

# Dr. Reynold C. K. Cheng

Associate Professor, Department of Computer Science, the University of Hong Kong

Department of Computer Science  
The University of Hong Kong  
Pokfulam Road, Hong Kong  
Tel: (852) 2219-4778  
Fax: (852) 2559-8447  
URL: <http://www.cs.hku.hk/~ckcheng/>  
Email: [ckcheng@cs.hku.hk](mailto:ckcheng@cs.hku.hk)

## Research Interests

Database and Uncertainty Management

## Education

### **Purdue University**

*West Lafayette, Indiana, USA*

Ph.D., Computer Science, Sep 2001 - May 2005

- *Dissertation Topic:* “Managing Uncertainty in Constantly-Evolving Environments”  
In database systems that monitor continuously-changing entities like temperature values and locations of moving objects, data are obtained from sensors continuously and streamed to the database. Due to limited bandwidth and battery power, it is infeasible for the system to keep track of the actual values of the entities. Queries that use these database values can produce incorrect answers. This dissertation studies uncertainty inherent to dynamic data, based on which *probabilistic queries* – query answers augmented with probabilistic guarantees – are proposed. A classification scheme of queries is presented, where we study evaluation of probabilistic queries and quality of probabilistic answers. We also investigate how to improve the computational and I/O performance of probabilistic queries. As a case study, we apply the idea of uncertainty handling to moving-object databases. We also study examine applications of probabilistic queries, including selection of sensors in a sensor network and control of location privacy.
- *Advisor:* Sunil Prabhakar

M.S., Computer Science, Sep 2001 - Dec 2003

### **University of Hong Kong**

*Hong Kong, China*

M.Phil., Computer Science, Sep 1998 - Sep 2000

- *Dissertation Topic:* “View Updates and Temporal Correctness in Real-Time Databases”  
This thesis is about the design of a set of concurrency control protocols, applied to a real-time database environment handling update streams and derived data. The proposed methods not only guarantee serializability, but also consider the trade-off between data freshness and transaction correctness.
- *Advisor:* Ben C. M. Kao

B. Eng., Computer Engineering, Sep 1995 - Jun 1998 (1<sup>st</sup> Class Honor)

- *Dissertation Topic:* “High Performance Generic Search Engine”  
This is a team project where we implemented a C toolbox that allows programmers to customize a search engine with minimal effort. It supports various data types and languages. I was the project leader, where my main role was to coordinate among the five members, and assure that the software development process adhered to strict software engineering disciplines. The system was demonstrated to IBM in Hong Kong.
- *Advisor:* David W. L. Cheung

## **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.

### **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. Ferragine(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. **S. Luo\***, B. Kao, G. Li, **J. Hu\***, **R. Cheng**, and **Y. Zheng\***. *TOAIN: A Throughput Optimizing Adaptive Index for Answering Dynamic kNN Queries on Road Networks*. In Proceedings of the VLDB Endowment (**PVLDB**). To be presented in the Very Large Databases Conf. (**VLDB 2018**), Rio De Janeiro, Brazil, Aug 27-31, 2018.

16. **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**), 55(2), pp. 305-331, 2018.
17. **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.
18. **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.
19. **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.
20. **Y. Fang\***, **R. Cheng**, **S. Luo\***, **J. Hu\***, and **K. Huang\***. *C-Explorer: Browsing Communities in Large Graphs*, 10(11), pp. 1885-1888, Aug 2017. Also presented in the 43<sup>rd</sup> Intl. Conf. on Very Large Data Bases (**VLDB**), Munich, Germany, August 2017.
21. **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.
22. **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.
23. 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.
24. 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.
25. 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, 29(2), pp. 286-299, Feb 2017.
26. **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.
27. Z. Li, **Y. Fang\***, Q. Liu, J. Cheng, **R. Cheng**, and J. C.S. Lui. *Walking in the Cloud: 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.
28. 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.
29. **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.
30. 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.

31. **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.
32. **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.
33. **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.
34. **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.
35. **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.
36. 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.
37. 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)*.
38. **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.
39. **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.
40. 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.
41. 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*)
42. 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*)
43. **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.*
44. **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*)
45. **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)
46. 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>
  47. **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*)
  48. 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.
  49. 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*)
  50. **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*)
  51. **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*)
  52. 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*)
  53. **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*)
  54. **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*)
  55. 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*)
  56. 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**

57. **Y. Fang\***, **R. Cheng**, G. Cong, N. Mamoulis, and Y. Li. *On Spatial Pattern Matching*. In the 34<sup>th</sup> IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2018**), Paris, France, Apr 2018.
58. **Y. Fang\***, **R. Cheng**, **J. Wang\***, **Budiman\***, G. Cong, and N. Mamoulis. *SpaceKey: Exploring Patterns in Spatial Databases (Demo)*. In the 34<sup>th</sup> IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2018**), Paris, France, Apr 2018.

59. **C. Shan\***, N. Mamoulis, G. Li, **R. Cheng**, **Z. Huang\***, and **Y. Zheng\***. *T-Crowd: Effective Crowdsourcing for Tabula Data*. In the 34<sup>th</sup> IEEE Intl. Conf. on Data Engineering (**IEEE ICDE 2018**), Paris, France, Apr 2018.
60. **X. Li\***, **R. Cheng**, **Y. Fang\***, **J. Hu\***, and S. Maniu. *Scalable Evaluation of k-NN Queries on Large Uncertain Graphs*. In the 21<sup>st</sup> Intl. Conf. on Extending Database Technology (**EDBT 2018**), Vienna, Austria, Mar 2018.
61. **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)*.
62. **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)*.
63. **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.
64. **Y. Fang\*** and **R. Cheng**. *On Attributed Community Search*. In **MATES@VLDB**, 2017, pp. 1-21.
65. **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.
66. **Z. Huang\***, B. Cautis, **R. Cheng**, and **Y. Zheng\***. *KB-Enabled Query Recommendation for Long-Tail Queries*. In **CIKM 2016**, Indianapolis, October 2016.
67. **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.
68. 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.
69. **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.
70. 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.
71. 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.
72. **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.
73. **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.
74. 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.*)

75. **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.
76. **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)*.
77. **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.
78. **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.
79. **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.
80. **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.
81. 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%*.
82. 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.
83. **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%*.
84. 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%*.
85. **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.
86. **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.
87. **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.
88. **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%*.
89. R. Li, B. Kao, B. Bi, **R. Cheng**, and E. Lo. *DQR: 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%*.
90. 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.

91. **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%*.
92. **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.
93. 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%*.
94. 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.
95. **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%*.
96. 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%*.
97. 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%*.
98. **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.
99. **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%*.
100. **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%*.
101. 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%*.
102. 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%*.
103. **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*)
104. **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.
105. **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.



106. 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.
107. **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%*.
108. **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*)
109. 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*)
110. **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.
111. 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*)
112. 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%*).
113. 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.
114. **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*)
115. 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*)
116. **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*)
117. **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*)
118. 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*)
119. 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.
120. **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*)

121. 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)
122. **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)
123. 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.
124. **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)
125. 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%.
126. 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.
127. **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).
128. **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.
129. **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).
130. **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%.
131. **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.
132. 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)
133. 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**

134. **R. Cheng** and **Y. Zheng\***. Crowdsourcing: Managing Crowd Wisdom in the Age of Big Data. Invited article for the Faculty of Engineering eNews. Submitted in September 2015.

135. **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

136. 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>

## Tutorials

1. **R. Cheng**. *Meta Paths and Meta Structures: Analyzing Large Heterogeneous Information Networks*. In **APWeb-WAIM Conference**, Beijing, 9<sup>th</sup> July, 2017.
2. G. Li, **Y. Zheng\***, J. Fan, J. Wang, and **R. Cheng**. *Crowdsourced Data Management: Overview and Challenges*. In **SIGMOD Conference 2017**, pp. 1711-1716.
3. **R. Cheng**. *Managing the Quality of Crowdsourced Databases*. In The 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining (**PAKDD-2014**), Tainan, Taiwan, 13-16 May, 2014.
4. **R. Cheng**, T. Emrich, H. Kriegel, N. Mamoulis, M. Renz, G. Trajcevski, and A. Zuefle. *Managing Uncertainty in Spatial and Spatio-temporal Data*. In Intl. Conf. on Data Engineering (**ICDE 2014**), Chicago, March 31-April 4, 2014.
5. M. Renz, **R. Cheng**, and H. Kriegel. *Similarity Search and Mining in Uncertain Databases*. In Very Large Databases Conf. (**VLDB 2010**), Singapore, 13-17 September, 2010.

## Research Projects

### As a Principal Investigator (PI):

#### [External grants]

1. **PI** (Co-I: B. Cautis and S. Maniu) – *Discovering and Querying Meta-Graphs in Large Heterogeneous Information Networks* (RGC **GRF**, Ref: 17229116, 2016-18). Amount: HKD 675,647.
2. **PI** (Co-I: W. Fan and P. Senellart) – *Efficient Query Algorithms for Uncertain Graph Databases* (RGC **GRF**, Ref: 17205115, 2015-17). Amount: HKD 462,528.
3. **PI** (Co-I: K. Rothermel and M. L. Yiu) – *UV-Diagram: A Voronoi Diagram for Uncertain Spatial Databases* (RGC **GRF**, Ref: 711110, 2011-13). Amount: HKD 746,400.
4. **PI** – Universitas 21 Fellowship 2011-12 (for visiting Prof. Anthony Tung, School of Computing, the National University of Singapore in July-August 2011. Amount: HKD 50,000. (*Completed*))
5. **PI** (Co-I: K. Rothermel and D. Olteanu) – *Scalable Cleaning of Probabilistic Databases with Quality Guarantees* (RGC **GRF**, Ref: 711309E, 2010-12). Amount: HKD 696,500 (*Completed*).
6. **PI** (Co-I: K. Rothermel) – *Scalable Continuous Query Processing on Imprecise Location Data* (RGC **GRF**, Ref: 513508, 2009-11). Amount: HKD 645,950 (*Completed*).
7. **PI** (Co-I: K. Rothermel) – *Adaptive Filters for Continuous Queries over Constantly-Evolving Data Streams* (RGC **CERG**, Ref: 513307, 2008-10). Amount: HKD 391,512 (*Completed*).
8. **PI** (Co-I: E. Bertino and S. Prabhakar) – *Privacy Protection in Location-based Services with Location Cloaking* (RGC **CERG**, Ref: 513806, 2007-09). Amount: HKD 356,000 (*Completed*).
9. **PI** (project partner: K. Rothermel) – *Efficient Protocols for Quality-Aware Querying of Sensor Data in Pervasive Environments*, RGC Germany/Hong Kong Joint Research Scheme 2006/2007. Ref: G\_HK013/06, Amount: HKD 59,600 (*Completed*)

## **[Internal grants]**

10. **PI** – *Managing Uncertainty of Heterogeneous Big Multimedia Data* (Seed Funding Programme for Basic Research), HKU. Ref: 104004572, 2017-18). Amount: HKD 44,320.
11. **PI** – *Managing Quality of Big Traffic Data*. HKU 37<sup>th</sup> Round PDF/RAP Scheme (for recruitment of a postdoctoral researcher in years 2017-18).
12. **PI** – *Scalable, Effective and Privacy-Aware Location-based Services for Big Traffic Data* (Seed Funding for Incubating Group-based Collaborative Research - CRF), HKU. Ref: 102009508, 2016-18). Amount: HKD 100,000.
13. **PI** – *LAPS: A Scalable Location-Aware Publish/Subscribe System* (Seed Funding Programme for Basic Research), HKU. Ref: 104004129, 2016-17). Amount: HKD 34,490.
14. **PI** – *Managing Crowdsourced Databases: A Probabilistic Approach* (Seed Funding Programme for Basic Research), HKU. Ref: 201411159171, 2015-16). Amount: HKD 44,925.
15. **PI** – *Effective Noise Removal Algorithms for Ambiguous Databases* (Seed Funding Programme for Basic Research), HKU. Ref: 201311159095, 2014-15). Amount: HKD 47,900.
16. **PI** – *Query-Efficient Representations of Uncertain Graphs* (Seed Funding Programme for Basic Research), HKU. Ref: 201211159083, 2013-14). Amount: HKD 154,200.
17. **PI** – *Travel Grants* for University Academic Staff to Visit Universities in Taiwan (National Chiao Tung University, Prof. Wen-Chih Peng, July 28 – August 4, 2013, with an award amount of HKD 9,600)
18. **PI** – *Management of Large-Scale Uncertain Databases*. HKU Engineering Postdoctoral Fellow Programme (for recruitment of a postdoctoral researcher in years 2010-12).
19. **PI** – *Probing Imprecise Sensor Data with Quality Guarantees* (Seed Funding Programme for Basic Research), HKU. Ref: 200808159002, 2008-10). Amount: HKD 240,000. (*Completed*)
20. **PI** (Co-I: D. Yau) - *Efficiency of Privacy Preservation Mechanisms in Routing over the Internet* (Internal Research Grant (**ICRG**), PolyU. Ref: A-PH09, 2006-08). Amount: HKD 120,000. (*Completed*)
21. **PI** (Co-I: D. Yau) – *Protecting Network Privacy with Spatial and Temporal Cloaking* (Internal Research Grant (**ICRG**), PolyU. Ref: A-PH39, 2007-08). Amount: HKD 120,000. (*Completed*)

## **As a Co-Investigator (co-I):**

22. **Co-I** (PI: N. Mamoulis) – *Query Suggestion for Geo-Textual Data* (RGC CERG. Ref: 17205015, 2016-18). Amount: HKD 496,028.
23. **Co-I** (PI: Lo Chi Lik Eric) – *Online Analytical Processing on Big Sequence Data* (RGC CERG. Ref: 521012E, 2012-13). Amount: HKD 700,000.
24. **Co-I** (PI: J. Cheng (SIAT)) – *Fast Algorithms for Analyzing Large Uncertain Graphs* (NSF China, Project for Young Scientists Fund, 2012-14. Amount: RMB 220,000.
25. **Co-I** (PI: W. Wu (SYSU)) -- *Cooperative Caching in Wireless Mesh Networks* (NSF China, Project for Young Scientists Fund, General Program, 2009-11. Amount: RMB 200,000.
26. **Co-I** (PI: D. Wang) – *Proactive Data Sweeping and Adaptive Communication for Efficient Data Collection in Sensor Networks* (RGC CERG. Ref: 530508, 2009-10). Amount: HKD 346,060.
27. **Co-I** (PI: Y. Tao) – *Query Processing on Historical Uncertain Spatiotemporal Data* (RGC CERG. Ref: 120206, 2006-07). Amount: HKD 961,920. (*Completed*)
28. **Co-I** (PI: Y. Liu) – *Hierarchical Multimedia Database for Semantic-based Multimedia Retrieval* (Internal Research Grant, PolyU, 2006-08. Ref: A-PH11). Amount: HKD 120,000.
29. **Affiliated Member** (PI: J. Ng) – *A Research Center for Ubiquitous Computing* (Central Allocation Group Research Projects, RGC, 2006-09, HKBU 1/05C)
30. **Member**, StruFus: Infrastructure for Information Fusion (joint project with U. Skövde, HKU, HKBU, IIT Bombay, U. Wuhan, 2008-09)

## **Teaching-Related Projects**

1. **PI** – *The Development of the Common Core Course “The Age of Big Data”* (Teaching Development Grant (TDG, HKU, 2015. Ref: **101000560**). Amount: **HKD 50,000**).

2. **Course Coordinator** (teacher: Dr. Ben Kao, Dr. Nikos Mamoulis, Dr. W.S. Ho) – *Handling Information in Everyday Life*, **University Common Core Course Stage I teaching proposal**, HKU, selected to proceed to Stage II under the AoI of “Scientific & Technological Literacy”. *Acceptance rate: 59.4%, 57/96, 2010.*
3. **Member** (PC: C. K. Chui) – *Kids can code* (Knowledge Exchange (KE) Funding Exercise, HKU, 2015-16, KE-IP-2015/16-35). Amount: HKD 100,000.
4. **Co-I** (PI: M. Wang) – *Support Self-Directed Learning in Interdisciplinary Studies by Developing a Knowledge Visualization and Navigation System* (Teaching Development Grant (TDG), HKU, 2009-11). Amount: HKD 268,240.
5. **Co-I** (PI: G. Ngai) – *PANDA: A Multi-User Platform for Teaching Computer Programming* (Learning and Teaching Development Grants, PolyU, 2006-07). Amount: HKD 100,000. (*Completed*)

### **Current Students**

- Chenhao Ma (PhD candidate, HKU, 2017 – current)
- Zichen Zhu (MPhil candidate, HKU, 2017 – current)
- Xiaodong Li (PhD candidate, HKU, 2016 – current)
- Nathan Yan (MPhil candidate, HKU, 2016 – current)
- Zhipeng Huang (PhD candidate, HKU, 2015 – current, co-supervised with Prof. Nikos Mamoulis)
- Caihua Shan (PhD candidate, HKU, 2015 – current, co-supervised with Prof. Nikos Mamoulis)
- Siqiang Luo (PhD candidate, HKU, 2015 – current, co-supervised with Prof. Ben Kao)
- Haiqi Sun (MPhil candidate, HKU, 2015 – current)
- Jiafeng Hu (PhD candidate, HKU, 2014 – current)
- Yixiang Fang (PhD candidate, HKU, 2013 – current)
- Yudian Zheng (PhD candidate, HKU, 2013 – current)
- Yanpeng Yang (PhD candidate, HKU, 2014 – current, co-supervised with Dr. B. Oliveira)

### **Graduated Students**

- Jinchuan Chen (PhD, HKPU, Jan 2006 – May 2009; co-supervised with Prof. Jiannong Cao; now associate professor in Renmin University of China. Thesis title: Querying and Cleaning Uncertain Data.)
- Yinuo Zhang (M. Phil., HKU, 2008 – Dec 2010; now PhD candidate in the University of Southern California (USC). Thesis title: Evaluating Continuous Probabilistic Queries over Sensor and Location Data.)
- Liwen Sun (M. Phil., HKU, 2008 – Dec 2010, co-supervised with Prof. D. Cheung; now PhD candidate in UC Berkeley. Thesis title: Mining Uncertain Data with Probabilistic Guarantees.)
- Jim Gong (PhD, HKU, 2007 – Apr 2011, co-supervised with Prof. D. Cheung; now Associate Consultant of Calypso Technology Limited. Thesis title: Managing Uncertainty in Schema Matchings.)
- Xiang Li (M. Phil., HKU, 2009 – Feb 2012, co-supervised with Prof. D. Cheung; now in Merrill Lynch, Hong Kong. Thesis title: Managing Query Quality in Probabilistic Databases.)
- Xike Xie (PhD, HKU, Jul 2008 – Mar 2012; now professor in USTC (University of Science and Technology of China). Thesis title: Evaluating Nearest Neighbor Queries over Uncertain Databases.)
- Yifan Jin (M. Phil. HKU, 2010 – Jul 2012, co-supervised with Prof. B. Kao. Thesis title: A Filter-based Protocol for Continuous Queries over Imprecise Location Data.)
- Peiwu Zhang (M. Phil. HKU, Sep 2010 – Dec 2012, co-supervised with Prof. N. Mamoulis. Thesis title: Voronoi-based Nearest Neighbor Search for Multi-Dimensional Uncertain Databases.)
- Xuan Yang (PhD, HKU, Sep 2009 – Dec 2013, co-supervised with Prof. D. Cheung; now in MSRA. Thesis title: Budget-Limited Data Disambiguation.)
- Wangda Zhang (M. Phil. candidate, HKU, Sep 2011 – Jan 2014, co-supervised with Prof. B. Kao; now PhD candidate in Columbia University. Thesis title: Evaluating Multi-Way Joins over Discounted Hitting Time.)
- Changpeng Meng (M. Phil. candidate, HKU, Sep 2012 – Jul 2014; now PhD candidate in Purdue University. Thesis title: Discovering Meta-Paths in Large Knowledge Bases.)
- Siyu Lei (M. Phil. Candidate, Sep 2012 – Feb 2015. Thesis title: Online Influence Maximization.)
- Yu Tang (M. Phil. candidate, HKU, 2011 – 2013, co-supervised with Dr. N. Mamoulis. Thesis title: Similarity Search with Earth Mover’s Distance at Scale.)
- Luyi Mo (PhD candidate, HKU, 2011 – 2015, primary supervisor, co-supervised with Prof. D. Cheung. Thesis title: Cleaning Algorithms for Novel Applications.)
- Xuhui Jia (PhD candidate, HKU, 2011 – 2016, co-supervised with Dr. K. P. Chan)

- Yong Xu (MPhil, HKU, Sep 2014 – Jun 2017. Thesis title: Top-*k* SAS: On Reliable Retrieval of Top-*k* Tags. Now in Microsoft Research Asia (MSRA))

### **Research Assistants**

- Silviu Maniu (Ph.D., Telecom ParisTech, *Postdoctoral Research Fellow*: Oct 2012 – Oct 2014)
- Yuxiang Wu (CS, Sun Yat Sen University; *Summer Intern Student*: Jul – Aug 2013)
- Caihua Shan (CS, Zhejiang University; *Summer Intern Student*: Jul – Aug 2013)
- Xu Yong (CS, Huazhong University of Science and Technology; *Summer Intern Student*: Jul – Aug 2013)
- Bo Liang (CS, Zhejiang University; *Student Intern Student*: Jul – Aug 2013)
- Zewei Chu (BEng (CS); *Student Research Assistant*: May – Nov 2012)
- Jiefeng Cheng (Ph.D., CUHK, *Postdoctoral Research Fellow*: Apr 2010 – Aug 2011; *Senior Research Assistant*: Oct 2009 – Mar 2010)
- James Siu (BEng (CS); *Student Research Assistant*: Jul – Sep 2011)
- Du Wei (BBAIS; *Student Research Assistant*: Dec 2009 – Jan 2010; Jul – Aug 2010)
- Chan Chi Wai (BBAIS; *Student Research Assistant*: Jul – Aug 2010)
- Bryan Hui (BEng (CE); *Student Research Assistant*: Jul – Aug 2009)
- Li Man Ho (BEng (CS); *Student Research Assistant*: Jul – Aug 2009)
- Jinchuan Chen (Jan 2009 – May 2009)
- Jin Jian (Jan 2009 – Mar 2009)
- Jianyun Fu (Sept 2006 – Oct 2007, HKPU)
- Kitty Yuen (summer 2007, HKPU)

### **MSc Students**

- Namm Andres, Zhao Yanhui, Lin Yu, Liang Hao, and Fang Cheng (Project: A Large Incident Management System (IMS) for Intelligent City Development, HKU, 2017-18).
- Wang Zheng (Project: Continuous Spatial-aware Community Search, HKU, 2017-18).
- Chen Yankai (Project: Effective and Efficient Community Search for Large Attributed Graphs, HKU, 2017-18).
- Ng Kwong Hei and Wong Ka Yu (Project: Discovering and Querying Meta-graphs in Large Heterogeneous Information Networks, HKU, 2017-18).
- Guan Ruixiang, Ma Mutian, Qiang Shuya, Weng Lili and Wu Mengying, (Project: i-tag: Collaborative Tagging System, HKU, 2014-15, completed).
- Jeff Smith, Zuyao Wang, and Lu Liu (Project: An Uncertain column-oriented databases, HKU, 2012-13, completed)
- Clement Ha (Project: Location-based services with uncertain database support, HKU, 2008-09, completed)

### **Undergraduate Final Year Project Student Supervision**

- Xu Fangyuan and Wu You (2016-17)
- Lau Piu Lam, Lo Ming Fai, Vinny Wong, and Dongguang You (2015-16)
- Denny Lin and Juntao Huang (2013-14)
- James Siu and Yuen Chi Leung (2011-12)
- Du Wei and Chan Chi Wai (2010-11)
- Lau Kai Wai Elvis, Lau Si Kit Alex, Li Tang Tommy, Chau Kei Man (2010-11)
- Bryan Hui, Li Man Ho, Michael Lau, and Hao Chan (2009-10)
- Chan Hong Ming, Chong Wing Tong, Wong Kar Man, Chan Hong Lap, Sin Ho Lun, Yau Siu Long, Jenny Lam, Wong Ka Ho, Wong Ping Hang (HKPU, 2007-08)
- Alfred Cheng, Gary Ching, Terry Wong, Victor Yu, Kitty Yuen (HKPU, 2006-07)

### **Awards Won by Supervised Staff and Students**

- Yunfan Kang was awarded the HKU Undergraduate Research Fellow Programme, for doing a project on *community search*, supervised by Dr. Reynold Cheng, in June-August 2017
- Fung Yuet was awarded the HKU Undergraduate Research Fellow Programme, for doing a project on *meta-path analysis*, supervised by Dr. Reynold Cheng in June-August 2016

- Xu Fangyuan and Wu You won the “Best 50 Final Year Projects” by Hong Kong X-Tech Startup Platform, on “Mining HKUCS Graduate Student Data: Extraction, Analysis, and Prediction”, 2016-17.
- Lau Piu Lam, Lo Ming Fai, Vinny Wong, and Dongguang You won the Bronze Award in the China Pan-Pearl River Delta Region University IT Project Competition 2016.
- Yudian Zheng, my PhD student, won the Hong Kong and China Gas Company Limited Postgraduate Scholarship 2015-16.
- Yudian Zheng, my PhD student, won the SIGKDD 2016 Student Travel Grant.
- Zhipeng Huang, my PhD student, won the SIGKDD 2016 Student Travel Grant.
- Yudian Zheng, my PhD student, won the SIGMOD 2015 Travel Grant, \$400 AUD.
- Jason Meng, my MPhil student, won the WWW 2015 Student Travel Grant.
- Lau Pui Lam, Lo Ming Fai, Wong Vinny, and You Dongguang won the **Third runner-up** prize for the HKU Final-Year Project Competition, on *A Mobile and Intelligent Student Interview System for HKUCS*, 2015-16.
- Luyi Mo, my PhD student, won the Hong Kong and China Gas Company Limited Postgraduate Scholarship 2012-13.
- Luyi Mo, my PhD student, won the Hong Kong PhD Fellowship awarded by the Hong Kong RGC, 2011
- Lau Kai Wai Elvis, Lau Si Kit Alex, Li Tang Tommy, Chau Kei Man won the **Champion** prize for the HKU Final-Year Project Competition, on *An Intelligent Mobile Transport System*, 2010-11.
- Bryan Hui, Li Man Ho, Michael Lau, and Hao Chan won the **First runner-up prize** for the HKU Final-Year Project Competition, on *Energy-aware location-based services*, 2009-10
- The HKU programming team, named “Far” and **coached by Dr. Reynold Cheng**, obtained an honorable mention in the **World Finals of the ACM International Collegiate Programming Contest**, in Stockholm in April 2009. Our team is among the 100 teams in this prestigious competition, which are chosen from 7,109 teams from 1,838 universities in 88 countries on six continents.
- The HKU team, **coached by Dr. Reynold Cheng**, won the 2<sup>nd</sup> Runner-Up award in the IBM Inter-University Programming Contest 2009, which consists of 37 strong teams from universities in Hong Kong and Macau.
- Three HKU programming teams, coached **by Dr. Reynold Cheng**, won the Gold, Silver, and Bronze Awards in the ACM Collegiate Programming Regional Contest (Indonesia), Binus University, Indonesia, Oct 20-21, 2008. The teams were among the top ten, with “Far” obtaining the highest ranking as the 4th among 49 strong teams. HKU was also the 2nd runner-up university in this prestigious competition.
- Chan Hong Ming won the Merit Awards of the Charles Baggage Final Year Project Awards, HKPU, 2007/08
- Kitty Yuen won the champion of the Amway Pan-Pearl IT Project Competition, China, for the final year project in 2007. The goal of the project is design a two-dimensional uncertain database, based on the ORION database system that I developed earlier. The Competition is organized by Guizhou Computer Federation, and the Internet Professional Association is the co-organizer in Hong Kong Region. After a local competition held in Hong Kong, 4 teams from 9 local tertiary institutions are selected to represent Hong Kong in *final judging* held in Guiyang, Guizhou. The final judging involves outstanding students from Macau and 9 provinces in the Pan-PRD region: Guangdong, Guangxi, Fujian, Hunan, Hainan, Sichuan, Guizhou, Yunan and Jiangxi.
- Kitty Yuen and Terry Wong won the Merit Awards of the Charles Baggage Final Year Project Awards, HKPU, 2006/07.

### **Invited Talks and Refereed Posters** (not including conference talks)

1. “Meta Paths and Meta Structures: Analyzing Large Heterogeneous Information Networks”, in **Jing Dong Data Science Lab**, Beijing, 13<sup>th</sup> July, 2017.
2. “Meta Paths and Meta Structures: Analyzing Large Heterogeneous Information Networks”, in Department of Computer Science, **Tsinghua University**, 11<sup>th</sup> July, 2017.
3. “Meta Paths and Meta Structures: Analyzing Large Heterogeneous Information Networks”, **ICT, Chinese Academy of Science**, Beijing, 7<sup>th</sup> July, 2017.
4. “On Attributed Community Search”, in the Department of Computer Science, **Harbin Institute of Technology (HIT) Shenzhen Graduate School**, 15<sup>th</sup> December, 2016.

5. “Meta Paths and Meta Structures: Computing Relevance in Large Heterogeneous Information Networks”, in **Shenzhen Institute of Advanced Technology (SIAT), Chinese Academy of Sciences**, 14<sup>th</sup> December, 2016.
6. “Meta Paths and Meta Structures: Computing Relevance in Large Heterogeneous Information Networks”, in the Department of Computer Science, **Fudan University**, 10<sup>th</sup> September, 2016.
7. “Meta Paths and Meta Structures: Computing Relevance in Large Heterogeneous Information Networks”, in the Department of Computer Science, **Macao University**, 29<sup>th</sup> July, 2016.
8. “Meta Paths and Meta Structures: Computing Relevance in Large Heterogeneous Information Networks”, in the Department of Computer Science, **City University of Hong Kong**, 28<sup>th</sup> July, 2016.
9. “Meta Paths and Meta Structures: Computing Relevance in Large Heterogeneous Information Networks”, in the Department of Computer Science, **INRIA/LIX**, Paris, 1<sup>st</sup> July, 2016.
10. “Enriching Knowledge Series: (20) The Uses and Applications of Big Data in Daily Life”, ICT training seminar for secondary school teachers, **Education Bureau**, Hong Kong, 2016.
11. “Big Data Applications”, Trans-disciplinary Research Workshops, **Graduate School**, University of Hong Kong 2015-16. Dec 2 and 16, 2015 (6 hours)
12. “The Age of Big Data”, in the Big Data and Privacy Workshop, **Law and Technology Centre**, University of Hong Kong, 30<sup>th</sup> November, 2015. URL: <http://www.lawtech.hk/big-data-and-privacy-workshop/>
13. “On Crowdsourcing Tags”, in the Department of Computer Science and Engineering, **Shanghai Jiao Tong University**, 7<sup>th</sup> May, 2015.
14. “On Incentive-based Tagging” (**Keynote Speaker**), in the 7<sup>th</sup> Intl. Symposium on Computational Intelligence and Design (**ISCID 2014**), 13-14<sup>th</sup> December, 2014.
15. “Managing Uncertain Databases”, in the College of Computer Science, **Zhejiang University**, 12<sup>th</sup> December, 2014.
16. “On Incentive-based Tagging”, in the **Summer School**, College of Computer Science and Software Engineering, Shenzhen University, 14<sup>th</sup> July, 2014.
17. “iTag: Incentive-based Tagging”, in the **2014 Big Data Workshop**, Chinese University of Hong Kong, 22<sup>nd</sup> April 2014.
18. “Managing Uncertain Databases”, in **Xi’an Jiaotong University**, 11<sup>th</sup> March 2014.
19. “Managing Uncertain Databases”, in the **Advanced Database Lab (adalab)**, Soochow University, 6<sup>th</sup> March 2014.
20. “Publishing Research Papers”, in the **Advanced Database Lab (adalab)**, Soochow University, 3<sup>rd</sup> March, 2014.
21. “On Incentive-based Tagging”, in the **Advanced Database Lab (adalab)**, Soochow University, 26<sup>th</sup> February 2014.
22. “Evaluating Multi-way Joins over Discounted Hitting Time”, research seminar at the **National Donghua University (NDHU)**, Hualien, Taiwan, 15<sup>th</sup> January, 2014.
23. “On Incentive-based Tagging”, research seminar at **Academic Sinica (AS)**, Taipei, Taiwan, 13<sup>th</sup> January, 2014.
24. “Evaluating Multi-way Joins over Discounted Hitting Time”, research seminar at the **National Tsinghua University (NTHU)**, Hsinchu, Taiwan, 10<sup>th</sup> January, 2014.
25. “On Incentive-based Tagging”, research seminar at the **National Chiao Tung University (NCTU)**, Hsinchu, Taiwan, 9<sup>th</sup> January, 2014.
26. “Writing Research Papers”, research seminar at the **National Chiao Tung University (NCTU)**, Hsinchu, Taiwan, 8<sup>th</sup> January, 2014.
27. “On Incentive-based Tagging”, research seminar at the **University of Oxford**, Oxford, UK, 30<sup>th</sup> July, 2013.
28. “On Incentive-based Tagging”, research seminar at the **Microsoft Research Asia (MSRA)**, Beijing, China, 7<sup>th</sup> June, 2013.
29. “On Incentive-based Tagging”, research seminar at the **East China Normal University**, Shanghai, China, 4<sup>th</sup> June, 2013.
30. “Filter-based Procotols for Continuous Queries over Imprecise Location Data”, research seminar at the **INQUEST workshop (INovative QUerying of Streams)**, Oxford University, 25-27<sup>th</sup> September, 2012.
31. “Explore or Exploit? Effective Strategies for Disambiguating Large Databases”, research seminar at the **Microsoft Research Asia (MSRA)**, Beijing, China, 13<sup>th</sup> September, 2012.
32. “Managing Uncertain Databases”, invited talk in the 2012 **International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)**, 26-27 August, 2012.
33. “Evaluating Probabilistic Queries over Uncertain Matching”, research seminar at the **Shandong University**, Jinan, China, 31<sup>st</sup> July, 2012.



34. "Explore or Exploit? Effective Strategies for Disambiguating Large Databases", research seminar at the **Shandong University**, Jinan, China, 27<sup>th</sup> July, 2012.
35. "Evaluating Probabilistic Queries over Uncertain Matching", research seminar at the **University of Munich (LMU)**, Munich, Germany, 19<sup>th</sup> June, 2012.
36. "Evaluating Probabilistic Queries over Uncertain Matching", in the 2011 **International Workshop on Advanced Data Management (ADM)**, December 2011.
37. "Explore or Exploit? Effective Strategies for Disambiguating Large Databases", in the **Nanyang Technological University (NTU)**, August 2011.
38. "Explore or Exploit? Effective Strategies for Disambiguating Large Databases", in the **National University of Singapore (NUS)**, August 2011.
39. "Managing Uncertain Databases", in the **National University of Singapore (NUS)**, July 2011.
40. "Managing Uncertain Databases", in the **Advanced Digital Sciences Center (ADSC)**, University of Illinois at Urbana-Champaign, Singapore, July 2011.
41. "Managing Uncertain Databases", at the Summer Research Institute seminar series, School of Computer and Communication Sciences, **EPFL**, Lausanne, 15<sup>th</sup> June, 2011.
42. "Managing Uncertain Databases", at the DKE seminar, **University of Queensland**, 31<sup>st</sup> May, 2011.
43. "Managing Uncertain Databases", at the Australian e-Health Research Centre, **CSIRO**, Australia, 3<sup>rd</sup> June, 2011.
44. "Explore or Exploit? Effective Strategies for Disambiguating Large Databases", at the ITEE seminar, **University of Queensland**, 1<sup>st</sup> June, 2011.
45. "Explore or Exploit? Effective Strategies for Disambiguating Large Databases", **Invited Talk**, in the 2<sup>nd</sup> StrucFus meeting, Sweden, Nov 30-Dec 2, 2010.
46. "UV-diagram: A Voronoi Diagram for Uncertain Data", at the **Singapore Management University**, 13 September, 2010.
47. "Managing Uncertain Databases", **Invited Talk**, in the **PhD Summer School** (Data Intensive Computing and Unstructured Data Management, Part 2: Unstructured Data Management), 19-31 July 2010, Renmin University of China.
48. "UV-diagram: A Voronoi Diagram for Uncertain Data", at the **Renmin University of China**, China, 27 July, 2010.
49. "UV-diagram: A Voronoi Diagram for Uncertain Data", at **Peking University**, China, 29, July, 2010.
50. "Cleaning Uncertain Data with Quality Guarantees", **Invited Talk**, in the 1<sup>st</sup> StrucFus meeting, Sweden, Aug 12-13, 2009.
51. "Querying and Cleaning Uncertain Data", **Invited Talk**, in the 1<sup>st</sup> Intl. Workshop on Quality of Context (**QuaCon**), Stuttgart, Germany, June 2009.
52. "Cleaning Uncertain Data with Quality Guarantees", the Informatics Colloquium, at the **University of Stuttgart**, Germany, 11 August 2008.
53. "Managing Spatial Uncertainty in a Database System", *Keynote speaker*, at the POLYNET IT Security Symposium 2008, **Hong Kong Polytechnic University (HKPU)**, 5 June 2008.
54. "Managing Attribute Uncertainty in a Database System", at the **Chinese University of Hong Kong (CUHK)**, Hong Kong, 19 November, 2007.
55. "Efficient Evaluation of Imprecise Location-Dependent Queries", at the **Nanyang Technological University (NTU)**, Singapore, 7 August 2007.
56. "Efficient Evaluation of Imprecise Location-Dependent Queries", at the **National University of Singapore (NUS)**, Singapore, 3 August 2007.
57. "Efficient Evaluation of Imprecise Location-Dependent Queries", the Informatics Colloquium, at the **University of Stuttgart**, Germany, 3 July 2007.
58. "Querying Uncertainty of Spatial Data", at the Hong Kong Database Research Workshop, the **University of Hong Kong (HKU)**, 31 May, 2007.
59. "Efficient Evaluation of Imprecise Location-Dependent Queries", at the CS departmental seminar of the **University of Hong Kong (HKU)**, 13 April, 2007.
60. "Preserving User Location Privacy in Mobile Data Management Infrastructures", the research seminar at the **Shandong University**, Jinan, China, 21 August 2006.
61. "Uncertainty Management in Moving-Object and Sensor Databases", the research seminar at the **Shandong University**, Jinan, China, 17 August 2006.
62. "Uncertainty Management in Moving-Object and Sensor Databases", the Informatics Colloquium, at the **University of Stuttgart**, Germany, 13 July 2006.

63. "Change Tolerant Indexing on Constantly Evolving Data", the database research seminar meeting, at the **Hong Kong University of Science and Technology**, Hong Kong, 25 November, 2005.
64. "U-DBMS: A Database System for Managing Constantly-Evolving Data", Software demo, in Very Large Databases Conference (**VLDB 2005**), Norway, Aug 2005.
65. "Adaptive Stream Filters for Entity-based Queries with Non-value Tolerance", at the departmental seminar of the **University of Hong Kong**, Hong Kong, August 17, 2005.
66. "Change Tolerant Indexing on Constantly Evolving Data", the annual research meeting, at the **Hong Kong Polytechnic University**, Hong Kong, July 15, 2005.
67. "Managing Uncertainty in Sensor-based Applications", in *Wired for Wireless: Poster Session* of Purdue's Center for Wireless Systems and Applications (**CWSA**) Workshop, Purdue University, USA, Nov 18-19, 2004.
68. "Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data", at the seminar of CIMIC, **Rutgers University**, New Jersey, USA, Oct 13, 2004.
69. "Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data", at the CS seminar of **University of Illinois at Urbana-Champaign (UIUC)**, Illinois, USA, Sep 30, 2004.
70. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the **University of Edinburgh**, Edinburgh, Scotland, Sep 14, 2004.
71. "Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data", at the database group seminar of **Purdue University**, Aug 25, 2004.
72. "Managing Uncertainty in Moving-Object and Sensor Databases", at the 1st Mobile Computing Research Forum of the **Hong Kong University of Science and Technology (HKUST)**, Jul 28, 2004.
73. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the **Hong Kong University of Science and Technology (HKUST)**, Jul 7, 2004.
74. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the **Hong Kong Baptist University (HKBU)**, Jun 30, 2004.
75. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the **Hong Kong Polytechnic University (HKPU)**, Jun 28, 2004.
76. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the **University of Hong Kong (HKU)**, Jun 8, 2004.
77. "Managing Uncertainty in Moving-Object and Sensor Databases", at the CS seminar of the Department of **Chinese University of Hong Kong (CUHK)**, May 27, 2004.

## Selected Work Experience

### **The University of Hong Kong**

**Jul 2008 – current**

- Associate professor in the Department of Computer Science, Mar 2012 – current
- Assistant professor in the Department of Computer Science, Jul 2008 – Feb 2012
- Course Instructor (score = course effectiveness):
  - Computer Organization (CSIS1120A & B), Spring 2012, Spring 2013
  - Topic in Computer Science: Uncertainty Management (CSIS7801), Fall 2011
  - Advanced Topics in Data Engineering (CSIS8101), Fall 2012
  - Introduction to Database Management Systems (CSIS0278), Spring 2009, Spring 2010, Fall 2010
  - Data Analysis and Machine Learning (CSIS9101), Spring 2009
  - Topic in Computer Science: Workflow Management (CSIS7801), Fall 2008, Summer 2010, Spring 2011
- MPhil/PhD Programme Director, 2012-13
- Chair of the Departmental Research Postgraduate Committee (DRPC), 2012-13
- Deputy Chair of the Departmental Research Postgraduate Committee (DRPC), 2011-12
- Chair of the China Programme Development, 2009-10, 2010-11, 2011-12
- Chair of the Committee on Mainland Recruitment (UG), 2009-10, 2010-11, 2011-12
- Chair of the Mainland UG Student Recruitment, 2009-10, 2010-11, 2011-12
- External Examiner/Academic Assessor for Higher Diploma in Logistics, Information Technology, Financial Information Management, and Business Systems, **HKUSPACE** (2008/09 – 2010/11)
- Member of the Board of the Faculty of Engineering, 2008 - current
- Member of the Admissions Committee (China), 2009-10, 2010-11, 2011-12
- Member of the Programme Committee for the Computer Engineering Course, 2009-10, 2010-11, 2011-12
- Member of the Committee on Advanced Undergraduate Training, 2009-10, 2010-11

- Member of the Departmental Curriculum Review and Development Committee (DCRDC), 2009-10, 2010-11, 2011-12
- Member of the Department Executive Committee (China Development), 2009-10
- Member of the Departmental Research Postgraduate Committee (DRPC), 2009-10, 2010-11
- Coach of three HKU teams for ACM Programming Competition, 2008-09
- Coach of five HKU teams for the IBM Inter-University Programming Contest 2009
- Presenter of mini-workshop in LOGIN 2010 and 2009 (a department activity for secondary school students)

### **The Hong Kong Polytechnic University**

**May 2005 – Jun 2008**

- Assistant professor in the Department of Computing.
- Course Instructor (score = average Student Feedback Questionnaire score, out of 5):
  - Advanced Database Systems (COMP 5151), Spring 08 [score: 4.3]
  - Workflow Management and Collaborative Systems (COMP 5524), Spring 2006-08 [score: 3.9, 4.3, 4.0]
  - Wireless Computing Systems and Applications (COMP 5326), Fall 2005-07 [score: 3.9, 4.2, 4.3]
  - Principles of Programming (COMP 201), Fall 2006 [score: 3.7]
- Member of the **Scholarship Assessment Board**, 2005-08

### **Oxford University**

- Visiting Researcher

**24 – 31 Jul 2013**

### **National University of Singapore**

- Visiting Researcher

**1-8 Aug 2007**

### **University of Stuttgart**

- Visiting Scientist

**Jun – Jul 2007, Jun – Aug 2006**

### **Shandong University**

- Visiting Researcher

**Jun 2013, Jul 2012, Aug 2006**

### **Purdue University**

**Aug 2001 – May 2005**

- Research assistant in the Department of Computer Sciences.
- Teaching assistant for *Computer Architecture* (CS 250), *Information Systems* (CS 348) and *Computer Networks* (CS 422). Work involved instructing recitation sessions, grading and assisting students.

### **Database Group, University of Hong Kong**

**May – Aug 2002, May – Aug 2004**

- Summer Intern as research assistant for Prof. Ben C. M. Kao. Work on database streams, sensor stream mining, real-time data stream query processing and querying uncertainty.

### **University of Hong Kong**

**Aug 1998 – 2001**

- Teaching assistant for Introduction to Data Structures and Algorithms (CSIS1119A) and graduate level Advanced Topics in Database Systems (CSIS 6954).

### **The HongKong Electric Company**

**Jun – Aug 1997**

- Summer intern as database programmer: Oracle database cleansing on IBM DB2 and VAX platforms.

## **Professional Activities**

1. **Member, Selection Committee for IEEE TKDE Editor-in-Chief**, 2016.
2. **Editorial Board Member, IEEE TKDE** (Transactions on Knowledge and Data Engineering), May 2015 – current.
3. **Editorial Board Member, DAPD** (Distributed and Parallel Databases Journal), March 2012 – current.
4. **Editorial Advisory Board, Information Systems** (Journal), July 2012 – current.
5. **Guest Editor, Journal of Internet Technology, Special Issue on Advances in Social Computing**, 2016
6. **Guest Editor, GeoInformatica, Special Issue on Spatial and Temporal Database**, 2013.
7. **Guest Editor, Distributed and Parallel Databases, Special Issue on Databases and Crowdsourcing**, 2013.
8. **Member, the SSTD Endowment**, March 2014 – current.
9. **General Chair, ACM-Hong Kong Programming Contest**, HKU, June 22<sup>nd</sup> 2013.
10. **Publicity Chair, ACM-Hong Kong Student Research and Career Day**, HKUST, November 19<sup>th</sup>, 2013.
11. **Vice Chairperson, ACM Hong Kong Chapter**, 2012-13.

12. **Head judge**, Hong Kong ACM Programming Contest, 2012.
13. **Keynote Speaker** for the Mobility Analytics for Spatio-temporal and Social Data (MATES 2017), co-located with VLDB 2017, 1<sup>st</sup> September, 2017.
14. **Invited Speaker** for the **2014 Big Data Workshop**, Chinese University of Hong Kong, 22<sup>nd</sup> April 2014.
15. **Invited Speaker** for the 2011 International Workshop on Advanced Data Management (**ADM 2011**), Shanghai, China, December 10-11, 2011.
16. **Speaker** for the tutorial, *Similarity Search and Mining in Uncertain Databases*, in **VLDB 2010**, September 13-17, 2010
17. **Invited Speaker** for the China NFS Summer School: Unstructured Data Management (2010 年全國研究生暑期學校, 國家自然科學基金委員會), 3 hours, 28 July, 2010. URL: <http://deke.ruc.edu.cn/summerschool/course.htm>
18. **Guest Editor, TKDE**, *Special Issues on Mining Large Uncertain and Probabilistic Databases*, published in September 2010
19. **Keynote Speaker** for the First International Workshop on Quality of Context (QuaCon'09)
20. **Speaker** for the Nanjing University Summer Advanced Seminar, July 9-10, 2009
21. **Senior PC member** for **APWeb-WAIM 2018**
22. **Area Chair** in **IEEE ICDE 2017** (IEEE Intl. Conf. on Data Engineering)
23. **Program Co-Chair** for **APWeb 2015** (The 17<sup>th</sup> Asia-Pacific Web Conference)
24. **Senior PC member** for **DASFAA 2015** (The 20<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications)
25. **Area Chair** for DB track (*Uncertain, Probabilistic, and Scientific Data Management*) in **CIKM 2014** (ACM Intl. Conference on Information and Knowledge Management)
26. **Area Chair** for *Crowdsourcing* in the 2<sup>nd</sup> edition of the Encyclopedia of Database Systems, edited by T. Oszu and L. Liu.
27. **Workshop Co-chair** for **ICDE 2014** (The 29<sup>th</sup> Intl. Conf. on Data Engineering)
28. **Publicity Chair** for **APWeb 2014** (The 16<sup>th</sup> Asia-Pacific Web Conference)
29. **Program Co-Chair** for **SSTD 2013** (The 13<sup>th</sup> Intl. Symposium on Spatial and Temporal Databases)
30. **Program Co-chair** for **1<sup>st</sup> PhD Symposium, PAKDD 2011** (Methodologies for Knowledge Discovery and Data Mining, Pacific-Asia Conference)
31. **Tutorial Co-chair** for **DASFAA 2011** (16th Intl. Conf. on Database Systems for Advanced Applications)
32. **Program Vice Chair** for **IWKDEWL 2010** (The Intl. Workshop on Knowledge and Data Engineering in Web-based Learning, in conjunction with the 9<sup>th</sup> Intl. Conf. on Web-based Learning (ICWL 2010))
33. **Registration Chair** for **CIKM 2009** (ACM 18th Conference on Information and Knowledge Management)
34. **Steering committee** for **MOUND 2010** (2nd Workshop on Management and Mining of Uncertain Data, with **ICDE 2010**)
35. **Program Co-chair** for **UMMM 2010** (1st Intl. Workshop on Uncertain Mobile Data Management and Mining, with **MDM 2010**)
36. **Program Co-chair** for **MOUND 2009** (1st Workshop on Management of Uncertain Data, with **ICDE 2009**)
37. **Program Co-chair** for **DUNE 2007** (1st Workshop on Data Mining of Uncertain Data, with **ICDM 2007**)
38. **Program Co-chair** for **WINPEN 2006** (Intl. Workshop on Information Processing over Evolving Networks, with **WAIM 2006**)
39. **Local Arrangement Chair** for **CIC 2006** (4<sup>th</sup> Intl. Conf. on Cooperative Internet Computing)
40. **Session Chair** for **VLDB 2010** (The 36<sup>th</sup> Intl. Conf. on Very Large Data Bases)
41. **Session Chair** for **ICDE 2010** (The 26<sup>th</sup> Intl. Conf. on Data Engineering)
42. **Session Chair** for **VLDB 2008** (The 34<sup>th</sup> Intl. Conf. on Very Large Data Bases)
43. **Session Chair** for **NPsec 2007** (The 3<sup>rd</sup> Workshop on Secure Network Protocols)
44. **Session Chair** for **WAIM 2006** (The 7<sup>th</sup> Intl. Conference on Web-Age Information Management)
45. **Session Chair** for **CIC 2006** (The 4<sup>th</sup> Intl. Conf. on Cooperative Internet Computing)
46. Regular member of **ACM**
47. Regular member of **IEEE**
48. Member of ACM Special Interest Group on Management of Data (**ACM SIGMOD**)
49. Member of the IEEE Computer Society, Hong Kong section
50. Member of the Hong Kong Information Technology Joint Council (**HKITJC**)
51. **Reviewer** of funding proposal for the Netherlands Organization for Scientific Research (NWO), 2010

52. **Reviewer** of funding proposal for the Natural Sciences and Engineering, Research Council of Canada (NSERC), 2009, 2010, 2011, 2016
53. **Reviewer** of funding proposal for the Israel Science Foundation in 2017
54. **External Reviewer**, tenure-track Assistant Professor search for Nanyang Technological University, Singapore, 2018
55. **Thesis exam committee chair** for PhD defence for Ms. Xiaoyan Xiong on “Computational Electromagnetic Methods for Multiphysics Simulations”, University of Hong Kong (2015)
56. **Thesis exam committee chair** for PhD defence for Mr. Haojie Zhao on “Auction-Based Resource Allocation in Selfish Networks”, University of Hong Kong (2014)
57. **Internal examiner** for PhD defence for Mr. Liu Yu on “Latent Topic Mining in Social Media with Probabilistic Graphical Model”, University of Hong Kong (2017)
58. **Internal examiner** for PhD defence for Mr. Shuyao Qi on “Advanced Ranking Queries on Composite Data”, University of Hong Kong (2016)
59. **Internal examiner** for PhD defence for Ms. Wenting Tu on “User-generated Content based Recommendation Systems for Investment and E-commerce”, University of Hong Kong (2016)
60. **Internal examiner** for PhD defence for Mr. Yilun Cai on “Ranking and Similarity Queries on Complex Data Types”, University of Hong Kong (2015)
61. **Internal examiner** for PhD defence for Mr. Xinjie Zhu on “STAP: A parallel signal track analytical processing system for flexible and efficient analysis of genomic data”, University of Hong Kong (2015)
62. **External examiner** for PhD defence for Mr. Saket Sathe on "Statistical Models for Querying and Managing Time-series Data", EPFL, Switzerland (2013)
63. **Internal examiner** for PhD defence for Mr. Chenghui Ren on “Algorithms for Evolving Graph Analysis”, University of Hong Kong (2013)
64. **Internal examiner** for PhD defence for Mr. Ge Shen on “Advanced Analysis and Join Queries in Multidimensional Spaces”, University of Hong Kong (2012)
65. **Internal examiner** for MPhil defence for Mr. Fei Wang on “Complex Stock Trading Strategy Based on Parallel Particle Swarm Optimization”, University of Hong Kong (2012)
66. **Internal examiner** for MPhil defence for Mr. Ruirui Li on “A Probabilistic Approach to Diversified Query Recommendation”, University of Hong Kong (2012)
67. **External examiner** for PhD defence for Mr. D. Šidlauskason on “Processing Massive Update and Query Workloads in Main Memory”, Aalborg University, Denmark (2012)
68. **External examiner** for PhD defence for Mr. Kai Zheng on "Uncertainty Management in Spatio-temporal Database", the University of Queensland, Australia (2012)
69. **Internal examiner** for PhD defence for Mr. Haisheng Tang on “Minimizing Interference in Wireless Sensor Networks”, University of Hong Kong (2011)
70. **Internal examiner** for PhD defence for Mr. Chun Kit Chui on “OLAP on Sequence Data”, University of Hong Kong (2010)
71. **Internal examiner** for MPhil defence for Mr. Xinjie Zhu on “Mining Order-Preserving Submatrices from Data with Repeated Measurements”, University of Hong Kong (2010)
72. **Internal examiner** for MPhil defence for Mr. Bin Bi on “Third-order Tensor Decomposition for Search in Social Tagging Systems”, University of Hong Kong (2010)
73. **Internal examiner** for MPhil defence for Mr. Wang Liang, University of Hong Kong (2010)
74. **Internal examiner** for PhD defence for Mr. Thomas Yau-tat Lee on “Formalisms on Semi-Structured and Unstructured Data Schema Computations”, University of Hong Kong (2010)
75. **Internal examiner** for PhD defence for Mr. Leong Hou U on “Matching Problems in Large Databases”, University of Hong Kong (2010)
76. **Internal examiner** for MPhil defence for Mr. Yiwei Zhang on “SIGBOT: Signature-based Multiple-Bug Localization”, University of Hong Kong (2010)
77. **Internal examiner** for MPhil defence for Mr. Lee King For on “Clustering Uncertain Data Using Voronoi Diagram”, University of Hong Kong (2009)
78. **Internal examiner** for MPhil defence for Mr. Lin Zhifeng on “Advanced Spatial Queries in Wireless Ad Hoc Networks”, University of Hong Kong (2009)

79. **Internal examiner** for PhD defence for Mr. Ngai Wang Kay on “Cluster Analysis on Uncertain Data”, University of Hong Kong (2008)
80. **Internal examiner** for MPhil defence for Mr. Smith Tsang on "Efficient Decision Tree Building Algorithms for Uncertain Data", University of Hong Kong (2008)
81. **External examiner** for MPhil defence for Mr. Alan Kwan on "Adaptive Stream Filters for Entity-based Queries with Non-Value Tolerance", University of Hong Kong (2007)
82. **External examiner** for MPhil defence for Mr. Dai Xaingyuan on "Spatial Preference Queries based on Non-spatial Constraints", University of Hong Kong (2006)
83. **Judge** for IT competition May 2006, co-organized with the PC Weekly magazine.
84. **Question Designer** for the ACM Programming Contest, Hong Kong region, 2006.
85. **Program Committee Member for:**

#### Year 2018

- **VLDB** (2017-18 review board for 44<sup>nd</sup> Intl. Conf. on Very Large Data Base)
- **SIGMOD** (ACM SIGMOD Intl. Conf. on Management of Data)
- **IEEE ICDE** (34<sup>nd</sup> Intl. Conf. on Data Engineering)
- **KDD** (24<sup>th</sup> ACM SIGKDD Conf. on Knowledge, Discovery, and Data Mining)
- **CIKM** (ACM 27<sup>th</sup> Conference on Information and Knowledge Management)
- **EDBT** (20<sup>th</sup> Intl. Conf. on Extending Database Technology)
- **IEEE BigData 2018** (Senior PC) (2018 IEEE Intl. Conf. on Big Data)
- **SUM** (12<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **DSAA** (5<sup>th</sup> IEEE Intl. Conf. on Data Science and Advanced Analytics)
- **APWeb-WAIM** (Senior PC) (Asia-Pacific Web Conference)

#### Year 2017

- **VLDB** (2016-17 review board for 43<sup>nd</sup> Intl. Conf. on Very Large Data Base)
- **SIGMOD** (ACM SIGMOD Intl. Conf. on Management of Data)
- **CIKM** (ACM 26<sup>th</sup> Conference on Information and Knowledge Management, Long papers track)
- **WWW** (26<sup>th</sup> Intl. World Wide Web Conf.)
- **IEEE BigData 2017** (Senior PC) (2017 IEEE Intl. Conf. on Big Data)
- **SUM** (11<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **HMDData** (1<sup>st</sup> IEEE Workshop on Human-Machine Collaboration in BigData, with IEEE Bigdata)

#### Year 2016

- **VLDB** (2015-16 review board for 42<sup>nd</sup> Intl. Conf. on Very Large Data Base)
- **SIGMOD** (ACM SIGMOD Intl. Conf. on Management of Data, demo track)
- **IEEE ICDE** (32<sup>nd</sup> Intl. Conf. on Data Engineering)
- **IJCAI** (Intl. Joint Conf. on Artificial Intelligence)
- **EDBT** (19<sup>th</sup> Intl. Conf. on Extending Database Technology)
- **KDD** (22<sup>nd</sup> ACM SIGKDD Conf. on Knowledge, Discovery, and Data Mining)
- **CIKM** (ACM 25<sup>th</sup> Conference on Information and Knowledge Management, Database Track)
- **IEEE BigData** (2016 IEEE Intl. Conf. on Big Data)
- **HCOMP** (4<sup>th</sup> AAAI Conf. on Human Computation and Crowdsourcing)
- **GeoRich** (2<sup>nd</sup> Intl. ACM Workshop on Managing and Mining Enriched Geo-Spatial Data, with **SIGMOD 2016**)

#### Year 2015

- **VLDB** (2015-16 review board for 42<sup>nd</sup> Intl. Conf. on Very Large Data Base)
- **KDD** (The 21<sup>st</sup> ACM SIGKDD Conf. on Knowledge, Discovery, and Data Mining)
- **CIKM** (ACM 24<sup>th</sup> Conference on Information and Knowledge Management, Database Track)
- **DASFAA** (Senior PC member; 20th Intl. Conf. on Database Systems for Advanced Applications)
- **SUM** (9<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **SSTD** (The 14<sup>th</sup> Intl. Symposium on Spatial and Temporal Databases)

- **ACM SIGSPATIAL GIS** (23<sup>rd</sup> ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **IEEE BigData** (2015 IEEE Intl. Conf. on Big Data)
- **WAIM** (16<sup>th</sup> Intl. Conf. on Web-Age Information Management)
- **SocialInformatics** (2015 ASE Intl. Conf. on Social Informatics)
- **BDSI** (Intl. Workshop on Big Data Security Intelligence)

### Year 2014

- **VLDB** (2013-14 review board for 40<sup>th</sup> Intl. Conf. on Very Large Data Base)
- **IEEE ICDE** (Intl. Conf. on Data Engineering)
- **DASFAA** (19<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications)
- **ACM SIGSPATIAL GIS** (22<sup>nd</sup> ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **WAIM** (15<sup>th</sup> Intl. Conf. on Web-Age Information Management)
- **BDSE** (The 3<sup>rd</sup> IEEE Intl. Conf. on Big Data Science and Engineering)
- **IEEE BigData** (2014 IEEE Intl. Conf. on Big Data)
- **BigData** (The 3<sup>rd</sup> IEEE Intl. Congress on BigData)
- **SISAP** (7th Intl. Conference on Similarity Search and Applications)
- **GeoRich** (1<sup>st</sup> Intl. ACM Workshop on Managing and Mining Enriched Geo-Spatial Data, with **SIGMOD 2014**)
- **HuMoComp** (The Intl. Workshop on Human Mobility Computing, with **MDM 2014**)
- **MobiGIS** (3<sup>rd</sup> ACM SIGSPATIAL Intl. Workshop on Mobile Geographic Information Systems)

### Year 2013

- **SIGMOD** (ACM SIGMOD Intl. Conf. on Management of Data)
- **VLDB** (2012-13 review board for 39<sup>th</sup> Intl. Conf. on Very Large Data Base)
- **IEEE ICDE** (Intl. Conf. on Data Engineering, research program)
- **IEEE ICDE** (Intl. Conf. on Data Engineering, demo track)
- **KDD** (The 19<sup>th</sup> ACM SIGKDD Conf. on Knowledge, Discovery, and Data Mining)
- **DASFAA** (18<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications)
- **DEXA** (The 24<sup>th</sup> Conf. on Database and Expert Systems Applications)
- **ACM SIGSPATIAL GIS** (21<sup>st</sup> ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **BDDS** (The 2<sup>nd</sup> IEEE Intl. Conf. on Big Data and Distributed Systems)
- **MDM** (Intl. Conf. on Mobile Data Management)
- **WAIM** (14<sup>th</sup> Intl. Conf. on Web-Age Information Management)
- **AICCSA** (“Big Data, Business, Analytics” track, ACS/IEEE Intl. Conf. on Computer Systems and Applications)
- **HuMoComp** (The Intl. Workshop on Human Mobility Computing, with **MDM 2013**)
- **MoDA** (The 1<sup>st</sup> Intl. Workshop on Mobile Data Analytics, with **ICDE 2013**)
- **MobiGIS** (2<sup>nd</sup> ACM SIGSPATIAL Intl. Workshop on Mobile Geographic Information Systems)

### Year 2012

- **VLDB** (2012-13 review board for 39<sup>th</sup> Intl. Conf. on Very Large Data Base)
- **VLDB** (2011-12 review board for 38<sup>th</sup> Intl. Conf. on Very Large Data Base, Istanbul, Turkey, August 2012)
- **CIKM** (ACM 19<sup>th</sup> Conference on Information and Knowledge Management)
- **DASFAA** (17<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications)
- **DEXA** (The 23<sup>rd</sup> Conf. on Database and Expert Systems Applications)
- **ACM SIGSPATIAL GIS** (20th ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **SUM** (6<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **APWeb** (The 14<sup>th</sup> Asia-Pacific Web Conference)
- **KMIS** (Intl. Conf. on Knowledge Management and Information Sharing)
- **AINA** (The 26<sup>th</sup> IEEE Intl. Conf. on Advanced Information Networking and Applications, Pervasive/Ubiquitous Computing and Services Track, 2012)
- **QuesT** (3<sup>rd</sup> ACM SIGSPATIAL Intl. Workshop on Querying and Mining Uncertain Spatio-Temporal Data)

- **LBSN** (4<sup>th</sup> ACM SIGSPATIAL Intl. Workshop on Location-Based Social Networks)
- **MobiGIS** (1<sup>st</sup> ACM SIGSPATIAL Intl. Workshop on Mobile Geographic Information Systems)
- **PAISI** (The Pacific-Asia Workshop on Intelligence and security Informatics)

### Year 2011

- **VLDB** (2011-12 review board for 38<sup>th</sup> Intl. Conf. on Very Large Data Base, Istanbul, Turkey, August 2012)
- **IEEE ICDE** (Intl. Conf. on Data Engineering, 2011, text, uncertain and probabilistic data track)
- **DASFAA** (16<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications)
- **DEXA** (The 22<sup>nd</sup> Conf. on Database and Expert Systems Applications)
- **ACM SIGSPATIAL GIS** (19th ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **WWW** (World Wide Web 2011 Poster Track)
- **KMIS** (Intl. Conf. on Knowledge Management and Information Sharing)
- **SUM** (5<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **SPRINGL** (The 4th Intl. Workshop on Security and Privacy in GIS and LBS, with **ACM SIGSPATIAL 2011**)
- **QeST** (2<sup>nd</sup> ACM SIGSPATIAL Intl. Workshop on Querying and Mining Uncertain Spatio-Temporal Data)
- **APWeb** (The 13<sup>th</sup> Asia-Pacific Web Conference)
- **PAISI** (The Pacific-Asia Workshop on Intelligence and security Informatics)
- **DaMEN** (The 1<sup>st</sup> Workshop on Data Management for Emerging Network Infrastructures, with **DASFAA11**)

### Year 2010

- **VLDB** (36<sup>th</sup> Intl. Conf. on Very Large Data Base)
- **IEEE ICDE** (Intl. Conf. on Data Engineering, 2010, text and uncertain data track)
- **ACM SIGMOD** (ACM Conf. on Management of Data, tutorial committee)
- **CIKM** (ACM 19<sup>th</sup> Conference on Information and Knowledge Management)
- **ACM SIGSPATIAL GIS** (18th ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **DASFAA** (15<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications, Demonstration Track)
- **DEXA** (The 21<sup>th</sup> Conf. on Database and Expert Systems Applications)
- **WAIM** (10<sup>th</sup> Intl. Conf. on Web-Age Information Management)
- **KMIS** (Intl. Conf. on Knowledge Management and Information Sharing)
- **EUC** (The IFIP Intl. Conf. on Embedded and Ubiquitous Computing, Multimedia and data management)
- **SUM** (4<sup>th</sup> Intl. Conf. on Scalable Uncertainty Management)
- **SPRINGL** (The 3rd Intl. Workshop on Security and Privacy in GIS and LBS, with **ACM GIS 2010**)
- **PAISI** (The Pacific-Asia Conf. on Intelligence and security Informatics)

### Year 2009

- **CIKM** (ACM 18<sup>th</sup> Conference on Information and Knowledge Management)
- **DEXA** (The 20<sup>th</sup> Conf. on Database and Expert Systems Applications)
- **SSTD** (The 11<sup>th</sup> Intl. Symposium on Spatial and Temporal Databases)
- **ACM GIS** (The 17<sup>th</sup> ACM SIGSPATIAL Intl. Conf. on Advances in Geographic Information Systems)
- **C3S2E** (The 2<sup>nd</sup> Conference on Computer Science & Software Engineering)
- **CSE** (The 12<sup>th</sup> IEEE Intl. Conference on Computational Science and Engineering)
- **APWeb-WAIM** (The Joint Intl. Conf. on Asia-Pacific Web Conference (APWeb) and Web-Age Information Management (WAIM))
- **EUC** (The IFIP Intl. Conf. on Embedded and Ubiquitous Computing)
- **WISM** (The 2009 Intl. Conf. on Web Information Systems and Mining)
- **Quacon** (1<sup>st</sup> Workshop on Quality of Context, Stuttgart University, 2009)
- **QeST** (1<sup>st</sup> Intl. Workshop on Querying and Mining Uncertain Spatio-Temporal Data, with **ACM GIS**)
- **SSPS** (3<sup>rd</sup> Intl. Workshop on Scalable Stream Processing Systems, co-located with **IPDPS 2008**)
- **SPRINGL** (The 2nd Intl. Workshop on Security and Privacy in GIS and LBS, with **ACM GIS**)
- **U'09** (1<sup>st</sup> ACM SIGKDD Workshop on Knowledge Discovery from Uncertain Data)
- **PAISI** (The Pacific-Asia Conf. on Intelligence and security Informatics)



## Year 2008

- **VLDB** (34<sup>th</sup> Intl. Conf. on Very Large Data Base)
- **DEXA** (The 19<sup>th</sup> Conf. on Database and Expert Systems Applications)
- **DASFAA** (13<sup>th</sup> Intl. Conf. on Database Systems for Advanced Applications, Demonstration Track)
- **IDEAS** (12<sup>th</sup> Intl. Database Engineering & Applications Symposium)
- **ACM GIS** (The 16<sup>th</sup> ACM Intl. Conf. on Advances in Geographic Information Systems 2008)
- **ACM GIS PhD Showcase Workshop** (for 16<sup>th</sup> ACM GIS 2008)
- **SSDBM** (20<sup>th</sup> Intl. Conf. on Scientific and Statistical Database Management)
- **WAIM** (9<sup>th</sup> Intl. Conf. on Web-Age Information Management)
- **APWeb** (10<sup>th</sup> Asia Pacific Web Conference)
- **ECDM** (The 2<sup>nd</sup> IADIS European Conference on Data Mining)
- **IEEE EDOC** (The 12<sup>th</sup> Enterprise Computing Conference)
- **EUC** (The IFIP Intl. Conf. on Embedded and Ubiquitous Computing)
- **PAISI** (The Pacific-Asia Conf. on Intelligence and security Informatics)
- **CSE** (9<sup>th</sup> Intl. Conf. on Computational Science and Engineering, Distributed Database and Data Mining)
- **SPRINGL2008** (The 1st Intl. Workshop on Security and Privacy in GIS and LBS, with **ACM GIS**)
- **MMSDE** (1<sup>st</sup> Intl. Workshop on Mining Multimedia Streams in a Large-scale Distributed Environments, co-located with **ICDE 2008**)
- **SSTDM** (Workshop on Mining Spatial and Spatio-temporal Data, in conjunction with the IEEE Intl. Conference on Data Mining (**ICDM 2008**))
- **SSPS** (2<sup>nd</sup> Intl. Workshop on Scalable Stream Processing Systems, co-located with **EDBT 2008**)
- **PALMS** (The 2<sup>nd</sup> Intl. Workshop on Privacy-Aware Location-based Mobile Services)
- **IDAR** (SIGMOD PhD Workshop on Innovative Database Research)
- **PiBA** (Intl. Workshop on Privacy in Location-Based Applications, co-located with the 13th European Symposium on Research in Computer Security (**ESORICS 2008**))

## Year 2007

- **IDEAS** (11<sup>th</sup> Intl. Database Engineering & Applications Symposium)
- **APWeb/WAIM** (Joint Conf. of the 9th Asia-Pacific Web Conf. and the 8th Intl. Conf. on Web-Age Information Management)
- **DBMAN** (Intl. Workshop on Database Management and Application over Networks in WAIM 2007)
- **AICCSA** (ACS/IEEE Intl. Conference on Computer Systems and Applications)
- **ECDM** (The 1<sup>st</sup> IADIS European Conference on Data Mining)
- **IEEE EDOC** (The 11<sup>th</sup> Enterprise Computing Conference)
- **SSTD** (The Intl. Symposium on Spatial and Temporal Databases)
- **SSPS** (1st Intl. Workshop on Scalable Stream Processing Systems, co-located with ICDE)
- **EUC** (The IFIP Intl. Conf. on Embedded and Ubiquitous Computing)
- **PAISI** (The Pacific-Asia Conf. on Intelligence and security Informatics)
- **PALMS** (The 1<sup>st</sup> Intl. Workshop on Privacy-Aware Location-based Mobile Services)
- **HICA-07** (Hybrid Intelligent Computing and its Application, co-located with ICNC '07)
- **SSTDM** (Workshop on Mining Spatial and Spatio-temporal Data, in conjunction with the IEEE Intl. Conference on Data Mining (**ICDM 2007**))

## Year 2006

- **SSTDM** (Workshop on Mining Spatial and Spatio-temporal Data, in conjunction with the IEEE Intl. Conference on Data Mining (**ICDM 2006**))
- **COMAD** (The 13<sup>th</sup> Intl. Conf. on Management of Data)
- **CIC** (The 4<sup>th</sup> Intl. Conf. on Cooperative Internet Computing)
- **IEEE TENCON** (IEEE Intl. Region 10 Conference)
- **IEEE EDOC** (The 10<sup>th</sup> Enterprise Computing Conference)
- **WAIM** (The 7<sup>th</sup> Intl. Conference on Web-Age Information Management)
- **XWICT** (the 1<sup>st</sup> Intl. Workshop on XML, Web, and Internet Contents Technologies, in **WAIM 2006**)

86. **Reviewer for the following journals/books:**

- VLDBJ (Very Large Databases Journal, 2008, 2009,2011, 2013-15)
- IEEE TKDE (Transactions on Knowledge and Data Engineering, 2007-17)
- DAPD (Distributed and Parallel Databases Journal, 2012, 2015)
- DKE (Data and Knowledge Engineering, 2009, 2011)
- IS (Information Systems Journal, 2007, 2009-12, 2015, 2017)
- IEEE TC (Transactions on Computer, 2007, 2011, 2012, 2014, 2015)
- Geoinformatica (2015)
- SpringerPlus (2016)
- IJSS (International Journal of Systems Science, 2011)
- FCS (Frontiers of Computer Science, 2011)
- ACM TODS (Transactions on Database Systems, 2008-10, 2012, 2015)
- TCS (Theoretical Computer Science, 2015)
- ACM TSAS (Transactions on Spatial Algorithms and Systems, 2014)
- IJBDSI (International Journal of Big Data Security Intelligence, 2015)
  
- TDP (Transactions on Data Privacy: Foundations and Technologies, 2012-13)
- JISE (Journal of Information Science and Engineering, 2010)
- JCST (Journal of Computer Science and Technology, 2008, 2009, 2010)
- IJSS (International Journal of Systems Science, 2011)
- PEVA (Performance Evaluation, 2009)
- DAPD – Special Issue on Ranking Databases (Distributed and Parallel Databases Journal 2008, 2010)
- KAIS (Knowledge and Information System Journal 2008, 2013)
- IJA (Intl. Journal of Approximate Reasoning, 2010)
- EIS (Enterprise Information Systems, 2008)
- IEEE TSC (Transactions on Service Computing, 2008)
- IEEE TMC (Transactions on Mobile Computing, 2006)
- IEEE ToN (Transactions on Networking, 2007)
- IPL (Information Processing Letters)
- WWWJ (World Wide Web Journal)
- Encyclopedia of Computer Science & Engineering, Prof. Benjamin Wah (Ed), John Wiley & Sons, 2007

87. **External Reviewer for the following conferences:**

- PODS 2012 (31<sup>st</sup> ACM Principles of Database Systems)
- KDD 2011 (17<sup>th</sup> ACM SIGKDD Conference on Knowledge Discovery and Data Mining)
- SIGMOD 2010 (ACM Special Interest Group on Management of Data)
- ICPP 2008 (37<sup>th</sup> Intl. Conf. on Parallel Processing)
- Percom 2008 (6<sup>th</sup> IEEE Intl. Conf. on Pervasive Computing and Communications)
- ICITM 2007 (Intl. Conf. on Information Technology and Management)
- VLDB 2006 (Very Large DataBases Conf.)
- IEEE ICDE 2004 - 06 (Intl. Conf. on Data Engineering)
- ICDM 2006 (Intl. Conf. on Data Mining)
- 7<sup>th</sup> ACM Postgraduate Research Day (Hong Kong)
- MDM 2005 (Intl. Conf. on Mobile Data Management)
- DASFAA 2001, 05, 06 (Database Systems for Advanced Applications)
- Book chapter review for DB Encyclopedia 2005
- DMSN 2004 (Intl. Workshop on Data Management for Sensor Networks)
- IEEE SCC 2004 (Intl. Conf. on Services Computing)
- PDCAT 2001 (Intl. Conf. on Parallel and Distributed Computing, Applications, and Techniques)
- RTCSA 2000 (Intl. Workshop on Real-Time Computing Systems and Applications)

88. Elected as **Faculty Search Committee Representative** of Graduate Student Board, responsible for assisting faculty search process, holding job-search seminars, and discussions with faculty candidates (2004-05).
89. Elected as **Database Group Representative** for communicating between the database group, the department, and other research groups, publishing research newsletter and hosting seminars for visiting speakers (2002-04).

90. Elected as **Graduate Committee Representative** of Graduate Student Board, for communicating with faculties about student affairs. Organizing orientation activities for prospective and new students (2003-04).
91. Elected as the **Secretary** of the Upsilon Pi Epsilon Honor Society for Computer Science, Purdue Chapter for helping in activities like induction ceremony and discussion with honor students (2003-04).
92. Elected as **Head Demonstrator** in the University of Hong Kong for planning teaching assistants' teaching assignments, organizing departmental seminars and extracurricular activities (1998-01).
93. **Student helper** for the ACM Programming Contest, Hong Kong in 2001.
94. **Student helper** for the 10<sup>th</sup> World Wide Web Conference, Hong Kong.

## Honors and Awards

1. **William Mong Visiting Research Fellowship 2015-16**, with Prof. H. V. Jagadish from University of Michigan, on “*Data Ethics: A Socio-Technical Framework for Doing Big Data Right*”
2. **Outstanding Young Researcher Award 2011-12**, the University of Hong Kong (granted a funding of HKD 300,000 and a research postgraduate student position), awarded on 27<sup>th</sup> March, 2013.
3. **William Mong Visiting Research Fellowship 2011-12**, with Prof. Divyakant Agrawal from University of California, Santa Barbara, on “*An Uncertain Database System for Cloud Environments*”.
4. **Prize Presentation Ceremony 2010-11**, Faculty of Engineering, November 2011 (for recognition of excellence in research activities).
5. **Universitas 21 Fellowship 2011-12** (for visiting the National University of Singapore in July-August 2011, with an award amount of HKD 50,000).
6. **Research Output Prize** in Department of Computer Science, on the journal article entitled “*Scalable processing of snapshot and continuous nearest-neighbor queries over one-dimensional uncertain data*”, with a monetary award of HK\$60,000, awarded by the Faculty of Engineering, HKU, 2010.
7. **Outstanding Service Award** (as Registration Chair), in the ACM 18<sup>th</sup> Conference on Information and Knowledge Management (CIKM), November 2-6, 2009, Hong Kong.
8. **Performance Reward**, Department of Computing, Hong Kong Polytechnic University, 2007.
9. **Performance Reward**, Department of Computing, Hong Kong Polytechnic University, 2006.
10. Microsoft sponsorship for attendance to the 6th Workshop on Privacy Enhancing Technologies in PET 2006.
11. Member, Upsilon Pi Epsilon (**UPE**) Honor Society for Computer Science, Purdue Chapter, 2003
12. Travel Grant Award for 1<sup>st</sup> Intl. Workshop on Data Management for Sensor Networks (**DMSN**) in VLDB 2004
13. **Dean’s List**, Purdue University for outstanding academic achievements, 2001 – 2004
14. Runner-up for **IT Innovation Award**, University of Hong Kong, for a study project about the use of IT for education development in Singapore, 2001
15. Runner-up for Open Category in the **Web Design Competition** in Hong Kong, for designing a website about free economy of Hong Kong, 2000
16. Graduate of the **Leadership and Life Skills** Class in the University of Hong Kong, a one-year course for leadership and social skill development, 2000 – 2001
17. **Best Paper Award** for the paper “*Updates and View Maintenance in Soft Real-Time Database Systems*” in the 2nd ACM Hong Kong Postgraduate Research Day, 1999
18. **Dean’s List**, the University of Hong Kong for outstanding achievements, 1995-1998
19. **First Class Honor** (Bachelor of Engineering), the University of Hong Kong, 1998

11-May-18