

7117 Wood Hollow Drive Apt 1935,  
Austin,  
TX, 78731

cjhsieh@cs.utexas.edu  
Cell Phone: 5129645633  
<http://www.cs.utexas.edu/~cjhsieh>

## Education

- **Ph.D. Dept. of Computer Science, University of Texas at Austin** 8/2010 – present  
Advisor: Inderjit Dhillon, GPA: 3.93/4.0
- **M.S. Dept. of Computer Science, National Taiwan University** 9/2007 – 6/2009  
Advisor: Chih-Jen Lin, GPA: 4.0/4.0
- **B.S., Dept. of Computer Science, National Taiwan University** 9/2003 – 6/2007  
Major GPA: 4.0/4.0
- **B.S., Dept. of Mathematics, National Taiwan University** 9/2004 – 6/2007  
Major GPA: 3.96/4.0

## Work Experience

- **Internship in IBM Research, Yorktown Heights** 6/2013–8/2013  
*Intern in Business Analysis and Math Sciences department*  
Mentor: Peder Olsen and Naoki Abe
  1. Built an efficient auto differentiation software.
  2. Developed an efficient nuclear norm minimization solver (published in ICML 2014).
- **Internship in LinkedIn** 6/2012–8/2012  
*Intern in SNA (Social Network Analysis) team*  
Mentor: Mitul Tiwari and Sam Shah
  1. Built a mathematical model using organizational time overlap to predict the probability of connections in social networks (published in WWW 2013).
- **Internship in Google** 5/2008–9/2008  
*Research Intern, Google Beijing Research Team*
  1. Studied parallel algorithms for training SVM and logistic regression.
  2. Improved the training and prediction speed of SVM in dependency parsing (published in JMLR).

## Research Interests

My main research focus is developing new algorithms and optimization techniques for large-scale machine learning problems. Currently, I am working on the following problems:

- Efficient solvers for large-scale linear and nonlinear SVMs.
- Scalable matrix completion algorithms in multi-core and distributed setting.
- Optimization techniques for high dimensional problems such as learning the structure of Gaussian Markov Random Fields (GMRF).
- Application areas including recommender systems and social networks.
- Theoretical analysis of optimization algorithms.

## Selected Honors

- IBM PhD Fellowships: 2013–2014, 2014–2015.
- NIPS Oral Presentation (top 1.4% submissions), 2013.
- Best Paper Award, ICDM 2012.
- Best Research Paper Award, KDD 2010.
- Best Master's Thesis Award, Institute of Information and Computing Machinery, 2009.

## Publications

Google Scholar Profile: Number of Citations=4100+; h-index = 14, i10-index = 19. Details available at <http://scholar.google.com/citations?user=Wy89g4IAAAAJ&hl=en&oi=ao>

## Journal Publications

1. **C.-J. Hsieh**, M. A. Sustik, I. S. Dhillon and P. Ravikumar. QUIC: Quadratic Approximation for Sparse Inverse Covariance Matrix Estimation. *Journal of Machine Learning Research(JMLR)*, 15:2911–2947, 2014.
2. K.-Y. Chiang, **C.-J. Hsieh**, N. Natarajan, A. Tewari, and I. S. Dhillon. Prediction and Clustering in Signed Networks: A Local to Global Perspective. *Journal of Machine Learning Research(JMLR)*, 15:1177–1213, 2014.
3. H. Yun, H.-F. Yu, **C.-J. Hsieh**, S. Vishwanathan, I. S. Dhillon. NOMAD: Non-locking, stOchastic Multi-machine algorithm for Asynchronous and Decentralized matrix completion. *Proceedings of the VLDB Endowment*, 7:11:975–986, 2014.
4. H.-F. Yu, **C.-J. Hsieh**, S. Si, I. S. Dhillon. Parallel Matrix Factorization for Recommender Systems. *Knowledge and Information Systems(KAIS)*, 2013.
5. H.-F. Yu, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. *ACM Transactions on Knowledge Discovery from Data(TKDD)*, 5:23:1–23, 2012.
6. G.-X. Yuan, K.-W. Chang, **C.-J. Hsieh**, and C.-J. Lin. A comparison of optimization methods for large-scale L1-regularized linear classification. *Journal of Machine Learning Research(JMLR)*, 11:3183–3234, 2010. **Cited 123 times.**
7. Y.-W. Chang, **C.-J. Hsieh**, K.-W. Chang, Michael Ringgaard, and C.-J. Lin. Low-Degree Polynomial Mapping of Data for SVM. *Journal of Machine Learning Research(JMLR)*, 11:1471–1490, 2010.
8. F.-L. Huang, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Iterative scaling and coordinate descent method for maximum entropy models. *Journal of Machine Learning Research(JMLR)*, 11:581–614, 2010.
9. R.-E. Fan, K.-W. Chang, **C.-J. Hsieh**, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research(JMLR)*, 9:1871–1874, 2008. **Cited 2664 times.**
10. K.-W. Chang, **C.-J. Hsieh**, and C.-J. Lin. Coordinate Descent Method for Large-scale L2-loss Linear SVM. *Journal of Machine Learning Research(JMLR)*, 9:1369–1398, 2008.

## Conference Publications

1. **C.-J. Hsieh**, H.-F. Yu, I. S. Dhillon. PASSCoDe: Parallel ASynchronous Stochastic dual Co-ordinate Descent. In *International Conference on Machine Learning(ICML)*, 2015.
2. **C.-J. Hsieh**, N. Natarajan, I. S. Dhillon. PU Learning for Matrix Completion. In *International Conference on Machine Learning(ICML)*, 2015.
3. H.-F. Yu, **C.-J. Hsieh**, H. Yun, S. Vishwanathan, I. S. Dhillon. A Scalable Asynchronous Distributed Algorithm for Topic Modeling. In *ACM WWW International conference on World Wide Web(WWW)*, 2015.

4. **C.-J. Hsieh**, I. S. Dhillon, P. Ravikumar, S. Becker, P. A. Olsen. QUIC & DIRTY: A Quadratic Approximation Approach for Dirty Statistical Models. In *Neural Information Processing Systems(NIPS)*, 2014.
5. **C.-J. Hsieh**, S. Si, I. S. Dhillon. Fast Prediction for Large-Scale Kernel Machines. In *Neural Information Processing Systems(NIPS)*, 2014.
6. E.-H. Yen, **C.-J. Hsieh**, P. Ravikumar, I. S. Dhillon. Constant Nullspace Strong Convexity and Fast Convergence of Proximal Methods under High-Dimensional Settings. In *Neural Information Processing Systems (NIPS)*, 2014.
7. **C.-J. Hsieh**, S. Si, I. S. Dhillon. A Divide-and-Conquer Solver for Kernel Support Vector Machines. In *International Conference on Machine Learning(ICML)*, 2014.
8. S. Si, **C.-J. Hsieh**, I. S. Dhillon. Memory Efficient Kernel Approximation. In *International Conference on Machine Learning(ICML)*, 2014. **Recommended for JMLR Fast Track, 18 out of 1260+**.
9. **C.-J. Hsieh**, P. A. Olsen. Nuclear Norm Minimization via Active Subspace Selection. In *International Conference on Machine Learning(ICML)*, 2014.
10. **C.-J. Hsieh**, M. A. Sustik, I. S. Dhillon, P. Ravikumar, R. A. Poldrack. BIG & QUIC: Sparse Inverse Covariance Estimation for a Million Variables. In *Neural Information Processing Systems(NIPS)*, 2013. **Oral presentation, 1.5% acceptance rate.**
11. H. Wang, A. Banerjee, **C.-J. Hsieh**, P. Ravikumar, I. S. Dhillon. Large Scale Distributed Sparse Precision Estimation. In *Neural Information Processing Systems (NIPS)*, 2013.
12. **C.-J. Hsieh**, M. Tiwari, S. Shah, D. Agarwal. Organizational Overlap on Social Networks and its Applications. In *ACM WWW International conference on World Wide Web(WWW)*, 2013.
13. H.-F. Yu, **C.-J. Hsieh**, S. Si, and I. S. Dhillon. Scalable Coordinate Descent Approaches to Parallel Matrix Factorization for Recommender Systems. In *IEEE International Conference on Data Mining(ICDM)*, 2012. **Best Paper Award.**
14. **C.-J. Hsieh**, I. S. Dhillon, P. Ravikumar and A. Banerjee. A Divide-and-Conquer Method for Sparse Inverse Covariance Estimation. In *Neural Information Processing Systems(NIPS)*, 2012.
15. **C.-J. Hsieh**, K.-Y. Chiang and I. S. Dhillon, Low-Rank Modeling of Signed Networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2012.
16. **C.-J. Hsieh**, M. A. Sustik, I. S. Dhillon and P. Ravikumar. Sparse Inverse Covariance Matrix Estimation Using Quadratic Approximation. In *Neural Information Processing Systems(NIPS)*, 2011. **Cited 130 times.**
17. **C.-J. Hsieh** and I. S. Dhillon. Fast Coordinate Descent Methods with Variable Selection for Non-negative Matrix Factorization. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2011.
18. H.-F. Yu, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. In *International Joint Conference on Artificial Intelligence(IJCAI)*, 2011. **The Best Paper Track.**
19. H.-F. Yu, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2010. **Best Research Paper Award.**
20. F.-L. Huang, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Iterative scaling and coordinate descent method for maximum entropy models. In *Association for Computational Linguistics(ACL)*, 2009. Short paper.
21. **C.-J. Hsieh**, K.-W. Chang, C.-J. Lin, S. Sathya Keerthi, and S. Sundararajan. A Dual Coordinate Descent Method for Large-scale Linear SVM. In *International Conference on Machine Learning(ICML)*, 2008. **Cited 409 times.**
22. S. S. Keerthi, S. Sundararajan. K.-W. Chang, **C.-J. Hsieh**, and C.-J. Lin. A sequential dual method for large scale multi-class linear SVMs. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2008.

## Other Publications

1. H.-Y. Lo, K.-W. Chang, S.-T. Chen, T.-H. Chiang, **C.-J. Hsieh**, Y.-K. Ko, T.-T. Kuo, H.-C. Lai, K.-Y. Lin, C.-H. Wang, H.-F. Yu, C.-J. Lin, H.-T. Lin and S.-D. Lin. An Ensemble of Three Classifiers for KDD Cup 2009: Expanded Linear Model, Heterogeneous Boosting, and Selective Naive Bayes. KDDCup '09, 2009, *JMLR Workshop and Conference Proceedings*, V.7, 57-64, 2009. Third Place of the Slow Track.
2. H.-F. Yu, **C.-J. Hsieh**, K.-W. Chang, and C.-J. Lin. Pascal Challenge: Linear Support Vector Machines. *Pascal Large Scale Learning Challenge in ICML Workshop*, 2008.

## Selected Software Packages

- LIBLINEAR – A Library for Large-scale Linear Classification
  1. <http://www.csie.ntu.edu.tw/~cjlin/liblinear>
  2. One of the main contributors.
  3. A comprehensive package containing several efficient linear classification and regression solvers.
- DC-SVM – A Divide-and-Conquer solver for kernel SVM
  1. <http://www.cs.utexas.edu/~cjhsieh/dcsvm>
  2. Solve classification problems with 0.5 million samples in 3 minutes.
- QUIC – QUadratic Inverse Covariance algorithm
  1. <http://www.cs.utexas.edu/user/sustik/quic>
  2. Proximal Newton method for sparse inverse covariance estimation.
  3. The extension–BIGQUIC can solve 1 million dimensional problems (with 1 trillion parameters) in one day using a single machine.
- LIBPMF – A parallel matrix factorization library.
  1. <http://www.cs.utexas.edu/~rofuyu/libpmf>
  2. Fast and scalable matrix completion solver (on multi-core platforms).
- NMF-CD – Coordinate descent methods for non-negative matrix factorization
  1. <http://www.cs.utexas.edu/~cjhsieh/nmf>
  2. Coordinate descent algorithms for least squares NMF and KL-NMF.
- AMD – An automatic matrix differentiation library.
  1. <https://github.com/pkambadu/AMD>
  2. One of the main contributors during my internship in IBM research.
  3. Efficient automatic differentiation computation for matrix functions.
- LIBSVM – A Library for Support Vector Machines
  1. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
  2. Assisted Professor Chih-Jen Lin in maintaining the library and answering questions from users.

## Talks

1. “Automatic Differentiation for Matrix Functions”. Machine Learning Symposium, National Taiwan University, Jan, 2015.
2. “Exploiting Structure in Large-scale Optimization for Machine Learning”. Guest lecture at UT Austin. Nov, 2014.
3. “Matrix Completion – Theories, Applications, and Scalable Solvers”. Appier, Taipei. July, 2014.

4. “Sparse Inverse Covariance Estimation for a Million Variables”. ICML workshop on Covariance Selection. Jun, 2014.
5. “Nuclear Norm Minimization via Active Subspace Selection”. International Conference on Machine Learning, Beijing. Jun, 2014.
6. “A Divide-and-Conquer Solver for Kernel Support Vector Machines”. International Conference on Machine Learning, Beijing. Jun, 2014.
7. “BIG & QUIC: Sparse Inverse Covariance Estimation for a Million Variables”. Machine Learning Symposium, National Taiwan University, Dec, 2013.
8. “BIG & QUIC: Sparse Inverse Covariance Estimation for a Million Variables”. Neural Information Processing Systems, Dec, 2013.
9. “Automatic differentiation for matrix functions and Nuclear Norm Minimization Solvers”. IBM Research, Yorktown Heights, NY. Aug, 2013.
10. “Sparse Inverse Covariance Estimation Using Quadratic Approximation”. MOPTA, Aug, 2013.
11. “Sparse Inverse Covariance Estimation Using Quadratic Approximation”. Machine Learning Symposium, National Taiwan University, Dec, 2012.
12. “Organizational Overlap on Social Networks”. LinkedIn, Aug, 2012.

## Other Awards

- Third Prize in the Slow Track of KDDCUP 2009.
- Winner of SVM Track of Pascal Large Scale Learning Challenge in ICML 2008 Workshop.
- International Olympiad in Informatics, 2nd place in Taiwan.
- National College Programming Contest, 2nd place.
- ACM ICPC Asia Regional programming contest, 6th place.
- Honorary Member of the Phi Tau Phi Scholastic Honor Society, National Taiwan University.
- Presidential Award (3 times, given to the top 5% undergraduate students each semester), National Taiwan University.

## Teaching Experience

1. TA: CS395T Scalable Machine Learning, Fall 2014, UT Austin.
2. TA: Introduction to the Theory of Computation, Fall 2008, National Taiwan University.
3. TA: Introduction to the Theory of Computation, Fall 2007, National Taiwan University.

## Professional Activities

1. Workshop organizer: ICML 2014 workshop on Covariance Estimation and Graphical Model Structure Learning.
2. Paper reviewer & Programming Committee: IEEE TNN, JMLR, IJPRAI, Neural Computation, Neural Computing, TKDD, IEEE TIT, TKDE, DAMI, IEEE WCCI '08, NIPS '11, NIPS '12, NIPS '13, IJCNN '13, ICML '14, AISTATS '14, KDD '14, NIPS '14, TAAI '14, AAAI '15, AISTATS '15, KDD '15, ICML '15.
3. Student volunteer: ICML '08, KDD '08, KDD '10, IJCAI '11, KDD '11, ICML '14, NIPS '14.

## Skills

Programming: C/C++, Python, MATLAB, Latex