005.
o **Program Member:** DASFAA 2005.
o **Program Member:** The Third International Semantic Web Conference, ISWC 2004.
o **Program Member**, CIKM 2004.
o **Program Member**: The Sixth Asia Pacific Web Conference, 2004.
o **Program Member**: the 9th International Conference on Database Systems for Advanced Applications (DASFAA 2004), Seogwipo KAL Hotel, Jeju Island, Korea.
o **Program Member**: The Fourth International Conference on Web-Age Information Management (WAIM 2003) China, 2003.
o **Program Member:** the first workshop on Semantics in Peer-to-Peer and Grid Computing. at the Twelfth International World Wide Web Conference 20 May 2003, Budapest, Hungary.
o **Panelist** of a review panel of the NSF Division of Information and Intelligent Systems, 2003.
o **Demo/Industry Chair, committee member, and Session Chair** of The Third International Conference on Web-Age Information Management (WAIM 2002), August, Beijing, China.

| | |
|---|---|
| | o **Panelist** of Advances on Web-age Database Technologies: An International Forum (WDBT'02), Beijing, China |
| | o **Organizer** of Cal-It(2) Workshop for Crisis Response, Tuesday, March 19th, 2002, Beckman Conference Center, UC Irvine. |
| **Affiliated Ph.D. students and postdocs** | o Inci Cetindil, PhD student, ICS, UCI |
| | o Jamshid Esmaelnezhad, PhD student, ICS, UCI |
| | o Jianfeng Jia, PhD student, ICS, UCI |
| | o Taewoo Kim, PhD student, ICS, UCI |
| | o |
| | o Young-Seok Kim, PhD student, ICS, UCI, 2016, Employer: Samsung |
| | o Sattam Mubark Alsubaiee, PhD, ICS, UCI, 2014, Research Assistant Professor at KACST, Saudi Arabia |
| | o Alex Behm, ICS, UCI, PhD 2013, Cloudera |
| | o Rares Vernica, ICS, UCI, Ph.D 2012, researcher at HP Labs |
| | o Shengyue Ji, ICS, UCI, PhD 2011, Google |
| | o Michal Shmueli-Scheuer: PhD 2009, researcher at IBM Haifa Research Lab |
| | o Jiaheng Lu, Postdoc, 2006~2008, Professor at Renmin University, China |
| **Awards and Funding** | o 2016, Research fund from ARL (Army Research Lab) |
| | o 2013, NSF grant 1305430 titled "CI-ADDO-NEW: Asterix: A Community Software Platform for Big Data Research, Analysis, and Management." |
| | o 2013, Gift funds from HTC. |
| | o 2013, DASFAA 10-year Best Paper Award. |
| | o 2012, ACM SIGMOD 10-year Test-of-Time award. |
| | o 2012, UCI ICS Dean's Award for Graduate Student Mentoring. |
| | o 2011, the third place at the Microsoft Speller Challenge competition. |
| | o 2010, NIH grant ($372,360) to support research on the iPubmed system |
| | o 2010, Intel grant on compression of human genome data. |
| | o 2010, NSF research award IIS 1030002 on powerful keyword search in a cloud-computing infrastructure. |
| | o 2009, NSF research award 0910989 titled "DC: Large: Collaborative Research: ASTERIX: A Highly Scalable Parallel Platform for Semistructured Data Management and Analysis." |
| | o 2009, NSF grant 0844574 titled "Supporting efficient fuzzy queries on large text repositories using Hadoop." |
| | o 2009, UCI ICS Dean's Award for excellence in Mid-Career Research, 2009. |
| | o 2008, Research Funding Award of the "Research Funds for Oversea Scholars" program of the National Natural Science Foundation of China, September. |
| | o 2007, NSF Grant titled "SGER: Answering Approximate String Queries Using Variable-Length Grams" |
| | o 2006, ICS Ted & Janice Smith Faculty Seed Fund |
| | o 2006, Google Research Award, $37.5K, renewable for a $2^{nd}$ year. |
| | o 2006: Gift Fund from Microsoft Research (7K). |
| | o 2005: UCI Career Development Award. |
| | o 2003 NSF CAREER Award, $400K, PI, starting 9/1/2003, for 5 years. |
| | o 2003 NSF ITR Award 0331707, "Responding to the Unexpected," $12.5M for 5 years, Senior Investigator, starting 10/1/2003. |
| | o Single Investigator Innovation Grant: CORCLR, UC Irvine, May 2003. |
| | o Stanford Graduate Fellowship, Stanford University, 1997 – 2001. |
| | o Merit-Based Scholarships, Tsinghua University, China, 1994 – 1996. |
| | o Outstanding Student Award, Tsinghua University, June 1994. |

- o Entrance exams waived, Tsinghua University, 1994 and 1989.