

## **Research Output (journal articles, conference papers, book chapters, software)**

Note: 1) A “\*” before the author name means he/she is my postdoc. fellow, graduate student, or research assistant.  
2) The citation numbers are obtained from Google scholar on 7<sup>th</sup> December, 2010. Only the top 25 citations are shown.

### **Ph.D. Thesis**

1. **R. Cheng**. *Managing Uncertainty in Constantly-Evolving Environments*. Ph.D thesis, Purdue University, 2005.

### **Scholarly Books, Monographs, and Chapters**

2. **R. Cheng**. *Uncertain Spatial Data Management*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer New York, pp. 1-6, 2017 (online). ISBN: 978-1-4899-7993-3.
3. S. Prabhakar and **R. Cheng**. *Indexing Uncertain Data*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer New York, pp. 1-6, 2017 (online). ISBN: 978-1-4899-7993-3.
4. **R. Cheng** and J. Chen. *Probabilistic Spatial Queries*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer New York, pp. 1-6, 2017 (online). ISBN: 978-1-4899-7993-3.
5. S. Prabhakar and **R. Cheng**. *Data Uncertainty Management in Sensor Networks*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer New York, pp. 1-6, 2017 (online). ISBN: 978-1-4899-7993-3.
6. **R. Cheng**. *Spatial Data, Indexing Techniques*. In the Encyclopedia of Geographical Information Science, S. Shekhar, H. Xiong, and X. Zhou (eds.), Springer, 2017, pp. 1992-2002. ISBN: 978-3-319-17884-4.
7. **R. Cheng**, **Y. Fang\***, and M. Renz. *Uncertain Data Classification*. In Data Classification: Algorithms and Applications, C. C. Aggarwal (eds.), Chapman & Hall / CRC Data Mining and Knowledge Discovery Series, May 2014, ISBN: 978-1466586741.
8. **R. Cheng**. *Managing Quality of Probabilistic Databases*. In the Handbook on Research and Practice in Data Quality, S. Sadiq (eds.), Springer, 2012.
9. **R. Cheng** and **J. Chen\***. *Probabilistic Spatial Queries*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer-Verlag, Sep 2009, ISBN: 978-0-387-35544-3.
10. S. Prabhakar and **R. Cheng**. *Data Uncertainty Management in Sensor Networks*. In the Encyclopedia of Database Systems, L. Liu and T. Ozsu (eds.), Springer-Verlag, Sep 2009. ISBN: 978-0-387-35544-3.
11. **R. Cheng**. *Spatial Data, Indexing Techniques*. In the Encyclopedia of Geographical Information Science, S. Shekhar and H. Xiong (eds.), Springer, 2008. ISBN: 978-0-387-30858-6.
12. **R. Cheng**, E. Chan and K. Y. Lam. *Quality Assurance of Probabilistic Queries*. In Sensor Network and Configuration: Fundamentals, Techniques, Platforms, and Experiments, edited by N. Mahalik, Springer-Verlag, Germany, 2007, XX, 510 p., 205 illus., Hardcover. ISBN: 978-3-540-37364-3.
13. **R. Cheng** and S. Prabhakar. *Sensors, Uncertainty Models and Probabilistic Queries*. In Encyclopedia of Database Technologies and Applications, L. Rivero, J. Doorn and E. Ferraggine(ed.), Idea Group Publishing, 2005.
14. B. Kao and **R. Cheng**. *Disk I/O Scheduling*. In Real-Time Database Systems: Architecture and Issues, edited by K. Y. Lam and T. W. Kuo, Kluwer Academic Publishers, pp. 97-107, Boston, Dec 2001.

### **Journal Papers**

15. **Y. Fang\***, **R. Cheng**, **S. Luo\***, **J. Hu\***, **X. Li\***. *Effective Community Search over Large Spatial Graphs*. In Proceedings of the VLDB Endowment (PVLDB), 10(6), pp. 709-720, Feb 2017. Also presented in the Very Large Databases Conf. (VLDB 2017), Munich, Germany, Aug 28-Sep 1, 2017.

16. **Y. Zheng\***, G. Li, Y. Li, **C. Shan\***, and **R. Cheng\***. *Truth Inference in Crowdsourcing: Is the Problem Solved? [Experiments and Analyses]*. In Proceedings of the VLDB Endowment (**PVLDB**), 10(5), pp. 541-552, Jan 2017. To be presented in the Very Large Databases Conf. (**VLDB 2017**), Munich, Germany, Aug 28-Sep 1, 2017.
17. **Y. Zheng\***, G. Li, and **R. Cheng\***. *DOCS: Domain-Aware Crowdsourcing System*. In Proceedings of the VLDB Endowment (**PVLDB**), 10(4), pp. 361-372, Dec 2016. To be presented in the Very Large Databases Conf. (**VLDB 2017**), Munich, Germany, Aug 28-Sep 1, 2017.
18. **J. Hu\***, X. Wu, **R. Cheng**, **S. Luo\***, and **Y. Fang\***. *On Minimal Steiner Maximum-Connected Subgraphs*. Accepted in the Transactions on Knowledge and Data Engineering (**IEEE TKDE**), July 2017.
19. **S. Maniu\***, **R. Cheng**, and P. Senellart. *An Indexing Framework for Queries on Probabilistic Graphs*. In ACM Transactions on Database Systems (**TODS**), 42(2), pp. 1-34, June 2017, ISSN: 0362-5915.
20. C. Ren, E. Lo, B. Kao, X. Zhu, **R. Cheng**, and D. Cheung. *Efficient Processing of Shortest Path Queries in Evolving Graph Sequences*. In Information Systems (**IS**), in press, accepted in May 2017.
21. **Y. Fang\***, X. Xie, X. Zhang, **R. Cheng**, and Z. Zhang. *STEM: a suffix tree-based method for web data records extraction*. In Knowledge and Information Systems (**KAIS**), pp. 1-27, May 2017.
22. K. A. Schmid, A. Zuefle, T. Emrich, M. Renz, and **R. Cheng**. *Uncertain UV cell computation based on space decomposition. (Selected as one of the best papers from SSTD for consideration in Geoinformatica.)* In Geoinformatica, Springer, ISSN 1573-7624, pp. 1-31, Feb 2017.
23. Z. He, P. Wong, B. Kao, E. Lo, **R. Cheng**, and Z. Feng. *Efficient Pattern-Based Aggregation on Sequence Data*. In the Transactions on Knowledge and Data Engineering (**IEEE TKDE**), ISSN: 1041-4347, PP(99), Oct 2016.
24. **Y. Fang\***, **R. Cheng**, **S. Luo\***, and **J. Hu\***. *Effective Community Search for Large Attributed Graphs*. In Proceedings of the VLDB Endowment (**PVLDB**), 9(12), pp. 1233-1244, Aug 2016. Also presented in the Very Large Databases Conf. (**VLDB 2016**), New Delhi, India, Sep 5-9, 2016.
25. Z. Li, **Y. Fang\***, Q. Liu, J. Cheng, **R. Cheng**, and J. C.S. Lui. *PASCO: Parallel SimRank Computation at Scale*. In Proceedings of the VLDB Endowment (**PVLDB**), 9(1), pp. 24-35, Sep 2015. Also presented in the Very Large Databases Conf. (**VLDB 2016**), New Delhi, India, Sep 5-9, 2016.
26. C. Dai, S. Nutanong, C.Y. Chow, and **R. Cheng**. *Entropy-based Scheduling Policy for Cross Aggregate Ranking Workloads*. In the IEEE Transactions on Services Computing (**IEEE TSC**), ISSN: 1939-1374, June 2016.
27. **Y. Fang\***, **R. Cheng**, W. Tang, **S. Maniu\***, and **X. Yang\***. *Scalable Algorithms for Nearest-Neighbor Joins on Big Trajectory Data*. In the Transactions on Knowledge and Data Engineering (**IEEE TKDE**), pp. 785-800, ISSN: 1041-4347, 28(3), March 2016.
28. Z. Wang, B. Yao, **R. Cheng**, X. Gao, L. Zou, H. Guan, and M. Guo. *SMe: Explicit and Implicit Constrained-Space Probabilistic Threshold Range Queries for Moving Objects*. Jan 2016, 20(1), pp. 19-58, ISSN 1384-6175, **Geoinformatica**, Springer.
29. **C. J. Zhu\***, K. Y. Lam, **R. Cheng**, and C. K. Poon. *On Using Broadcast Index for Efficient Execution of Shortest Path Continuous Queries*. In Information Systems (**IS**), Volume 49, April 2015, pp. 142-162.
30. **Y. Tang\***, L. H. U, Y. Cai, N. Mamoulis, and **R. Cheng**. *Earth Mover's Distance based Similarity Search at Scale*. In Proceedings of the VLDB Endowment (**PVLDB**), 7(4), pp. 313-324. Also presented in the Very Large Databases Conf. (**VLDB 2014**), Hangzhou, China, Sep 1-5, 2014.
31. **X. Xie\***, M. L. Yiu, **R. Cheng**, and L. Hua. *Scalable Evaluation of Trajectory Queries over Imprecise Location Data*. In the Transactions on Knowledge and Data Engineering (**TKDE**), pp. 2029-2044, ISSN: 1041-4347, May 2013.

32. **X. Xie\***, **R. Cheng**, M. L. Yiu, **L. Sun\***, and **J. Chen**. *UV-Diagram: A Voronoi Diagram for Uncertain Spatial Databases*. In the Very Large Databases Journal (**VLDBJ**), 22(3), pp. 319-344, June 2013.
33. **Y. Zhang\*** and **R. Cheng**. *Probabilistic Filters: A Stream Protocol for Continuous Probabilistic Queries*. In Information Systems (**IS**), 38(1), pp. 132-154, March 2013.
34. L. Wang, D. W. Cheung, **R. Cheng**, S. D. Lee, and **X. Yang\***. *Efficient Mining of Frequent Itemsets on Large Uncertain Databases*. In the IEEE Transactions on Knowledge and Data Engineering (**IEEE TKDE**), 24(12), pp. 2170-2183, Dec 2012.
35. T. Bernecker, **R. Cheng**, D. Cheung, H. Kriegel, S. D. Lee, M. Renz, F. Verhein, L. Wang, and A. Zuefle. *Model-based Probabilistic Frequent Itemset Mining*. In the Knowledge and Information Systems Journal (**KAIS**), October 2013, Volume 37, Issue 1, pp 181-217, Springer, *impact factor: 2 (Selected among the best papers in CIKM'10 conference for journal extension)*.
36. **J. Gong\***, **R. Cheng**, and D. W. Cheung. *Efficient Management of Uncertainty in XML Schema Matching*. In the Very Large Databases Journal (**VLDBJ**). 21(3), pp. 385-409, 2012.
37. **L. Sun\***, **R. Cheng**, **X. Li\***, D. Cheung, and J. Han. *On Link-based Similarity Join*. In Proceedings of the VLDB Endowment (**PVLDB**), 4(11), pp. 714-725. Also presented in the Very Large Databases Conf. (**VLDB 2011**), Seattle, Aug, 2011.
38. C. Ren, E. Lo, B. Kao, X. Zhu, and **R. Cheng**. *On Querying Historical Evolving Graph Sequences*. In Proceedings of the VLDB Endowment (**PVLDB**), 4(11), pp. 726-737. Also presented in the Very Large Databases Conf. (**VLDB 2011**), Seattle, Aug, 2011.
39. W. K. Ngai, B. Kao, **R. Cheng**, M. Chau, S. D. Lee, D. W. Cheung, and K. Y. Yip. *Metric and Trigonometric Pruning for Clustering of Uncertain Data in 2D Geometric Space*. In Information Systems (**IS**), 36(2), pp. 476-497, 2011. (*Impact factor: 1.966*)
40. T. Farrell, K. Rothermel, and **R. Cheng**. *Processing Continuous Range Queries with Spatio-Temporal Tolerance*. In IEEE Transactions on Mobile Computing (**IEEE TMC**), 10(3), pp. 320-334, Mar 2011. (*Impact factor: 2.647*)
41. **R. Cheng**, E. Lo, **X. Yang\***, M. Luk, **X. Li\***, and **X. Xie\***. *Explore or Exploit? Effective Strategies for Disambiguating Large Databases*. In Very Large Databases Conf. (**VLDB 2010**), Singapore, Sep, 2010. *Acceptance rate: 17.3%; also appeared in Volume 3 of the Journal "Proceedings of the VLDB Endowment" (PVLDB), a new Database journal.*
42. **R. Cheng**, B. Kao, A. Kwan, S. Prabhakar and Y. Tu. *Filtering Data Streams for Entity-based Continuous Queries*. In IEEE Transactions on Knowledge and Data Engineering (**IEEE TKDE**), Vol. 22, No. 2, pp. 234-248, Feb 2010. (*Impact factor: 2.285*)
43. **J. Chen\***, **R. Cheng**, M. Mokbel, and C. Chow. *Scalable Processing of Snapshot and Continuous Nearest-Neighbor Queries over One-Dimensional Uncertain Data*. In Very Large Databases Journal (**VLDBJ**), Special Issue on Uncertain and Probabilistic Databases, Vol. 18, No. 5, pp. 1219-1240, Oct 2009. (*Impact factor: 4.517*) (Awarded the **Research Output Prize** in Department of Computer Science, Faculty of Engineering, HKU, 2010)
44. D. Lin, E. Bertino, **R. Cheng**, and S. Prabhakar. *Position Transformation: A Location Privacy Protection Method for Moving Objects*. In Transactions on Data Privacy: Foundations and Technologies (**TDP**), 2(1): 21-46 (April 2009). URL: <http://www.tdp.cat>
45. **R. Cheng**, **J. Chen\***, and **X. Xie\***. *Cleaning Uncertain Data with Quality Guarantees*. In Very Large Databases Conf. (**VLDB 2008**), New Zealand, Aug 2008. *Also appeared in Volume 1 of the Journal "Proceedings of the VLDB Endowment" (PVLDB), a new Database journal. (Acceptance rate: 16.5%; cited 15 times)*

46. Y. Xia, **R. Cheng**, S. Prabhakar, S. Lei, and R. Shah. *Indexing Continuously Changing Data with Mean-Variance Tree*. In the Intl. Journal of High Performance Computing and Networking (**IJHPCN**): A Special Issue on Recent Advances in Collaborative Internet Computing, Vol. 5, No. 4, pp. 263-272, Inderscience Publishers, 27 Dec 2008.
47. Y. Tao, X. Xiao, and **R. Cheng**. *Range Search on Multidimensional Uncertain Data*. In ACM Transactions on Database Systems (**TODS**). 32(3):15, Aug 2007. (*Impact factor: 1.245; cited 46 times*)
48. **R. Cheng**, K.Y. Lam, S. Prabhakar, and B. Liang. *An Efficient Location Update Mechanism for Continuous Queries over Moving Objects*. In Information Systems (**IS**), Vol. 32, No. 4, pp. 593-620, Jun 2007. URL: <http://dx.doi.org/10.1016/j.is.2006.03.002> (*Impact factor: 1.966; cited 12 times*)
49. **R. Cheng**, D. Kalashnikov, and S. Prabhakar. *Evaluation of Probabilistic Queries over Imprecise Data in Constantly-Evolving Environments*. In Information Systems (**IS**), Vol. 32, No. 1, pp. 104-130, Mar 2007. URL: <http://dx.doi.org/10.1016/j.is.2005.06.002> (*Impact factor: 1.966; cited 23 times*)
50. S. Han, E. Chan, **R. Cheng**, and K. Y. Lam. *A Statistics-Based Sensor Selection Scheme for Continuous Probabilistic Queries in Sensor Network*. In Real Time Systems Journal (**RTS**), Vol. 35, No. 1, pp. 33-58, Jan 2007. URL: <http://www.springerlink.com/content/4635361u56167401> (*Impact factor: 1.200; cited 13 times*)
51. **R. Cheng**, D. Kalashnikov, and S. Prabhakar. *Querying Imprecise Data in Moving Object Environments*. In IEEE Transactions on Knowledge and Data Engineering (**IEEE TKDE**), Vol. 16, No. 9, pp. 1112-1127, Sep 2004. (*Impact factor: 2.285; cited 208 times*)
52. **R. Cheng** and S. Prabhakar. *Managing Uncertainty in Sensor Databases*. In Special Section on Sensor Network Technology and Sensor Data Management, **SIGMOD Record**, Vol.32, No.4, pp.41-46, Dec 2003. (*Impact factor: 0.642; cited 55 times*)
53. B. Kao, K. Y. Lam, B. Adelberg, **R. Cheng**, and T. Lee. *Maintaining Temporal Consistency of Discrete Objects in Soft Real-Time Database Systems*. In IEEE Transactions on Computers (**IEEE TC**), Vol. 52, No. 3, pp. 373-389, 2003. (*Impact factor: 1.822; cited 24 times*)
54. K. Y. Lam, T. W. Kuo, B. Kao, T. Lee, and **R. Cheng**. *Evaluation of Concurrency Control Strategies for Mixed Soft Real-Time Database Systems*. In Information Systems Journal (**IS**), Vol. 27, No. 2, pp. 123-149, Elsevier Science, 2002. (*Impact factor: 1.966; cited 26 times*)

### **Conference and Workshop Papers**

55. **J. Hu\***, **R. Cheng**, **Z. Huang\***, **Y. Fang\***, and **S. Luo\***. *On Embedding Uncertain Graphs*. In the 26th ACM Conf. on Information and Knowledge Management (**ACM CIKM 2017**), Singapore, Nov 6 – Nov 10, 2017. *Acceptance rate: 21% (171/820)*.
56. **S. Luo\***, **J. Hu\***, **R. Cheng**, and B. Kao. *SEQ: Example-based Query for Spatial Objects*. (Short Paper). In the 26th ACM Conf. on Information and Knowledge Management (**ACM CIKM 2017**), Singapore, Nov 6 – Nov 10, 2017. *Acceptance rate: 30% (119/398)*.
57. **Y. Xu\***, **R. Cheng**, and **Y. Zheng\***. *Reliable Retrieval of Top-k Tags*. In the 18<sup>th</sup> Intl. Conf. on Web Information Systems Engineering (**WISE**), Moscow, Russia, October 2017.
58. **Y. Fang\***, **R. Cheng**, **S. Luo\***, **J. Hu\***, and **K. Huang\***. *C-Explorer: Browsing Communities in Large Graphs (demo)*. In the 43<sup>rd</sup> Intl. Conf. on Very Large Data Bases (**VLDB**), Munich, Germany, August 2017.
59. **J. Hu\***, X. Wu, **R. Cheng**, **S. Luo\***, and **Y. Fang\***. *Querying Minimal Steiner Maximum-Connected Subgraphs in Large Graphs*. In **CIKM 2016**, pp. 1241-1250, Indianapolis, October 2016.

60. **Z. Huang\***, B. Cautis, **R. Cheng**, and **Y. Zheng\***. *KB-Enabled Query Recommendation for Long-Tail Queries*. In **CIKM 2016**, Indianapolis, October 2016.
61. **Z. Huang\***, **Y. Zheng\***, **R. Cheng**, Y. Sun, N. Mamoulis, and X. Li. *Meta Structure: Computing Relevance in Large Heterogeneous Information Networks*. In the 22nd ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (**KDD 2016**), San Francisco, US, August 2016.
62. H. Hu, **Y. Zheng\***, Z. Bao, G. Li, J. Feng, and **R. Cheng**. *Crowdsourced POI Labelling: Location-Aware Result Inference and Task Assignment*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2016**), Helsinki, Finland, May 2016.
63. **Y. Fang\***, **R. Cheng**, W. Tang, **S. Maniu\***, and **X. Yang\***. *Scalable Algorithms for Nearest-Neighbor Joins on Big Trajectory Data*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2016**), *TKDE Extended Abstract*, Helsinki, Finland, May 2016.
64. Y. Sun, **H. Sun\***, and **R. Cheng**. *Fast and Semantic Measurements on Collaborative Tagging Quality*. In the Advances in Knowledge Discovery and Data Mining – 20<sup>th</sup> Pacific-Asia Conference (**PAKDD 2016**), New Zealand, April 2016, pp. 363-375.
65. Z. Li, **Y. Fang\***, Q. Liu, J. Cheng, **R. Cheng**, and J. C.S. Lui. *PASCO: Parallel SimRank Computation at Scale (Poster)*. In the ACM Symposium on Cloud Computing (**SoCC 2015**), Hawaii, August 27-29, 2015.
66. **S. Lei\***, **S. Maniu\***, **L. Mo\***, **R. Cheng**, and P. Senellart. *Online Influence Maximization*. In the 21th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (**KDD 2015**), Sydney, Australia, August 2015.
67. **J. Hu\***, **R. Cheng**, D. Wu, and B. Jin. *Efficient Top-k Subscription Matching for Location-Aware Publish/Subscribe*. In the 14<sup>th</sup> Intl. Symposium on Spatial and Temporal Databases (**SSTD 2015**), Seoul, South Korea, August 2015.
68. T. Emrich, K. A. Schmid, A. Zuefle, M. Renz, and **R. Cheng**. *Uncertain Voronoi Cell Computation based on Space Decomposition*. In the 14<sup>th</sup> Intl. Symposium on Spatial and Temporal Databases (**SSTD 2015**), Seoul, South Korea, August 2015. (*Selected as one of the best papers for nomination to publication in Geoinformatica.*)
69. **Y. Zheng\***, J. Wang, G. Li, **R. Cheng**, and J. Feng. *QASCA: A Quality-Aware Task Assignment System for Crowdsourcing Applications*. In ACM SIGMOD Intl. Conf. on Management of Data (**SIGMOD 2015**), Melbourne, Victoria, Australia, May 2015.
70. **C. Meng\***, **R. Cheng**, **S. Maniu\***, P. Senellart, and **W. Zhang\***. *Discovering Meta-Paths in Large Heterogeneous Information Networks*. In the 24<sup>th</sup> Intl. World Wide Web Conf. (**WWW 2015**), Florence, Italy, May 2015. *Acceptance rate: 14.1% (131 out of 929 papers).*
71. **Y. Zheng\***, **R. Cheng**, **L. Mo\***, and **S. Maniu\***. *On Optimality of Jury Selection in Crowdsourcing*. In the 18<sup>th</sup> Intl. Conf. on Extending Database Technology (**EDBT 2015**), Brussels, Belgium, Mar 2015.
72. **S. Maniu\***, **R. Cheng**, and P. Senellart. *ProbTree: A Query-Efficient Representation of Probabilistic Graphs*. In the 1<sup>st</sup> Intl. Workshop on Big Uncertain Data (**BUDA**), associated with **SIGMOD**, Utah, US, Jun 2014.
73. **W. Zhang\***, **R. Cheng**, and B. Kao. *Evaluating Multi-way Joins over Discounted Hitting Time*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2014**), Chicago, US, Apr 2014.
74. **S. Lei\***, **X. Yang\***, **L. Mo\***, **S. Maniu\***, and **R. Cheng**. *iTag: Incentive-Based Tagging*. Demo paper, in Intl. Conf. on Data Engineering (**IEEE ICDE 2014**), Chicago, US, Apr 2014.
75. C. Ren, **L. Mo\***, B. Kao, **R. Cheng**, and D. Cheung. *CLUDE: An Efficient Algorithm for LU Decomposition over a Sequence of Evolving Graphs*. In the 17<sup>th</sup> Intl. Conf. on Extending Database Technology (**EDBT 2014**), Athens, Greece, Mar 2014. *Acceptance rate: 20%.*

76. Y. Sun, H. Xu, and **R. Cheng**. *Privacy Preserving Path Recommendation for Moving User on Location Based Service*. In the 10<sup>th</sup> IEEE Intl. Conference on Ubiquitous Intelligence and Computing (**UIC 2013**), Italy, Dec 2013.
77. **L. Mo\***, **R. Cheng**, B. Kao, **X. Yang\***, C. Ren, **S. Lei\***, and **E. Lo\***. *Optimizing Plurality for Human Intelligence Tasks*. In the 22<sup>nd</sup> ACM Conf. on Information and Knowledge Management (**ACM CIKM 2013**), San Francisco, Oct 27 – Nov 1, 2013. *Acceptance rate: 16.9%*.
78. Z. He, P. Wong, B. Kao, E. Lo, and **R. Cheng\***. *Fast Evaluation of Iceberg Pattern-Based Aggregate Queries*. In the 22<sup>nd</sup> ACM Conf. on Information and Knowledge Management (**ACM CIKM 2013**), San Francisco, Oct 27 – Nov 1, 2013. *Acceptance rate: 16.9%*.
79. **X. Yang\***, **R. Cheng**, **L. Mo\***, B. Kao, and D. Cheung. *On Incentive-based Tagging*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2013**), Brisbane, Apr 2013.
80. **L. Mo\***, **R. Cheng**, **X. Li\***, D. Cheung, and **X. Yang\***. *Cleaning Uncertain Data for Top-k Queries*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2013**), Brisbane, Apr 2013.
81. **P. Zhang\***, **R. Cheng**, N. Mamoulis, M. Renz, A. Zuefle, **Y. Tang\***, and T. Emrich. *Voronoi-based Nearest Neighbor Search for Multi-Dimensional Uncertain Databases*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2013**), Brisbane, Apr 2013.
82. **Y. Jin\***, **R. Cheng**, B. Kao, K. Y. Lam, and **Y. Zhang\***. *A Filter-based Protocol for Continuous Queries over Imprecise Location Data*. In the 21<sup>st</sup> ACM Conf. on Information and Knowledge Management (**ACM CIKM 2012**), Hawaii, Oct 2012. *Acceptance rate: 13.4%*.
83. R. Li, B. Kao, B. Bi, **R. Cheng**, and E. Lo. *DRQ: A Probabilistic Approach to Diversified Query Recommendation*. In the 21<sup>st</sup> ACM Conf. on Information and Knowledge Management (**ACM CIKM 2012**), Hawaii, Oct 2012. *Acceptance rate: 13.4%*.
84. R. Tang, **R. Cheng**, H. Wu, and S. Bressan. *A Framework for Conditioning Uncertain Relational Data*. In the 23<sup>rd</sup> Conf. on Database and Expert Systems Applications (**DEXA 2012**), Austria, September 2012.
85. **X. Xie\***, **R. Cheng**, and M. L. Yiu. *Evaluating Trajectory Queries Over Imprecise Location Data*. In the 24th Intl. Conf. on Scientific and Statistical Database Management (**SSDBM 2012**), Greece, June 2012. *Acceptance rate: 51.6%*.
86. **J. Gong\***, **R. Cheng**, D. W. Cheung, and J. Cheng. *Evaluating Probabilistic Queries over Uncertain Schema Matching*. In the IEEE Intl. Conf. Data Engineering (**IEEE ICDE 2012**), Washington D.C., April 2012.
87. C. K. Chui, **R. Cheng**, B. Kao, and E. Lo. *I/O Efficient Algorithms for Evaluating Pattern-based Aggregate Queries in a Sequence OLAP System*. In the 20<sup>th</sup> ACM Conf. on Information and Knowledge Management (**ACM CIKM 2011**), Glasgow, Oct 2011. *Acceptance rate: 15%*.
88. B. Bi, S. D. Lee, B. Kao, and **R. Cheng**. *CubeLSI: An Effective and Efficient Method for Searching Resources in Social Tagging Systems*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2011**), Hannover, Germany, Apr, 2011.
89. **L. Sun\***, **R. Cheng**, D. W. Cheung, and **J. Cheng\***. *Mining Uncertain Data with Probabilistic Guarantees*. In the 16<sup>th</sup> ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (**ACM SIGKDD 2010**), Washington D.C., USA, Jul, 2010 (Full paper). *Acceptance rate: 17%*.
90. L. Wang, D. W. Cheung, **R. Cheng**, and S. D. Lee. *Accelerating Probabilistic Frequent Itemset Mining: A Model-Based Approach*. In the ACM 19<sup>th</sup> Conf. on Information and Knowledge Management (**ACM CIKM 2010**), Toronto, Canada, Oct 2010. **Selected as one of the best papers in CIKM'10**. *Acceptance rate: 13.4%*.

91. Y. Zhang\*, **R. Cheng**, and **J. Chen\***. *Evaluating Continuous Probabilistic Queries over Imprecise Sensor Data*. In the Database Systems for Advanced Applications (**DASFAA**), Apr, 2010 (Full paper). *Acceptance rate: 23.2%*.
92. **J. Cheng\***, J. X. Yu, and **R. Cheng**. *On-Line Preferential Nearest Neighbor Browsing in Large Attributed Graphs (Invited paper)*, 1<sup>st</sup> Intl. Workshop on Graph Data Management: Techniques and Applications (**GDM 2010**), associated with **DASFAA**, Tsukuba, Japan, Apr 2010.
93. **R. Cheng**, **X. Xie\***, M. L. Yiu, **J. Chen\***, and **L. Sun\***. *UV-diagram: A Voronoi Diagram for Uncertain Data*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2010**), Long Beach, USA, Mar, 2010 (Full paper). *Acceptance rate: 12.5%*.
94. **R. Cheng**, **J. Gong\***, and D. Cheung. *Managing Uncertainty of XML Schema Matching*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2010**), Long Beach, USA, Mar, 2010 (Full paper). *Acceptance rate: 12.5%*.
95. J. Ren, S. D. Lee, X. Chen, B. Kao, **R. Cheng**, and D. Cheung. *Naïve Bayes Classification of Uncertain Data*. In the IEEE Intl. Conf. on Data Mining (**IEEE ICDM 2009**), Miami, USA, Dec, 2009. *Acceptance rate: 17.8%*.
96. Z. Zhang, **R. Cheng**, D. Papadias, and A. Tung. *Minimizing the Communication Cost for Continuous Skyline Maintenance*. In Proc. ACM Conf. on Management of Data (**SIGMOD 2009**), Providence, RI, USA, July 2009. *Acceptance rate: 15.9%*.
97. **R. Cheng**, L. Chen, **J. Chen\***, and **X. Xie\***. *Evaluating Probability Threshold  $k$ -Nearest-Neighbor Queries over Uncertain Data*. In the 12<sup>th</sup> Intl. Conf. on Extending Database Technology (**EDBT 2009**), St. Petersburg, Russia, Mar 2009. (*Acceptance rate: 32.5%; cited 13 times*)
98. **R. Cheng**. *Querying and Cleaning Uncertain Data (Invited Paper)*. In the 1<sup>st</sup> Intl. Workshop on Quality of Context (**QuaCon**), LNCS, Springer, Stuttgart, Germany, June 2009.
99. **J. Chen\***, **R. Cheng**, **Y. Zhang**, and **J. Jian**. *A Probabilistic Filter Protocol for Continuous Queries*. In the 1<sup>st</sup> Intl. Workshop on Quality of Context (**QuaCon**), LNCS, Springer, Stuttgart, Germany, June 2009.
100. D. Lin, E. Bertino, **R. Cheng**, and S. Prabhakar. *Position Transformation: A Location Privacy Protection Method for Moving Objects*. In SIGSPATIAL ACM GIS 2008 Intl. Workshop on Security and Privacy in GIS and LBS (**SPRINGL 2008**), Irvine, CA, USA, November 2008.
101. **J. Chen\*** and **R. Cheng**. *Quality-Aware Probing of Uncertain Data with Resource Constraints*. In 20th Intl. Conf. on Scientific and Statistical Database Management (**SSDBM 2008**), Hong Kong, Jul 2008. *Acceptance rate: 34.5%*.
102. **R. Cheng**, **J. Chen\***, M. Mokbel, and C. Chow. *Probabilistic Verifiers: Evaluating Constrained Probabilistic Nearest-Neighbor Queries over Uncertain Data*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2008**), Cancun, Mexico, Apr, 2008 (Full paper, full talk). (*Acceptance rate: 12.1%; cited 62 times*)
103. S. Singh, R. Shah, S. Prabhakar, S. Hambrusch, C. Mayfield, J. Neville, and **R. Cheng**. *Database Support for Probabilistic Attributes and Tuples*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2008**), Cancun, Mexico, Apr, 2008 (Full paper, short talk). (*Acceptance rate: 7.1%; cited 29 times*)
104. **R. Cheng**, D. Yau, and J. Fu (**my research assistant**). *Packet Cloaking: Protecting Receiver Privacy Against Traffic Analysis*. In the 3<sup>rd</sup> Workshop on Secure Network Protocols (**NPsec**), with the 15<sup>th</sup> IEEE Intl. Conf. on Network Protocols (**IEEE ICNP 2007**), Beijing, China, Oct 2007.
105. S. Lee, B. Kao, and **R. Cheng**. *Reducing  $UK$ -means to  $K$ -means*. In the 1<sup>st</sup> Workshop on Data Mining of Uncertain Data (**DUNE**), co-located with the IEEE Conf. on Data Mining (**IEEE ICDM 2007**), USA, Oct, 2007. (*cited 20 times*)



106. T. Farrell, **R. Cheng**, and K. Rothermel. *Energy-Efficient Monitoring of Mobile Objects with Uncertainty-Aware Tolerances*. In the 11<sup>th</sup> Intl. Database Engineering & Applications Symposium (**IDEAS 2007**), Banff, Canada, Sept, 2007. (*Acceptance rate: 31.5%*).
107. Y. Liu and **R. Cheng**. *Intelligent Tutoring System Based on Semantic Multimedia Browsing and Retrieval*. In the 5<sup>th</sup> Intl. Conf. on Intelligent Multimedia & Ambient Intelligence (**IMAI 2007**), Information Sciences, 10<sup>th</sup> Joint Conference, Salt Lake City, Utah, USA, July 2007.
108. **J. Chen\*** and **R. Cheng**. *Efficient Evaluation of Imprecise Location-Dependent Queries*. In the IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2007**), Istanbul, Turkey, Apr, 2007. (*Acceptance rate: 18.5%; cited 54 times*)
109. J. Ngai, B. Kao, C. Chui, **R. Cheng**, M. Chau, and K. Yip. *Efficient Clustering of Uncertain Data*. In the IEEE Intl. Conf. on Data Mining (**IEEE ICDM 2006**), Hong Kong, Dec, 2006. (*Acceptance rate: 10%; cited 65 times*)
110. **R. Cheng**, S. Singh, S. Prabhakar, R. Shah, J. Vitter, and Y. Xia. *Efficient Join Processing over Uncertain Data*. In the ACM 15<sup>th</sup> Conf. on Information and Knowledge Management (**ACM CIKM 2006**), Arlington, USA, Nov 2006. (*Acceptance rate: 15%; cited 35 times*)
111. **R. Cheng**, Y. Zhang, E. Bertino, and S. Prabhakar. *Preserving User Location Privacy in Mobile Data Management Infrastructures*. In the Lecture Notes in Computer Science (LNCS), Privacy Enhancing Technology Workshop (**PET 2006**), Cambridge, UK, June 2006, pp. 393-412. (*Acceptance rate: 26%; cited 94 times*)
112. M. Chau, **R. Cheng**, B. Kao, and J. Ng. *Uncertain Data Mining: An Example in Clustering Location Data*. In the Methodologies for Knowledge Discovery and Data Mining, Pacific-Asia Conference (**PAKDD 2006**), Singapore, April 2006, pp. 199-204. (*Acceptance rate: 13.4%; cited 42 times*)
113. M. Chau, **R. Cheng**, and B. Kao. *Uncertain Data Mining: A New Research Direction*. Invited Paper, in the Workshop on the Sciences of The Artificial (**WSA**) 2005, National Dong Hwa University, Taiwan, Dec 2005.
114. **R. Cheng**, B. Kao, S. Prabhakar, A. Kwan, and Y. Tu. *Adaptive Stream Filters for Entity-based Queries with Non-value Tolerance*. In Very Large Databases Conf. (**VLDB 2005**), Norway, Aug 2005, pp. 37-48. (*Acceptance rate: 16.5%; cited 50 times*)
115. Y. Tao, **R. Cheng**, X. Xiao, W. Ngai, B. Kao, and S. Prabhakar. *Indexing Multi-Dimensional Uncertain Data with Arbitrary Probability Density Functions*. In Very Large Databases Conf. (**VLDB 2005**), Norway, Aug 2005, pp. 922-933. (*Acceptance rate: 16.5%; cited 148 times*)
116. **R. Cheng**, S. Singh, and S. Prabhakar. *U-DBMS: A Database System for Managing Constantly-Evolving Data (Software Demonstration)*. In Very Large Databases Conf. (**VLDB 2005**), Norway, Aug 2005, pp.1271-1274. (*Acceptance Rate 29/69=42%; cited 59 times*)
117. S. Han, E. Chan, **R. Cheng**, and K.Y. Lam. *A Statistics-Based Sensor Selection Scheme for Continuous Probabilistic Queries in Sensor Networks*. In 11th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (**RTCSA 2005**), Hong Kong, Aug 2005, pp. 331-336.
118. **R. Cheng**, Y. Xia, S. Prabhakar, and R. Shah. *Change Tolerant Indexing on Constantly Evolving Data*. In Intl. Conf. on Data Engineering (**IEEE ICDE 2005**), Tokyo, Japan, Apr 2005. (*Acceptance rate: 12.9%, 67/521; cited 24 times*)
119. Y. Xia, S. Prabhakar, S. Lei, **R. Cheng**, and R. Shah. *Indexing Continuously Changing Data with Mean-Variance Tree*. In the 20th Annual ACM Symposium on Applied Computing (**ACM SAC 2005**), Mar 2005. *Acceptance rate:30%*.
120. K. Y. Lam, **R. Cheng**, B. Liang, and J. Chau. *Sensor Node Selection for Execution of Continuous Probabilistic Queries in Wireless Sensor Networks*. In ACM 2nd Intl. Workshop on Video Surveillance and Sensor Networks (**ACM VSSN 2004**), in conjunction with 12th ACM Intl Conf. on Multimedia, pp. 63-71, New York, Oct 2004.



121. **R. Cheng**, Y. Xia, S. Prabhakar, R. Shah, and J. S. Vitter. *Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data*. In Very Large Databases Conf. (**VLDB 2004**), pp. 876-887, Toronto, Canada, Sep 2004. (Acceptance rate: 16%; cited 165 times).
122. **R. Cheng** and S. Prabhakar. *Using Uncertainty to Provide Privacy-Preserving and High-Quality Location-Based Service*. In the **Mobile HCI 2004** workshop on Location Systems Privacy and Control, Glasgow, Sep 2004.
123. **R. Cheng**, D. Kalashnikov and S. Prabhakar. *Evaluating Probabilistic Queries over Imprecise Data*. In Proc. of the ACM Special Interest Group on Management of Data (**ACM SIGMOD 2003**), pp. 551-562, June 2003. (Acceptance rate: 15.2%, 52/342; cited 368 times).
124. **R. Cheng**, S. Prabhakar, and D. Kalashnikov. *Querying Imprecise Data in Moving Object Environments*. In Proc. of the Intl. Conf. on Data Engineering (**IEEE ICDE 2003**), pp. 723-725, Bangalore, India, March 2003. Acceptance rate: 13.5%.
125. **R. Cheng**. *Updates and View Maintenance in Soft Real-Time Database Systems*. In the 2nd ACM Hong Kong Postgraduate Research Day (**Best Paper Award**), University of Hong Kong, 1999.
126. B. Kao, K.Y. Lam, B. Adelberg, **R. Cheng** and T. Lee. *Updates and View Maintenance in Soft Real-Time Database Systems*. In the Intl. Conf. on Information and Knowledge Management (**ACM CIKM 1999**), pp. 300-307, Kansas City, Missouri, USA, Nov 1999. (Acceptance rate: 39%; cited 17 times)
127. C. Yip, K. Loo, B. Kao, D. Cheung, and **R. Cheng**. *LGen - A Lattice-Based Candidate Set Generation Algorithm for I/O Efficient Association Rule Mining*. In Methodologies for Knowledge Discovery and Data Mining, Third Pacific-Asia Conference (**PAKDD 1999**), pp. 54-63, Beijing, Apr 1999. Acceptance Rate: 18.3%.

### **Knowledge Exchange Articles**

128. **R. Cheng**. Scalable Continuous Query Processing on Imprecise Location Data (移動位置資料庫的連續性查詢處理). Research Frontiers, the Research Grants Council (RGC) of the University Grants Committee (UGC), Issue 23, Oct 2012. URL: <http://www.ugc.edu.hk/rgc>

### **Software**

129. C. Mayfield, S. Singh, **R. Cheng** and S. Prabhakar. *ORION: A Database System for Managing Uncertain Data*, version 0.1, March 2006, copyrighted by Purdue University. URL: <http://orion.cs.purdue.edu>

9-Aug-17