

Curriculum Vitae

Name: Prabhakar Raghavan
Title: Chief Strategy Officer and Head, Yahoo! Labs
Office address: Yahoo!
701 First Avenue,
Sunnyvale, CA 94089, USA.
Office phone: (408) 349 2326
Internet: pragh@yahoo-inc.com
<http://theory.stanford.edu/~pragh>
Home address: 13636 Deer Trail Court
Saratoga, CA 95070, USA.

Education

<u>Years</u>	<u>University</u>	<u>Department</u>	<u>Degree</u>
1976-81	IIT, Madras	Electrical Engineering	B.Tech
1981-82	U.C. Santa Barbara	Electrical & Computer Engineering	M.S.
1982-86	U.C. Berkeley	Computer Science	Ph.D.

Employment

1981-1986 Graduate Assistantships and Fellowships, UC Berkeley and Santa Barbara.
1986-1994 Research Staff Member, IBM T.J. Watson Research Center.
1994-1995 Manager, Theory of Computing, IBM T.J. Watson Research Center.
1995-2000 Sr. Manager, Computer Science Principles, IBM Almaden Research Center.
2000-2004 Vice-President and Chief Technology Officer, Verity, Inc.
2004-2005 Senior Vice-President and Chief Technology Officer, Verity, Inc.
2005- VP/SVP and Head, Yahoo! Labs.

Additional Appointments

1989 Visiting Instructor, Department of Computer Science, Yale University.
1991 and 1994 Visiting Scientist, IBM Tokyo Research Laboratory.
1995-1997 Consulting Associate Professor, Dept. of Computer Science, Stanford University.
1997 - Consulting Professor, Dept. of Computer Science, Stanford University.

Awards and honors

- 1984 Outstanding Teaching Assistant, Academic Senate, U.C. Berkeley.
- 1985-1986 IBM Doctoral Fellowship
- 1986 Machtey Award for the Best Student Paper, 27th IEEE Symposium on Foundations of Computer Science.
- 1989 IBM Research Division Award for PLA partitioning.
- 2000 Fellow of the IEEE.
- 2000 Best Paper, ACM Symposium on Principles of Database Systems.
- 2000 Best Paper, Ninth International World-Wide Web Conference (WWW9).
- 2002 Fellow of the ACM.
- 2006 Distinguished Alumnus Award, UC Berkeley Division of Computer Science.
- 2008 Member, National Academy of Engineering.
- 2009 *Laurea Honoris Causa*. Honorary degree from the University of Bologna, Italy.

Research Interests

Randomized algorithms, information retrieval, web and data mining, combinatorial optimization.

Publications

Books

1. Randomized Algorithms (with R. Motwani). *Cambridge University Press*, 1995.
2. Introduction to Information Retrieval (with C. D. Manning and H. Schütze). *Cambridge University Press*, 2008.

Book Chapters

1. Randomized Algorithms (with R. Motwani). *The Computer Science and Engineering Handbook*, A.B. Tucker Jr., editor. CRC Press, 1996, 141-160.
2. Randomized approximation algorithms in combinatorial optimization (with R. Motwani and S. Naor). Chapter in *Approximation Algorithms*, D.S. Hochbaum, ed. PWS publishing company, 1996.
3. Randomized Algorithms (with R. Motwani). *Handbook of Discrete and Combinatorial Mathematics*, K.H. Rosen, editor. CRC Press, 1997.
4. An overview of randomized algorithms (with R. Motwani). In *Probabilistic Methods for Algorithmic Discrete Mathematics*, M. Habib, C. McDiarmid, R. Ramirez-Alfonsin and B. Reed (Eds.), pp. 93-115. Springer, 1998.

Full-length papers

1. A Class of Graphs for Processor Interconnection (with S.M. Reddy and J.G. Kuhl). *Proceedings of the IEEE International Conference on Parallel Processing*, May 1983, 154-157.

2. On Estimating the Performance of VLSI Circuits (with C.D. Thompson). *Proceedings of the 1984 MIT Conference on Advanced Research in VLSI*, Cambridge, MA, January 1984, 34-44.
3. Randomized Rounding: Provably Good Algorithms and Algorithmic Proofs (with C.D. Thompson).
 - (a) *Proceedings of the 17th Annual ACM Symposium on Theory of Computing*, May 1985, 79-87. (Appears under the title *Provably Good Routing in Graphs: Regular Arrays*.)
 - (b) *Combinatorica*, vol. 7, no. 4, 1987, 365-374 (invited to a special issue of selected papers from the conference).
4. A Specification Language for Steiner Trees (with A.P-C. Ng and C.D. Thompson). *Proceedings of the 1986 Design Automation Conference*, June 1986, 659-662.
5. Experimental Results for a Linear Program Global Router (with A.P-C. Ng and C.D. Thompson). *Computers and Artificial Intelligence*, vol. 6, no. 3, 1987, 229-242.
6. Deferred Data-Structuring.
 - (a) (with R. Motwani). *Proceedings of the 2nd Annual ACM Symposium on Computational Geometry*, June 1986, 303-312.
 - (b) (with R.M. Karp and R. Motwani). *SIAM Journal on Computing*, vol. 17, no. 5, October 1988, 883-902.
7. Probabilistic Construction of Deterministic Algorithms: Approximating Packing Integer Programs.
 - (a) *Proceedings of the 27th Annual IEEE Symposium on Foundations of Computer Science*, October 1986, 10-18. (This paper was awarded the Machtey Prize for the Best Student Paper).
 - (b) *Journal of Computer and System Sciences*, vol. 37, 1988, 130-143 (invited to a special issue of selected papers from the conference).
8. Multiterminal Global Routing: A Deterministic Approximation Scheme (with C.D. Thompson). *Algorithmica*, vol. 6, 1991, 73-82.
9. Randomized Rounding and Discrete Ham-Sandwich Theorems: Provably Good Algorithms for Routing and Packing Problems. PhD. Thesis, Computer Science Division, U.C. Berkeley, August 1986; Technical Report UCB/CSD 87/312.
10. Parallel Graph Algorithms that are Efficient on Average (with D. Coppersmith and M. Tompa).
 - (a) *Proceedings of the 28th Annual IEEE Symposium on Foundations of Compute Science*, October 1987, 260-270.
 - (b) *Information and Computation*, vol. 81, 1989, 318-333.
11. Randomized Algorithms and Pseudorandom Numbers (with H.J. Karloff).
 - (a) *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, May 1988, 310-321.
 - (b) *Journal of the ACM*, vol. 40, July 1993, 421-453.
12. Energy Consumption in VLSI Circuits (with A. Aggarwal and A.K. Chandra). *Proceedings of the 20th ACM Annual Symposium on Theory of Computing*, Chicago, IL, May 1988, 205-216.
13. Learning in Threshold Networks: A Computational Model and its Applications. *Proceedings of the First Workshop on Computational Learning Theory*, Morgan-Kaufmann, Inc., 1988, 19-27.
14. Integer Programming in VLSI Design.

- (a) Extended abstract of invited lecture in the *Proceedings of the 13th Symposium uber Operations Research*, Paderborn, September 1988.
- (b) *Discrete Applied Mathematics*, vol. 40, 1992, 29-43.
- 15. Program Correctness: Can one test for it? (with M. Blum). *Proceedings of the 11th IFIP World Computer Congress*, August 1989, 127-34. North Holland.
- 16. The Electrical Resistance of a Graph captures its Commute and Cover Times (with A.K. Chandra, W.L. Ruzzo, R. Smolensky and P. Tiwari).
 - (a) *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, May 1989, 574-586.
 - (b) *Computational Complexity*, vol. 6, 312-340, 1997.
- 17. Trading Space for Time in Undirected s-t Connectivity (with A.Z. Broder, A.R. Karlin and E. Upfal).
 - (a) *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, May 1989, 543-549.
 - (b) *SIAM Journal on Computing*, vol. 23, 1994, 324-334.
- 18. Memory Versus Randomization in On-line Algorithms (with M. Snir).
 - (a) *Proceedings of the 16th International Colloquium on Automata, Languages and Programming*, July 1989. Springer-Verlag Lecture Notes in Computer Science 372,687-703.
 - (b) *IBM Journal of Research and Development*, vol. 38, 1994, 683-707.
- 19. Fast Geometric Approximation Techniques and Geometric Embedding Problems (with M.W. Bern, H.J. Karloff and B. Schieber).
 - (a) *Proceedings of the 5th Annual ACM Symposium on Computational Geometry*, June 1989, 292-301.
 - (b) *Theoretical Computer Science*, vol. 106, no.2, Dec. 1992, 265-281.
- 20. Robust Algorithms for Packet Routing on a Mesh.
 - (a) *Proceedings of the 1st ACM Symposium on Parallel Algorithms and Architectures*, June 1989, 344-350.
 - (b) *Mathematical Systems Theory*, 28(1):1-11, January/February 1995.
- 21. Random Walks on Weighted Graphs, and Applications to On-line Algorithms (with D. Coppersmith, P.G. Doyle and M. Snir).
 - (a) *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, May 1990, 369-378.
 - (b) *Journal of the ACM*, vol. 40, July 1993, 454-476.
- 22. Computing with Noisy Information (with U. Feige, D. Peleg and E. Upfal).
 - (a) *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing*, May 1990, 128-137.
 - (b) *SIAM Journal on Computing*, vol. 23, Oct. 1994, 1001-1018.
- 23. Lecture Notes on Randomized Algorithms. Technical Report based on a course taught at the Computer Science Department, Yale University, Fall 1989. *IBM Research Report RC 15340*, IBM Yorktown, January 1990, 170 pp.; also available as Technical Report YALEU/DCS/TR-757, Computer Science Dept., Yale University.
- 24. Random Broadcast in Networks (with U. Feige, D. Peleg and E. Upfal).
 - (a) *Proceedings of the SIGAL International Symposium on Algorithms*, August 1990. Springer-Verlag Lecture Notes in Computer Science 450, 128-137.

- (b) *Random Structures and Algorithms*, vol. 1, No. 4, 1990, 447-460.
25. Time-Space Tradeoffs for Undirected Graph Connectivity (with P. Beame, A. Borodin, W. L. Ruzzo and M. Tompa).
- (a) *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, October 1990, 429-438.
- (b) A Time-Space Tradeoff for Undirected Graph Traversal by Walking Automata. *SIAM Journal on Computing* Volume 28, Number 3, 1051-1072, 1999.
- (c) Time-space tradeoffs for undirected graph traversal by graph automata." *Information and Computation*, vol 130, 101-129, 1996.
26. Asymptotically Tight Bounds for Computing with Faulty Arrays of Processors (with C. Kaklamanis, A.R. Karlin, F.T. Leighton, V. Milenkovic, S. Rao, C.D. Thomborson and A. Tsantilas). *Proceedings of the 31st Annual IEEE Symposium on Foundations of Computer Science*, October 1990, 285-296.
27. On the Parallel Complexity of Game-Tree Evaluation (with A.Z. Broder, A.R. Karlin and E. Upfal). *Proceedings of the 2nd Annual ACM-SIAM Symposium on Discrete Algorithms*, San Francisco, January 1991, 404-413.
28. Navigating in Unfamiliar Geometric Terrain (with A. Blum and B. Schieber).
- (a) *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, May 1991, 494-504.
- (b) *SIAM Journal on Computing*, vol. 26, 110-137, 1997.
29. Competitive Paging with Locality of Reference (with A. Borodin, S. Irani and B. Schieber).
- (a) *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, May 1991, 249-259.
- (b) *Journal of Computer and System Sciences* (special issue containing selected papers from the conference), vol. 50, 244-258, 1995.
30. A Statistical Adversary for On-line Algorithms. *Proceedings of the DIMACS Workshop on On-line Algorithms*, Feb 1991. L.A. McGeoch and D.D.K. Sleator, eds., American Mathematical Society.
31. The robot localization problem in two dimensions (with L. Guibas and R. Motwani).
- (a) *Proceedings of the 3rd Annual ACM-SIAM Symposium on Discrete Algorithms*, Jan. 1992, pp. 259-268.
- (b) *SIAM Journal on Computing*, 26(4), 1120-1138, 1997.
32. A theory of wormhole routing in parallel computers (with S. Felperin and E. Upfal).
- (a) *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 563-572, 1992.
- (b) *IEEE Transactions on Computers*, Jun. 1996, 704-713.
33. Exact Analysis of Hot Potato Routing (with U. Feige). *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 553-562, 1992.
34. Markov Paging (with A. Karlin and S. Phillips).
- (a) *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 208-217, 1992.
- (b) *SIAM Journal on Computing*, 1998.
35. An Experimental Study of Wormhole Routing in Parallel Computers (with S. Felperin and E. Upfal). *Parallel Architectures and Their Efficient Use*. F. Meyer auf der Heide, B.

- Monien, A.L. Rosenberg (Eds.). Lecture Notes in Computer Science 678, Springer-Verlag 1993, pages 156-165.
36. How much can hardware help routing? (with A. Borodin, B. Schieber and E. Upfal).
 - (a) *Proceedings of the 1993 ACM Symposium on Theory of Computing*, 573-582 (1993).
 - (b) *Journal of the ACM*, 44(5):726-741, (1997).
 37. Fast deflection routing for packets and worms (with A. Bar-Noy, B. Schieber and H. Tamaki). *Proceedings of the 1993 ACM Symposium on Principles of Distributed Computing*, 75-86, 1993.
 38. Efficient Routing in All-Optical Networks (with E. Upfal). *Proceedings of the 26th ACM Symposium on Theory of Computing*, 134-143, 1994.
 39. The minimum latency problem (with A. Blum, P. Chalasani, D. Coppersmith, W.R. Pulleyblank, and M. Sudan). *Proceedings of the 26th ACM Symposium on Theory of Computing*, 163-171, 1994.
 40. The Traveling Cameraman Problem, and Applications to Automatic Optical Inspection (with K. Iwano and H. Tamaki). *Proceedings of the 5th International Symposium on Algorithms and Computation*, Springer-Verlag LNCS 834, 29-37, 1994.
 41. Random Walks and Undirected Graph Connectivity: A Survey (with A. Karlin). *Finite Markov Chain Renaissance* (D. Aldous and P. Diaconis, eds.), IMA volumes in mathematics and its applications 72, Springer-Verlag, 95-102, 1995.
 42. Motion Planning on a Graph (with C.H. Papadimitriou, M. Sudan and H. Tamaki). *Proceedings of the 35th IEEE Symposium on Foundations of Computer Science*, 511-520, 1994.
 43. Randomized approximation algorithms in combinatorial optimization. Invited paper in *Proceedings of the 14th Conference on the Foundations of Software Technology and Theoretical Computer Science*, Springer-Verlag LNCS 880, 300-317, 1994.
 44. Motion planning for a steering-constrained robot through moderate obstacles (with P. Agarwal and H. Tamaki). *Proceedings of the 27th ACM Symposium on Theory of Computing*, 343-352, 1995.
 45. Stochastic contention resolution with short delays (with E. Upfal).
 - (a) *Proceedings of the 27th ACM Symposium on Theory of Computing*, 229-237, 1995.
 - (b) *SIAM Journal on Computing*, Volume 28, Number 2, pp. 709-719, 1998.
 46. Randomized query processing for robot motion planning (with L. Kavraki, J-C. Latombe and R. Motwani).
 - (a) *Proceedings of the 27th ACM Symposium on Theory of Computing*, 353-362, 1995.
 - (b) *Journal of Computer and System Sciences* 57(1), 50-60, 1998.
 47. Randomized path planning in large configuration spaces (with J. Barraquand, L. Kavraki, J-C. Latombe, T-Y. Li and R. Motwani).
 - (a) *Proceedings of the 7th International Symposium on Robotics Research*, G. Giralt and G. Hirzinger (eds.), Springer, New York, NY, 249-264, 1996.
 - (b) *International Journal of Robotics Research*, 16(6), 759-774, 1997.
 48. Combinatorial and experimental results for point-matching algorithms (with S. Irani).
 - (a) *Proceedings of the 12th Annual ACM Symposium on Computational Geometry*, 68-77, 1996.
 - (b) *Computational Geometry*, 12, 17-31, 1999.
 49. Randomized Algorithms (with R. Motwani). *ACM Computing Surveys*, 33-37, 1996.

50. Adversarial queuing theory (with A. Borodin, J. Kleinberg, M. Sudan and D.P. Williamson).
 - (a) *Proceedings of the 28th ACM Symposium on Theory of Computing*, 376-385, 1996.
 - (b) *Journal of the ACM*, [48](#)(1): 13-38, 2001.
51. A linear algorithm for deviation detection (with A. Arning and R. Agrawal). *Proceedings of the 2nd Conference on Knowledge Discovery and Data Mining*, 164-169, 1996.
52. Information retrieval algorithms: a survey. Invited paper in *Proceedings of the 7th ACM-SIAM Symposium on Discrete Algorithms*, 1997.
53. Locality-preserving multidimensional hashing (with P. Indyk, R. Motwani and S. Vempala.) *Proceedings of the 29th ACM Symposium on Theory of Computing*, 618-625, 1997.
54. Keyword detection, navigation and annotation in hierarchical text databases (with S. Chakrabarti, B. Dom and R. Agrawal).
 - (a) *Proceedings of the VLDB conference*, 446-455, 1997.
 - (b) *VLDB Journal* 7(3): 163-178 (1998) (appears under the title “Scalable Feature Selection, Classification and Signature Generation for Organizing Large Text Databases into Hierarchical Topic Taxonomies”).
55. Motion via pushing (with P. Agarwal, J-C. Latombe and R. Motwani). *Proceedings of the 1997 International Conference on Robotics and Automation*.
56. Constrained TSP and low-power computing (with M. Charikar, R. Motwani and C. Silverstein). *Proceedings of WADS 97*.
57. Maintaining mortal archives (with J. Kleinberg, R. Motwani and S. Venkatasubramanian). *Proceedings of the 38th IEEE Symposium on Foundations of Computer Science*, 1997.
58. Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text (with S. Chakrabarti, B. Dom, D. Gibson, J. Kleinberg and S. Rajagopalan). *Proceedings of the 7th International World-Wide Web Conference*, 1998; *Computer Networks* 30(1-7): 65-74.
59. Dynamic Schemes for Speculative Execution of Code (with H. Shachnai and M. Yaniv.).
 - (a) Proc. of the 6th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, Montreal, July 1998.
 - (b) *Performance Evaluation*, [53](#)(2), July 2003, 125-142.
60. Automatic Subspace Clustering of High Dimensional Data for Data Mining Applications (with Rakesh Agrawal, Johannes Gehrke, Dimitrios Gunopulos). *Proceedings of ACM SIGMOD*, 1998.
61. A polynomial-time approximation scheme for the Euclidean k -medians problem (with S. Arora and S. Rao). *Proceedings of the ACM Symposium on Theory of Computing*, 1998.
62. Approximation algorithms for segmentation problems (with J. Kleinberg and C.H. Papadimitriou).
 - (a) *Proceedings of the ACM Symposium on Theory of Computing*, 1998.
 - (b) *Journal of the ACM*, March 2004.
63. A probabilistic analysis of latent semantic indexing (with C.H. Papadimitriou, H. Tamaki and S. Vempala).
 - (a) *Proceedings of the ACM Symposium on Principles of Database Systems*, 1998.
 - (b) *Journal of Computer and System Sciences*. Special issue of selected papers from the conference, 2001.

64. Inferring web communities from link topology (with D. Gibson and J. Kleinberg). *Proceedings of the ACM Symposium on Hypertext and Hypermedia*, 1998.
65. Spectral filtering for resource discovery (with S. Chakrabarti, B. Dom, S. Ravi Kumar, S. Rajagopalan and A. Tomkins). *Proceedings of the ACM SIGIR workshop on hypertext information retrieval*, July 1998.
66. Clustering categorical data: an approach based on dynamical systems (with D. Gibson and J. Kleinberg).
 - (a) *Proceedings of the VLDB conference*, 1998.
 - (b) *VLDB Journal* 8(3-4): 222-236, 2000.
67. Recommendation systems: a probabilistic analysis (with S.R. Kumar, S. Rajagopalan and A. Tomkins).
 - (a) *Proceedings of the 39th IEEE Symposium on Foundations of Computer Science*, 1998.
 - (b) *Journal of Computer and System Sciences*, 63(1):42-61, 2001.
68. A microeconomic view of data mining (with J. Kleinberg and C.H. Papadimitriou). *Data Mining and Knowledge Discovery*, 2:4, 311-324, 1998.
69. On targeting Markov segments (with M. Charikar, S.R. Kumar, S. Rajagopalan and A. Tomkins). *Proceedings of the ACM Symposium on Theory of Computing*, 1999.
70. Trawling emerging cyber-communities automatically (with S.R. Kumar, S. Rajagopalan and A. Tomkins). *Proceedings of the 8th International World-Wide Web Conference*, 1999.
71. Hypersearching the web (with S. Chakrabarti, B. Dom, D. Gibson, J. Kleinberg, S.R. Kumar, S. Rajagopalan and A. Tomkins). *Scientific American*, June 1999.
72. The Web as a graph: measurements, models, and methods (with J.M. Kleinberg, S.R. Kumar, S. Rajagopalan and A. S. Tomkins). *Proceedings of the Fifth Annual International Computing and Combinatorics Conference*, 1999. Springer-Verlag Lecture Notes in Computer Science 1627, 1-17, 1999.
73. Computing on data streams (with M.R. Henzinger and S. Rajagopalan). In *External Memory Algorithms*, DIMACS Series in Discrete Mathematics 50, American Mathematical Society, 107-118, 1999.
74. Extracting large-scale knowledge bases from the web (with S.R. Kumar, S.Rajagopalan and A. Tomkins). *Proceedings of the 25th International Conference on Very Large Databases*, VLDB1999.
75. Mining the Link Structure of the World Wide Web (with S.Chakrabarti, B.Dom, S.R. Kumar, D.Gibson, J.Kleinberg, S.Rajagopalan and A.S. Tomkins). *IEEE Computer*, August 1999.
76. Random walks with “back buttons” (with R. Fagin, A. Karlin, J. Kleinberg, S. Rajagopalan, R. Rubinfeld, M. Sudan and A. Tomkins).
 - (a) *Proceedings of the ACM Symposium on Theory of Computing*, 2000.
 - (b) *Annals of Applied Probability*, 2001.
77. Query strategies for priced information (with M. Charikar, R. Fagin, V. Guruswami, J.M. Kleinberg and A. Sahai).
 - (a) *Proceedings of the ACM Symposium on Theory of Computing*, 2000.
 - (b) *Journal of Computer and System Sciences* 64(4):785-819 (2002).
78. Graph structure of the Web: a survey. Invited paper in *Proceedings of Latin Informatics*, 2000, Springer-Verlag Lecture Notes in Computer Science, 123-125.
79. Auditing Discrete Attributes (with J. Kleinberg and C. Papadimitriou).

- (a) *Proceedings of the ACM/SIGMOD Symposium on Principles of Database Systems, 2000.*
This paper won the Best Paper award at the Conference.
- (b) *Journal of Computer and System Sciences* 66(1): 244-253, 2003.
80. Topic distillation and spectral filtering (with S. Chakrabarti, B. Dom, D. Gibson, S. Ravi Kumar, S. Rajagopalan and A. Tomkins). *Artificial Intelligence Review*, 13:409-435, 1999.
81. Graph structure in the Web (with A.Z. Broder, R. Kumar, F. Maghoul, S. Rajagopalan, R. Stata, A. Tomkins and J. Wiener). *Proceedings of the Ninth International World-Wide Web Conference WWW9*, May 2000. This paper won the Best Paper award at the Conference.
82. Stochastic models for the web graph (with S. Ravi Kumar, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal). *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science, 2000.*
83. On semi-automated Web taxonomy construction (with S. Ravi Kumar, A. Tomkins and S. Rajagopalan). *Proceedings of the Fourth International Workshop on the Web and Databases, WebDB 2001.*
84. On the value of private information (with J. Kleinberg and C.H. Papadimitriou). *Theoretical Aspects of Rationality and Knowledge (TARK) 2001.*
85. Multidimensional Cube Packing (with Y. Kohayakawa, F.K. Miyazawa and Y. Wakabayashi). *Brazilian Symposium on Graphs, Algorithms and Combinatorics, 2001.*
86. Building low-diameter P2P networks (with G. Pandurangan and E. Upfal). *Proceedings of the IEEE Symposium on Foundations of Computer Science, 492-499, 2001.*
87. Competitive recommendation systems (with P. Drineas and I. Kerenidis). *Proceedings of the ACM Symposium on Theory of Computing, 82-90, 2002.*
88. Social Networks: from the Web to the Enterprise. *IEEE Internet Computing*, pp. 91-94, Jan/Feb 2002.
89. Structured and unstructured search in enterprises. *IEEE TCDE Bulletin, 24(4) 2001, 15-18.*
90. Using PageRank to characterize web structure (with G. Pandurangan and E. Upfal).
(a) *Proceedings of COCOON*, Springer-Verlag Lecture Notes in Computer Science 2387, 330-339, 2002.
(b) *IEEE Journal on Selected Areas in Communication, 2003.*
91. Mining significant associations in large-scale text corpora (with P. Tsaparas). *Proceedings of the IEEE International Conference on Data Mining, 2002.*
92. Symphony: distributed hashing in a small world (with G. Manku and M. Bawa). *Proceedings of the Fourth USENIX Symposium on Internet Technologies and Systems, 2003.*
93. The web and social networks (with R. Kumar, S. Rajagopalan and A. Tomkins). *IEEE Computer*, November 2002.
94. A Deterministic $(2-2/(k+1))^n$ Algorithm for k-SAT Based on Local Search (with E. Dantson, A. Goerd, E.A. Hirsch, R. Kannan, J. Kleinberg, C.H. Papadimitriou and U. Schöning). *Theoretical Computer Science* 289/1, 2002, pp. 69-83.
95. On the bursty evolution of Blogspace (with R. Kumar, J. Novak and A. Tomkins). *Proceedings of the 2003 International Conference on the WWW.*
96. SETS: Search enhanced by Topic Segmentation (with G. Manku and M. Bawa). *Proceedings of the 26th Annual ACM SIGIR Conference, 2003.*
97. Propagation of Trust and Distrust (with R.V. Guha, R. Kumar and A. Tomkins). *Proceedings of the 2004 International Conference on the WWW, 403-412.*

98. Anti-Aliasing on the Web (with J. Novak and A. Tomkins). *Proceedings of the 2004 International Conference on the WWW*, 30-39.
99. Efficiency-Quality Tradeoffs for Vector Score Aggregation (with P. Singitham and M. Mahabashyam). *Proceedings of VLDB 2004*.
100. Structure and evolution of Blogspace (with R. Kumar, J. Novak and A. Tomkins). *Communications of the ACM*, December 2004.
101. Encoding XML in vector spaces (with V. Kakade). *Proceedings of the European Conference on Information Retrieval*, March 2005.
102. Geographic Routing in Social Networks (with D. Liben-Nowell, J. Novak, R. Kumar and A. Tomkins). In [*Proceedings of the National Academy of Sciences*](#), 102(33):11623-11628, August 2005.
103. Query Incentive Networks (with J.M. Kleinberg). *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2005.
104. Visualizing tags over time (with M. Dubinko, J. Novak, R. Kumar and A. Tomkins). *Proceedings of WWW*, 2006.
105. Finding Near Neighbors Through Cluster Pruning (with F. Chierichetti, A. Panconesi, M. Sozio, A. Tiberi and E. Upfal). *Proceedings of the ACM Symposium on Principles of Database Systems*, 2007.
106. Online story scheduling in web advertising (with A. Dasgupta, A. Ghosh, H. Nazerzadeh). *Proceedings of the ACM Symposium on Discrete Algorithms*, 2009.
107. Compressed web indexes (with F. Chierichetti and R. Kumar). *Proceedings of WWW*, 2009.
108. On compressing social networks (with F. Chierichetti, R. Kumar, S. Lattanzi, A. Panconesi and M. Mitzenmacher). *Proceedings of ACM KDD*, 2009.
109. Models for the compressible Web (with F. Chierichetti, R. Kumar, S. Lattanzi, M. Mitzenmacher and A. Panconesi). *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2009.
110. Optimizing Two-Dimensional Search Results Presentation (with F. Chierichetti and R. Kumar). *Proceedings of Web Search and Data Mining*, 2011.
111. An Algorithmic Treatment of Strong Queries (with R. Kumar and S. Lattanzi). *Proceedings of Web Search and Data Mining*, 2011.
112. Markov Layout (with F. Chierichetti and R. Kumar). *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 2011.

Short publications

1. Optimal Time Bounds for Some Proximity Problems (with A. Aggarwal, H. Edelsbrunner and P. Tiwari). *Information Processing Letters*, vol. 42, 1992, 55-60.
2. Timing-Driven Partitioning of PLAs (with G.S. Ditlow). *IBM Technical Disclosure Bulletin*, vol. 31, no. 9, February 1989. This work won an IBM Research Division Award.
3. Multidimensional On-line Bin Packing: Algorithms and Worst-Case Analysis (with D. Coppersmith). *Operations Research Letters*, vol. 8, no. 1, 1989, 17-20.
4. Review of the book *Complexity in Information Theory*, Y.S. Abu-Mostafa, ed., in the *IEEE Communications Magazine*, vol. 27, no. 6, June 1989.
5. Deferred data structure for the nearest neighbor problem (with A. Aggarwal). *Information Processing Letters* 40, pp. 119-122, 1991.

6. The worst-case running time of the random simplex algorithm is exponential in the height (with A.Z. Broder, M.E. Dyer, A.M. Frieze and E. Upfal). *Information Processing Letters* 56, pp. 79-81, 1995.
7. Technology 2006: A vision of the electric utility in ten years (with F. LeGoues). *Proceedings of the 11th Conference of the Electric Power Supply Industry*, 1996.
8. A theory of computing symposia (with A. Blum). ACM SIGACT News, 1998.
9. Web search using automated classification (with C. Chekuri, M. Goldwasser and E. Upfal). Poster paper in the 6th World-Wide Web conference, Santa Clara, 1997.
10. Information technology for the utilities in a new century (with F. LeGoues). *Proceedings of the 12th Conference of the Electric Power Supply Industry*, 1998.
11. Finding anything in the billion-page web: are algorithms the key? Panel Moderator's abstract, *Proceedings of the 8th World-Wide Web conference*, 1999.
12. Structural Analysis of the World Wide Web (with D. Gibson and J. Kleinberg). Position paper at the *WWW Consortium Web Characterization Workshop*, November 1998.
13. Search: beyond the keyword interface. Panel Moderator's abstract, *Proceedings of the 10th World-Wide Web conference*, 2001.
14. Navigating large-scale semi-structured data in business portals (with M. Abrol, N. Latache, U. Mahadevan, J. Mao, R. Mukherjee, P. Raghavan, M. Tourn, J. Wang, G. Zhang). *Proceedings of VLDB (Industrial Track)*, 2001.
15. Review of the book *The Discrepancy Method* by Bernard Chazelle, in *SIAM Review* 43/3, pp. 570-571, 2001.
16. Information retrieval for enterprise content. *Upgrade* 8(3), June 2002. Also appears in *Upgrade's* Spanish edition *Novatica* under the title "Recuperacion de Informacion de contenidos empresariales".
17. Lord of the Files – Social Networks, Portals and Knowledge Management. *Journal of Public Inquiry*, Spring-Summer 2002, 20-23.
18. Thematic Mapping – from unstructured documents to taxonomies (with C. Chung, R. Lieu, J. Liu, A. Luk and J. Mao). *Proceedings of the ACM International Conference on Information and Knowledge Management (Industrial Track)*, 2002.
19. More on random walks, electrical networks and the harmonic k -server algorithm (with Y. Bartal, M. Chrobak and J. Noga). *Information Processing Letters* 84 (2002), 271-276.
20. Review of the book *Nexus: Small Worlds and the Ground-breaking Science of Networks* by Mark Buchanan, in *American Scientist* 91, Jan-Feb 2003.
21. Review of the book *Six Degrees: The Science of a Connected Age* by Duncan J. Watts, in *American Scientist* 94, Jul-Aug 2003.
22. Industrial session paper *The new frontiers of Web search: going beyond the 10 blue links* (with R. Baeza-Yates, A. Broder and Y. Maarek), ACM SIGIR 2010.

Issued U.S. patents

1. US5796827: System and method for near-field human-body coupling for encrypted communication with identification cards (with D. Coppersmith and T.G. Zimmerman), Aug 1998.
2. US5884305: System and method for data mining from relational data by sieving through iterated relational reinforcement (with J.M. Kleinberg), March 1999.

3. US6003029: Automatic subspace clustering of high dimensional data for data mining applications (with R. Agrawal, J. Gehrke and D. Gunopulos), Dec 1999.
4. US6233575: Multilevel taxonomy based on features derived from training documents classification using Fisher values as discrimination values (with R. Agrawal, S. Chakrabarti and B. Dom), May 2001.
5. US6334131: Method for cataloging, filtering and relevance ranking frame-based hierarchical information structures (with S. Chakrabarti, B. Dom, D. Gibson, S. Rajagopalan, S. Ravikumar and A. Tomkins), Dec 2001.
6. US6336112: Method for interactively creating an information database including preferred information elements, such as, preferred-authority, world wide web pages (with S. Chakrabarti, B. Dom, D. Gibson, S. Rajagopalan, S. Ravikumar and A. Tomkins), Jan 2002.
7. US6356899: Method for interactively creating an information database including preferred information elements, such as, preferred-authority, world wide web pages (with S. Chakrabarti, B. Dom, D. Gibson, S. Rajagopalan, S. Ravikumar and A. Tomkins), Mar 2002.
8. US6457047: Application caching system and method (with A. Chandra, N. Latache and J. Mao), Sep 2002.
9. US6728704: Method and apparatus for merging results lists from multiple search engines (with J. Mao, R. Mukherjee and P. Tsaparas), April 2004.
10. US6757686: Method and apparatus for representing database and query information using interval hash tree (with T. Sayeda-Mahmood and N. Megiddo), June 2004.
11. US6792419. System and method for ranking hyperlinked documents based on a stochastic backoff processes, Sep 2004.
12. US6738764: Apparatus and method for adaptively ranking search results (with JC. Mao, M. Abrol, R. Mukherjee, M. Tourn), May 2004.
13. US6886129: Method and system for trawling the World-wide Web to identify implicitly-defined communities of web pages (with R. Kumar, S. Rajagopalan and A. Tomkins), April 2005.
14. US6996572: Method and system for filtering of information entities (with S. Chakrabarti, B. Dom, D. Gibson, J.M. Kleinberg and S. Rajagopalan), February 2006.
15. [US7581184: System and method for visualizing the temporal evolution of object metadata](#) (with A. Tomkins, J. Novak. R. Kumar, J. Magnini and M. Dubinko), May 2006.
16. US7624104: User-sensitive pagerank (with P. Berkhin, U. Fayyad and A. Tomkins), June 2006.
17. [US7739275](#): System and method for selecting object metadata evolving over time (with M. Dubinko, S. Ravikumar, J. Magnani, J. Novak and A. Tomkins). May 2006.

Professional and University service

Teaching Interests

Design and analysis of algorithms, information retrieval and data mining, computer architecture and hardware, combinatorics, probabilistic analysis.

Teaching experience

Graduate courses

CS 661: Randomized Algorithms, FALL 1989 (Computer Science Department, Yale University).
CS 347: Information retrieval and distributed databases, SPRING 2001 (Computer Science Department, Stanford University).
CS 276A-B / SYMBSYS 239IRA-B / LING 239IRA-B: Text Information Retrieval, Mining, and Exploitation, AUTUMN 2002 - WINTER 2003 (Computer Science Department, Stanford University).
CS 276: Text retrieval and mining. AUTUMN 2003 (Computer Science Department, Stanford University).
CS 276A-B: Information Retrieval, Mining, and Exploitation, AUTUMN 2004 - WINTER 2005 (Computer Science Department, Stanford University).
CS 276, AUTUMN 2005-2009 (Computer Science Department, Stanford University).

Undergraduate course

CS99E: Great Ideas in Computer Science, AUTUMN 1998 Freshman Honors course, (Computer Science Department, Stanford University).

Thesis committees

Yiftach Ravid, PhD, Tel-Aviv University, 1992.
Ran El-Yaniv, PhD, University of Toronto, 1994.
Prasad Chalasani, PhD, Carnegie-Mellon University, 1994.
Alberto Pedrotti, PhD, Scuola Normale Superiore, Pisa, 1998.
Vinay Kakade, MS, Stanford University, 2004.
Mayank Bawa, PhD, Stanford University, Spring 2005.
Zoltan Gyongyi, PhD, Stanford University, 2007.

Professional activities

Selected listing of invited and plenary lectures, lecture series

- Integer Programming in VLSI Design. *13th Symposium uber Operations Research*. Paderborn, Germany, September 1988.
- Random Walks, Electric Networks and Algorithmic Applications. *15th Theory Day at Columbia University*. March 1990.
- Overview of Randomized Algorithms.
 - *Annual Meeting of the American Association for the Advancement of Science*. Boston, 1993.
 - *Annual meeting of the American Mathematical Society*. Baltimore, January 1998.
 - *First Pacific Rim Conference on Mathematics*. Hongkong, January 1998.
- Lecture Series on Randomized Algorithms.
 - Melbourne, Australia, 1993.
 - Newcastle-upon-Tyne, England, 1996.
 - Technical University of Eindhoven, Netherlands, May 1998.
 - *Summer School on Probabilistic Methods in Algorithms*. Montpellier, France, August 1998.

- 3rd Max-Planck Advanced Course on the Foundations of Computer Science, Sep 2002, Saarbrücken, Germany.
- Randomized Approximation Algorithms for Combinatorial Problems. *14th Conference on the Foundations of Software Technology and Theoretical Computer Science*. Madras, India, 1994.
- Information retrieval algorithms: a survey. *ACM-SIAM Symposium on Discrete Algorithms*. New Orleans, Louisiana, 1997.
- Nine favorite randomized algorithms from the 90's. *Workshop in honor of Michael Rabin's 65th birthday*. Jerusalem, Israel, June 1997.
- Data mining: problems and algorithms. *Israel Symposium on Theory of Computing and Systems*. Ramat-Gan, Israel, June 1997.
- The design and analysis of algorithms. Lecture series at the International Centre for Theoretical Physics. Trieste, Italy. 1999.
- The web as a graph.
 - *International Conference on Combinatorics and Computing (COCOON)*. Tokyo, July 1999.
 - *Latin American Theoretical Informatics 2000*. Uruguay, April 2000.
 - *ACM/SIGMOD Symposium on Principles of Database Systems*. Dallas, May 2000.
 - *CISM Summer School*, Udine, Italy, 2002.
 - *Lipari Summer School*, Lipari, Italy, 2003.
- Mining social networks.
 - *Web Intelligence*. Maebashi, Japan, October 2001.
 - *Dertouzos Distinguished Lecture Series*, MIT Laboratory of Computer Science. Cambridge, MA, April 2002.
 - *Italian Association for Artificial Intelligence*, Siena, Italy, Sep 2002.
 - *World Knowledge Forum*, Seoul, Korea, October 2002.
 - *UT Austin Distinguished Lecture Series on Grid and Internet Computing*, February 2003.
 - *University of Pittsburgh Distinguished Lecture Series*, February 2003.
 - *Center of Nonlinear Studies Annual Conference*, Santa Fe, NM, 2003.
 - *Onassis Foundation Lecture Series*, Crete, July 2003.
- Tutorial on Web Information Retrieval (with A.Z. Broder and either R. Baeza-Yates, or K. Bharat and D. Hawking)
 - *ACM SIGIR*, New Orleans, September 2001.
 - *ACM SIGIR*, Tampere, Finland, August 2002.
 - *ACM SIGIR*, August 2005-2007.
 - *WWW*, May 2006, 2007.
- Lecture Series on Information Retrieval.
 - *WEBBAR Summer school on Web information retrieval*, Italy, August 2004, 2005 and 2007.
 - *5th Max-Planck Advanced Course on the Foundations of Computer Science*, Sep 2004, Saarbrücken, Germany.
 - *IPAM Graduate Summer School*, UCLA, July 2005.
- Incentive Networks.
 - *ACM Annual Symposium on Knowledge Discovery and Data Mining KDD-2005*, August 2005.

- The changing face of web search.
 - *Pacific-Asia Symposium on Data mining (PAKDD)*, Singapore, April 2006.
 - *Mobile Data Management (MDM)*, Nara, Japan, May 2006.
 - *ACM Symposium on Theory of Computing*, Seattle, May 2006.
 - *Flexible Query Answering Systems (FQAS)*, Milan, June 2006.
 - *Norman E. Friedmann Distinguished Lecture*, UCLA, March 2007.
 - *Bruce Nelson Memorial Lecture*, CMU, April 2007.
 - *University of Illinois, Chicago*, April 2007.
 - *High-Performance Computing*, Goa, December 2007.
 - *Mueller-Thuns lecture*, University of Illinois, Urbana-Champaign, April 2008.
 - University of Maryland, College Park, November 2008.
 - University of Toronto, Canada, November 2008.
- Web *n.0*: what sciences will it take?
 - *WWW2007 plenary lecture*, Banff, May 2007.
 - *Rosser lecture*, University of Wisconsin, Madison, April 2008.
- Heavy tails, and models for the Web and social networks.
 - Karlsruhe Institute of Technology, November 2009.
 - Tel Aviv University, November 2009.
- The quantitative study of user behavior online.
 - AT&T Labs Distinguished Speaker Series, May 2010.
 - Symposium on Computer Science in Russia, June 2010.

Editorial Activities

- Editor-in-Chief, *Journal of the ACM*, 2003–2009
- Member, Editorial Board, *Communications of the ACM*, 2008–
- Member, Editorial Board, *Journal of the ACM*, 1998-2003.
- Member, Editorial Board, *Internet Mathematics*, 2003–
- Member, Editorial Board, *ACM Transactions on Internet Technology*, 2000-2003.
- Associate Editor, Web Intelligence and Agent Systems, 2002-2003.
- Member, Editorial Board, *SIAM Journal On Computing*, 1992-2003.
- Member, Editorial Board, *Random Structures and Algorithms*, 1992-2000.
- Member, Editorial Board, *Journal of Combinatorial Optimization*, 1995-2000.

Professional Service

- 2000-2002 Chair, IEEE Technical Committee on Mathematical Foundations of Computing. This included responsibility for the IEEE Annual Symposia on Foundations of Computing, Logic in Computer Science, and Computational Complexity.
- 2000-2004 Awards Chair, ACM SIGACT.
- 1999-2001 Member, ACM SIGACT Executive Committee.
- 2001-2004 Member, ACM Kanellakis Award Committee. Chair for the 2004 award.
- 2003-2006 Member of the Board on Mathematical Sciences and Applications of the National Academy of Sciences.

- 2006- Member of the Computer Science and Telecommunications Board of the National Academy of Sciences.
- 2009- Chairman, Board of Trustees, International Computer Science Institute, Berkeley.
- 2009- Industrial Curatory Board, Schloss Dagstuhl, Germany.
- 2007- Member, Board of Trustees, Mathematical Sciences Research Institute, Berkeley.

Conference Program Committees

Program Chair

- 4th International Symposium on Algorithms and Computation, Hong Kong, 1993.
- 36th IEEE Symposium on Foundations of Computer Science, Milwaukee, 1995.
- WWW2005, 14th Annual Conference on the World-Wide Web, Tokyo, May 2005.

Program Committees (selected listing only)

- 32nd IEEE Symposium on Foundations of Computer Science, Puerto Rico, 1991.
- 34th IEEE Symposium on Foundations of Computer Science, Palo Alto, 1993.
- 6th ACM-SIAM Symposium on Discrete Algorithms, San Francisco, 1995.
- 2nd Workshop on Algorithmic Foundations of Robotics, Toulouse, 1996.
- 2nd Annual International Computing and Combinatorics Conference, Hong Kong, 1996.
- ACM-SIGMOD Symposium on Principles of Database Systems, 1999.
- 10th ACM-SIAM Symposium on Discrete Algorithms, 1999.
- International Colloquium on Automata, Languages and Programming (ICALP), 1999.
- International Symposium on Algorithms and Computation (ISAAC), 1999.
- 33rd ACM Symposium on Theory of Computing, 2000.
- WWW2002, 11th Annual Conference on the World-Wide Web, May 2002.
- 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2002.
- Deputy Vice-Chair for Search and Data Mining, WWW2003, 12th Annual Conference on the World-Wide Web, May 2003.
- Vice-Chair for Search, WWW2004, 13th Annual Conference on the World-Wide Web, May 2004.
- International Colloquium on Automata, Languages and Programming (ICALP), 2004.
- 23rd ACM SIGMOD International Conference on Management of Data SIGMOD, 2004.
- ACM Symposium on Principles of Distributed Systems, PODS 2005.
- 25th ACM SIGMOD International Conference on Management of Data SIGMOD, 2006.
- 21st AAAI, 2006.

Corporate and management experience

- 1994-2000 IBM Research, TJ Watson and Almaden Research Centers.
- Responsible for first- and second-line management of up to 15 researchers.
 - Formulation of overall group research directions.

- Assisting with IBM Research division strategy for Mathematical sciences and for the Utilities vertical industry.
- Mentoring group members and establishing connections across peer research groups.

2000-2004 Vice President/SVP and Chief Technology Officer, Verity, Inc.

- Technical strategy for the corporation including partnership and acquisition evaluations and due diligence; screening led to one successful acquisition (Inktomi's Enterprise Search business) that was accretive within 6 months.
- Line management of over 30 developers; this included a newly-formed Emerging Technologies department of over 20 scientists and engineers responsible for upgrading the company's technology.
- Product delivery responsibility for one product line (Verity Ultraseek, integrated from Inktomi) plus a significant portion of a second (Verity K2), together accounting for over \$100M a year in revenue.
- Technical analyst relations with analysts of all the major firms: Gartner, Delphi, IDC, Forrester, META and Yankee groups in the US as well as Ovum and AMR in the UK. During this period, Verity went from being viewed as a no-vision company with old technology to defining the cutting edge in its space.
- Technical spokesperson to the media for the corporation, doing over 20 media interviews a year, directly leading to an average of over 10 news stories a year.
- Sales support through direct customer engagement or indirect field support in most sizeable direct and OEM transactions in the US and in Europe.
- After-sales field support for customer escalations and patches on products in my responsibility.

2005- Senior Vice President and Head, Yahoo! Labs. From 2011, Chief Strategy Officer.

- Growing a world-class science and technology organization, hiring talent and establishing a number of research centers across the world. <http://labs.yahoo.com>
- Ensuring the impact of the labs' research on the corporation and on the scientific community. Yahoo! Labs is renowned for award-winning scientific output, as well as significant business impact on the company.