1 PageRank Papers

The original Brin and Page papers are [37] and [38]. Some others on the topic of PageRank are [14], [29], [55], [125], [75], [76], [124], [28], [142], [114], [91], [85], [52], [79]. A good survey of web searching with PageRank summary is [13]. Modeling web surfing with the 'back button': [58].

Updating PageRank papers: [43], [116], [115], [97], [98].

2 HITS and its modified versions Papers

The original HITS paper is [93]. Others are [100], [55], [54], [26], [79], [46], [47], [7], [48], [41], [61], [62], [121], [42], [92], [49]. Good survey paper: [33]. HITS for bibliometrics: [36].

3 LSI Papers

The first LSI paper [56] was followed by [20], [21], [23], [53], [50], [140], [141], [101], [87], [138], [89], [90], [94], [82], [88], [31], [30].

4 Similarity Measures

Some new similarity measures proposed are: [86], [68], [121].

5 Stats and Structure of Web

Several papers study the structure of the web and statistics regarding the web: [39], [3], [2], [80], [24], [1], [4], [5], [25], [60], [32], [12], [120], [119].

6 Laplacian and Graph Theory

The eigenvalues of the Laplacian of an undirected graph provide a surprising amount of information about the structure of the associated matrix. Do such findings extend to directed graphs, like those underlying a Markov chain? See [113], [112], [63], [64].

7 Clustering

Clustering is a difficult and important issue for IR. See [136], [34], [135], [51], [50], [77], [17], [67], [22], [78], [63], [64], [111], [45], [133].

8 Markov Chains

Classic Markov chain papers are [66], [72], [44], [107], [108], [110], [129], [57], [73], [134], [106], [81], [16], [84], [83], [128].

9 Books

Useful books for IR researchers are [40], [109], [132], [137], [131], [130], [118], [15], [18], [19], [65], [71], [95], [139].

10 Miscellaneous

Some papers that are harder to classify: [27], [102], [123, 99] deal with relevance feedback and user-tailored queries. [104] and [59] combine search engine results into one ranked list and [105] post-processes major search engine results annotating retrieved documents with topic information [103]. Spammer's guide to PageRank: [126]. Meta-search engines: [35]. TREC-5 conference: [74]. Medline data set: [6]. An ML info. retrieval model: [117]. IR and genetic algorithm: [122]. Inquirus 2 metasearch engine [70, 99]. Sampling pages uniformly from the Web [127]. Using maximum likelihood and training sets to detect "hidden" links between groups, such as for government intelligence applications: [96]. Using a training set to classify documents and incorporating text content and link information: [69]. Several papers by same authors on matrix perturbation and spectral decomposition: [11], [8], [9], [10].

References

[1] Search engine watch, March 1998. `https://www.searchenginewatch.com/`.

[2] Cyveillance: sizing the internet. webpage, July 2000. `https://www.cyveillance.com/web/us/corporate/white-paper.htm`.

[3] *Books in Print*. R.R. Bowker, New York, 2001.

[4] *Ulrich's International Periodicals Directory*. Bowker, New York, 2001.

[5] How much information? webpage, 2002. `https://www.sims.berkeley.edu/research/projects/how-much-info/internet.html#1`.

[6] Medlars test collection, January 2002. Available at `https://www.cs.utk.edu/ lsi/`.

[7] Dimitris Achlioptas, Amos Fiat, Anna R. Karlin, and Frank McSherry. Web search via hub synthesis. In *IEEE Symposium on Foundations of Computer Science*, pages 500–509, 2001.

[8] Rafikul Alam and Shreemayee Bora. Effect of linear perturbation on spectra of matrices.

[9] Rafikul Alam and Shreemayee Bora. A geometric approach to sensitivity analysis of eigenvalues and eigendecompositions.

[10] Rafikul Alam and Shreemayee Bora. On stable eigendecompositions of matrices.

[11] Rafikul Alam and Shreemayee Bora. Stability of eigenvalues and spectral decompositions under linear perturbation.

[12] Brian Amento, Loren Terveen, and Will Hill. Does authority mean quality? predicting expert quality ratings of web documents. In *The Twenty-Third Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2000.

[13] Arvind Arasu, Junghoo Cho, Hector Garcia-Molina, Andreas Paepcke, and Sriram Raghavan. Searching the web. *ACM Transactions on Internet Technology*, 2001.

[14] Arvind Arasu, Jasmine Novak, Andrew Tomkins, and John Tomlin. PageRank computation and the structure of the web: experiments and algorithms. In *The Eleventh International WWW Conference*, May 2002.

[15] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. ACM Press, New York, 1999.

[16] Michele Benzi and Miroslav Tuma. A parallel solver for large-scale Markov chains. *Applied Numerical Mathematics*, 41:135–153, 2002.

[17] Pavel Berkhin. Survey of clustering data mining techniques. Technical report, Accrue Software, San Jose, CA, 2002.

[18] Michael W. Berry, editor. *Computational Information Retrieval*. SIAM, Philadelphia, 2001.

[19] Michael W. Berry and Murray Browne. *Understanding Search Engines: Mathematical Modeling and Text Retrieval*. SIAM, Philadelphia, 1999.

[20] Michael W. Berry, Z. Drmac, and Elizabeth R. Jessup. Matrices, vector spaces and information retrieval. *SIAM Review*, 41:335–362, 1999.

[21] Michael W. Berry and R. D. Fierro. Low-rank orthogonal decompositions for information retrieval applications. *Journal of Numerical Linear Algebra with Applications*, 1(1):1–27, 1996.

[22] Michael W. Berry, Bruce Hendrickson, and Padma Raghavan. Sparse matrix reordering schemes for browsing hypertext. In J. Renegar, M. Shub, , and S. Smale, editors, *Lectures in Applied Mathematics (LAM)*, volume 36, pages 99–123. American Mathematical Society, 1996.

[23] Michael W. Berry and Gavin W. O'Brien. Using linear algebra for intelligent information retrieval. *SIAM Review*, 37:573–595, 1998.

[24] Michael W. Berry, P. Wang, and J. Bownas. Website query analysis: trend and behavior detection. In *Second SIAM Conference on Data Mining*, April 2002.

[25] Krishna Bharat and Andrei Broder. Estimating the relative size and overlap of public web search engines. In *The Seventh International WWW Conference*, New York, 1998. Elsevier Science.

[26] Krishna Bharat and Monika R. Henzinger. Improved algorithms for topic distillation in hyperlinked environments. In *21st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 104–111, 1998.

[27] Krishna Bharat, Farzin Maghoul, and Raymie Stata. The term vector database: fast access to indexing terms for webpages. *Computer Networks*, 33:247–255, 2000.

[28] Krishna Bharat and George A. Mihaila. When experts agree: using non-affiliated experts to rank popular topics. *ACM Transactions on Information Systems*, 20(1):47–58, 2002.

[29] Monica Bianchini, Marco Gori, and Franco Scarselli. PageRank: A circuital analysis. In *The Eleventh International WWW Conference*, May 2002.

[30] Katarina Blom. *Information retrieval using the singular value decomposition and Krylov subspaces*. PhD thesis, University of Chalmers, January 1999.

[31] Katarina Blom and Axel Ruhe. Information retrieval using very short Krylov sequences. In *Computational Information Retrieval*, pages 41–56, 2001.

[32] Paolo Boldi, Bruno Codenotti, Massimo Santini, and Sebastiano Vigna. Structural properties of the African web. In *The Eleventh International WWW Conference*, May 2002.

[33] Allan Borodin, Gareth O. Roberts, Jeffrey S. Rosenthal, and Panayiotis Tsaparas. Finding authorities and hubs from link structures on the world wide web. *World Wide Web*, pages 415–429, 2001.

[34] Rodrigo A. Botafogo and Ben Shneiderman. Identifying aggregates in hypertext structures. In *UK Conference on Hypertext*, pages 63–74, 1991.

[35] P. Bradley. Multi-search engines - a comparison. webpage, January 2002. `https://www.philb.com/msengine.htm`.

[36] U. Brandes and T. Willhalm. Visualizing bibliographic networks with a reshaped landscape metaphor. In *4th Joint Eurographics - IEEE TVCG Symposium on Visualization (VisSym '02)*, pages 159–164. ACM Press, 2002.

[37] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 33:107–117, 1998.

[38] Sergey Brin, Lawrence Page, R. Motwami, and Terry Winograd. The PageRank citation ranking: bringing order to the web. Technical report, Computer Science Department, Stanford University, 1998.

[39] Andrei Broder, Ravi Kumar, and Marzin Maghoul. Graph structure in the web. In *The Ninth International WWW Conference*, May 2000.

[40] Steven Campbell and Carl D. Meyer. *Generalized Inverses of Linear Transformations*. Pitman, San Francisco, 1979.

[41] Soumen Chakrabarti, Byron Dom, David Gibson, Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, and Andrew Tomkins. Spectral filtering for resource discovery. In *ACM SIGIR workshop on Hypertext Information Retrieval on the Web*, 1998.

[42] Zheng Chen, Jidong Wang, Liu Wenyin, and Wei-Ying Ma. A unified framework for web link analysis. 2002.

[43] Steve Chien, Cynthia Dwork, Ravi Kumar, and D. Sivakumar. Towards exploiting link evolution.

[44] Grace E. Cho and Carl D. Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335(1–3):137–150, 2001.

[45] Hwajeong Choi and Daniel B. Szyld. Application of threshold partitioning of sparse matrices to markov chains. In *IEEE International Computer Performance and Dependability Symposium IPDS'96*, pages 158–165, September 1996.

[46] David Cohn and Huan Chang. Learning to probabilistically identify authoritative documents. In *Proceedings of the 17th International Conference on Machine Learning*, pages 167–174, Stanford, CA, 2000.

[47] David Cohn, Huan Chang, and Andrew McCallum. Creating customized authority lists. In *Proceedings of the 17th International Conference on Machine Learning*, Stanford, CA, 2000.

[48] David Cohn and Thomas Hofmann. The missing link: a probabilistic model of document content and hyperlink connectivity. *Advances in Neural Information Processing Systems*, 13, 2001.

[49] Brian D. Davison, Apostolos Gerasoulis, Konstantinos Kleisouris, Yingfang Lu, Hyun ju Seo, Wei Wang, and Baohua Wu. Discoweb: applying link analysis to web search. May 1999.

[50] Inderjit S. Dhillon. Concept decompositions for large sparse text data using clustering. *Machine Learning*, 42(1/2):143–175, 2001.

[51] Inderjit S. Dhillon. Efficient clustering of very large document collections. In Robert L. Grossman, editor, *Data Mining for Scientific and Engineering Applications*, pages 357–388. Kluwer Academic Publishers, 2001.

[52] Michelangelo Diligenti, Marco Gori, and Marco Maggini. Web page scoring systems for horizontal and vertical search. In *WWW2002*, Honolulu, Hawaii, USA, 2002.

[53] Chris Ding. A similarity-based probability model for LSI. In *The Twenty-Second ACM SIGIR '99 Conference*, pages 59–65, 1999.

[54] Chris Ding, Xiaofeng He, Parry Husbands, Hongyuan Zha, and Horst Simon. Link analysis: hubs and authorities on the World Wide Web. Technical Report 47847, Lawrence Berkeley National Laboratory, May 2001.

[55] Chris Ding, Xiaofeng He, Hongyuan Zha, and Horst Simon. PageRank, HITS and a unified framework for link analysis. In *Proceedings of the 25th ACM SIGIR Conference*, pages 353–354, Tampere, Finland, August 2002.

[56] Susan T. DuMais. Improving the retrieval of information from external sources. *Behavior Research Methods, Instruments and Computers*, 23:229–236, 1991.

[57] Dundee proceedings. *A numerical analyst looks at the 'cutoff phenomenon' in card shuffling and other Markov chains*, 1997.

[58] Ronald Fagin, Anna R. Karlin, Jon Kleinberg, Prabhakar Raghavan, Sridhar Rajagopalan, Ronitt Rubinfeld, Madhu Sudan, and Andrew Tomkins. Random walks with 'back buttons'. In *32nd ACM Symposium on Theory of Computing*, 2000.

[59] Ronald Fagin, Ravi Kumar, and D. Sivakumar. Comparing top k lists. In *ACM SIAM Symposium on Discrete Algorithms*, 2003.

[60] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[61] Ayman Farahat, Thomas Lofaro, Joel C. Miller, Gregory Rae, F. Schaefer, and Lesley A. Ward. Modifications of kleinberg's HITS algorithm using matrix exponentiation and web log records. In *ACM SIGIR Conference*, pages 444–445, September 2001.

[62] Ayman Farahat, Thomas Lofaro, Joel C. Miller, Gregory Rae, and Lesley A. Ward. Existence and uniqueness of ranking vectors for linear link analysis. In *ACM SIGIR Conference*, September 2001.

[63] Miroslav Fielder. Algebraic connectivity of graphs. *Czechoslovak Mathematics Journal*, 23(98):298–305, 1973.

[64] Miroslav Fielder. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Mathematics Journal*, 25(100):619–633, 1975.

[65] W.B. Frakes and Ricardo Baeza-Yates. *Information Retrieval: Data Structures and Algorithms*. Prentice Hall, Englewood Cliffs, NJ, 1992.

[66] Robert E. Funderlic and Carl D. Meyer. Sensitivity of the stationary distribution vector for an ergodic Markov chain. *Linear Algebra and its Applications*, 76:1–17, 1986.

[67] Glenn Fung. A comprehensive overview of basic clustering algorithms. May 2001.

[68] Prasanna Ganesan, Hector Garcia-Molina, and Jennifer Widom. Exploiting hierarchical domain structure to compute similarity. Technical Report 2001-26, Stanford University Database Group, June 2001.

[69] Lise Getoor, Eran Segal, Ben Taskar, and Daphne Koller. Probabilistic models or text and link structure for hypertext classification. In *IJCAI Workshop on Text Learning: Beyond Supervision*, August 2001.

[70] Eric J. Glover, Steve Lawrence, Michael D. Gordon, William P. Birmingham, and C. Lee Giles. Web search – your way. *Communications of the ACM*, 1999. accepted for publication.

[71] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 1996.

[72] Gene H. Golub and Carl D. Meyer. Using the QR factorization and group inverse to compute, differentiate and estimate the sensitivity of stationary probabilities for Markov chains. *SIAM Journal on Algebraic and Discrete Methods*, 17:273–281, 1986.

[73] Winfried K. Grassmann, M. I. Taskar, and Daniel P. Heyman. Regenerative analysis and steady state distributions for Markov chains. *Operations Research*, 33(5):1107–1116, 1985.

[74] D. Harman and Ellen Voorhees. Overview of the fifth Text REtrieval Conference (TREC-5). In *Information Technology: The Fifth Text REtrieval Conference (TREC-5)*, pages 1–28, Gaithersburg, MD: NIST, November 1996.

[75] Taher H. Haveliwala. Efficient computation of PageRank. Technical report, Computer Science Department, Stanford University, 1999.

[76] Taher H. Haveliwala. Topic-sensitive PageRank. In *The Eleventh International WWW Conference*, May 2002.

[77] Xiaofeng He, Hongyuan Zha, Chris Ding, and Horst Simon. Web document clustering using hyperlink structures. *Computational Statistics and Data Analysis*, 45:19–45, 2002.

[78] Bruce Hendrickson and Tamara G. Kolda. Partitioning sparse rectangular and structurally nonsymmetric matrices for parallel computation. *SIAM Journal of Scientific Computing*, 21(6):2048–2072, 2000.

[79] Monika Henzinger. Link analysis in web information retrieval. *IEEE Data Engineering Bulletin*, 23(3):3–8, 2000.

[80] Monika R. Henzinger, Hannes Marais, Michael Moricz, and Craig Silverstein. Analysis of a very large altavista query log. Technical Report 1998-014, Digital SRC, October 1998.

[81] Daniel P. Heyman and Dianne P. O'Leary. What is fundamental for Markov chains: First passage times, fundamental matrices, and group generalized inverses. In William J. Stewart, editor, *Numerical Solution of Markov Chains (NSMC'95)*, pages 151–161. Kluwer Academic, 1995.

[82] M. K. Hughey and Michael W. Berry. Improved query matching using kd-trees, a latent semantic indexing enhancement. *Information Retrieval*, 2:287–302, 2000.

[83] Jeffrey J. Hunter. Generalized inverses, stationary distributions and mean first passage times with applications to perturbed Markov chains. *Research Letters in the Information and Mathematical Sciences*, 3:99–116, 2002.

[84] Jeffrey J. Hunter. Stationary distributions and mean first passage times of perturbed Markov chains. *Research Letters in the Information and Mathematical Sciences*, 3:85–98, 2002.

[85] Glen Jeh and Jennifer Widom. Scaling personalized web search. Technical report, Stanford University, 2002.

[86] Glen Jeh and Jennifer Widom. SimRank: A measure of structural-context similarity. In *Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, volume 11, Edmonton, Alberta, Canada, July 2002.

[87] Eric P. Jiang and Michael W. Berry. Solving total least squares problems in information retrieval. *Linear Algebra and its Applications*, 316:137–156, 2000.

[88] Fan Jiang and Michael L. Littman. Approximate dimension equalization in vector-based information retrieval. In *The Seventeenth International Conference on Machine Learning*, pages 423–430, 2000.

[89] K. Sparck Jones. A statistical interpretation of term specificity and its applications in retrieval. *J. Documentation*, 28:11–21, 1972.

[90] William P. Jones and George W. Furnas. Pictures of relevance: A geometric analysis of similarity measures. *Journal of American Society for Information Retrieval*, 38:420–442, 1987.

[91] Sepandar D. Kamvar, Taher H. Haveliwala, Christopher D. Manning, and Gene H. Golub. Extrapolation methods for accelerating PageRank computations. In *WWW 2003*, Budapest, Hungary, May 2003.

[92] Hung-Yu Kao, Ming-Syan Chen, Shian-Hua Lin, and Jan-Ming Ho. Entropy-based link analysis for mining web informative structures. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 574–581. ACM Press, 2002.

[93] Jon Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46, 1999.

[94] Tamara G. Kolda and Dianne P. O'Leary. A semi-discrete matrix decomposition for latent semantic indexing in information retrieval. *ACM Transactions on Information Systems*, 16:322–346, 1998.

[95] Robert R. Korfhage. *Information Storage and Retrieval*. Wiley, New York, 1997.

[96] Jeremy Kubica, Andrew Moore, Jeff Schneider, and Yiming Yang. Stochastic link and group detection. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, pages 798–804. AAAI Press/MIT Press, July 2002.

[97] Amy N. Langville and Carl D. Meyer. Updating PageRank using the group inverse and stochastic complementation. Technical Report crsc02-tr32, North Carolina State University, Mathematics Department, CRSC, 2002.

[98] Amy N. Langville and Carl D. Meyer. Updating the stationary vector of an irreducible Markov chain. Technical Report crsc02-tr33, North Carolina State University, Mathematics Department, CRSC, 2002.

[99] Steve Lawrence. Context in web search. *IEEE Data Engineering Bulletin*, 23(3):25–32, 2000.

[100] Ronny Lempel and S. Moran. The stochastic approach for link-structure analysis (SALSA) and the TKC effect. In *The Ninth International WWW Conference*, May 2000.

[101] Todd A. Letsche and Michael W. Berry. Large-scale information retrieval with LSI. *Informatics and Computer Science*, pages 105–137, 1997.

[102] Massimo Marchiori. The quest for correct information of the web: hyper search engines. In *The Sixth International WWW Conference*, April 1997.

[103] Alberto O. Mendelzon and Davood Rafiei. What do the neighbours think? computing web page reputations. *IEEE Data Engineering Bulletin*, 23(3):9–16, 2000.

[104] Alberto O. Mendelzon and Davood Rafiei. An autonomous page ranking method for metasearch engines. In *The Eleventh International WWW Conference*, May 2002.

[105] Alberto O. Mendelzon and Davood Rafiei. Topic: measuring webpage reputation. webpage, September 2002. `https://www.cs.toronto.edu/db/topic/about.html`.

[106] Carl D. Meyer. An alternative expression for the mean first passage matrix. *Linear Algebra and its Applications*, 22:41–47, 1978.

[107] Carl D. Meyer. Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems. *SIAM Review*, 31(2):240–272, 1989.

[108] Carl D. Meyer. Uncoupling the Perron eigenvector problem. *Linear Algebra and its Applications*, 114/115:69–74, 1989.

[109] Carl D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000.

[110] Carl D. Meyer and James M. Shoaf. Updating finite Markov chains by using techniques of group matrix inversion. *Journal of Statistical Computation and Simulation*, 11:163–181, 1980.

[111] Violeta Migallon, Jose Penades, and Daniel B. Szyld. Experimental study of parallel iterative solutions of Markov chains with block partitions. In Brigitte Plateau, William J. Stewart, and Manuel Silva, editors, *Numerical Solutions of Markov Chains (NSMC'99)*, pages 96–110. Prensas Universitarias de Zaragoza, 1999.

[112] Bojan Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, pages 871–898. Wiley, 1991.

[113] Bojan Mohar. Some applications of Laplace eigenvalues of graphs. In G. Hahn and G. Sabidussi, editors, *Graph Symmetry: Algebraic Methods and Applications*, pages 225–275. Kluwer Academic, 1997.

[114] Cleve Moler. The world's largest matrix computation. *Matlab News and Notes*, pages 12–13, October 2002.

[115] Andrew Y. Ng, Alice X. Zheng, and Michael I. Jordan. Link analysis, eigenvectors and stability. In *Seventh International Joint Conference on Artificial Intelligence*, 2001.

[116] Andrew Y. Ng, Alice X. Zheng, and Michael I. Jordan. Stable algorithms for link analysis. In *Proceedings of the 24th Annual International ACM SIGIR Conference*. ACM, 2001.

[117] Kenney Ng. A maximum likelihood ratio information retrieval model. In *The Eighth Text REtrieval Conference (TREC-8)*, 1999.

[118] J. R. Norris. *Markov Chains.* Cambridge University Press, 1997.

[119] Gopal Pandurangan, Prabhakara Raghavan, and Eli Upfal. Using PageRank to characterize web structure. In *8th Annual International Computing and Combinatorics Conference (COCOON)*, 2002.

[120] David M. Pennock, Gary W. Flake, Steve Lawrence, Eric J. Glover, and C. Lee Giles. Winners don't take all: characterizing the competition for links on the web. *Proceedings of the National Academy of Sciences*, 99(8):5207–5211, 2002.

[121] Donnacha Phelan and Nicholas Kushmerick. A descendant-based link analysis algorithm for web search. 2002.

[122] F. Picarougne, N. Monmarch, A. Oliver, and G. Venturini. Web mining with a genetic algorithm. In *The Eleventh International WWW Conference*, May 2002.

[123] Jesus U. Quevedo, Ana G. Covarrubias, and S. H. Huang. Improving retrieval by querying and examining prestige. In *The Eleventh International WWW Conference*, May 2002.

[124] Davood Rafiei and Alberto O. Mendelzon. What is this page known for? computing webpage reputations. In *The Ninth International WWW Conference*, pages 823–835. Elsevier Science, May 2000.

[125] Matthew Richardson and Petro Domingos. The intelligent surfer: probabilistic combination of link and content information in PageRank. *Advances in Neural Information Processing Systems*, 14, 2002.

[126] Chris Ridings. PageRank explained: everything you've always wanted to know about PageRank. Black Box Group: Rank Write Roundtable. `https://www.rankwrite.com/`, May 2002.

[127] Paat Rusmevichientong, David M. Pennock, Steve Lawrence, and C. Lee Giles. Methods for sampling pages uniformly from the World Wide Web. In *AAAI Fall Symposium on Using Uncertainty Within Computation*, pages 121–128, 2001.

[128] Kimberly Ann Flagg Sellers. Iterative methods for computing mean first passage times of Markov chains. Master's thesis, University of Maryland, 1998.

[129] Herbert A. Simon and Albert Ando. Aggregation of variables in dynamic systems. *Econometrica*, 29:111–138, 1961.

[130] G. W. Stewart. *Introduction to Matrix Computations.* Academic Press, Inc., 1973.

[131] G. W. Stewart and Ji-guang Sun. *Matrix Perturbation Theory.* Academic Press, Inc., 1990.

[132] William J. Stewart. *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, 1994.

[133] William J. Stewart and Tugrul Dayar. Comparison of partitioning techniques for two-level iterative solvers on large, sparse Markov chains. *SIAM Journal on Scientific Computing*, 21(5):1691–1705, 2000.

[134] William J. Stewart and W. Wu. Numerical experiments with iteration and aggregation for Markov chains. *ORSA Journal on Computing*, 4(3):336–350, 1992.

[135] Stijn Marinus van Dongen. *Graph Clustering by Flow Simulation.* PhD thesis, University of Utrecht, May 2000.

[136] Ron Weiss, Bienvenido Velez, Mark A. Sheldon, Chanathip Namprempre, Peter Szilagyi, Andrzej Duda, and David K. Gifford. Hypursuit: A hierarchical network search engine that exploits content-link hypertext clustering. In *Seventh ACM Conference on Hypertext*, pages 180–193, 1996.

[137] James H. Wilkinson. *The Algebraic Eigenvalue Problem.* Claredon Press, 1965.

[138] Dian I. Witter and Michael W. Berry. Downdating the latent semantic indexing model for conceptual information retrieval. *The Computer Journal*, 41(1):589–601, 1998.

[139] Richard S. Wurman. *Information Anxiety.* Doubleday, New York, 1989.

[140] Hongyuan Zha, Osni Marques, and Horst D. Simon. A subspace-based model for information retrieval with applications in latent semantic indexing. *Lecture Notes in Computer Science*, 1457:29–42, 1998.

[141] Hongyuan Zha and Horst D. Simon. On updating problems in latent semantic indexing. *SIAM Journal on Scientific Computing*, 21(2):782–791, 1999.

[142] Dell Zhang and Yisheng Dong. An efficient algorithm to rank web resources. *Computer Networks*, 33:449–455, 2000.