

David Gleich

dgleich@stanford.edu

74 Barnes Ct. #100
Stanford, CA 94305
801-949-0331

Research

My broad research interests are scientific computing, numerical analysis, and data mining; particularly, they include parallel computing, large-scale graph algorithms, visualization, ranking, PageRank, and parameterized matrix equations.

Awards

- Microsoft Live Labs Fellowship, 2007.
- Best speaker, SCREAM 2005.
- National Science Foundation Graduate Fellowship, Honorable Mention, Harvey Mudd College.
- High distinction, Harvey Mudd College.
- Honors in Computer Science, Harvey Mudd College.
- Borrelli Prize, Harvey Mudd College.
- High school valedictorian, Lourdes High School.
- US Army Excellence in Computer Science, Lourdes High School.
- Intel Award of Excellence in Computer Science, Lourdes High School.

Education

Stanford University

Stanford, CA

- Received Doctor of Philosophy from the Institute of Computational and Mathematical Engineering, September 2009.
- Thesis: Models and Algorithms for PageRank Sensitivity.
- Adviser: Michael Saunders (Former Adviser: Professor Gene Golub deceased Nov. 2007.)
- Received Master of Science degree in Computational and Mathematical Engineering, June 2006.
- Microsoft Live Labs fellow from August 2007, Overall 4.08 GPA.
- Courses: Partial Differential Equations of Applied Mathematics, Matrix Computations, Introduction to Combinatorics and its Applications, Numerical Optimization, Approximation Algorithms, Computational Genomics, Advanced Matrix Computations (Eigenvalue Problems), Numerical Solutions to Partial Differential Equations, Stochastic Methods in Engineering, Information Networks, Algorithmic Game Theory and Auction Design, Metric Embeddings and Algorithmic Applications, Stochastic Calculus, Iterative Methods, Large-Scale Numerical Optimization, Advanced Optimization, Discrete Data Analysis.

Harvey Mudd College

Claremont, CA

- Received Bachelor of Science degree, Computer Science and Mathematics joint major, May 2004.
- Graduated with high distinction and honors in computer science.
- Overall 3.864 GPA; 3.894 Major GPA.
- Computer Science Courses: Principles of Computer Science, C++ and Data Structures, Logic, Architecture and Operating Systems, Software Development, Programming Languages, Algorithms, Theory of Computing, Artificial Neural Networks, Independent Study in Computer Chess, Advanced Algorithms.
- Math Courses: Discrete Mathematics, Probability and Statistics, Differential Equations II, Real Analysis I, Abstract Algebra I, Numerical Analysis, Complex Analysis, Dynamical Systems.
- Concentration in Psychology: Introduction to Psychology, Theories of Personality, Foundations of Neuroscience, Physiological Psychology, Language and the Brain.
- Clinic: Overture Services (2002-2003) - Improved Relevance Ordering in Web Search.
 - used micro information units to improve relevance ordering
 - Erin Bodine, David Gleich, Cathy Kurata, Jordan Kwan, Lesley Ward, and Daniel Fain. Three Methods for Improving Relevance in Web Search. Clinic Report, Harvey Mudd College, 102 pages, May 9, 2003. Includes fully documented program code on accompanying CD.
- Dean's list in Spring 2001, Fall 2001, Spring 2002, Fall 2002, and Spring 2003 semesters.
- Borrelli Prize for Outstanding Contribution to a Mathematics Clinic.

dgleich@stanford.edu

University of Minnesota	Minneapolis, MN	1997-2000
-------------------------	-----------------	-----------

- Courses through the University of Minnesota Talented Youth Mathematics Program: Single Variable Calculus, Multivariable Calculus, Differential Equations, and Linear Algebra.

Lourdes High School	Rochester, MN	1996-2000
---------------------	---------------	-----------

- 4.0 GPA and class valedictorian.
- National Honors Society, 1998-2000.
- US Army Excellence and Intel Excellence Awards in Computer Science for sophomore science fair project on data compression; 3rd place in Hiawatha Division of Math League; Spanish student of the year 2000.

Teaching/ Research

Stanford University	Stanford, CA
---------------------	--------------

Teaching Assistantships

- Fall 2004: Linear Algebra with Applications to Engineering Computations. (Professor Margot Gerritsen)
- Winter 2005: Matrix Computations with Applications to Data Mining and IT. (Professor Gene Golub)
- Fall 2005: Linear Algebra with Applications to Engineering Computations. (Professor Margot Gerritsen)

Research Assistantships

- Spring 2009: Experiments with the random alpha PageRank model and spam identification; preparation for oral defense. (Professor Michael Saunders)
- Winter 2009: Developed new heuristics for the network alignment problem. (Professor Michael Saunders, and Mohsen Bayati) Began working on network analysis algorithms on epilepsy propagation networks. (Professor Josef Parvizi)
- Fall 2008: Updated MatlabBGL library with new routines; developed theory on PageRank with a matrix as the teleportation parameter (Professor Michael Saunders).
- Spring 2008: Developed theory to prove fast convergence of the inner-outer iteration for PageRank (UBC Professor Chen Greif); helped develop theory of matrix equations with random parameters (with Paul Constantine).
- Winter 2008: Improved work on the random alpha PageRank model with new algorithms and provable runtime and approximation bounds for all algorithms
- Fall 2007: Developed a quadratic programming framework for the overlap graph matching or network alignment problem for large graphs; studied direct methods for sparse matrices. (Professors Gene Golub and Amin Saberi)
- Spring 2007: Continued work on constrained least squares. Visited numerical analysis group at Oxford University. (Professor Gene Golub)
- Winter 2007: Worked on PageRank sensitivity analysis and the polynomial chaos expansion. Continued work on constrained least squares problems. (Professor Gene Golub)
- Fall 2006: Continued work on constrained least squares problems and iterative methods for total least squares and data least squares. Worked on algorithms to compute the Drazin inverse of a matrix. Worked on PageRank sensitivity analysis. (Professor Gene Golub)
- Summer 2006: Constrained least squares problems with two quadratic constraints. Special case software for total least squares and data least squares problems. (Professor Gene Golub)
- Spring 2006: Continued work on Monte Carlo approximations to the adjoint equation. Developed software for large-scale graph analysis in Matlab (MatlabBGL). (Professor Amin Saberi)
- Winter 2006: Monte Carlo approximations to the adjoint equation with applications to control. (Professor Amin Saberi)

Selected Publications

Paul G. Constantine, David F. Gleich, and Gianluca Iaccarino. Spectral methods for parameterized matrix equations. Submitted.

dgleich@stanford.edu

Mohsen Bayati, Margot Gerritsen, David F. Gleich, Amin Saberi, Ying Wang. Algorithms for large, sparse network alignment problems. ICDM 2009. (Includes all software.)

David F. Gleich, Andrew P. Gray, Chen Greif, Tracy Lau. An inner-outer iteration for computing PageRank. SIAM Journal of Scientific Computing, in press. (Includes all software.)

Paul G. Constantine, David F. Gleich. Using polynomial chaos to compute the influence of multiple random surfers in the PageRank model, Proceedings of the 5th Workshop on Algorithms and Models for the Web Graph (WAW2007), 82-95. (Published December 2007.)

Selected Software

MatlabBGL. Version 4.0. A large scale graph algorithm package for Matlab based on the Boost Graph Library. Downloaded over 7000 times and cited in more than 5 publications, including in the Proceedings of the National Academy of Sciences.

Experience

Microsoft Corporation Redmond, WA Summer 2008

Research Intern, Live Labs

- Investigated sensitivity of PageRank to large changes in web-graph structure (with Dennis Decoste).
- Assisted writing large scale codes for computing PageRank with overlapping information on large distributed graphs (with Reid Andersen).

Flixbee Inc. Palo Alto, CA Aug. 2007 – June 2008,
Oct. 2008 to present

- Developed search engines to aid movie discovery for a small start-up company; assisted with developing a recommendation engine. (Website: hellomovies.com)

Stanford University Stanford, CA Summer 2007

Computational Approaches to Digital Stewardship Intern

- Investigated properties of the Library of Congress subject headings.
- Helped develop a multi-lingual search engine using latent semantic indexing.
- Developed visualization software for relationships between subject headings.
- Compared the Library of Congress subject headings and the category system in Wikipedia.

Yahoo! Santa Clara, CA Sept. 2005 – May 2006

Part-Time Research Associate

- Visualization of flickr social network.
- Linear algebra based information propagation techniques for directed graphs with applications to labeling unknown nodes in social networks and web-graphs.

Intel Santa Clara, CA Summer 2005

Research Intern with Applications Research Lab

- Developed and proved theoretical bounds on an approximation to personalized PageRank computable on a desktop computing without disclosing personalized information.
- Investigated privacy issues in personalized web-search.

Publications

David Gleich and Marzia Polito. Approximating Personalized PageRank with Minimal Use of Web Graph Data. *Internet Mathematics*. (Published December 2007).

Yahoo! Sunnyvale, CA 2004-May 2005

Part-Time Research Associate

- Theoretical work on web search including extensive research into PageRank and multi-grid algorithms for PageRank computing.

dgleich@stanford.edu

Publications

- Paul G. Constantine, David F. Gleich. Random Teleportation Parameters in the PageRank model of a Random Surfer. In preparation, extended draft available on website. (Includes all software.)
- Paul G. Constantine, David F. Gleich, and Gianluca Iaccarino. Spectral methods for parameterized matrix equations. Submitted.
- Mohsen Bayati, Margot Gerritsen, David F. Gleich, Amin Saberi, Ying Wang. Algorithms for large, sparse network alignment problems. In Proceedings of the 9th International Conference on Data Mining (ICDM2009). Miami, FL, December
- David F. Gleich, Andrew P. Gray, Chen Greif, Tracy Lau. An inner-outer iteration for computing PageRank. SIAM Journal of Scientific Computing, in press. (Includes all software.)
- Qiqi Wang, David Gleich, Amin Saberi, Nasrollah Etemadi, Parviz Moin. A Monte Carlo Method for Solving the Unsteady Adjoint Equations. *Journal of Computational Physics* 227:6184-6205, 2008. (Published June 2008.)
- Paul G. Constantine, David F. Gleich. Using polynomial chaos to compute the influence of multiple random surfers in the PageRank model, Proceedings of the 5th Workshop on Algorithms and Models for the Web Graph (WAW2007), 82-95. (Published December 2007.)
- David Gleich and Marzia Polito. Approximating Personalized PageRank with Minimal Use of Web Graph Data. *Internet Mathematics* 3:257-294. (Published December 2007.)
- David F. Gleich, Peter Glynn, Gene H. Golub, Chen Greif. Three results on the PageRank vector: eigenstructure, sensitivity, and the derivative, Web Information Retrieval and Linear Algebra Algorithms, Dagstuhl Seminar Proceedings. (Published online June 2007.)
- David Gleich, Leonid Zhukov, Matt Rasmussen, Kevin Lang. The World of Music: SDP layout of high dimensional data. Interactive Poster, Information Visualization 2005.
- David Gleich and Leonid Zhukov. Scalable Computing with Power-Law Graphs: Experience with Parallel PageRank. Poster, SuperComputing 2005.
- Dennis Decoste, David Gleich, Tejaswi Kasturi, Sathiya Keerthi, Omid Madani, Seung-Taek Park, David M. Pennock, Corey Porter, Sumit Sanghai, Farial Shahnaz, and Leonid Zhukov. Recommender Systems Research at Yahoo! Research Labs. Position Statement, Beyond Personalization 2005. San Diego, CA, Jan 9-12, 2005.
- David Gleich and Leonid Zhukov. An SVD based term suggestion and ranking system. In Proceedings of the 4th International Conference on Data Mining. Brighton, UK, Nov 1-4th, 2004.
- Erin Bodine, David Gleich, Cathy Kurata, Jordan Kwan, Lesley Ward, and Daniel Fain. Three Methods for Improving Relevance in Web Search. Clinic Report, Harvey Mudd College, 102 pages, May 9, 2003. Includes full documented program code on accompanying CD.

Conference Organization

- (With Amy Langville.) Mini-Symposium on Matrix Methods for Sparse Text and Data Mining, SIAM Annual Meeting 2008. Speakers: David F. Gleich (Random parameters in PageRank), Michael Saunders (Basis Pursuit for Sparse Personalized PageRank vectors), Yehuda Koren (Matrix Methods in the Netflix competition), and Brett W. Bader (Methods for Multilingual Text Analysis).
- Symposium on current research in engineering and applied mathematics (SCREAM 2007). Stanford University, organized by the SIAM Student Chapter.

Presentations

- Building the Library of Congress Subject Headings galaxy. Stanford presentation to Delft University Students, September 21st, 2009, Stanford, CA.
- Building the Library of Congress Subject Headings galaxy. Joint Stanford-Library of Congress report meeting, September 10th, 2009, Washington DC.
- Algorithms for sparse network alignment problems. Keynote presentation at the Southeastern ranking and clustering workshop, August 2009.
- A tutorial on my research software for graph and sparse matrix computations. Southeastern ranking and clustering workshop, August 2009.
- New Heuristics for the Large-Scale Network Alignment Problem. 23rd Biennial Conference on Numerical Analysis, University of Strathclyde, June 24th, 2009.

dgleich@stanford.edu

Data visualization and prior research in the CADS Group. CADS meeting with the Library of Congress, Stanford University, June 9th, 2009.

Models and algorithms for PageRank sensitivity. Uncertainty quantification research group meeting, Stanford University, June 3, 2009.

Models and algorithms for PageRank sensitivity. Ph.D. oral defense, Stanford University, May 26, 2009.

The Network Alignment Problem and Algorithms. Art Owen's Research Group Seminar, February 12, 2009.

Inner Outer Iterations for PageRank. Linear Algebra and Optimization Seminar, February 11, 2009.

Uncertainty quantification for PageRank and friends. SIAM 2008 Annual Meeting; mini-symposium on Matrix Methods for Sparse Text and Data Mining, San Diego CA, July 6, 2008.

Uncertainty quantification for PageRank. University of Washington Applied Math Seminar, Seattle WA, May 6, 2008.

Finding Gems at the Library of Congress. College of Charleston, Charleston SC, January 11, 2008.

Really random surfers in PageRank. College of Charleston, Charleston SC, January 17, 2008.

(With Paul G. Constantine.) Really random surfers in PageRank. Workshop on Algorithms for the Web Graph, San Diego, CA, December 11, 2007.

Random PageRank. Linear Algebra and Optimization seminar, Stanford University, October 31, 2007.

Data Visualization. Library of Congress, Washington DC, September 21, 2007. (With CADS Presentation Group Professor Margot Gerritsen, Professor Amin Saberi, Jeremy Kozdon, Ying Wang, Vinayak Ganeshan.)

The PageRank problem and what happens when changing the probability of teleporting around the web. OCIAM, Oxford UK, May 11, 2007.

The PageRank problem and what happens when changing the probability of teleporting around the web. . Numerical analysis group seminar, Oxford UK, May 1, 2007.

Some results on the sensitivity of the PageRank vector. Dagstuhl Seminar on Web Information Retrieval and Linear Algebra, Schloss Dagstuhl, Germany, February 12, 2007.

(With Margot Gerritsen.) Linear Algebra and Search Engines. Library of Congress, Washington DC, January 5, 2007.

Fast Parallel PageRank. SCREAM 2005, Stanford University, February 26, 2005.

Fast Parallel PageRank: Methods and Evaluations. Seminar presentation. Intel Corporation, February 17th, 2005.

An SVD based term suggestion and ranking system. ICDM 2004, Brighton, UK.

Orthogonal Subspace Projection and SVD for Search Term Suggestion. Southern California Applied Mathematics Symposium, April 24, 2004.

(With Erin Bodine, Cathy Kurata, Jordan Kwan.) Three Methods for Improving Relevance in Web Search. Harve Mudd College Presentation Days. May 6, 2003.

Academic Visits

Chen Grief. University of British Columbia, April 30th-May 5th, 2008. Developed theory to analyze the inner-outer iteration for PageRank

Amy Langville. College of Charleston, January 6th-19th, 2008. Worked on sports ranking, dangling node models for PageRank, and large-scale rank aggregation.

Chen Greif. University of British Columbia, May 31st-June 12th, 2007. Investigated stationary iterations for PageRank.

Gene Golub. Oxford University, April 30-May 28th, 2007.

dgleich@stanford.edu

- Participation** SIAM Student Chapter President, Stanford University, October 2006-March 2008.
- Organized a conference (SCREAM 2007), a lecture series on historical research in numerical analysis at Stanford, and assisted with a memorial service for Gene Golub.
- ICME Social Chair, Stanford University, September 2006-September 2007.
- Planned annual BBQ for new students and awards celebration. Organized frequent social outings.
- Patents** Methods for Ranking Graphs using Random Parameters, with Paul Constantine. Pending.
- Refereeing** Refereeing articles for SIAM Journal on Scientific Computing, Information Visualization Conference, SIAM Symposium on Discrete Algorithms, Transactions on Parallel and Distributed Computing, Mathematics of Computation, Physics Review Letters, Mathematics and Computation, Physics Review E, Transactions on Knowledge and Data Engineering.
- Interviews** Richard Giles. *How to Use Flickr*. Course Technology PTR, 2006. Interview about visualizations of the Flickr social network.
- “A Visual Exploration of Complex Networks.” Seed Magazine Online, 24 July 2006. Accessed via http://www.seedmagazine.com/news/2006/07/look_around_you.php. Interview about visualizations of the LAUNCHcast recommendation network.
- Software** MatlabBGL. Version 4.0. A large scale graph algorithm package for Matlab based on the Boost Graph Library. Downloaded over 7000 times and cited in more than 5 publications, including in the Proceedings of the National Academy of Sciences.
- vismatrix. Version 1.01. Software to view sparse matrices interactively with labeled data. Downloaded over 375 times and used in presentations to the NSA and Library of Congress.
- gaimc. Version 1.0. Graph algorithms implemented in pure Matlab code that are only 2-4 times slower than MatlabBGL but extremely portable.
- Computer Skills** C/C++, Win32 API, COM, ATL, C#, .NET, Perl, PHP, Java, SQL, HTML, Linux, Unix, LaTeX, Matlab, MPI, Python, R
- Interests** Digital Photography, Skiing, Hiking, Cooking, Wine Appreciation