

Dawei Yin

CONTACT INFORMATION	JD.com 8 Beichen West Rd. Beijing, China, 10000	<i>Voice:</i> +1 (610) 739-9810 <i>E-mail:</i> yindawei@acm.org <i>Web:</i> www.yindawei.com
RESEARCH INTERESTS	Data Mining, Applied Machine learning, Recommender Systems, Online Advertising	
EDUCATION	Lehigh University , Bethlehem, PA, USA	
	Doctor of Philosophy in Computer Science	Sep. 2013
	<ul style="list-style-type: none">• Topic: Prediction and recommendation in online media.• Advisor: Prof. Brian D. Davison• Rossin Doctoral Fellow	
	Master of Science in Computer Science	Sep. 2010
	<ul style="list-style-type: none">• GPA: 3.95/4.0• Advisor: Prof. Brian D. Davison and Prof. Henry S. Baird	
	The University of Hong Kong , Hong Kong, China	
	M.Phil. student in Computer Science	Jun. 2007 - Jun. 2008
	<ul style="list-style-type: none">• GPA: 4.0/4.0 (Transfer to Lehigh University)	
	Shandong University , Jinan, Shandong, China	
	Bachelor of Engineering, Computer Science	Jun. 2006
	<ul style="list-style-type: none">• GPA: 89.41/100, Major GPA: 92.0/100, Rank: 1st/53• Outstanding Graduate	
WORK EXPERIENCE	JD.com, Beijing, China	
	<i>Senior Director</i>	Jun. 2016 - present
	<ul style="list-style-type: none">• Leading the science efforts of JD's search and recommendation.• Manage over 100 talented scientists/engineers.• Founding head of Data Science Lab at JD.com (http://datascience.jd.com)• WSDM 2018 Best Student Paper Award	
	Yahoo Labs, Sunnyvale, CA, USA	
	<i>Senior Research Manager</i>	Aug. 2013 - Jun. 2016
	<ul style="list-style-type: none">• Manage the core ranking team, which includes about 10 top talented scientists and engineers.• Lead the science efforts on Yahoo search core relevance, and the projects include: ranking function, click modeling, active learning, deep learning, query rewriting, real time features etc..• As a leading author, publish "Ranking Relevance in Yahoo Search" in KDD 2016 and Yahoo Tech Pulse 2015 (internal tech conference).• LEAP Award (Labs Excellence Award Program) for significantly improving Yahoo search relevance.• WSDM 2016 Best Paper Award• KDD 2016 Best Paper Award	

Machine Learning Group, Microsoft Research Asia, Beijing, China

Research Intern **Feb. 2012 - Aug. 2012**

- Perform research focused on Internet Advertising, Sponsored Search, and Recommender Systems.
- Analyze several factors influencing the CTR from the perspective of context, including the number of displayed ads, the content of the ads, the relationship between the query and ads, and the mutual influences between ads. Based on this data analysis, we designed a novel Context-Aware Click Model for sponsored search.

Web Understanding, Modeling, and Evaluation Lab, Lehigh University

Research Assistant **Jan. 2009 - May 2013**

- Research under Prof. Davison's supervision, mainly focusing on web mining, machine learning and social computing.

Machine Learning Group, XEROX Research Center Europe, Grenoble, France

Research Intern **May. 2011 - Sep. 2011**

- Work on Bayesian Recommender System, Bayesian multi-relational analysis.

Pattern Recognition Research Lab, Lehigh University

Research Assistant **Jan. 2009 - Jan. 2011**

- Research under Prof. Baird's supervision on DARPA project - Document analysis and exploitation project (<http://dae.cse.lehigh.edu>)
- Investigate algorithm enhancements for fast approximate nearest-neighbor classifiers.

Computer Science and Engineering Department, Lehigh University

Teaching Assistant **Jan. 2013 - present**

- Undergraduate semester course "Computer Organization Architecture".
- Gave tutorials (once a week), and deployed experimental system for 81 students.

Computer Science Department, The University of Hong Kong

Teaching Assistant **Sep. 2007 - Jan. 2008**

- Undergraduate semester course "Discrete Mathematics".
- Gave tutorials (once a week), designed and graded the homework and final exams for 80 students.

REFEREED
PUBLICATIONS

M. Yamada, J. Tang, J. Lugo-Martinez, E. Hodzic, R. Shrestha, A. Saha, H. Ouyang, D. Yin, H. Mamitsuka, C. Sahinalp, P. Radivojac, F. Menczer, Y. Chang. **Ultra High-Dimensional Nonlinear Feature Selection for Big Biological Data**. In *IEEE Transactions on Knowledge and Data Engineering*, 2018.

Hongshen Chen, Zhaochun Ren, Jiliang Tang, Yihong Eric Zhao and Dawei Yin. **Hierarchical Variational Memory Network for Dialogue Generation**. In *Proceedings of the 27th International World Wide Web Conference (WWW 2018)*, Lyon, France, 2018.

Hongshen Chen, Xiaorui Liu, Dawei Yin, Jiliang Tang. **A Survey on Dialogue Systems: Recent Advances and New Frontiers**. In *SIGKDD Explorations*, 2018

Meizi Zhou, Zhuoye Ding, Jiliang Tang, Dawei Yin. **Micro Behaviors: A New Perspective in E-commerce Recommender Systems**, In *Proceedings of the 11th ACM Conference on Web Search and Data Mining (WSDM 2018)*, Los Angeles, California, USA, 2018.

Yao Ma, Zhaochun Ren, Ziheng Jiang, Jiliang Tang, Dawei Yin. **Multidimensional network embedding with hierarchical structures**, In *Proceedings of the 11th ACM Conference on Web Search and Data Mining (WSDM 2018)*, Los Angeles, California, USA, 2018.

Zihan Wang, Ziheng Jiang, Zhaochun Ren, Jiliang Tang, Dawei Yin. **A Path-constrained Framework for Discriminating Substitutable and Complementary Products in E-commerce**, In *Proceedings of the 11th ACM Conference on Web Search and Data Mining (WSDM 2018)*, Los Angeles, California, USA, 2018. [**Best Student Paper Award**]

Zhiwei Wang, Tyler Derr, Dawei Yin and Jiliang Tang. **Understanding and Predicting Weight Loss with Mobile Social Networking Data**. In *Proceedings of the 26th ACM International Conference on Information and Knowledge Management (CIKM2017)*, Singapore, Nov. 2017.

Yuening Hu, Changsung Kang, Jiliang Tang, Dawei Yin, and Yi Chang. **Large-scale Location Prediction for Web Pages**. In *IEEE Trans on Knowledge and Data Engineering (TKDE)*, 2017

Shiyu Chang, Yang Zhang, Jiliang Tang, Dawei Yin, Yi Chang, Mark Hasegawa-Johnson, Thomas S. Huang. **Streaming Recommender Systems**. In *Proceedings of the 26th International World Wide Web Conference (WWW 2017)*, Perth, Australia, Apr. 2017.

Yunlong He, Jiliang Tang, Hua Ouyang, Changsung Kang, Dawei Yin and Yi Chang. **Learning to rewrite queries**. In *Proceedings of 25th ACM Conference on Information and Knowledge Management (CIKM 2016)*, Indianapolis, IN, USA, Oct. 2016.

Dawei Yin, Yuening Hu, Jiliang Tang, Tim Daly Jr., Mianwei Zhou, Hua Ouyang, Jianhui Chen, Changsung Kang, Hongbo Deng, Chikashi Nobata, Jean-Marc Langlois, Yi Chang. **Ranking Relevance in Yahoo Search**. In *Proceedings of the 22nd annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2016)*, San Francisco, CA, Aug. 2016 [**Best Paper Award**]

Shiyu Chang, Yang Zhang, Jiliang Tang, Dawei Yin, Yi Chang, Mark Hasegawa-Johnson, Thomas S. Huang. **Positive-Unlabeled Learning in Streaming Networks**. In *Proceedings of the 22nd annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD2016)*, San Francisco, CA, Aug. 2016

Yi Chang, Jiliang Tang, Dawei Yin, Makoto Yamada, and Yan Liu. **Timeline Summarization with Life Cycle Models**. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI 2016)*, New York, USA, July, 2016.

Neil O'Hare, Paloma De Juan, Rossano Schifanella, Yunlong He, Dawei Yin, Yi Chang. **Leveraging User Interaction Signals for Web Image Search**. In *Proceedings of*

the 39th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2016), Pisa, Tuscany, Italy, July, 2016.

Shan Jiang, Yuening Hu, Changsung Kang, Tim Daly, Dawei Yin, Yi Chang. **Learning Query and Document Relevance from Web-scale Click Graph**. In *Proceedings of the 39th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2016)*, Pisa, Tuscany, Italy, July, 2016.

Jiliang Tang , SuhangWang, Xia Hu, Dawei Yin, Yingzhou Bi, Yi Chang and Huan Liu. **Recommendation with Social Dimensions**. In *the 30th AAAI Conference on Artificial Intelligence (AAAI)*, Phoenix, Arizona USA, Feb 2016.

Yue Wang, Dawei Yin, Jie Luo, Pengyuan Wang, Makoto Yamada, Yi Chang, Qiaozhu Mei. **Beyond Ranking: Optimizing Whole-Page Presentation**, In *Proceedings of the 9th ACM Conference on Web Search and Data Mining (WSDM 2016)*, San Francisco, US, Feb 2016. [**Best Paper Award**]

Changsung Kang, Dawei Yin, Ruiqiang Zhang, Nicolas Torzec, Jianzhang He, Yi Chang. **Learning to rank related entities in web search**. In *Neurocomputing*.

Zaihan Yang, Liangjie Hong, Dawei Yin, and Brian D. Davison. **Ranking Authors with Learning-to-Rank Topic Modeling**. In *International Journal of Innovative Computing, Information and Control*.

Pengyuan Wang, Wei Sun and Dawei Yin. **What Size Should A Mobile Ad Be?**. In *the 1st Workshop on Offline and Online Evaluation of Web-based Services at WWW 2015*. Florence, Italy, May 2015.

Pengyuan Wang, Dawei Yin, Marsha Meytlis, Jian Yang, Yi Chang. **Rethink Targeting: Detect ‘Smart Cheating’ in Online Advertising through Causal Inference**. In *Proceedings of the 24th International World Wide Web Conference (WWW 2015)*, Florence, Italy, May 2015.

Suriya Gunasekar, Makoto Yamada, Dawei Yin, Yi Chang. **Consistent Collective Matrix Completion under Joint Low Rank Structure**. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS2015)*, San Diego, CA, USA, April 2015.

Wei Sun, Pengyuan Wang, Dawei Yin, Jian Yang and Yi Chang. **Causal Inference via Sparse Additive Models with Application to Online Advertising**. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*, Austin, Texas, USA, Jan 2015.

Pengyuan Wang, Wei Sun, Dawei Yin, Jian Yang and Yi Chang. **Robust Tree-based Causal Inference for Complex Ad Effectiveness Analysis**. In *Proceedings of the 8th ACM Conference on Web Search and Data Mining (WSDM 2015)*, Shanghai, China, Feb 2015.

Zaihan Yang, Dawei Yin, Brian D. Davison. **Recommendations in Academia: a Joint Multi-Relational Model**. In *Proceedings of 2014 IEEE/ACM International Conference on Advances in Social Network Analysis and Mining*, Beijing, China, 2014.

Dawei Yin, Shike Mei, Bin Cao, Jian-Tao Sun, Brian D. Davison. **Exploiting Contextual Factors for Click Modeling in Sponsored Search**. In *Proceedings of the*

7th ACM Conference on Web Search and Data Mining (WSDM 2014), New York, USA, Feb 2014.

Dawei Yin, Bin Cao, Jian-Tao Sun, Brian D. Davison. **Estimating Ad Group Performance in Sponsored Search**. In *Proceedings of the 7th ACM Conference on Web Search and Data Mining (WSDM 2014)*, New York, USA, Feb 2014.

Guillaume Bouchard, Shengbo Guo and Dawei Yin. **Convex Collective Matrix Factorization**. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2013)*, Scottsdale, AZ, USA, April 2013.

Dawei Yin, Shenbo Guo, Boris Chidlovskii, Brian D. Davison, Cedric Archambeau and Guillaume Bouchard. **Connecting Comments and Tags: Improved Modeling of Social Tagging Systems**. In *Proceedings of the 6th ACM Conference on Web Search and Data Mining (WSDM 2013)*, Rome, Italy, Feb 2013.

Dawei Yin, Liangjie Hong and Brian D. Davison. **Structural Link Analysis and Prediction in MicroBlogs**. In *Proceedings of the 20th ACM Conference on Information and Knowledge Management (CIKM 2011)*, Glasgow, Scotland, UK, October 2011.

Liangjie Hong, Dawei Yin, Jian Guo and Brian D. Davison. **Tracking Trends: Incorporate Volume into Temporal Topic Models**. In *Proceedings of the 17th Annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2011)*, San Diego, CA, 2011.

Dawei Yin, Liangjie Hong, Zhenzhen Xue and Brian D. Davison. **Temporal Dynamics of User Interests in Tagging Systems**. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence (AAAI 2011)*, San Francisco, USA, August, 2011.

Zhenzhen Xue, Dawei Yin and Brian D. Davison. **Normalizing Microtext**. In *Proceedings of the AAAI-11 Workshop on Analyzing Microtext*, San Francisco, USA, August, 2011.

Dawei Yin, Liangjie Hong, Xiong Xiong and Brian D. Davison. **Link Formation Analysis in MicroBlogs**. In *Proceedings of the 34th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2011)*, Beijing, China, July, 2011.

Zaihan Yang, Dawei Yin and Brian Davison. **Award Prediction with Temporal Citation Network Analysis**. In *Proceedings of the 34th Annual ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2011)*, Beijing, China, July, 2011.

Dawei Yin, Liangjie Hong and Brian D. Davison. **Exploiting Session-like Behavior in Tag Prediction**. In *Proceedings of the 20th International World Wide Web Conference (WWW 2011)*, Hyderabad, India, March 2011.

Xiaoguang Qi, Dawei Yin, Zhenzhen Xue and Brian D. Davison. **Choosing Your Own Adventure: Automatic Taxonomy Expansion to Permit Many Paths**. In *Proceedings of 19th ACM Conference on Information and Knowledge Management (CIKM 2010)*, Toronto, Canada, October 2010.

Dawei Yin, Zhenzhen Xue, Liangjie Hong and Brian D. Davison. **A Probabilistic Model for Personalized Tag Prediction.** In *Proceedings of the 16th Annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010)*, Washington, DC, July 2010.

Dawei Yin, Chang An, Henry S. Baird. **Imbalance and Concentration in k-NN Classification.** In *Proceedings of IAPR 20th International Conference on Pattern Recognition (ICPR 2010)*, Istanbul, Turkey, August 2010.

Chang An, Dawei Yin, Henry S. Baird. **Document Segmentation using Pixel-Accurate Ground Truth.** In *Proceedings of IAPR 20th International Conference on Pattern Recognition (ICPR 2010)*, Istanbul, Turkey, August 2010.

Dawei Yin, Chang An and Henry S. Baird. **Safely Selecting Subsets of Training Data.** In *Proceedings of IAPR 9th International Workshop on Document Analysis Systems (DAS 2010)*, Boston, MA, June 2010.

Dawei Yin, Henry S. Baird and Chang An. **Time and Space Optimization of Document Content Classifiers.** In *Proceedings of IS&T/SPIE Document Recognition and Retrieval Conf. (DR&R XVII)*, San Jose, CA, January 17-21, 2010.

Dawei Yin, Zhenzhen Xue, Liangjie Hong, B. D. Davison, A. Kontostathis, and L. Edwards. **Detection of Harassment on Web 2.0.** In *Proceedings of the Content Analysis in the WEB 2.0 (CAW2.0) Workshop at WWW2009*, Madrid, Spain, April 2009.

Xin Han, He Guo, Dawei Yin and Yong Zhang. **A note on on-line broadcast scheduling with deadlines.** *Information Processing Letters*, Volume 109 Issue 3, January 2009

OTHER
PUBLICATIONS

Dawei Yin, Changsung Kang, Chikashi Nobata, Hongbo Deng, Hua Ouyang, Jianhui Chen, Jiliang Tang, Mianwei Zhou, Tim Daly, Yuening Hu, Jean-Marc Langlois, Yi Chang. **Ranking in Yahoo Search Technology**, *Yahoo Tech Pulse (internal publication)*, 2015.

Zaihan Yang, Dawei Yin and Brian D. Davison. **Optimizing MAP in Joint Multi-Relational Models for Recommendations in the Academic Network.** *Technical Report Technical Report LU-CSE-14-002*, Dept. of Computer Science and Engineering, Lehigh, 2014.

Xiaoguang Qi, Dawei Yin, Zhenzhen Xue and Brian D. Davison. **Enhancing Taxonomies by Providing Many Paths.** *Technical Report LU-CSE-10-005*, Dept. of Computer Science and Engineering, Lehigh, 2010.

Dawei Yin, Zhenzhen Xue, Xiaoguang Qi and Brian D. Davison. **Diversifying Search Results with Popular Subtopics.** In *Proceedings of the Eighteenth Text REtrieval Conference (TREC 2009)*, Gaithersburg, MD: NIST, November 2009.

PATENTS

US Patent **Automatic Social Media Content Timeline Summarization Method and Apparatus**, US Patent App. 15/200,922.

US Patent **Method and System for Rewriting a Query.** US Patent App. 15/130,217.

US Patent **Computerized system and method for high-quality and high-ranking digital content discovery**. US Patent App. 15/074,028

US Patent **Method and system for recommending content**. US Patent App. 14/983,663.

US Patent **Location-sensitive ranking for search and related techniques**. US Patent App. 14/868,135

US Patent **Predicting locations for web pages and related techniques**. US Patent App. 14/868,154.

US Patent **Method and system for ranking search content**. US Patent App. 14/959,122.

US Patent **Systems and methods for measuring complex online strategy effectiveness**. US Patent App. 14/587,328

US Patent **Probabilistic Relational Data Analysis**. US Patent App. 13/690,071.

US Patent **Convex Collective Matrix Factorization**. US Patent 9,058,303.

INVITED TALKS

Ranking Relevance in Web Search at *China Big Data Technology Conference*, Beijing, China, Dec. 2016.

Ranking and Beyond-ranking in Web Search at *Qihoo 360 Technology Co Ltd.*, Beijing, China, Sep. 2016.

Ranking and Beyond-ranking in Web Search at *Institute of Information Engineering, Chinese Academy of Science*, Beijing, China, Sep. 2016.

Ranking and Beyond-Ranking in Web Search, at *Facebook*, Menlo Park, CA, Aug. 2016.

Prediction and Recommendation in Online Media at *Shenzhen Institutes of Advanced Technology, Chinese Academy of Science*, Shenzhen, China, July 2013.

Prediction and Recommendation in Online Media at *Yahoo! Labs*, Sunnyvale, CA, May 2013.

Prediction and Optimization in Online Advertising at *Thomson Reuters*, New York, NY, April 2013.

Prediction and Optimization in Online Advertising at *Samsung Research America*, San Jose, CA, April 2013.

Temporal Dynamics of User Interests in Tagging Systems at *XEROX Research Center Europe*, Grenoble, France, June 2011.

HONORS AND AWARDS

Best Student Paper Award, ACM WSDM, February 2018

Best Paper Award, ACM SIGKDD, August 2016

Yahoo Excellence Award for the highest level performer, Yahoo Inc, March 2016.

Best Paper Award, ACM WSDM, February 2016

Yahoo Patent Milestone Award, Yahoo Inc, 2016
Yahoo Invention Award, Yahoo Inc., January 2015.
LEAP Award (Labs Excellence Award Program), Yahoo Labs, September 2014.
WSDM Student Travel Grant, ACM WSDM, February 2013.
Rossin Doctoral Fellow, Lehigh University, 2012
Donald B. Crouch Travel Grant, ACM SIGIR, 2011.
SIGIR Student Travel Grant, ACM SIGIR, 2011.
Gotshal Fellowship, Lehigh University, 2008.
Graduate Studentship, the University of Hong Kong, 2007.
Honorable Mention Award of Mathematical Contest in Modeling, COMAP, 2006.
Outstanding Undergraduate Award (3%), Shandong University, June 2006.
China Mobile Scholarship, China Mobile, December 2005.
Top-class Scholarship (5%), Shandong University, 2004, 2005.
Excellent Students Award (5%), Shandong University, 2004, 2005.
Excellent Student Leader Award, Shandong University, 2003, 2004, 2005.

MENTORSHIP

JD Data Science Lab:

Ziyi Guo (Lehigh University)

Meizi Zhou (The University of Minnesota, Twin Cities)

Yahoo Labs:

Shiyu Chang (The University of Illinois at Urbana-Champaign)

Yue Wang (The University of Michigan)

Wei Sun (Purdue University)

Suriya Gunasekar (The University of Texas, Austin)

OTHER
EXPERIENCES

Chinese Scholar and Student Association, Lehigh University

Vice President

Sep. 2008 - Jun. 2009

- Organizes social activities for Chinese scholars and students in Lehigh University.

Material Science and Engineering Department, Shandong University

Student Counselor

Sep. 2006 - Jun. 2007

- Managing about 200 undergraduates and guiding their social and academic activities.

Computer Science Department, Shandong University

Class President

Sep. 2002 - Jun. 2006

- Continuously elected as class president from Sep. 2002 to Jun. 2006. Leading 52 classmates and organizing both social and academic activities.

SERVICE

Organizer:

Industry Track Chair at SIGIR 2018

Sponsorship Chair at WSDM 2018

Co-Chair of the Second Vertical Search Relevance Workshop at WSDM 2015

Track Chair/Senior Program committee:

SIGIR 2018

WSDM 2018

Program committee:

SDM 2015, 2016
SIGIR 2014, 2015, 2016, 2017
WSDM 2014
AAAI 2014, 2016, 2017, 2018
IJCAI 2016
CIKM 2016
WWW 2017, 2018

Journal Reviewer:

IEEE Transactions on Knowledge and Data Engineering
ACM Transactions on Knowledge Discovery from Data
Data Mining and Knowledge Discovery
ACM Transactions on Intelligent Systems and Technology
Neurocomputing
Pattern Recognition Letters

MEMBERSHIP ACM/SIG member
 AAAI member

SKILLS Proficient with: ANSI C, C++, Python, Matlab
 Familiar with: Perl, Java, Pig, Hadoop
 Solid math background
 Excellent teamwork
 Strong problem-solving abilities

REFERENCES **Available upon request**