Rahul Shah

Roy Paul Daniels Distinguished Professor
Division of Computer Science and Engg, Louisiana State University,
Patrick Taylor Hall, LA 70803
http://www.csc.lsu.edu/~rahul/ e-mail: rahul@csc.lsu.edu

RESEARCH INTERESTS

Area: Algorithms, Data Structures, Theoretical Computer Science
Applied Areas: Database Systems, Information Retrieval, Data Compression, Bioinformatics.
Interests: developing theory and designing algorithms for emerging applications**.**
Specific Research Topics: Indexing Data Structures for Compressed Text/Sequence (String) Data, Uncertain Data, (Parallel) Disk-Based BigData.

EDUCATION

Doctor of Philosophy (Computer Science) May 2002
Rutgers University, New Brunswick, NJ

Master of Science (Computer Science) January 1999
Rutgers University, New Brunswick, NJ

Bachelor of Technology (Comp. Sci. & Engg.) May 1997
Indian Institute of Technology, Bombay, India

WORK EXPERIENCE

Position	University/Company	Start Date	End Date
Professor	Louisiana State University	August 2018	-
Program Director	National Science Foundation	July 2015	June 2019
Associate Professor	Louisiana State University	August 2013	July 2018
Assistant Professor	Louisiana State University	August 2007	July 2012
Research Assistant Professor	Purdue University	May 2005	July 2007
Research Staff Member	IBM India Research Lab	August 2004	May 2005
Visiting Assistant Professor	Purdue University	Oct. 2002	July 2004
Summer Intern	Telcordia Technologies (BellCore)	May 2000	Sept. 2000
Summer Intern	Bell Communications Research (BellCore)	May 1998	August 1998
Instructor/Teaching Assistant	Rutgers University	Fall 1997	Spring 2002

PROFESSIONAL SERVICES

- Senior Member, ACM
- Program Director, Algorithmic Foundations, CCF, CISE
    a. Responsible for AF Core programs, CAREER, CRII
    b. Cross-cutting programs:
        1. BIGDATA
        2. Research Experience for Undergraduates (REU) Sites
        3. Smart and Connected Communities (SCC)
        4. Algorithms in the Field (AitF)
        5. Transdisciplinary Research in Prin. of Data Science (TRIPODS)
        6. Education: NSF Research traineeship (NRT) / Innovations in Graduate Edu.(IGE)

- NSF Panelist 2011, 2012, 2013, 2020
- DARPA Reviewer 2016, 2017
- Associate Editor, PeerJ Computer Science (Algorithms Area), 2015 – date
- Program Committees
    - ACM-SIAM Symposium on Discrete Algorithms SODA 2018
    - Combinatorial Pattern Matching CPM 2018, 2017, 2012, 2010
    - Data Compression Conference DCC 2023, 2022, 2021, 2020, 2019, 2018, 2017, 2016
    - International workshop on Combinatorial Algorithms IWOCA 2019
    - International Symposium on Algorithms and Computation ISAAC 2015
    - String Processing and Information Retrieval SPIRE 2020, 2014, 2013, 2012, 2007
    - ACM SIGSPATIAL Workshop on Querying Uncertain and Spatio-Temporal Data 2012, 2009
    - Int. Conf. on Algorithms and Architectures in Parallel Processing ICA3PP 2013, 2013, 2011, 2010
    - International Conference on Parallel and Distributed Systems ICPADS 2011
- External Reviewer for Conferences: SODA, ESA, ISAAC, LATIN, FSTTCS, WADS, ICDE, SIGMOD, VLDB (over multiple years for many of these conferences)
- External Reviewer for various Journals (including):
  Algorithmica, Journal of Discrete Algorithms, Journal of Classification, TPDS, VLDB Journal, TKDE, Information and Computation, ACM Transactions on Algorithms (TALG), IEEE Transactions on Systems, Man and Cybernetics, Theoretical Computer Science, Information Processing Letters, ACM Trans. On Spatial Algorithms and Systems
- Graduate Student Supervision: 2 **academic** placements + 3 students placed at **Facebook/Google**
    - Sharma Thankachan (PhD 2013, assistant professor at University of Central Florida)
    - Manish Patil (PhD 2014, research scientist at Facebook)
    - Sudip Biswas (PhD 2015, software engineer at Facebook)
    - Arnab Ganguly (PhD 2017, assistant professor at University of Wisconsin-White Water)
    - Xuanting Cai (MS Thesis 2014, Google)
    - Sabrina Chandrasekaran (MS Thesis 2010, Amazon)
    - Emran Chowdhury (MS Thesis, Amazon)
    - Ambujam Krishnan (MS Thesis, 2015)
    - Phani Siddapareddy (MS Project, 2009)
    - Rasagnya Ryali (MS Project, 2011)

GRANTS

Funded continuously by NSF programs since 2006

- EAGER: CCF: AF: Sublinear Data Structures for Approximate Queries
  Funding Agency: National Science Foundation CCF – 2137057
  Duration: October 2021 – September 2023
  Role: PI
  Amount: $225,768

- IPA Award (Non-research grant supporting only 20% research effort)
  Funding Agency: National Science Foundation 1547616
  Duration: July 2015 – June 2019
  Role: PI
  Amount: ~$800,000

- AF: Small: Compact Data Structures for String Matching and Retrieval
  Funding Agency: National Science Foundation CCF – 1527435
  Duration: June 2015 – May 2021
  Role: PI
  Amount: $225,787

- AF: III: Small: Space-Efficient Frameworks for Multi-pattern Matching in Text Streams
  Funding Agency: National Science Foundation CCF - 1218904
  Duration: October 2012 – September 2016
  Role: PI
  Amount: $184,261

- AF:DC: Collaborative Research: Pattern Matching for Massive Data Sets
  Funding Agency: National Science Foundation CCF – 1017623
  Duration: August 2010 – July 2013
  Role: PI
  Total amount: $500,000

- Performance Models and Systems Optimization for Disk-Bound Applications
  Funding Agency: National Science Foundation CCF – 0621457
  Duration: October 2006 – September 2010
  Role: co-PI
  Mithuna Thottethodi (PI), Vijay Pai (co-PI), Rahul Shah (co-PI), TN Vijaykumar (co-PI), Jeffrey Vitter (co-PI)
  Total amount: $889,788
  Co-PI Shah's budget: $325,000

- Space Saving String Retrieval System
  Agency: Louisiana State University Summer Research Grant
  Duration: July 2008
  Role: PI
  Amount: $5,000

PUBLICATIONS

Notes/Highlights:

- For most papers, the authors are listed in alphabetical order as is the norm in the theoretical computer science community.
- H-Index 27; Citations ~ 3000; > 100 publications ; > 50% in A* or A venues
- Best student paper award at top conference (ESA 2001)
- Major Venues (Theory): JACM, FOCS, SODA, ICALP, ESA, ISAAC, SPAA, Algorithmica, TCS
- Major Venues (Applied): INFOCOM, BMC Genomics, SIGIR, SIGMOD, VLDB, ICDE, PODS, ICDT
- Frequent area specific venues: DCC, CPM, SPIRE

Journal Papers

1. Sharma Thankachan, Rahul Shah, Jeffrey Scott Vitter, Cheng Sheng, *Ranked Document Retrieval in External Memory*, ACM Transactions on Algorithms, 2022

2. Arnab Ganguly, Daniel Gibney, Sharma V. Thankachan, Rahul Shah,
   I/O-optimal categorical 3-sided skyline queries. Theoretical Computer Science, 896*, 132-144, 2021*

3. Arnab Ganguly, Wing-Kai Hon, Kunihiko Sadakane, Rahul Shah, Sharma V. Thankachan, Yiling Yang,
   A framework for designing space-efficient dictionaries for parameterized and order-preserving matching, Theoretical Computer Science, 854, 52-62, 2021

4. Arnab Ganguly, Wing-Kai Hon, Rahul Shah, Kunihiko Sadakane, Sharma V. Thankachan, Yilin Yang,
   Succinct Dictionaries for Parameterized and Order-Preserving Matching, accepted pending minor revisions, Theoretical Computer Science, 2020.

5. Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Succinct Non-overlapping Indexing*, Algorithmica, 82(1), 107-117, 2020.

6. J. Ian Munro, Gonzalo Navarro, Rahul Shah, Sharma V. Thankachan, *Ranked document selection,* Theoretical Computer Science 812: 149-159, 2020.

7. Paniz Abedin, Arnab Ganguly, Wing-Kai Hon, Kotaro Matsuda, Yakov Nekrich, Kunihiko Sadakane, Rahul Shah, Sharma V. Thankachan,*A linear-space data structure for range-LCP queries in poly-logarithmic time.* Theoretical Computer Science 822*:* 15-22 2020.

8. Arnab Ganguly, Manish Patil, Rahul Shah, Sharma V. Thankachan, *A Linear Space Data Structure for Range LCP Queries,* Fundam. Inform. 163(3), 245-251, 2018.

9. Wing-Kai Hon, Tak-Wah Lam, Rahul Shah, Sharma V. Thankachan, Hing-Fung Ting, Yilin Yang, Dictionary Matching with a Bounded Gap in Pattern or in Text, in Algorithmica 80(2), 698-713, 2018.

10. Sudip Biswas, Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Space-efficient indexes for forbidden extension queries.* Journal of Discrete Algorithms 50, 23-35, 2018.

11. Sudip Biswas, Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Ranked document retrieval for multiple patterns,* Theoreretical Computer Science. 746, 98-111, 2018.

12. Sudip Biswas, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, *Position-Restricted Substring Searching Over Small Alphabets,* Journal of Discrete Algorithms, 2017.

13. Arnab Ganguly; Wing-Kai Hon; Rahul Shah; Sharma V. Thankachan, *Space-Time Trade-offs for Finding Shortest Unique Substrings and Maximal Unique Matches*, Theoretical Computer Science (TCS), 700:75-88, 2017.

14. J. Ian Munro, Gonzalo Navarro, Jesper Sindahl Nielsen, Rahul Shah, Sharma V. Thankachan, *Top-k Term-Proximity in Succinct Space*, Algorithmica 78(2), 379-393, 2017.

15. Stephane Durocher, Rahul Shah, Matthew Skala, Sharma V. Thankachan, *Linear-Space Data Structures for Range Frequency Queries on Arrays and Trees*, Algorithmica 74(1), 344-366, 2016.

16. Sudip Biswas, Manish Patil, Rahul Shah, Sharma V. Thankachan, *Succinct indexes for reporting discriminating and generic words*, Theoretical. Computer Science, 593, 165-173, 2015.

17. Wing-Kai Hon, Tsung-Han Ku, Tak Wah Lam, Rahul Shah, Siu-Lung Tam, Sharma V. Thankachan, Jeffrey Scott Vitter, *Compressing Dictionary Matching Index via Sparsification Technique*, Algorithmica 72(2), 515-538, 2015.

18. Yu-Feng Chien, Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *Geometric BWT: Compressed Text Indexing via Sparse Suffixes and Range Searching*, in Algorithmica 2015.

19. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *Space-Efficient Frameworks for Top-k String Retrieval*, Journal of ACM, 2014.

20. Wing-Kai Hon, Manish Patil, Rahul Shah, and Sharma V. Thankachan, *Compressed Property Suffix Trees*, in Information and Computation, 232, 10-18, 2013.

21. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *Faster Compressed Dictionary Matching*, Theoretical Computer Science, 475, 113-119, 2013.

22. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter**,** *Compressed text indexing with wildcards***,** Journal of Discrete Algorithms, 19, 23-29, 2013.

23. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *On Position Restricted Substring Searching in Succinct Space*, Journal of Discrete Algorithms, 17, 109-114, 2012.

24. Manish Patil, Rahul Shah, Sharma V. Thankachan, *Succinct Representations of Weighted Trees Supporting Path Queries*, Journal of Discrete Algorithms, 17, 103-108, 2012.

25. Wing-Kai Hon, Tak Wah Lam, Rahul Shah, Siu-Lung Tam, Jeffrey Scott Vitter, *Cache-oblivious index for approximate string matching*, In Theoretical Computer Science, 412(29), 2011, 3579-3588.

26. M. Oguzhan Külekci, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, Bojian Xu, *PSI-RA: A parallel sparse index for read alignment on genomes*, In BMC Genomics, 12:S2, 2011,1-7.

27. Wing-Kai Hon, Manish Patil, Rahul Shah, Shih-Bin Wu, *Efficient index for retrieving top-k most frequent documents*, In Journal of Discrete Algorithms, 8(4), 2010, 402-417.

28. Yuni Xia, Sunil Prabhakar, Shan Lei, Reynold Cheng, Rahul Shah, *Indexing Constantly Changing Data with Mean-Variance Tree*, In International Journal of High Performance Computing and Networking (IJHPCN), 5(4), 2008, 263-272.

29. Ankur Gupta, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Compressed Data Structures: Dictionaries and Data-Aware Measures,* In Theoretical Computer Science (TCS), 387(3), 2007*,* 313-331.

30. Ihab F. Ilyas, Hicham Elmongui, Walid G. Aref, Ahmed K. Elmagarmid, Rahul Shah, Jeffrey Scott Vitter, *Adaptive Rank-aware Query Optimization in Relational Databases*, In ACM Transactions on Database Systems (TODS) 31(4), 2006, 1257-1304.

31. Rahul Shah, Martin Farach-Colton, *On the Complexity of Ordinal Clustering*, In Journal of Classification, 23(1), 2006, 79-102.

32. Rahul Shah, Zulfikar Ramzan, Ravi Jain, Raghu Dendukuri, Farooq Anjum, *Efficient Dissemination of Personalized Information Using Content-Based Multicast,* In IEEE Transactions on Mobile Computing (TMC), 3(4), 2004, 394-408.

Conference Papers

1. Arnab Ganguly*,* Rahul Shah*,* Sharma V. Thankachan*, Fully Functional Parameterized Suffix Trees in Compact Space*. In International Conference on Automata, Languages and Programming (ICALP) 2022*, 65:1-65:18*

2. Arnab Ganguly, Dhrumil Patel, Rahul Shah, Sharma Thankachan, Order-Isomorphic Text Indexing in Compact Space, In International Conference on Automata, Languages and Programming (ICALP) 2021, *71:1-71:19*

3. Dhrumil Patel*,* Rahul Shah*,* Inverse Suffix Array Queries for 2-Dimensional Pattern Matching in Near-Compact Space. International Symposium on Algorithms and Computation (ISAAC)*, 2021, 60:1-60:14*

4. Daniel Gibney*,* Sharma V. Thankachan*,* Arnab Ganguly*,* Rahul Shah*, I/O Optimal Data Structures for Categorical Range Skyline Queries*. Canadian Conference on Computational Geometry (CCCG) 2019*, 9-16*

5. Arnab Ganguly, J. Ian Munro, Yakov Nekrich, Rahul Shah, Sharma V. Thankachan, *Categorical Range Reporting with Frequencies*, to appear in International Conference on Database Theory (ICDT) 2019.

6. Arnab Ganguly, Wing-Kai Hon, Solon Pissis, Rahul Shah, Sharma V Thankachan, *Parameterized Text Indexing with One Wildcard*, in Data Compression Conference, 2019.

7. Paniz Abedin, Arnab Ganguly, Wing-Kai Hon, Yakov Nekrich, Kunihiko Sadakane, Rahul Shah, Sharma V. Thankachan*, A Linear-Space Data Structure for Range-LCP Queries in Poly-Logarithmic Time*, International Computing and Combinatorics Conference (COCOON), 615-625, 2018

8. Arnab Ganguly, Rahul Shah, Sharma Thankachan, *Encoding Structural Suffix Trees – Succinctly*, International Symposium on Algorithms and Computation (ISAAC), 35:1-35:13, 2017.

9. Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *pBWT: Achieving Succinct Data Structures for Parameterized Pattern Matching and Related Problems*, In ACM-SIAM Symposium on Discrete Algorithms (SODA), 397-407, 2017.

10. Arnab Ganguly, Wing-Kai Hon, Rahul Shah, *Stabbing Colors in One Dimension,* Data Compression Conference (DCC), 2017.

11. Arnab Ganguly, Wing-Kai Hon, Kunihiko Sadakane, Rahul Shah, Sharma V. Thankachan, Yilin Yang, *Space-Efficient Dictionaries for Parameterized and Order-Preserving Pattern Matching,* In Combinatorial Pattern Matching (CPM), 2:1-2:12, 2016.

12. Sudip Biswas, Manish Patil, Sharma V. Thankachan, Rahul Shah, *Probabilistic Threshold Indexing for Uncertain Strings,* International Conference on Extending Database Technology (EDBT), 401-412, 2016.

13. Arnab Ganguly, Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, *Space-Time Trade-Offs for the Shortest Unique Substring Problem,* International Symposium on Algorithms and Computation (ISAAC), 34:1-34:13, 2016.

14. Arnab Ganguly, Wing-Kai Hon, Rahul Shah, *A Framework for Dynamic Parameterized Dictionary Matching,* Symposium on Algorithms Theory (SWAT), 10:1-10:1, 2016.

15. Sudip Biswas, Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Ranked Document Retrieval with Forbidden Pattern,* Combinatorial Pattern Matching (CPM), 77-88, 2015.

16. Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Succinct Non-overlapping Indexing,* Combinatorial Pattern Matching (CPM), 185-195, 2015.

17. Wing-Kai Hon, Tak Wah Lam, Rahul Shah, Sharma V. Thankachan, Hing-Fung Ting, Yilin Yang, *Dictionary Matching with Uneven Gaps,* Combinatorial Pattern Matching (CPM), 247-260, 2015.

18. Sudip Biswas, Arnab Ganguly, Rahul Shah, *Restricted Shortest Path in Temporal Graphs,* Database amd Expert Systems Applications (DEXA), (1), 13-27, 2015.

19. Sudip Biswas, Arnab Ganguly, Rahul Shah, Sharma V. Thankachan, *Forbidden Extension Queries*, Foundations of Software Technology and Theoretical Computer Science (FSTTCS), 320-335, 2015.

20. Sudip Biswas, Manish Patil, Rahul Shah, Sharma V. Thankachan, *Shared-Constraint Range Reporting,* International Conference on Database Theory (ICDT), 277-290, 2015.

21. Umesh Chandra Satish, Praveenkumar Kondikoppa, Seung-Jong Park, Manish Patil, Rahul Shah, *MapReduce based parallel suffix tree construction for human genome*, 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS), 664-670, 2014

22. J. Ian Munro, Gonzalo Navarro, Jesper Sindahl Nielsen, Rahul Shah, Sharma V. Thankachan, *Top- k Term-Proximity in Succinct Space,* International Symposium on Algorithms and Computation (ISAAC), 169-180, 2014.

23. Manish Patil, Sharma V. Thankachan, Rahul Shah, Yakov Nekrich, Jeffrey Scott Vitter, *Categorical range maxima queries,* Proceedings of the 33rd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, 266-277, 2014.

24. Manish Patil, Rahul Shah, *Similarity joins for uncertain strings,* International Conference on Management of Data (SIGMOD Conference), 1471-1482, 2014.

25. Sudip Biswas, Manish Patil, Rahul Shah, Sharma V. Thankachan, *Succinct Indexes for Reporting Discriminating and Generic Words*, String Processing and Information Retrieval (SPIRE), 89-100, 2014.

26. J. Ian Munro, Gonzalo Navarro, Rahul Shah, Sharma V. Thankachan, *Ranked Document Selection,* 15th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT), 344-356, 2014.

27. Rahul Shah, Cheng Sheng, Sharma V. Thankachan, Jeffrey Scott Vitter, *Top-k Document Retrieval in External Memory,* European Symposium on Algorithms (ESA) 2013, 803-814.

28. Stephane Durocher, Rahul Shah, Matthew Skala, Sharma V. Thankachan, *Linear-Space Data Structures for Range Frequency Queries on Arrays and Trees*, Mathematical Foundations of Computer Science (MFCS) 2013, 325-336.

29. Manish Patil, Xuanting Cai, Sharma V. Thankachan, Rahul Shah, Seung-Jong Park, David Foltz, *Approximate String Matching by Position Restricted Alignment,* EDBT/ICDT Workshops 2013, 384-391

30. Manish Patil, Rahul Shah, Sharma V. Thankachan, *Faster Range LCP Queries*, String Processing and Information Retrieval (SPIRE) 2013, 263-270

31. Manish Patil, Rahul Shah, Sharma V. Thankachan: *Top-k Join Queries: Overcoming the Curse of Anti-Correlation*, IDEAS 2013: 76-85

32. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan: *Space-Efficient Construction Algorithm for the Circular Suffix Tree*, Combinatorial Pattern Matching (CPM) 2013, 142-152

33. Stephane Durocher, Rahul Shah, Matthew Skala, Sharma V. Thankachan: *Top-k Color Queries on Tree Paths*, String Processing and Information Retrieval (SPIRE) 2013: 109-115

34. Sudip Biswas, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan: *Position-Restricted Substring Searching over Small Alphabets*, String Processing and Information Retrieval (SPIRE) 2013, 29-36.

35. Wing-Kai Hon, Manish Patil, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter*: Indexes for Document Retrieval with Relevance*, Space-Efficient Data Structures, Streams, and Algorithms 2013, 351-362.

36. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, *Space-Efficient Construction Algorithm for the Circular Suffix Tree*, Data Compression Conference (DCC) 2013, 496

37. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan , *Faster Compressed Document Retrieval*, Data Compression Conference (DCC) 2013, 341-350.

38. Wing-Kai Hon, Tsung-Han Ku, Chen-Hua Lu, Rahul Shah and Sharma V. Thankachan, *Efficient Algorithm for Circular Burrows-Wheeler Transform,* In Annual Symposium on Combinatorial Pattern Matching (CPM) 2012, 257-268.

39. Sharma V. Thankachan, Rahul Shah, Wing-Kai Hon and Jeffrey Scott Vitter, *Document Listing for Queries With Excluded Pattern*, In Annual Symposium on Combinatorial Pattern Matching (CPM) 2012, 185-195.

40. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, *Towards an Optimal Space-and-Query-Time Index for Top-k Document Retrieval*, In Annual Symposium on Combinatorial Pattern Matching (CPM) 2012, 173-184.

41. Manish Patil, Rahul Shah, Sharma V. Thankachan, *A Truly Dynamic Data Structure for Top-k Queries on Uncertain Data*, In International Conference on Scientific and Statistical Database Management (SSDBM) 2011, 91-108.

42. Wing-Kai Hon, C. H. Lu, Sharma V. Thankachan, and Rahul Shah, *Succinct Indexes for Circular Patterns*, In International Symposium on Algorithms and Computation (ISAAC) 2011, 673-682.

43. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *Compressed Text Indexing with Wildcards*, In International Conference on String Processing and Information Retrieval (SPIRE), 2011, 267-277.

44. Manish Patil, Sharma V. Thankachan, Rahul Shah, Wing-Kai Hon, Jeffrey Scott Vitter, Sabrina Chandrasekaran, *Inverted Indexes for Phrases and Strings*, In ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), 2011, 555-564.

45. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan and Jeffrey Scott Vitter, *Compressed Dictionary Matching with One Error*, In IEEE Data Compression Conference (DCC) 2011, 113-122.

46. Wing-Kai Hon, Manish Patil, Rahul Shah, and Sharma V. Thankachan, *Compressed Property Suffix Trees*, In IEEE Data Compression Conference (DCC) 2011, 123-132.

47. M. Oguzhan Külekci, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, Bojian Xu, *PSI-RA: A Parallel Sparse Index for Read Alignment on Genomes*, In IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2010, 663-668.

48. Wing-Kai Hon, Tsung-Han Ku, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *Faster Compressed Dictionary Matching*, In International Conference on String Processing and Information Retrieval (SPIRE) 2010, 191-200.

49. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *String Retrieval for Multi-pattern Queries*, In International Conference on String Processing and Information Retrieval (SPIRE) 2010, 55-66.

50. Sheng-Yuan Chiu, Wing-Kai Hon, Rahul Shah, and Jeffrey Scott Vitter, *I/O-Efficient Compressed Text Indexes: From Theory to Practice*, In IEEE Data Compression Conference (DCC) 2010, 426-434.

51. Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Space-Efficient Framework for Top-k String Retrieval Problems*, In IEEE Symposium on Foundations of Computer Science (FOCS) 2009, 713-722.

52. Wing-Kai Hon, Tak-Wah Lam, Rahul Shah, Siu-Lung Tam, Jeffrey Scott Vitter, *Succinct Index for Dynamic Dictionary Matching*, In International Symposium on Algorithms and Computation (ISAAC) 2009, 1034-1043.

53. Wing-Kai Hon, Rahul Shah, Shih-Bin Wu*, Efficient Index for Retrieving Top-k Most Frequent Documents*, In International Conference on String Processing and Information Retrieval, (SPIRE) 2009, 182-193.

54. Wing-Kai Hon, Rahul Shah, Sharma V. Thankachan, Jeffrey Scott Vitter, *On Entropy-Compressed Text Indexing in External Memory*, In International Conference on String Processing and Information Retrieval (SPIRE) 2009, 75-89.

55. Wing-Kai Hon, Tak Wah Lam, Rahul Shah, Siu-Lung Tam, Jeffrey Scott Vitter, *Compressed Index for Dictionary Matching*, IEEE Data Compression Conference (DCC) 2008, 23-32.

56. Yu-Feng Chien, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Geometric Burrows-Wheeler Transform: Linking Range Searching and Text Indexing*, In IEEE Data Compression Conference (DCC) 2008, 252-261.

57. Mohamed Y. Eltabakh, Wing-Kai Hon, Rahul Shah, Walid G. Aref, Jeffrey Scott Vitter, *The SBC-tree: an Index for Run-length Compressed Sequences*. In International Conference on Extending Database technology (EDBT) 2008, 523-534.

58. Sarvjeet Singh, Chris Mayfield, Rahul Shah, Sunil Prabhakar, Susanne E. Hambrusch, Jennifer Neville, Reynold Cheng, *Database Support for Probabilistic Attributes and Tuples*, In International Conference on Data Engineering, (ICDE) 2008, 1053-1061.

59. Paolo Ferragina, Roberto Grossi, Ankur Gupta, Rahul Shah, Jeffrey Scott Vitter, *On Searching Compressed String Collections Cache-obliviously*, In ACM Symposium on Principals of Database Systems (PODS) 2008, 181-190.

60. Wing-Kai Hon, Rahul Shah, Peter J. Varman, Jeffrey Scott Vitter, *Tight Competitive Ratios for Parallel Disk Prefetching and Caching*, In ACM Symposium on Parallelism in Algorithms and Architectures (SPAA) 2008, 352-361.

61. Sarvjeet Singh, Chris Mayfield, Rahul Shah, Sunil Prabhakar, Susanne E. Hambrusch, *Query Selectivity Estimation for Uncertain Data*, In International Conference on Scientific and Statistical Database Management (SSDBM) 2008, 61-78.

62. Yinian Qi, Sarvjeet Singh, Rahul Shah, Sunil Prabhakar, *Indexing Probabilistic Nearest-Neighbor Threshold Queries*, In Joint International Workshop on Quality in Databases and Management of Uncertain Data in conjunction with VLDB 2008, 87-102.

63. Sarvjeet Singh, Chris Mayfield, Sunil Prabhakar, Rahul Shah, Susanne Hambrusch, *Indexing Uncertain Categorical Data*, In International Conference on Data Engineering (ICDE) 2007, 616-625.

64. Wing-Kai Hon, Tak-Wah Lam, Alan Tam, Rahul Shah, Jeffrey Scott Vitter, *Cache Oblivious Index for Approximate String Matching*, In Combinatorial Pattern Matching Conference (CPM) 2007, 40-51.

65. Ankur Gupta, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *A Framework for Dynamizing Succinct Data Structures*, In International Conference on Automata, Languages and Programming (ICALP) 2007, 521-532.

66. Reynold Cheng, Sarvjeet Singh, Yuni Xia, Sunil Prabhakar, Rahul Shah, Jeffrey Scott Vitter, *Efficient Join-processing over Uncertain Data*, In International Conference on Information and Knowledge Management (CIKM) 2006, 738-747.

67. Ankur Gupta, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Compressed Dictionaries: Space Measures, Data Sets and Experiments*, In Workshop on Experimental and Efficient Algorithms (WEA) 2006, 158-169.

68. Ankur Gupta, Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Compressed Data Structures: Dictionaries and Data-aware Measures*, In IEEE Data Compression Conference (DCC) 2006, 213-222.

69. Rahul Shah, Peter Varman, Jeffrey Scott Vitter, *On the Online Read-many Parallel Disks Scheduling*, In ACM Symposium on Parallelism in Algorithms and Architecture (SPAA), 2005, 217.

70. Yuni Xia, Sunil Prabhakar, Shan Lei, Reynold Cheng, Rahul Shah, *Indexing Constantly Changing Data with Mean-Variance Tree*, In ACM Symposium on Applied Computing (SAC) 2005, 1125-1132.

71. Reynold Cheng, Yuni Xia, Sunil Prabhakar, Rahul Shah, *Change Tolerant Indexing for Constantly Evolving Data*, In international conference on Data Engineering (ICDE) 2005, 391-402.

72. Reynold Cheng, Yuni Xia, Sunil Prabhakar, Rahul Shah, Jeffrey Scott Vitter, *Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data*, In International Conference on Very Large Databases (VLDB) 2004, 876-887.

73. Rahul Shah, Peter J. Varman, Jeffrey Scott Vitter, *Online Algorithms for Prefetching and Caching on Parallel Disks*, In ACM Symposium on Parallelism in Algorithms and Architecture (SPAA) 2004, 255-264.

74. S. Muthukrishnan, Rahul Shah, Jeffrey Scott Vitter, *Mining Deviants in Time Series Data Streams*, In International Conference on Scientific and Statistical Database Management (SSDBM) 2004, 41-50.

75. Ihab F. Ilyas, Rahul Shah, Walid G. Aref, Jeffrey Scott Vitter, Ahmed K. Elmagarmid, *Rank Aware Query Optimization*, In International Conference on Management of Data (SIGMOD) 2004, 203-214.

76. Thanaa Ghanem, Rahul Shah, Mohammed F. Mokbel, Walid Aref, Jeffrey Scott Vitter, *Bulk Operations for Space Partitioning Trees*, In International Conference on Data Engineering (ICDE) 2004, 29-40.

77. Rahul Shah, Farooq Anjum, Ravi Jain, *Efficient Dissemination of Personalized Information Using Content-Based Multicast*, In 21st Joint Conference of IEEE Computer and Communications Societies (InfoCom) 2002.

78. Rahul Shah, Martin Farach-Colton, *Undiscretized Dynamic Programming: Faster Algorithms for Facility Location and Related Problems on Trees*, In ACM-SIAM symposium on Discrete Algorithms (SODA) 2002, 108-115.

79. Stefan Langerman, Sachin Lodha, Rahul Shah, *Algorithms for Efficient Filtering in Content-Based Multicast*, In European Symposium on Algorithms (ESA) 2001, **Winner of Best Student Paper award**, 428-439.

80. Rahul Shah, Martin Farach-Colton, *On the Midpath Tree Conjecture: a Counter-Example*, In ACM-SIAM Symposium on Discrete Algorithms (SODA) 2001, 208-209.

Other Publications (Non-Peer Reviewed): Book Chapters/Surveys/Plenary/Tech Reports/Thesis

1. Rahul Shah, Document Retrieval on String Collections, Encyclopedia of Algorithms, 2016. (Invited article – nominally reviewed)

2. Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Compression, Indexing, and Retrieval for Massive String Data*, in International Conference on Combinatorial Pattern Matching (CPM), 2010, 260-274. (Invited Plenary – nominally reviewed)

3. Sunil Prabhakar, Rahul Shah, and Sarvjeet Singh, *Indexing Uncertain Data*, In Managing and Mining Uncertain Data, Charu Aggarwal ed. (Springer), Book Chapter, 2009.

4. Wing-Kai Hon, Rahul Shah, Jeffrey Scott Vitter, *Ordered Pattern Matching: Towards Full-Text Retrieval*, CS TR #06-013, Dept. of Computer Science, Purdue University.

5. Rahul Shah, *Faster Algorithms for the k-Median Problem on Trees with Smaller Heights*, CS TR #03-030, Dept. of Computer Science, Purdue University.

6. Rahul Shah, *Undiscretized Dynamic Programming and Ordinal Embeddings*, PhD Thesis, Rutgers University, 2002. (Committee: Martin Farach-Colton (chair), S. Muthukrishnan, Sampath Kannan, Vasek Chvatal)

7. Rahul Shah, *Enumerating Independent Sets in Trees and Chordal Graphs*, Bachelor's Project, IIT Bombay, 1997. (Advisor: A. A. Diwan)

Demonstrations

1. Sarvjeet Singh, Chris Mayfield, Sagar Mittal, Sunil Prabhakar, Susanne E. Hambrusch, Rahul Shah, *Orion 2.0: Native Support for Uncertain Data*, ACM SIGMOD International Conference on Management

of Data (Demo), 2008, 1239-1242.

2. Sarvjeet Singh, Chris Mayfield, Sagar Mittal, Sunil Prabhakar, Susanne E. Hambrusch, Rahul Shah: *The Orion Uncertain Data Management System*. COMAD (Demo) (2008) 273-276.

REFERENCES

Available Upon Request