

Curriculum Vitae of

Yin “David” Yang

Ph.D., HKUST

Assistant Professor
College of Science and Engineering
Hamad Bin Khalifa University, Qatar

Email: yyang@hbku.edu.qa
Tel: +974-445-HKUST

Education

- | | |
|-----------|--|
| 2004-2009 | Ph.D. in Computer Science
Supervisor: Dimitris Papadias
Department of Computer Science and Engineering
Hong Kong University of Science and Technology (HKUST) |
| 2000-2004 | Bachelor of Engineering in Computer Science
Department of Computer Science and Engineering
Shanghai Jiao Tong University, Shanghai, China (SJTU) <ul style="list-style-type: none">– Ranked 1st among 130 students.– 6th place worldwide, ACM International Collegiate Programming Contest, World Finals, 2003. |

Current Research Interests

- *Big data*: interactive query processing and complex data analysis over terabytes of data.
 - *Fast data*: real-time processing of high-velocity data streams.
- *Data security and privacy*: privacy-preserving data collection, publication and analysis, database outsourcing and query authentication.

Professional Experience

- | | |
|-----------|---|
| 2014-now | Assistant Professor
College of Science and Engineering
Hamad Bin Khalifa University (HBKU), Qatar |
| 2011-2014 | Research Scientist
Advanced Digital Sciences Center (ADSC), Singapore
Adjunct positions: <ul style="list-style-type: none">▪ Adjunct Assistant Professor at Nanyang Technological University (NTU), Singapore |
| 2009-2011 | Postdoc
At HKUST (2009-10) and then at University of Hong Kong (2010-11) |

Publications

Google Scholar statistics: (as of July 2017)

- h-index: 21
 - i100-index: 3
 - i10-index: 31
- Number of citations: 1500+

Number of publications in top venues:

Data Science	<i>TODS</i>	<i>VLDB J.</i>	<i>TKDE</i>	<i>SIGMOD</i>	<i>VLDB</i>	<i>ICDE</i>	<i>KDD</i>
Number of papers	2	3	6	5	6	7	2

Cybersecurity	<i>CCS</i>	<i>WPES</i>
Number of papers	2	1

Others	<i>Bioinformatics</i>	<i>WWW</i>	<i>WSDM</i>	<i>TON</i>	<i>ICDCS</i>
Number of papers	1	1	1	1	1

Best Paper Award: *IEEE IC2E'13*.

Journal papers

- J1. Fu, T., Ding, J., Ma, R., Winslett, M., **Yang, Y.**, Zhang, Z. DRS: Dynamic Resource Scheduling for Real-Time Analytics over Fast Streams. *IEEE/ACM Transactions on Networking (TON)*, to appear.
- J2. Wang, S., Tang, Y., Xiao, X., **Yang, Y.**, Li, Z. HubPPR: Effective Indexing for Approximate Personalized PageRank. *Proceedings of the VLDB Endowment (PVLDB)*, 10(3): 205-216, 2016.
- J3. Wang, S., Xiao, X., **Yang, Y.**, Lin, W. Effective Indexing for Approximate Constrained Shortest Path Queries on Large Road Networks. *Proceedings of the VLDB Endowment (PVLDB)*, 10(2): 61-72, 2016.
- J4. Ding, J., Zhang, Z., Ma, R., **Yang, Y.** ABACUS: An Auction-Based Approach to Cloud Service Differentiation. *Elsevier Computer Networks*, 94: 231-249, 2016.
- J5. Cai, R., Hao, Z., Winslett, M., Xiao, X., **Yang, Y.**, Zhang, Z., Zhou, S. Deterministic Identification of Specific Individuals from GWAS Results. *Bioinformatics*, 31(11): 1701-1707, 2015.
- J6. Yuan, G., Zhang, Z., Winslett, M., Xiao, X., **Yang, Y.**, Hao, Z. Low-Rank Mechanism: Optimizing Batch Queries under Differential Privacy. *ACM Transactions on Database Systems (TODS)*, 40(2), Article 11, 2015.
- J7. Zhang, S., **Yang, Y.**, Fan, W., Winslett, M. Design and Implementation of a Real-Time Interactive Analytics System for Large Spatio-Temporal Data. *Proceedings of the VLDB Endowment (PVLDB)*, 7(13): 1754-1759, 2014.
- J8. Wang, H., Cai, Y., **Yang, Y.**, Zhang, S., Mamoulis, N. Durable Queries over Historical Time Series. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 26(3): 595-607, 2014.

- J9. Xu, J., Zhang, Z., Xiao, X., **Yang, Y.**, Yu, G., Winslett, M. Differentially Private Histogram Publication. *Very Large Data Bases Journal (VLDB J.)*, 22(6): 792-822, 2013.
- J10. Li, J., **Yang, Y.**, Mamoulis, N. Optimal Route Queries with Arbitrary Order Constraints. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 25(5):1097-1100, 2013.
- J11. Trimponias, G., Bartolini, I., Papadias, D., **Yang, Y.** Skyline Processing on Distributed Vertical Decompositions. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 25(4):850-862, 2013.
- J12. Zhang, J., Zhang, Z., Xiao, X., **Yang, Y.**, Winslett, M., Functional Mechanism: Regression Analysis under Differential Privacy. *Proceedings of the VLDB Endowment (PVLDB)*, 5(11): 1364-1375, 2012.
- J13. Yuan, G., Z. Zhang, Winslett, M., Xiao, X., **Yang, Y.**, Hao, Z. Low-Rank Mechanism: Optimizing Batch Queries under Differential Privacy. *Proceedings of the VLDB Endowment (PVLDB)*, 5(11):1352-1363, 2012.
- J14. Papadopoulos, S., Wang, L., **Yang, Y.**, Papadias, D., Karras, P. Authenticated Multi-Step Nearest Search. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 23(5): 641-654, 2011.
- J15. Papadopoulos, S., **Yang, Y.**, Papadias, D. Continuous Authentication on Relational Streams. *Very Large Data Bases Journal (VLDB J.)*, 19, 161-180, 2010.
- J16. Markowetz, A., **Yang, Y.**, Papadias, D. Keyword Search over Relational Tables and Streams. *ACM Transactions on Database Systems (TODS)*, 34(3), Article 17, 2009.
- J17. **Yang, Y.**, Papadopoulos, S., Papadias, D., Kollios, G. Authenticated Indexing for Outsourced Spatial Databases. *Very Large Data Bases Journal (VLDB J.)*, 18(3): 631-648, 2009.
- J18. Zhang, Z., **Yang, Y.**, Tung, A., Papadias, D. Continuous k -Means Monitoring over Moving Objects. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 20(9): 1205-1216, 2008.
- J19. **Yang, Y.**, Papadias, D., Krämer, J., Seeger, B. HybMig: A Hybrid Approach to Dynamic Plan Migration for Continuous Queries. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 19(3): 398-411, 2007.

Refereed conference papers

Full papers

- C1. Qin, Z., Yu, T., **Yang, Y.**, Khalil, I., Xiao, X., Ren, K. Generating Synthetic Decentralized Social Graphs with Local Differential Privacy. To appear in *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, Dallas, TX, USA, October 30-November 3, 2017.
- C2. Wang, S. Yang, R., Xiao, X., Wei, Z., **Yang, Y.** FORA: Simple and Effective Approximate Single-Source Personalized PageRank. To appear in *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, Halifax, Canada, August 13-17, 2017. Acceptance rate: 1/10 (full paper with oral presentation).
- C3. Wang, N., Xiao, X., **Yang, Y.**, Zhang, Z., Gu, Y., Yu, G. PrivSuper: a Superset-First Approach to Frequent Itemset Mining under Differential Privacy. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, San Diego, California, USA, April 29-22, 2017. Acceptance rate: 1/5.

- C4. Qin, Z., **Yang, Y.**, Yu, T., Khalil, I., Xiao, X., Ren, K. Heavy Hitter Estimation over Set-Valued Data with Local Differential Privacy. *Proceedings of the ACM Conference on Computer and Communications Security (CCS)*, Vienna, Austria, October 24-28, 2016. Acceptance rate: 1/6.
- C5. Yuan, G., **Yang, Y.**, Zhang, Z., Hao, Z. Convex Optimization for Linear Query Processing under Approximate Differential Privacy. *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, San Francisco, California, USA, August 13-17, 2016. Acceptance rate: 1/5.
- C6. Wang, L., Zhou, M., Zhang, Z., **Yang, Y.**, Zhou, A. Elastic Pipelining in an In-Memory Database Cluster. *Proceedings of the ACM Conference on the Management of Data (SIGMOD)*, San Francisco, California, USA. June 26-July 1, 2016. Acceptance rate: 1/5.
- C7. Fu, T., Ding, J., Ma, R., Winslett, M., **Yang, Y.**, Zhang, Z. DRS: Dynamic Resource Scheduling for Real-Time Analytics over Fast Streams. *Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS)*, Columbus, Ohio, USA, June 29-July 2, 2015. Acceptance rate: 1/7 (full paper with oral presentation).
- C8. Zhang, J., Xiao, X., **Yang, Y.**, Zhang, Z., Winslett, M. PrivGene: Differentially Private Model Fitting Using Genetic Algorithms. *Proceedings of the ACM Conference on the Management of Data (SIGMOD)*, New York City, New York, USA, June 22-27, 2013. Acceptance rate: 1/5.
- C9. Zhang, Z., Ma, R., Ding, J., **Yang, Y.** ABACUS: An Auction-Based Approach to Cloud Service Differentiation. *Proceedings of the IEEE International Conference on Cloud Engineering (IC2E)*, San Francisco, CA, USA, March 25-27, 2013. Acceptance rate: 1/5. **Best paper award (1 / 107).**
- C10. Zhang, Z., Shu, H., Chong, Z., Lu, H., **Yang, Y.** C-Cube: Elastic Continuous Clustering in the Cloud. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, Brisbane, Australia, April 8-11, 2013. Acceptance rate: 1/5.
- C11. Peng, S., **Yang, Y.**, Zhang, Z., Winslett, M., Yu, Y. Query Optimization for Differentially Private Data Management Systems. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, Brisbane, Australia, April 8-11, 2013. Acceptance rate: 1/5.
- C12. Xu, J., Zhang, Z., Xiao, X., **Yang, Y.**, Yu, G. Differentially Private Histogram Publication. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, Washington D. C., USA, April 1-5, 2012. Acceptance rate: 1/5.
- C13. Wang, L., **Yang, Y.**, Miao, X., Papadias, D., Liu, Y. Algorithms for Local Sensor Synchronization. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, Hannover, Germany, April 11-16, 2011. Acceptance rate: 1/6 (full paper with full presentation).
- C14. Papadopoulos, S., **Yang, Y.**, Bakiras, S., Papadias, D. Continuous Spatial Authentication. *Proceedings of the International Symposium on Spatial and Temporal Databases (SSTD)*, Aalborg, Denmark, July 8-10, 2009. Acceptance rate: 1/3.
- C15. **Yang, Y.**, Papadias, D., Papadopoulos, S., Kalnis, P. Authenticated Join Processing in Outsourced Databases. *Proceedings of the ACM Conference on the Management of Data (SIGMOD)*, Providence, RI, USA, June 29th - July 2, 2009. Acceptance rate: 1/7.
- C16. Zhang, Z., **Yang, Y.**, Cai, R., Papadias, D., Tung, A. Kernel-Based Skyline Cardinality Estimation. *Proceedings of the ACM Conference on the Management of Data (SIGMOD)*, Providence, RI, USA, June 29th - July 2, 2009. Acceptance rate: 1/7.

- C17. **Yang, Y.**, Bansal, N., Dakka, W., Ipeirotis, P., Koudas, N., Papadias, D. Query by Document. Proceedings of the *ACM International Conference on Web Search and Data Mining (WSDM)*, Barcelona, Spain, February 9-12, 2009. Acceptance rate: 1/6.
- C18. **Yang, Y.**, Papadopoulos, S., Papadias, D., Kollios, G. Spatial Outsourcing for Location-based Services. Proceedings of the *IEEE International Conference on Data Engineering (ICDE)*, Cancún, México, April 7-12, 2008. Acceptance rate: 1/7.
- C19. **Yang, Y.**, Papadias, D. Just-In-Time Processing of Continuous Queries. Proceedings of the *IEEE International Conference on Data Engineering (ICDE)*, Cancún, México, April 7-12, 2008. Acceptance rate: 1/7.
- C20. Papadopoulos, S., **Yang, Y.**, Papadias, D. CADS: Continuous Authentication on Data Streams. Proceedings of the *Very Large Data Bases Conference (VLDB)*, Vienna, Austria, September 23-28, 2007. Acceptance rate: 1/7.
- C21. Markowetz, A., **Yang, Y.**, Papadias, D. Keyword Search on Relational Data Streams. Proceedings of the *ACM Conference on the Management of Data (SIGMOD)*, Beijing, China, June 14-17, 2007. Acceptance rate: 1/7.
- C22. **Yang, Y.**, Ng, W., Lau, H.-L., Cheng, J. An Efficient Approach to Querying Secure Outsourced XML Information. Proceedings of the *International Conference on Advanced Information Systems Engineering (CAiSE)*, Luxembourg, 2006. Acceptance rate: 1/6.
- C23. Zhang, L., Yu, Y., Lin, C., Zhou, J., **Yang, Y.** An Enhanced Model for Searching in Semantic Portals. Proceedings of the *International World Wide Web Conference (WWW)*, Chiba, Japan, 2005. Acceptance rate: 1/7.

Short papers, demo papers, and workshop papers

- C24. Zhang, S., **Yang, Y.**, Fan, W., Lan, L., Yuan, M. OceanRT: Real-Time Analytics over Large Temporal Data. *Proceedings of the ACM Conference on the Management of Data (SIGMOD)*, Snowbird, Utah, USA, June 22-27, 2014. Acceptance rate: 1/3.
- C25. Li, Y., Zhang, Z., Winslett, M., **Yang, Y.** Compressive Mechanism: Utilizing Sparse Representation in Differential Privacy. Proceedings of the *Workshop on Privacy in the Electronic Society (WPES)*, Chicago, IL, USA, October 17, 2011. Acceptance rate: 1/4.
- C26. Markowetz, A., **Yang, Y.**, Papadias, D. Reachability Indexes for Relational Keyword Search. Proceedings of the *IEEE International Conference on Data Engineering (ICDE)*, Shanghai, China, March 29 - April 4, 2009. Acceptance rate: 1/4.
- C27. Krämer, J., **Yang, Y.**, Cammert, M., Seeger, B., Papadias, D. Dynamic Plan Migration for Snapshot-Equivalent Continuous Queries in Data Stream Systems. Proceedings of the *Workshops of Extending Database Technology (EDBT)*, Munich, Germany, March 26-31, 2006.

Tutorials and posters

- C28. Tan, T., Ma, R., Winslett, W., **Yang, Y.**, Zhang, Z. Realtime Elastic Streaming Analytics in the Cloud. Proceedings of the *ACM Conference on the Management of Data (SIGMOD)*, New York City, New York, USA, June 22-27, 2013 (poster).

- C29. Peng, S., **Yang, Y.**, Zhang, Z., Winslett, M., Yu, Y. DP-Tree: Indexing Multi-Dimensional Data under Differential Privacy. Proceedings of the *ACM Conference on the Management of Data (SIGMOD)*, Scottsdale, Arizona, USA, May 20-24, 2012 (poster).
- C30. **Yang Y.**, Zhang, Z., Miklau, G., Winslett, M., Xiao, X. Differential Privacy in Data Publication and Analysis. Proceedings of the *ACM Conference on the Management of Data (SIGMOD)*, Scottsdale, Arizona, USA, May 20-24, 2012 (tutorial).

Grants

Role	Funding Source	Program	Title	Amount	Duration
Lead PI	QNRF ¹	NPRP	Detecting Fraudulent Messages with Strong Privacy Protection	Pending (est. US\$ 800k)	2017-2019
Lead PI	QNRF	NPRP	Real-Time Analytics for Sports Video Streams	US\$717k	2016–2018
Co-PI	Samsung	GRO	Discovering Patterns from Smart Device Users with Differential Privacy	US\$100k	2016–2017
Co-PI	Samsung	GRO	Collecting and Analyzing Data from Smart Device Users with Strong Privacy Protection	US\$100k	2015–2016
Co-PI	A*STAR ²	HCCS	Scalable, Real-Time Analytics on Challenging Data	US\$1.3m	2014–2016

¹QNRF: Qatar National Research Fund

²A*STAR: Agency for Science, Technology and Research, Singapore

Teaching

COMP343: Multimedia Computing, HKUST, fall 2009

- Main contents: digital audio, image and video fundamentals, representation, storage, and processing.

CZ4033: Advanced Data Management, NTU, spring 2014

- Main contents: distributed data management, data management in the cloud, column-oriented DBMS, graph data management.

ICT605: Applied Data Analytics. HBKU, spring 2016 (as CSE505), spring 2017

- Main contents: classification, regression, outlier detection, principal component analysis.

CPEG280: Introduction to Multimedia Computing. HBKU, summer 2016

- Main contents: MIDI, digital audio, image and video processing.

DSEG633: Advanced Data Management Systems. HBKU, fall 2016, fall 2017

- Main contents: data warehouse, distributed databases, NoSQL.

CPEG110: Principles of Computing. HBKU, spring 2017

- Main contents: Python programming, introductory topics on data structures, algorithms, image processing, computer security and networking.

Current Team

Postdocs at HBKU

Shiping Wen Postdoc
Biao Luo Postdoc

Students at HBKU

Mohammed Amro MS Data Science
Ameema Zainab MS Data Science
Mansoor Al-Thani MS Data Science
Alanood Al-Kuwari MS Data Science
Abdulaziz Al-Homid MS Data Science (with Abdelkader Baggag@QCRI)
Shaymaa Khalifa MS Data Science (with Bhiksha Raj@CMU)
Wael Radwan MS Data Science (with Bhiksha Raj@CMU)

Selected Professional Services

Organizer: IEEE ICDE Workshop on Differentially Private Data Publication and Analysis (*PrivDB*), 2013. As co-chair.
Asia-Pacific Web Conference (*APWeb*), 2015. As publication co-chair.

PC member: IEEE International Conference on Data Engineering (*ICDE*), 2013, 2015, 2017.
ACM SIGMOD Conference on Management of Data (*SIGMOD*), 2013, 2018.
ACM Conference on Information and Knowledge Management (*CIKM*), 2017.

Journal referee: ACM Transactions on Database Systems (*TODS*).
ACM Transactions on Information and System Security (*TISSEC*).
VLDB Journal (*VLDB J.*).
ACM Transactions on the Web (*TWEB*).
IEEE Transactions on Dependable and Secure Computing (*TDSC*).
IEEE Transactions on Knowledge and Data Engineering (*TKDE*).
IEEE Transactions on Mobile Computing (*TMC*).

Other role(s): Elsevier Information and Processing Management (*IPM*), as Editorial Board member (since 2017).