

August 13-17, 2016
San Francisco, CA, USA



Association for
Computing Machinery

Advancing Computing as a Science & Profession



San Francisco, CA, USA

KDD 2016

KDD'16

Proceedings of the 22nd ACM SIGKDD International Conference on
Knowledge Discovery and Data Mining

Sponsored by:

ACM SIGKDD and ACM SIGMOD



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

**The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York, New York 10121-0701**

Copyright © 2016 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: permissions@acm.org or Fax +1 (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through www.copyright.com.

Notice to Past Authors of ACM-Published Articles

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that has been previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform permissions@acm.org, stating the title of the work, the author(s), and where and when published.

ISBN: 978-1-4503-4232-2

Additional copies may be ordered prepaid from:

ACM Order Department
PO Box 30777
New York, NY 10087-0777, USA

Phone: 1-800-342-6626 (USA and Canada)
+1-212-626-0500 (Global)
Fax: +1-212-944-1318
E-mail: acmhelp@acm.org
Hours of Operation: 8:30 am – 4:30 pm ET

Printed in the USA .

KDD 2016 Chairs' Welcome Message

It is our great pleasure to welcome you to the 2016 ACM Conference on Knowledge Discovery and Data Mining – KDD'16. We hope that the content and the professional network at KDD'16 will help you succeed professionally by enabling you to: identify technology trends early; make new/creative contributions; increase your productivity by using newer/better tools, processes or ways of organizing teams; identify new job opportunities; and hire new team members.

We are living in an exciting time for our profession. On the one hand, we are witnessing the *industrialization of data science*, and the emergence of the *industrial assembly line* processes characterized by the division of labor, integrated processes/pipelines of work, standards, automation, and repeatability. Data science practitioners are organizing themselves in more sophisticated ways, embedding themselves in larger teams in many industry verticals, improving their productivity substantially, and achieving a much larger scale of social impact. On the other hand we are also witnessing astonishing progress from research in algorithms and systems — for example the field of deep neural networks has revolutionized speech recognition, NLP, computer vision, image recognition, etc. By facilitating interaction between practitioners at large companies & startups on the one hand, and the algorithm development researchers including leading academics on the other, KDD'16 fosters technological and entrepreneurial innovation in the area of data science.

This year's conference continues its tradition of being the premier forum for presentation of results in the field of data mining, both in the form of cutting edge research, and in the form of insights from the development and deployment of real world applications. Further, the conference continues with its tradition of a strong tutorial and workshop program on leading edge issues of data mining. The mission of this conference has broadened in recent years even as we placed a significant amount of focus on both the research and applied aspects of data mining. As an example of this broadened focus, this year we have introduced a strong hands-on tutorial program during the conference in which participants will learn how to use practical tools for data mining. KDD'16 also gives researchers and practitioners a unique opportunity to form professional networks, and to share their perspectives with others interested in the various aspects of data mining. For example, we have introduced office hours for budding entrepreneurs from our community to meet leading Venture Capitalists investing in this area. We hope that KDD 2016 conference will serve as a meeting ground for researchers, practitioners, funding agencies, and investors to help create new algorithms and commercial products.

The call for papers attracted a significant number of submissions from countries all over the world. In particular, the research track attracted 784 submissions and the applied data science track attracted 331 submissions. Papers were accepted either as full papers or as posters. The overall acceptance rate either as full papers or posters was less than 20%. For full papers in the research track, the acceptance rate was lower than 10%. This is consistent with the fact that the KDD Conference is a premier conference in data mining and the acceptance rates historically tend to be low. It is noteworthy that the applied data science track received a larger number of submissions compared to previous years. We view this as an encouraging sign that research in data mining is increasingly becoming relevant to industrial applications. All papers were reviewed by at least three program committee members and then discussed by the PC members in a discussion moderated by a meta-reviewer. Borderline papers were thoroughly reviewed by the program chairs before final decisions were made.

In addition, KDD'16 continues to have a strong tutorial and workshop program, as is its tradition since its inception. The program committee reviewed and accepted the following:

<i>Venue or Track</i>	<i>Reviewed</i>	<i>Accepted</i>	<i>Acceptance Rate</i>
Research Track Papers	784	70+72	8.9%+9.2%
Applied Data Science Track Papers	331	40+26	12%+7.9%
Workshops	47	17	36.2%
Tutorials	37	11	29.7%
Applied Data Science Talks	Invited	12	Invited
Hands-on Tutorials	Invited	8	Invited
Regular Keynotes	Invited	4	Invited
Turing Award Winner Talk	Special	1	Special

In the table above the research track papers and applied data science track papers are broken up into the full presentations and poster presentations. We have four excellent keynote presentations from Jennifer Chayes, Joe Hellerstein, Greg Papadopoulos, and Nando de Freitas. In addition, we are honored to have a Turing Award winner, Whitfield Diffie, as one of the keynote speakers. We also encourage attendees to attend the keynote and invited talk presentations.

Putting together *KDD'16* was a team effort. We first thank the authors for providing the content of the program. We are grateful to the program committee and the senior program committee, who worked very hard in reviewing papers and providing feedback for authors. Finally, we thank the numerous sponsors and ACM SIGKDD for their support in hosting the conference.

We hope that you will find this program interesting and thought-provoking and that the conference will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

Charu Aggarwal
Alex Smola

KDD'16 Research Track Chairs

Rajeev Rastogi
Dou Shen

*KDD'16 Applied Data
Science Track Chairs*

Balaji Krishnapuram
Mohak Shah

KDD'16 General Chairs

Table of Contents

KDD 2016 Conference Organization	xx
---	-----------

KDD 2016 Sponsors & Supporters	xxiv
---	-------------

Keynote Talks

• Graphons and Machine Learning: Modeling and Estimation of Sparse Massive Networks	1
Jennifer Chayes (<i>Microsoft Research</i>),	
• Learning to Learn and Compositionality with Deep Recurrent Neural Networks: Learning to Learn and Compositionality	3
Nando de Freitas (<i>Google DeepMind & University of Oxford</i>),	
• The Evolving Meaning of Information Security	5
Whitfield Diffie (<i>Stanford</i>),	
• People, Computers, and The Hot Mess of Real Data	7
Joseph M. Hellerstein (<i>University of California, Berkeley & Trifacta</i>),	
• A VC View of Investing in ML	9
Greg Papadopoulos (<i>NEA</i>),	

Panel

• Big Data Needs Big Dreamers: Lessons from Successful Big Data Investors	11
Evangelos Simoudis (<i>Synapse Partners</i>), Mark Gorenberg (<i>Zetta Venture Partners</i>), Tim Guleri (<i>Sierra Ventures</i>), Matt Ocko (<i>Data Collectivee</i>), Greg Sands (<i>Costanoa Venture Capital</i>)	

Applied Data Science Track Oral Papers

• Designing Policy Recommendations to Reduce Home Abandonment in Mexico	13
Klaus Ackermann (<i>Monash University</i>), Eduardo Blancas Reyes (<i>University of Chicago</i>), Sue He (<i>University of Virginia</i>), Thomas Anderson Keller (<i>University of California, San Diego</i>), Paul van der Boor (<i>University of Chicago</i>), Romana Khan (<i>Northwestern University</i>), Rayid Ghani (<i>University of Chicago</i>), José Carlos González (<i>Infonavit</i>)	
• Aircraft Trajectory Prediction <i>Made Easy</i> with Predictive Analytics	21
Samet Ayhan, Hanan Samet (<i>University of Maryland</i>)	
• Matrix Computations and Optimization in Apache Spark	31
Reza Bosagh Zadeh (<i>Stanford & Matroid</i>), Xiangrui Meng (<i>Databricks</i>), Alexander Ulanov (<i>HP Labs</i>), Burak Yavuz (<i>Databricks</i>), Li Pu (<i>Twitter</i>), Shivaram Venkataraman, Evan Sparks (<i>University of California, Berkeley</i>), Aaron Staple (<i>Databricks</i>), Matei Zaharia (<i>Databricks & Massachusetts Institute of Technology</i>)	
• Predicting Disk Replacement towards Reliable Data Centers	39
Mirela Botezatu, Ioana Giurgiu, Jasmina Bogojenska, Dorothea Wiesmann (<i>IBM Research</i>)	
• Developing a Data-Driven Player Ranking in Soccer Using Predictive Model Weights	49
Joel Brooks, Matthew Kerr, John Gutttag (<i>Massachusetts Institute of Technology</i>)	
• The Legislative Influence Detector: Finding Text Reuse in State Legislation	57
Matthew Burgess (<i>University of Michigan</i>), Eugenia Giraudy (<i>University of California, Berkeley</i>), Julian Katz-Samuels (<i>University of Michigan</i>), Joe Walsh (<i>University of Chicago</i>), Derek Willis (<i>ProPublica</i>), Lauren Haynes, Rayid Ghani (<i>University of Chicago</i>)	
• Identifying Police Officers at Risk of Adverse Events	67
Samuel Carton (<i>University of Michigan</i>), Jennifer Helsby (<i>University of Chicago</i>), Kenneth Joseph (<i>Carnegie Mellon University</i>), Ayesha Mahmud (<i>Princeton University</i>), Youngsoo Park (<i>University of Arizona</i>), Joe Walsh (<i>University of Chicago</i>), Crystal Cody, CPT Estella Patterson (<i>Charlotte-Mecklenburg Police Department</i>), Lauren Haynes, Rayid Ghani (<i>University of Chicago</i>)	

• Data-Driven Metric Development for Online Controlled Experiments: Seven Lessons Learned	77
Alex Deng (<i>Microsoft</i>), Xiaolin Shi (<i>Yahoo Research</i>)	
• Catch Me If You Can: Detecting Pickpocket Suspects from Large-Scale Transit Records	87
Bowen Du (<i>Beihang University</i>), Chuanren Liu (<i>Drexel University</i>), Wenjun Zhou (<i>University of Tennessee</i>), Zhenshan Hou (<i>Beihang University</i>), Hui Xiong (<i>Rutgers</i>)	
• Email Volume Optimization at LinkedIn	97
Rupesh Gupta, Guanfeng Liang, Hsiao-Ping Tseng, Ravi Kiran Holur Vijay, Xiaoyu Chen, Romer Rosales (<i>LinkedIn</i>)	
• Large-Scale Item Categorization in e-Commerce Using Multiple Recurrent Neural Networks	107
Jung-Woo Ha, Hyuna Pyo, Jeonghee Kim (<i>NAVER Corp.</i>)	
• Online Dual Decomposition for Performance and Delivery-Based Distributed Ad Allocation	117
Jim C. Huang, Rodolphe Jenatton, Cedric Archambeau (<i>Amazon</i>)	
• Minimizing Legal Exposure of High-Tech Companies through Collaborative Filtering Methods	127
Bo Jin (<i>Dalian University of Technology</i>), Chao Che (<i>Dalian University</i>), Kuifei Yu (<i>Zhigu Tech</i>), Yue Qu, Li Guo, Cuili Yao (<i>Dalian University of Technology</i>), Ruiyun Yu (<i>Northeastern University</i>), Qiang Zhang (<i>Dalian University</i>)	
• Ranking Universities Based on Career Outcomes of Graduates	137
Navneet Kapur (<i>GoFundMe</i>), Nikita Lytkin, Bee-Chung Chen, Deepak Agarwal, Igor Perisic (<i>LinkedIn Corporation</i>)	
• Predictors without Borders: Behavioral Modeling of Product Adoption in Three Developing Countries	145
Muhammad Raza Khan, Joshua E. Blumenstock (<i>University of Washington</i>)	
• Repeat Buyer Prediction for E-Commerce	155
Guimei Liu, Tam T. Nguyen (<i>Institute for Infocomm Research</i>), Gang Zhao (<i>Development Bank of Singapore</i>), Wei Zha (<i>Institute for Infocomm Research</i>), Jianbo Yang (<i>General Electric</i>), Jianneng Cao, Min Wu, Peilin Zhao (<i>Institute for Infocomm Research</i>), Wei Chen (<i>Development Bank of Singapore</i>)	
• Audience Expansion for Online Social Network Advertising	165
Haishan Liu, David Pardoe, Kun Liu, Manoj Thakur, Frank Cao, Chongzhe Li (<i>LinkedIn Corporation</i>)	
• From Online Behaviors to Offline Retailing	175
Ping Luo, Su Yan (<i>Institute of Computing Technology, CAS</i>), Zhiqiang Liu, Zhiyong Shen, Shengwen Yang (<i>Baidu, Inc.</i>), Qing He (<i>Institute of Computing Technology, CAS</i>)	
• Firebird: Predicting Fire Risk and Prioritizing Fire Inspections in Atlanta	185
Michael Madaio (<i>Carnegie Mellon University</i>), Shang-Tse Chen (<i>Georgia Institute of Technology</i>), Oliver L. Haimson (<i>University of California, Irvine</i>), Wenwen Zhang (<i>Georgia Institute of Technology</i>), Xiang Cheng (<i>Emory University</i>), Matthew Hinds-Aldrich (<i>Atlanta Fire Rescue Department</i>), Duen Horng Chau, Bistra Dilkina (<i>Georgia Institute of Technology</i>)	
• DopeLearning: A Computational Approach to Rap Lyrics Generation	195
Eric Malmi (<i>Aalto University & HIIT</i>), Pyry Takala (<i>Aalto University</i>), Hannu Toivonen (<i>University of Helsinki & HIIT</i>), Tapani Raiko (<i>Aalto University</i>), Aristides Gionis (<i>Aalto University & HIIT</i>)	
• EMBERS at 4 years: Experiences operating an Open Source Indicators Forecasting System	205
Sathappan Muthiah, Patrick Butler, Rupinder Paul Khandpur, Parang Saraf, Nathan Self, Alla Rozovskaya, Liang Zhao, Jose Cadena, Chang-Tien Lu, Anil Vullikanti, Achla Marathe (<i>Virginia Tech</i>), Kristen Summers (<i>IBM Watson Group</i>), Graham Katz, Andy Doyle (<i>CACI</i>), Jaime Arredondo (<i>University of California, San Diego</i>), Dipak K. Gupta (<i>San Diego State University</i>), David Mares (<i>University of California, San Diego</i>), Naren Ramakrishnan (<i>Virginia Tech</i>)	
• Anomaly Detection Using Program Control Flow Graph Mining From Execution Logs	215
Animesh Nandi, Atri Mandal (<i>IBM Research</i>), Shubham Atreja (<i>IIT Kanpur</i>), Gargi B. Dasgupta, Subhrajit Bhattacharya (<i>IBM Research</i>)	

• Engagement Capacity and Engaging Team Formation for Reach Maximization of Online Social Media Platforms	225
Alexander Nikolaev, Shounak Gore, Venu Govindaraju (<i>University at Buffalo</i>)	
• Boosted Decision Tree Regression Adjustment for Variance Reduction in Online Controlled Experiments	235
Alexey Poyarkov, Alexey Drutsa, Andrey Khalyavin, Gleb Gusev, Pavel Serdyukov (<i>Yandex</i>)	
• Dynamic and Robust Wildfire Risk Prediction System: An Unsupervised Approach	245
Mahsa Salehi, Laura Irina Rusu, Timothy Lynar, Anna Phan (<i>IBM Research</i>)	
• Deep Crossing: Web-Scale Modeling without Manually Crafted Combinatorial Features	255
Ying Shan, T. Ryan Hoens, Jian Jiao, Haijing Wang, Dong Yu, JC Mao (<i>Microsoft Corporation</i>)	
• Question Independent Grading using Machine Learning: The Case of Computer Program Grading	263
Gursimran Singh, Shashank Srikant, Varun Aggarwal (<i>Aspiring Minds</i>)	
• Contextual Intent Tracking for Personal Assistants	273
Yu Sun (<i>University of Melbourne</i>), Nicholas Jing Yuan (<i>Microsoft Corporation</i>), Yingzi Wang (<i>University of Science and Technology of China & Microsoft Research</i>), Xing Xie (<i>Microsoft Research</i>), Kieran McDonald (<i>Microsoft Corporation</i>), Rui Zhang (<i>University of Melbourne</i>)	
• An Empirical Study on Recommendation with Multiple Types of Feedback	283
Liang Tang, Bo Long, Bee-Chung Chen, Deepak Agarwal (<i>LinkedIn</i>)	
• An Engagement-Based Customer Lifetime Value System for E-commerce	293
Ali Vanderveld, Addhyan Pandey (<i>Groupon</i>), Angela Han (<i>Google</i>), Rajesh Parekh (<i>Facebook</i>)	
• Identifying Earmarks in Congressional Bills	303
Ellery Wulczyn (<i>Stanford University</i>), Madian Khabza (<i>Microsoft</i>), Vrushank Vora (<i>University of Chicago</i>), Matthew Heston (<i>Northwestern</i>), Joe Walsh, Christopher Berry, Rayid Ghani (<i>University of Chicago</i>)	
• Evaluating Mobile Apps with A/B and Quasi A/B Tests	313
Ya Xu, Nanyu Chen (<i>LinkedIn Corporation</i>)	
• Ranking Relevance in Yahoo Search	323
Dawei Yin, Yuening Hu, Jiliang Tang, Tim Daly, Mianwei Zhou, Hua Ouyang, Jianhui Chen, Changsung Kang, Hongbo Deng, Chikashi Nobata, Jean-Marc Langlois, Yi Chang (<i>Yahoo! Inc.</i>)	
• Identifying Decision Makers from Professional Social Networks	333
Shipeng Yu (<i>LinkedIn</i>), Evangelia Christakopoulou (<i>University of Minnesota</i>), Abhishek Gupta (<i>LinkedIn</i>)	
• Batch Model for Batched Timestamps Data Analysis with Application to the SSA Disability Program	343
Qingqi Yue, Ao Yuan, Xuan Che (<i>National Institutes of Health</i>), Minh Huynh (<i>Impac International LLC & National Institutes of Health</i>), Chunxiao Zhou (<i>National Institutes of Health</i>)	
• Collaborative Knowledge Base Embedding for Recommender Systems	353
Fuzheng Zhang, Nicholas Jing Yuan (<i>Microsoft Research</i>), Defu Lian (<i>University of Electronic Science and Technology of China</i>), Xing Xie, Wei-Ying Ma (<i>Microsoft Research</i>)	
• GLMix: Generalized Linear Mixed Models For Large-Scale Response Prediction	363
XianXing Zhang, Yitong Zhou, Yiming Ma, Bee-Chung Chen, Liang Zhang, Deepak Agarwal (<i>LinkedIn</i>)	
• A Non-parametric Approach to Detect Epileptogenic Lesions using Restricted Boltzmann Machines	373
Yijun Zhao, Bilal Ahmed (<i>Tufts University</i>), Thomas Thesen, Karen E. Blackmon (<i>New York University</i>), Jennifer G. Dy, Carla E. Brodley (<i>Northeastern University</i>), Ruben Kuznieky, Orrin Devinsky (<i>New York University</i>)	
• Recruitment Market Trend Analysis with Sequential Latent Variable Models	383
Chen Zhu (<i>Baidu Inc.</i>), Hengshu Zhu (<i>Baidu Inc. & Baidu Research-Big Data Lab</i>), Hui Xiong (<i>Rutgers University</i>), Pengliang Ding, Fang Xie (<i>Baidu Inc.</i>)	

- **Days on Market: Measuring Liquidity in Real Estate Markets**393
Hengshu Zhu (*Baidu Research-Big Data Lab*), Hui Xiong (*Rutgers University*),
Fangshuang Tang, Qi Liu (*University of Science and Technology of China*), Yong Ge (*University of Arizona*),
Enhong Chen (*University of Science and Technology of China*),
Yanjie Fu (*Missouri University of Science and Technology*)

Applied Data Science Track Invited Talks

- **Can You Teach the Elephant to Dance? AKA: Culture Eats Data Science for Breakfast**403
Jonathan Becher (*SAP*)
- **How Machine Learning has Finally Solved Wanamaker's Dilemma**405
Oliver Downs (*Amplero*)
- **Learning Sparse Models at Scale**407
Ralf Herbrich (*Amazon*)
- **Profiling Users from Online Social Behaviors with Applications for Tencent Social Ads**409
Ching Law (*Tencent*)
- **The Wisdom of Crowds: Best Practices for Data Prep & Machine Learning derived from Millions of Data Science Workflows**411
Ingo Mierswa (*RapidMiner*)
- **Bayesian Optimization and Embedded Learning Systems**413
Jeff Schneider (*Uber*)
- **Accelerating the Race to Autonomous Cars**415
Danny Shapiro (*NVIDIA*)
- **Large-Scale Machine Learning at Verizon: Theory and Applications**417
Ashok Srivastava (*Verizon*)
- **Computational Social Science: Exciting Progress and Future Challenges**419
Duncan Watts (*Microsoft*)

Applied Data Science Track Posters

- **MAP: Frequency-Based Maximization of Airline Profits based on an Ensemble Forecasting Approach**421
Bo An, Haipeng Chen (*Nanyang Technological University*),
Noseong Park, V.S. Subrahmanian (*University of Maryland*)
- **Gemello: Creating a Detailed Energy Breakdown from Just the Monthly Electricity Bill**431
Nipun Batra, Amarjeet Singh (*IIT Delhi*), Kamin Whitehouse (*University of Virginia*)
- **CaSMoS: A Framework for Learning Candidate Selection Models over Structured Queries and Documents**441
Fedor Borisjuk, Krishnamurthy Kenthapadi, David Stein, Bo Zhao (*LinkedIn Corporation*)
- **Domain Adaptation in the Absence of Source Domain Data**451
Boris Chidlovskii, Stephane Clinchant, Gabriela Csurka (*Xerox Research Centre Europe*)
- **Kam1n0: MapReduce-based Assembly Clone Search for Reverse Engineering**461
Steven H. H. Ding, Benjamin C. M. Fung (*McGill University*),
Philippe Charland (*Defence R&D Canada - Valcartier*)
- **Joint Optimization of Multiple Performance Metrics in Online Video Advertising**471
Sahin Cem Geyik, Sergey Faleev, Jianqiang Shen, Sean O'Donnell, Santanu Kolay (*Turn Inc.*)
- **Convolutional Neural Networks for Steady Flow Approximation**481
Xiaoxiao Guo (*University of Michigan*), Wei Li (*Autodesk Research*), Francesco Iorio (*Autodesk Research*)

• Computational Drug Repositioning Using Continuous Self-Controlled Case Series	491
Zhaobin Kuang (<i>University of Wisconsin</i>), James Thomson (<i>Morgridge Institute</i>), Michael Caldwell, Peggy Peissig (<i>Marshfield Clinic</i>), Ron Stewart (<i>Morgridge Institute</i>), David Page (<i>University of Wisconsin</i>)	
• How to Get Them a Dream Job?: Entity-Aware Features for Personalized Job Search Ranking	501
Jia Li (<i>University of Illinois at Chicago</i>), Dhruv Arya, Viet Ha-Thuc, Shakti Sinha (<i>LinkedIn</i>)	
• Scalable Fast Rank-1 Dictionary Learning for fMRI Big Data Analysis	511
Xiang Li, Milad Makkie (<i>University of Georgia</i>), Binbin Lin (<i>University of Michigan</i>), Mojtaba Sedigh Fazli (<i>University of Georgia</i>), Ian Davidson (<i>University of California, Davis</i>), Jieping Ye (<i>University of Michigan</i>), Tianming Liu, Shannon Quinn (<i>University of Georgia</i>)	
• CompanyDepot: Employer Name Normalization in the Online Recruitment Industry	521
Qiaoling Liu, Faizan Javed, Matt McNair (<i>CareerBuilder LLC</i>)	
• Understanding Behaviors that Lead to Purchasing: A Case Study of Pinterest	531
Caroline Lo (<i>Stanford University</i>), Dan Frankowski (<i>Pinterest</i>), Jure Leskovec (<i>Pinterest & Stanford University</i>)	
• Images Don't Lie: Transferring Deep Visual Semantic Features to Large-Scale Multimodal Learning to Rank	541
Corey Lynch, Kamelia Aryafar, Josh Attenberg (<i>Etsy</i>)	
• Text Mining in Clinical Domain: Dealing with Noise	549
Hoang Nguyen (<i>Data61 - CSIRO</i>), Jon Patrick (<i>University of Sydney</i>)	
• Detecting Devastating Diseases in Search Logs	559
John Paparizos (<i>Columbia University</i>), Ryen W. White, Eric Horvitz (<i>Microsoft Research</i>)	
• When Recommendation Goes Wrong - Anomalous Link Discovery in Recommendation Networks	569
Bryan Perozzi (<i>Stony Brook University</i>), Michael Schueppert, Jack Saalweachter (<i>Google</i>), Mayur Thakur (<i>Work performed while at Google</i>)	
• Deploying Analytics with the Portable Format for Analytics (PFA)	579
Jim Pivarski, Collin Bennett, Robert L. Grossman (<i>Open Data Group</i>)	
• Singapore in Motion: Insights on Public Transport Service Level Through Farecard and Mobile Data Analytics	589
Hasan Poonawala (<i>IBM Research</i>), Vinay Kolar (<i>Cisco systems inc.</i>), Sebastien Blandin, Laura Wynter, Sambit Sahu (<i>IBM research</i>)	
• EMBERS AutoGSR: Automated Coding of Civil Unrest Events	599
Parang Saraf, Naren Ramakrishnan (<i>Virginia Tech</i>)	
• Compute Job Memory Recommender System Using Machine Learning	609
Taraneh Taghavi, Maria Lupetini, Yaron Kretchmer (<i>Qualcomm Technologies Inc.</i>)	
• Scalable Time-Decaying Adaptive Prediction Algorithm	617
Yinyan Tan, Zhe Fan, Guilin Li, Fangshan Wang, Zhengbing Li, Shikai Liu, Qiuling Pan (<i>Huawei Software Technologies CO. LTD</i>), Eric P. Xing, Qirong Ho (<i>Carnegie Mellon University</i>)	
• Analyzing Volleyball Match Data from the 2014 World Championships Using Machine Learning Techniques	627
Jan Van Haaren (<i>KU Leuven</i>), Horesh Ben Shitrit (<i>PlayfulVision</i>), Jesse Davis (<i>KU Leuven</i>), Pascal Fua (<i>École Polytechnique Fédérale de Lausanne</i>)	
• Crime Rate Inference with Big Data	635
Hongjian Wang, Daniel Kifer, Corina Graif, Zhenhui Li (<i>Pennsylvania State University</i>)	
• Improving the Sensitivity of Online Controlled Experiments: Case Studies at Netflix	645
Huizhi Xie, Juliette Aurisset (<i>Netflix</i>)	
• Talent Circle Detection in Job Transition Networks	655
Huang Xu, Zhiwen Yu (<i>Northwestern Polytechnical University</i>), Jingyuan Yang, Hui Xiong (<i>Rutgers University</i>), Hengshu Zhu (<i>Baidu Research-Big Data Lab</i>)	

- **Bid-aware Gradient Descent for Unbiased Learning with Censored Data in Display Advertising** 665
Weinan Zhang (*Shanghai Jiao Tong University*), Tianxiong Zhou (*TukMob Inc. & University College London*), Jun Wang (*University College London & Media Gamma Limited*), Jian Xu (*TouchPal Inc.*)

Research Track Oral Papers

- **Compact and Scalable Graph Neighborhood Sketching** 685
Takuya Akiba, Yosuke Yano (*National Institute of Informatics*)
- **Streaming-LDA: A Copula-based Approach to Modeling Topic Dependencies in Document Streams** 695
Hesam Amoualian, Marianne Clausel, Eric Gaussier, Massih-Reza Amini (*University of Grenoble Alps*)
- **Assessing Human Error Against a Benchmark of Perfection** 705
Ashton Anderson (*Microsoft Research*), Jon Kleinberg (*Cornell University*), Sendhil Mullainathan (*Harvard University*)
- **Inferring Network Effects from Observational Data** 715
David Arbour, Dan Garant, David Jensen (*UMass Amherst*)
- **Communication Efficient Distributed Kernel Principal Component Analysis** 725
Maria-Florina Balcan (*Carnegie Mellon University*), Yingyu Liang (*Princeton University*), Le Song (*Georgia Institute of Technology*), David Woodruff (*IBM Research*), Bo Xie (*Georgia Institute of Technology*)
- **Keeping it Short and Simple: Summarising Complex Event Sequences with Multivariate Patterns** 735
Roel Bertens (*Utrecht University*), Jilles Vreeken (*Max Planck Institute for Informatics and Saarland University*), Arno Siebes (*Utrecht University*)
- **The Limits of Popularity-Based Recommendations, and the Role of Social Ties** 745
Marco Bressan, Stefano Leucci, Alessandro Panconesi (*Sapienza University of Rome*), Prabhakar Raghavan (*Google*), Erisa Terolli (*Sapienza University of Rome*)
- **Positive-Unlabeled Learning in Streaming Networks** 755
Shiyu Chang, Yang Zhang (*University of Illinois at Urbana-Champaign*), Jiliang Tang (*Michigan State University*), Dawei Yin, Yi Chang (*Yahoo! Inc.*), Mark A. Hasegawa-Johnson, Thomas S. Huang (*University of Illinois at Urbana-Champaign*)
- **FASCINATE: Fast Cross-Layer Dependency Inference on Multi-layered Networks** 765
Chen Chen, Hanghang Tong (*Arizona State University*), Lei Xie (*City University of New York*), Lei Ying (*Arizona State University*), Qing He (*University at Buffalo*)
- **Predicting Matchups and Preferences in Context** 775
Shuo Chen, Thorsten Joachims (*Cornell University*)
- **XGBoost: A Scalable Tree Boosting System** 785
Tianqi Chen, Carlos Guestrin (*University of Washington*)
- **Robust Influence Maximization** 795
Wei Chen (*Microsoft Research*), Tian Lin, Zihan Tan, Mingfei Zhao (*Tsinghua University*), Xuren Zhou (*The Hong Kong University of Science and Technology*)
- **Ranking Causal Anomalies via Temporal and Dynamical Analysis on Vanishing Correlations** 805
Wei Cheng, Kai Zhang, Haifeng Chen, Guofei Jiang, Zhengzhang Chen (*NEC Laboratories America*), Wei Wang (*University of California, Los Angeles*)
- **Towards Conversational Recommender Systems** 815
Konstantina Christakopoulou (*University of Minnesota*), Filip Radlinski, Katja Hofmann (*Microsoft*)
- **TRIEST: Counting Local and Global Triangles in Fully-Dynamic Streams with Fixed Memory Size** 825
Lorenzo De Stefani (*Brown University*), Alessandro Epasto (*Google*), Matteo Riondato (*Two Sigma Investments*), Eli Upfal (*Brown University*)
- **A Subsequence Interleaving Model for Sequential Pattern Mining** 835
Jaroslav Fowkes, Charles Sutton (*University of Edinburgh*)

• Efficient Frequent Directions Algorithm for Sparse Matrices	845
Mina Ghashami (<i>University of Utah</i>), Edo Liberty (<i>Yahoo Research</i>), Jeff M. Phillips (<i>University of Utah</i>)	
• node2vec: Scalable Feature Learning for Networks	855
Aditya Grover, Jure Leskovec (<i>Stanford University</i>)	
• Generalized Hierarchical Sparse Model for Arbitrary-Order Interactive Antigenic Sites Identification in Flu Virus Data	865
Lei Han (<i>Rutgers University</i>), Yu Zhang (<i>Hong Kong University of Science and Technology</i>), Xiu-Feng Wan (<i>Mississippi State University</i>), Tong Zhang (<i>Rutgers University & Baidu Inc.</i>)	
• Joint Community and Structural Hole Spanner Detection via Harmonic Modularity	875
Lifang He (<i>Shenzhen University</i>), Chun-Ta Lu (<i>University of Illinois at Chicago</i>), Jiaqi Ma (<i>Tsinghua University</i>), Jianping Cao (<i>National University of Defense Technology</i>), Linlin Shen (<i>Shenzhen University</i>), Philip S. Yu (<i>University of Illinois at Chicago</i>)	
• Robust Influence Maximization	885
Xinran He, David Kempe (<i>University of Southern California</i>)	
• FRAUDAR: Bounding Graph Fraud in the Face of Camouflage	895
Bryan Hooi, Hyun Ah Song, Alex Beutel, Neil Shah, Kijung Shin, Christos Faloutsos (<i>Carnegie Mellon University</i>)	
• Temporal Order-based First-Take-All Hashing for Fast Attention-Deficit-Hyperactive-Disorder Detection	905
Hao Hu, Joey Velez-Ginorio, Guo-Jun Qi (<i>University of Central Florida</i>)	
• When Social Influence Meets Item Inference	915
Hui-Ju Hung (<i>Pennsylvania State University</i>), Hong-Han Shuai, De-Nian Yang, Liang-Hao Huang (<i>Academia Sinica</i>), Wang-Chien Lee (<i>The Pennsylvania State University</i>), Jian Pei (<i>Simon Fraser University</i>), Ming-Syan Chen (<i>National Taiwan University</i>)	
• Privacy-preserving Class Ratio Estimation	925
Arun Iyer, J. Saketha Nath, Sunita Sarawagi (<i>IIT Bombay</i>)	
• Extreme Multi-label Loss Functions for Recommendation, Tagging, Ranking & Other Missing Label Applications	935
Himanshu Jain, Yashoteja Prabhu (<i>Indian Institute of Technology Delhi</i>), Manik Varma (<i>Microsoft Research</i>)	
• CatchTartan: Representing and Summarizing Dynamic Multicontextual Behaviors	945
Meng Jiang (<i>University of Illinois at Urbana-Champaign</i>), Christos Faloutsos (<i>Carnegie Mellon University</i>), Jiawei Han (<i>University of Illinois at Urbana-Champaign</i>)	
• Smart Reply: Automated Response Suggestion for Email	955
Anjuli Kannan, Karol Kurach, Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado, Laszlo Lukacs, Marina Ganea, Peter Young, Vivek Ramavajjala (<i>Google</i>)	
• Mining Subgroups with Exceptional Transition Behavior	965
Florian Lemmerich (<i>GESIS - Leibniz Institute for the Social Sciences & University of Koblenz-Landau</i>), Martin Becker (<i>University of Würzburg</i>), Philipp Singer (<i>GESIS - Leibniz Institute for the Social Sciences & University of Koblenz-Landau</i>), Denis Helic (<i>Graz University of Technology</i>), Andreas Hotho (<i>University of Würzburg & L3S Hannover</i>), Markus Strohmaier (<i>GESIS - Leibniz Institute for the Social Sciences & University of Koblenz-Landau</i>)	
• Point-of-Interest Recommendations: Learning Potential Check-ins from Friends	975
Huayu Li (<i>The University of North Carolina at Charlotte</i>), Yong Ge (<i>University of Arizona</i>), Richang Hong (<i>Hefei University of Technology</i>), Hengshu Zhu (<i>Baidu Research-Big Data</i>)	
• QUINT: On Query-Specific Optimal Networks	985
Liangyue Li (<i>Arizona State University</i>), Yuan Yao (<i>Nanjing University</i>), Jie Tang (<i>Tsinghua University</i>), Wei Fan (<i>Baidu Big Data Lab</i>), Hanghang Tong (<i>Arizona State University</i>)	
• Dynamic Clustering of Streaming Short Documents	995
Shangsong Liang, Emine Yilmaz (<i>University College London</i>), Evangelos Kanoulas (<i>University of Amsterdam</i>)	
• Rebalancing Bike Sharing Systems: A Multi-source Data Smart Optimization	1005
Junming Liu (<i>Rutgers University</i>), Leilei Sun (<i>Dalian University of Technology</i>), Weiwei Chen, Hui Xiong (<i>Rutgers University</i>)	
• Unified Point-of-Interest Recommendation with Temporal Interval Assessment	1015
Yanchi Liu (<i>Rutgers University</i>), Chuanren Liu (<i>Drexel University</i>), Bin Liu, Meng Qu, Hui Xiong (<i>Rutgers University</i>)	

• AnyDBC: An Efficient Anytime Density-based Clustering Algorithm for Very Large Complex Datasets	1025
Son T. Mai, Ira Assent, Martin Storgaard (<i>Aarhus University</i>)	
• Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs	1035
Emaad Manzoor (<i>Stony Brook University</i>), Sadegh M. Milajerdi (<i>University of Illinois at Chicago</i>), Leman Akoglu (<i>Stony Brook University</i>)	
• Regime Shifts in Streams: Real-time Forecasting of Co-evolving Time Sequences	1045
Yasuko Matsubara, Yasushi Sakurai (<i>Kumamoto University</i>)	
• Skinny-dip: Clustering in a Sea of Noise	1055
Samuel Maurus (<i>Helmholtz Zentrum München & Technische Universität München</i>), Claudia Plant (<i>University of Vienna</i>)	
• Semi-Markov Switching Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems	1065
Igor Melnyk, Arindam Banerjee (<i>University of Minnesota</i>), Bryan Matthews, Nikunj Oza (<i>NASA Ames Research Center</i>)	
• Continuous Experience-aware Language Model	1075
Subhabrata Mukherjee (<i>Max Planck Institute for Informatics</i>), Stephan Günnemann (<i>Technical University of Munich</i>), Gerhard Weikum (<i>Max Planck Institute for Informatics</i>)	
• Structural Neighborhood Based Classification of Nodes in a Network	1085
Sharad Nandanwar, M. N. Murty (<i>Indian Institute of Science</i>)	
• Modeling Precursors for Event Forecasting via Nested Multi-Instance Learning	1095
Yue Ning, Sathappan Muthiah (<i>Virginia Tech</i>), Huzefa Rangwala (<i>George Mason University</i>), Naren Ramakrishnan (<i>Virginia Tech</i>)	
• Asymmetric Transitivity Preserving Graph Embedding	1105
Mingdong Ou, Peng Cui (<i>Tsinghua University</i>), Jian Pei (<i>Simon Fraser University</i>), Ziwei Zhang, Wenwu Zhu (<i>Tsinghua University</i>)	
• PTE: Enumerating Trillion Triangles On Distributed Systems	1115
Ha-Myung Park, Sung-Hyon Myaeng (<i>Korea Advanced Institute of Science and Technology</i>), U. Kang (<i>Seoul National University</i>)	
• Robust Large-Scale Machine Learning in the Cloud	1125
Steffen Rendle, Dennis Fetterly, Eugene J. Shekita, Bor-ying Su (<i>Google Inc.</i>)	
• Why Should I Trust You? Explaining the Predictions of Any Classifier	1135
Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin (<i>University of Washington</i>)	
• ABRA: Approximating Betweenness Centrality in Static and Dynamic Graphs with Rademacher Averages	1145
Matteo Riondato (<i>Two Sigma Investments</i>), Eli Upfal (<i>Brown University</i>)	
• Sampling of Attributed Networks from Hierarchical Generative Models	1155
Pablo Robles (<i>Purdue University</i>), Sebastian Moreno (<i>Universidad Adolfo Ibañez</i>), Jennifer Neville (<i>Purdue University</i>)	
• Goal-Directed Inductive Matrix Completion	1165
Si Si, Kai-Yang Chiang (<i>University of Texas at Austin</i>), Cho-Jui Hsieh (<i>University of California, Davis</i>), Nikhil Rao (<i>Technicolor R&I</i>), Inderjit S. Dhillon (<i>University of Texas at Austin</i>)	
• Graph Wavelets via Sparse Cuts	1175
Arlei Silva, Xuan Hong Dang (<i>University of California, Santa Barbara</i>), Prithwish Basu (<i>Raytheon BBN Technologies</i>), Ambuj Singh (<i>University of California, Santa Barbara</i>), Ananthram Swami (<i>Army Research Laboratory</i>)	
• Lexis: An Optimization Framework for Discovering the Hierarchical Structure of Sequential Data	1185
Payam Siyari, Bistra Dilkina, Constantine Dovrolis (<i>Georgia Institute of Technology</i>)	
• Towards Optimal Cardinality Estimation of Unions and Intersections with Sketches	1195
Daniel Ting (<i>Facebook</i>),	

• Overcoming Key Weaknesses of Distance-based Neighbourhood Methods using a Data Dependent Dissimilarity Measure	1205
Kai Ming Ting (<i>Federation University</i>), Ye Zhu, Mark Carman (<i>Monash University</i>), Yue Zhu, Zhi-Hua Zhou (<i>Nanjing University</i>)	
• Just One More: Modeling Binge Watching Behavior	1215
William Trouleau (<i>École Polytechnique Fédérale de Lausanne</i>), Azin Ashkan, Weicong Ding, Brian Eriksson (<i>Technicolor Research</i>)	
• Structural Deep Network Embedding	1225
Daixin Wang, Peng Cui, Wenwu Zhu (<i>Tsinghua University</i>)	
• Targeted Topic Modeling for Focused Analysis	1235
Shuai Wang, Zhiyuan Chen, Geli Fei, Bing Liu, Sherry Emery (<i>University of Illinois at Chicago</i>)	
• Structured Doubly Stochastic Matrix for Graph Based Clustering: Structured Doubly Stochastic Matrix	1245
Xiaoqian Wang, Feiping Nie, Heng Huang (<i>University of Texas at Arlington</i>)	
• A Multiple Test Correction for Streams and Cascades of Statistical Hypothesis Tests	1255
Geoffrey I. Webb, François Petitjean (<i>Monash University</i>)	
• Revisiting Random Binning Features: Fast Convergence and Strong Parallelizability	1265
Lingfei Wu (<i>College of William and Mary</i>), Ian E.H. Yen (<i>University of Texas at Austin</i>), Jie Chen (<i>IBM Research</i>), Rui Yan (<i>Baidu Inc.</i>)	
• Robust Extreme Multi-label Learning	1275
Chang Xu Dacheng Tao (<i>University of Technology Sydney</i>), Chao Xu (<i>Peking University</i>)	
• Taxi Driving Behavior Analysis in Latent Vehicle-to-Vehicle Networks: A Social Influence Perspective	1285
Tong Xu (<i>University of Science and Technology of China</i>), Hengshu Zhu (<i>Baidu Research-Big Data Lab</i>), Xiangyu Zhao, Qi Liu (<i>University of Science and Technology of China</i>), Hao Zhong (<i>Rutgers University</i>), Enhong Chen (<i>University of Science and Technology of China</i>), Hui Xiong (<i>Rutgers University</i>)	
• DeepIntent: Learning Attentions for Online Advertising with Recurrent Neural Networks	1295
Shuangfei Zhai (<i>Binghamton University</i>), Keng-hao Chang, Ruofei Zhang (<i>Microsoft</i>), Zhongfei Mark Zhang (<i>Binghamton University</i>)	
• GMove: Group-Level Mobility Modeling Using Geo-Tagged Social Media	1305
Chao Zhang, Keyang Zhang, Quan Yuan (<i>University of Illinois at Urbana-Champaign</i>), Luming Zhang (<i>Hefei University of Technology</i>), Tim Hanratty (<i>U.S. Army Research Laboratory</i>), Jiawei Han (<i>University of Illinois at Urbana-Champaign</i>)	
• Approximate Personalized PageRank on Dynamic Graphs	1315
Hongyang Zhang, Peter Lofgren, Ashish Goel (<i>Stanford University</i>)	
• Annealed Sparsity via Adaptive and Dynamic Shrinking	1325
Kai Zhang (<i>NEC Laboratories America</i>), Shandian Zhe (<i>Purdue University</i>), Chaoran Cheng, Zhi Wei (<i>New Jersey Institute of Technology</i>), Zhengzhang Chen, Haifeng Chen, Guofei Jiang (<i>NEC Laboratories America</i>), Yuan Qi (<i>Purdue University</i>), Jieping Ye (<i>University of Michigan, Ann Arbor</i>)	
• Partial Label Learning via Feature-Aware Disambiguation	1335
Min-Ling Zhang, Bin-Bin Zhou, Xu-Ying Liu (<i>Southeast University</i>)	
• FINAL: Fast Attributed Network Alignment	1345
Si Zhang, Hanghang Tong (<i>Arizona State University</i>)	
• Come-and-Go Patterns of Group Evolution: A Dynamic Model	1355
Tianyang Zhang, Peng Cui (<i>Tsinghua University</i>), Christos Faloutsos (<i>Carnegie Mellon University</i>), Yunfei Lu (<i>Tsinghua University</i>), Hao Ye (<i>Tencent Corporation</i>), Wenwu Zhu, Shiqiang Yang (<i>Tsinghua University</i>)	
• NetCycle: Collective Evolution Inference in Heterogeneous Information Networks	1365
Yizhou Zhang, Yun Xiong (<i>Fudan University</i>), Xiangnan Kong (<i>Worcester Polytechnic Institute</i>), Yangyong Zhu (<i>Fudan University</i>)	

- **Accelerating Online CP Decompositions for Higher Order Tensors** 1375
Shuo Zhou, Nguyen Xuan Vinh, James Bailey, Yunzhe Jia (*The University of Melbourne*),
Ian Davidson (*The University of California - Davis*)

Research Track Posters

- **Optimal Reserve Prices in Upstream Auctions: Empirical Application on Online Video Advertising** 1395
Miguel Angel Alcobendas Lisbona, Sheide Chammas, Kuang-chih Lee (*Yahoo! Inc.*)
- **Burstiness Scale: A Parsimonious Model for Characterizing Random Series of Events** 1405
Rodrigo A. S. Alves (*CEFET-MG*), Renato Assunção, Pedro O. S. Vaz de Melo (*UFMG*)
- **MANTRA: A Scalable Approach to Mining Temporally Anomalous Sub-trajectories** 1415
Prithu Banerjee (*University of British Columbia*), Pranali Yawalkar, Sayan Ranu (*IIT Madras*)
- **From Prediction to Action: A Closed-Loop Approach for Data-Guided Network Resource Allocation** 1425
Yanan Bao, Huasen Wu, Xin Liu (*University of California, Davis*)
- **Towards Robust and Versatile Causal Discovery for Business Applications** 1435
Giorgos Borboudakis, Ioannis Tsamardinos (*University of Crete; Gnosis Data Analysis IKE*)
- **Deep Visual-Semantic Hashing for Cross-Modal Retrieval** 1445
Yue Cao, Mingsheng Long, Jianmin Wang (*Tsinghua University*),
Qiang Yang (*Hong Kong University of Science and Technology*),
Philip S. Yu (*University of Illinois at Chicago & Tsinghua University*)
- **Predicting Socio-Economic Indicators using News Events** 1455
Sunandan Chakraborty, Ashwin Venkataraman, Srikanth Jagabathula, Lakshminarayanan Subramanian
(*New York University*)
- **City-Scale Map Creation and Updating using GPS Collections** 1465
Chen Chen, Cewu Lu (*Stanford University*), Qixing Huang (*Toyota Technological Institute at Chicago*),
Qiang Yang (*Hong Kong University of Science and Technology*), Dimitrios Gunopulos (*University of Athens*),
Leonidas Guibas (*Stanford University*)
- **Compressing Convolutional Neural Networks in the Frequency Domain** 1475
Wenlin Chen (*Washington University in St. Louis*), James Wilson (*University of Edinburgh*),
Stephen Tyree (*NVIDIA*), Kilian Q. Weinberger (*Cornell University*),
Yixin Chen (*Washington University in St. Louis*)
- **Parallel Dual Coordinate Descent Method for Large-scale Linear Classification in Multi-core Environments** 1485
Wei-Lin Chiang, Mu-Chu Lee, Chih-Jen Lin (*National Taiwan University*)
- **Multi-layer Representation Learning for Medical Concepts** 1495
Edward Choi, Mohammad Taha Bahadori (*Georgia Institute of Technology*),
Elizabeth Searles, Catherine Coffey, Michael Thompson, James Bost, Javier Tejedor-Sojo
(*Children's Healthcare of Atlanta*),
Jimeng Sun (*Georgia Institute of Technology*)
- **Finding Gangs in War from Signed Networks** 1505
Linyang Chu (*Simon Fraser University*), Zhefeng Wang (*University of Science and Technology of China*),
Jian Pei, Jiannan Wang, Zijin Zhao (*Simon Fraser University*),
Enhong Chen (*University of Science and Technology of China*)
- **Efficient Processing of Network Proximity Queries via Chebyshev Acceleration** 1515
Mustafa Coskun (*Case Western Reserve University*), Ananth Grama (*Purdue University*),
Mehmet Koyuturk (*Case Western Reserve University*)
- **Latent Space Model for Road Networks to Predict Time-Varying Traffic** 1525
Dingxiong Deng, Cyrus Shahabi, Ugur Demiryurek (*University of Southern California*),
Linhong Zhu (*Information Sciences Institute*), Rose Yu, Yan Liu (*University of Southern California*)

• Compressing Graphs and Indexes with Recursive Graph Bisection	1535
Laxman Dhulipala (<i>Carnegie Mellon University</i>), Igor Kabiljo, Brian Karrer, Giuseppe Ottaviano, Sergey Pupyrev, Alon Shalita (<i>Facebook</i>)	
• Fast Unsupervised Online Drift Detection Using Incremental Kolmogorov-Smirnov Test	1545
Denis dos Reis (<i>Universidade de São Paulo</i>), Peter Flach (<i>University of Bristol</i>), Stan Matwin (<i>Polish Academy of Sciences</i>), Gustavo Batista (<i>Universidade de São Paulo</i>)	
• Recurrent Marked Temporal Point Processes: Embedding Event History to Vector	1555
Nan Du, Hanjun Dai, Rakshit Trivedi (<i>Georgia Institute of Technology</i>), Utkarsh Upadhyay, Manuel Gomez-Rodriguez (<i>Max Planck Institute for Software Systems</i>), Le Song (<i>Georgia Institute of Technology</i>)	
• Learning Cumulatively to Become More Knowledgeable	1565
Geli Fei, Shuai Wang, Bing Liu (<i>University of Illinois at Chicago</i>)	
• Squish: Near-Optimal Compression for Archival of Relational Datasets	1575
Yihan Gao, Aditya Parameswaran (<i>University of Illinois at Urbana-Champaign</i>)	
• Fast Component Pursuit for Large-Scale Inverse Covariance Estimation	1585
Lei Han (<i>Rutgers University</i>), Yu Zhang (<i>Hong Kong University of Science and Technology</i>), Tong Zhang (<i>Rutgers University & Baidu Inc.</i>)	
• Meta Structure: Computing Relevance in Large Heterogeneous Information Networks	1595
Zhipeng Huang, Yudian Zheng, Reynold Cheng (<i>The University of Hong Kong</i>), Yizhou Sun (<i>Northeastern University</i>), Nikos Mamoulis, Xiang Li (<i>The University of Hong Kong</i>)	
• Robust and Effective Metric Learning Using Capped Trace Norm: Metric Learning via Capped Trace Norm	1605
Zhouyuan Huo, Feiping Nie, Heng Huang (<i>University of Texas at Arlington</i>)	
• Subjectively Interesting Component Analysis: Data Projections that Contrast with Prior Expectations	1615
Bo Kang (<i>Ghent University</i>), Jefrey Lijffijt (<i>Ghent University & University of Bristol</i>), Raúl Santos-Rodríguez (<i>University of Bristol</i>), Tijl De Bie (<i>Ghent University</i>)	
• Online Optimization Methods for the Quantification Problem	1625
Purushottam Kar (<i>Indian Institute of Technology Kanpur</i>), Shuai Li (<i>University of Insubria</i>), Harikrishna Narasimhan (<i>Harvard University</i>), Sanjay Chawla, Fabrizio Sebastiani (<i>Qatar Computing Research Institute, HBKU</i>)	
• Smart Broadcasting: Do You Want to be Seen?	1635
Mohammad Reza Karimi, Erfan Tavakoli (<i>Sharif University</i>), Mehrdad Farajtabar, Le Song (<i>Georgia Institute of Technology</i>), Manuel Gomez Rodriguez (<i>Max Planck Institute for Software Systems</i>)	
• How to Compete Online for News Audience: Modeling Words that Attract Clicks	1645
Joon Hee Kim (<i>Korea Advanced Institute of Science and Technology</i>), Amin Mantrach, Alejandro Jaimes (<i>Yahoo! Labs</i>), Alice Oh (<i>Korea Advanced Institute of Science and Technology</i>)	
• Causal Clustering for 1-Factor Measurement Models	1655
Erich Kummerfeld (<i>University of Pittsburgh</i>), Joseph Ramsey (<i>Carnegie Mellon University</i>)	
• Optimally Discriminative Choice Sets in Discrete Choice Models: Application to Data-Driven Test Design	1665
Igor Labutov, Frans Schalekamp (<i>Cornell University</i>), Kelvin Luu (<i>University of Washington</i>), Hod Lipson (<i>Columbia University</i>), Christoph Studer (<i>Cornell University</i>)	
• Interpretable Decision Sets: A Joint Framework for Description and Prediction	1675
Himabindu Lakkaraju, Stephen H. Bach, Jure Leskovec (<i>Stanford University</i>)	
• Lightweight Monitoring of Distributed Streams	1685
Amnon Lazerson (<i>Technion -- Israel Institute of Technology</i>), Daniel Keren (<i>Haifa University</i>), Assaf Schuster (<i>Technion -- Israel Institute of Technology</i>)	

• Bayesian Inference of Arrival Rate and Substitution Behavior from Sales Transaction Data with Stockouts	1695
Benjamin Letham, Lydia M. Letham (<i>Massachusetts Institute of Technology</i>), Cynthia Rudin (<i>Duke University</i>)	
• Parallel Lasso Screening for Big Data Optimization	1705
Qingyang Li (<i>Arizona State University</i>), Shuang Qiu (<i>University of Michigan</i>), Shuiwang Ji (<i>Washington State University</i>), Paul M. Thompson (<i>University of Southern California</i>), Jieping Ye, Jie Wang (<i>University of Michigan</i>)	
• A Multi-Task Learning Formulation for Survival Analysis	1715
Yan Li (<i>Wayne State University</i>), Jie Wang, Jieping Ye (<i>University of Michigan</i>), Chandan K. Reddy (<i>Wayne State University</i>)	
• A Real Linear and Parallel Multiple Longest Common Subsequences (MLCS) Algorithm	1725
Yanni Li, Hui Li (<i>Xidian University</i>), Tihua Duan (<i>Shanghai Finance University</i>), Sheng Wang (<i>Conventry University</i>), Zhi Wang, Yang Cheng (<i>Xidian University</i>)	
• Multi-Task Feature Interaction Learning	1735
Kaixiang Lin, Jianpeng Xu, Inci M. Baytas (<i>Michigan State University</i>), Shuiwang Ji (<i>Washington State University</i>), Jiayu Zhou (<i>Michigan State University</i>)	
• Infinite Ensemble for Image Clustering	1745
Hongfu Liu, Ming Shao, Sheng Li, Yun Fu (<i>Northeastern University</i>)	
• Scalable Pattern Matching over Compressed Graphs via Dedensification	1755
Antonio Maccioni (<i>Roma Tre University</i>), Daniel J. Abadi (<i>Yale University</i>)	
• Scalable Betweenness Centrality Maximization via Sampling	1765
Ahmad Mahmoody (<i>Brown University</i>), Charalampos E. Tsourakakis (<i>Harvard University</i>), Eli Upfal (<i>Brown University</i>)	
• User Identity Linkage by Latent User Space Modelling	1775
Xin Mu (<i>Nanjing University</i>), Feida Zhu, Ee-Peng Lim (<i>Singapore Management University</i>), Jing Xiao, Jianzong Wang (<i>Ping An Technology (Shenzhen) Co.,Ltd</i>), Zhi-Hua Zhou (<i>Nanjing University</i>)	
• Safe Pattern Pruning: An Efficient Approach for Predictive Pattern Mining	1785
Kazuya Nakagawa, Shinya Suzumura, Masayuki Karasuyama (<i>Nagoya Institute of Technology</i>), Koji Tsuda (<i>University of Tokyo</i>), Ichiro Takeuchi (<i>Nagoya Institute of Technology</i>)	
• Predict Risk of Relapse for Patients with Multiple Stages of Treatment of Depression	1795
Zhi Nie (<i>Arizona State University & University of Michigan, Ann Arbor</i>), Pinghua Gong, Jieping Ye (<i>University of Michigan, Ann Arbor</i>)	
• Lossless Separation of Web Pages into Layout Code and Data	1805
Adi Omari, Benny Kimelfeld, Eran Yahav (<i>Technion</i>), Sharon Shoham (<i>Tel Aviv University</i>)	
• Unbounded Human Learning: Optimal Scheduling for Spaced Repetition	1815
Siddharth Reddy, Igor Labutov, Siddhartha Banerjee, Thorsten Joachims (<i>Cornell University</i>)	
• Label Noise Reduction in Entity Typing by Heterogeneous Partial-Label Embedding	1825
Xiang Ren, Wenqi He, Meng Qu (<i>University of Illinois at Urbana-Champaign</i>), Clare R. Voss (<i>Army Research Laboratory</i>), Heng Ji (<i>Rensselaer Polytechnic Institute</i>), Jiawei Han (<i>University of Illinois at Urbana-Champaign</i>)	
• Reconstructing an Epidemic Over Time	1835
Polina Rozenshtein, Aristides Gionis (<i>Aalto University</i>), B. Aditya Prakash (<i>Virginia Tech</i>), Jilles Vreeken (<i>Saarland University & Max Planck Institute for Informatics</i>)	
• Improving Survey Aggregation with Sparsely Represented Signals	1845
Tianlin Shi (<i>Stanford University</i>), Forest Agostinelli (<i>University of California, Irvine</i>), Matthew Staib (<i>Massachusetts Institute of Technology</i>), David Wipf, Thomas Moscibroda (<i>Microsoft Research</i>)	
• Dynamics of Large Multi-View Social Networks: Synergy, Cannibalization and Cross-View Interplay	1855
Yu Shi (<i>University of Illinois at Urbana-Champaign</i>), Myunghwan Kim, Shaunak Chatterjee, Mitul Tiwari, Souvik Ghosh, Rómer Rosales (<i>LinkedIn Corporation</i>)	

- **Data-driven Automatic Treatment Regimen Development and Recommendation** 1865
Leilei Sun (*Dalian University of Technology*), Chuanren Liu (*Drexel University*),
Chonghui Guo (*Dalian University of Technology*), Hui Xiong (*Rutgers University*),
Yanming Xie (*China Academy of Chinese Medical Sciences*)
- **Scalable Partial Least Squares Regression on Grammar-Compressed Data Matrices** 1875
Yasuo Tabei (*Japan Science and Technology Agency*), Hiroto Saigo, Yoshihiro Yamanishi (*Kyushu University*),
Simon J. Puglisi (*University of Helsinki*)
- **From Truth Discovery to Trustworthy Opinion Discovery: An Uncertainty-Aware Quantitative Modeling Approach** 1885
Mengting Wan (*University of California, San Diego*), Xiangyu Chen (*University of Illinois, Urbana-Champaign*),
Lance Kaplan (*U.S. Army Research Laboratory*), Jiawei Han (*University of Illinois, Urbana-Champaign*),
Jing Gao (*SUNY Buffalo*), Bo Zhao (*LinkedIn*)
- **The Million Domain Challenge: Broadcast Email Prioritization by Cross-domain Recommendation** 1895
Beidou Wang (*Zhejiang University & Simon Fraser University*), Martin Ester (*Simon Fraser University*),
Yikang Liao, Jiajun Bu, Yu Zhu (*Zhejiang University*), Ziyu Guan (*Northwest University of China*),
Deng Cai (*Zhejiang University*)
- **Transfer Knowledge between Cities** 1905
Ying Wei (*Hong Kong University of Science and Technology*), Yu Zheng (*Microsoft Research*),
Qiang Yang (*Hong Kong University of Science and Technology*)
- **Probabilistic Robust Route Recovery with Spatio-Temporal Dynamics** 1915
Hao Wu, Jiangyun Mao, Weiwei Sun (*Fudan University*), Baihua Zheng (*Singapore Management University*),
Hanyuan Zhang, Ziyang Chen, Wei Wang (*Fudan University*)
- **A Truth Discovery Approach with Theoretical Guarantee** 1925
Houping Xiao, Jing Gao (*The State University of New York at Buffalo*), Zhaoran Wang (*Princeton University*),
Shiyu Wang (*University of Illinois at Urbana-Champaign*), Lu Su (*The State University of New York at Buffalo*),
Han Liu (*Princeton University*)
- **Towards Confidence in the Truth: A Bootstrapping based Truth Discovery Approach** 1935
Houping Xiao, Jing Gao, Qi Li, Fenglong Ma, Lu Su (*The State University of New York at Buffalo*),
Yunlong Feng (*Katholieke Universiteit Leuven*), Aidong Zhang (*The State University of New York at Buffalo*)
- **Online Feature Selection: A Limited-Memory Substitution Algorithm and Its Asynchronous Parallel Variation** 1945
Haichuan Yang (*University of Rochester*), Ryohei Fujimaki, Yukitaka Kusumura (*NEC*),
Ji Liu (*University of Rochester*)
- **Absolute Fused Lasso and Its Application to Genome-Wide Association Studies** 1955
Tao Yang (*Arizona State University*), Jun Liu (*SAS Institute Inc.*), Pinghua Gong (*University of Michigan*),
Ruiwen Zhang (*SAS Institute Inc.*), Xiaotong Shen, Jieping Ye (*University of Michigan*)
- **Diversified Temporal Subgraph Pattern Mining** 1965
Yi Yang (*Fudan University & The Chinese University of Hong Kong*),
Da Yan, Huanhuan Wu, James Cheng (*The Chinese University of Hong Kong*),
Shuigeng Zhou (*Fudan University*), John C.S. Lui (*The Chinese University of Hong Kong*)
- **Distributing the Stochastic Gradient Sampler for Large-Scale LDA** 1975
Yuan Yang (*Beihang University & Tsinghua University*), Jianfei Chen, Jun Zhu (*Tsinghua University*)
- **FUSE: Full Spectral Clustering** 1985
Wei Ye (*Ludwig-Maximilians-Universität München*), Sebastian Goebel (*Ludwig-Maximilians-Universität München*),
Claudia Plant (*University of Vienna*), Christian Böhm (*Ludwig-Maximilians-Universität München*)
- **A Text Clustering Algorithm Using an Online Clustering Scheme for Initialization** 1995
Jianhua Yin (*Tsinghua University*), Jianyong Wang (*Tsinghua University & Jiangsu Normal University*)
- **Convex Optimization for Linear Query Processing under Approximate Differential Privacy** 2005
Ganzhao Yuan (*South China University of Technology*), Yin Yang (*Hamad Bin Khalifa University*),
Zhenjie Zhang (*Illinois at Singapore Pte. Ltd.*), Zhifeng Hao (*Foshan University*)

• Beyond Sigmoids: The NetTide Model for Social Network Growth, and Its Applications	2015
Chengxi Zang, Peng Cui (<i>Tsinghua University</i>), , Christos Faloutsos (<i>Carnegie Mellon University</i>)	
• Online Context-Aware Recommendation with Time Varying Multi-Armed Bandit	2025
Chunqiu Zeng, Qing Wang, Shekoofeh Mokhtari, Tao Li (<i>Florida International University</i>)	
• Accelerated Stochastic Block Coordinate Descent with Optimal Sampling	2035
Aston Zhang (<i>University of Illinois at Urbana-Champaign</i>), Quanquan Gu (<i>University of Virginia</i>)	
• Collaborative Multi-View Denoising	2045
Lei Zhang, Shupeng Wang, Xiaoyu Zhang, Yong Wang, Binbin Li (<i>Chinese Academy of Sciences</i>), Dinggang Shen (<i>University of North Carolina</i>), Shuiwang Ji (<i>Washington State University</i>)	
• Online Asymmetric Active Learning with Imbalanced Data	2055
Xiaoxuan Zhang, Tianbao Yang, Padmini Srinivasan (<i>The University of Iowa</i>)	
• FLASH: Fast Bayesian Optimization for Data Analytic Pipelines	2065
Yuyu Zhang, Mohammad Taha Bahadori, Hang Su, Jimeng Sun (<i>Georgia Institute of Technology</i>)	
• Portfolio Selections in P2P Lending: A Multi-Objective Perspective	2075
Hongke Zhao, Qi Liu, Guifeng Wang (<i>University of Science and Technology of China</i>), Yong Ge (<i>University of Arizona</i>), Enhong Chen (<i>University of Science and Technology of China</i>)	
• Hierarchical Incomplete Multi-source Feature Learning for Spatiotemporal Event Forecasting	2085
Liang Zhao (<i>Virginia Tech</i>), Jieping Ye (<i>University of Michigan</i>), Feng Chen (<i>University at Albany, SUNY</i>), Chang-Tien Lu, Naren Ramakrishnan (<i>Virginia Tech</i>)	
• Efficient Shift-Invariant Dictionary Learning	2095
Guoqing Zheng, Yiming Yang, Jaime Carbonell (<i>Carnegie Mellon University</i>)	
• Topic Modeling of Short Texts: A Pseudo-Document View	2105
Yuan Zuo, Junjie Wu, Hui Zhang, Hao Lin, Fei Wang, Ke Xu (<i>Beihang University</i>), Hui Xiong (<i>Rutgers University</i>)	

Tutorials

• Scalable Data Analytics Using R: Single Machines to Hadoop Spark Clusters	2115
John-Mark Agosta, Debraj GuhaThakurta, Robert Horton, Mario Inchiosa, Srini Kumar, Mengyue Zhao (<i>Microsoft</i>)	
• Lifelong Machine Learning and Computer Reading the Web	2117
Zhiyuan Chen (<i>Google</i>), Estevam R. Hruschka Jr. (<i>Federal University of Sao Carlos & Carnegie Mellon University</i>), Bing Liu (<i>University of Illinois at Chicago</i>)	
• IoT Big Data Stream Mining	2119
Gianmarco De Francisci Morales (<i>Qatar Computing Research Institute</i>), Albert Bifet (<i>Université Paris-Saclay</i>), Latifur Khan (<i>University of Texas at Dallas</i>), Joao Gama (<i>University of Porto</i>), Wei Fan (<i>Baidu Research</i>)	
• Mining Reliable Information from Passively and Actively Crowdsourced Data	2121
Jing Gao, Qi Li (<i>SUNY Buffalo</i>), Bo Zhao (<i>LinkedIn</i>), Wei Fan (<i>Baidu Research Big Data Lab</i>), Jiawei Han (<i>University of Illinois at Urbana-Champaign</i>)	
• Streaming Analytics	2123
Ashish Gupta (<i>LinkedIn Corporation</i>), Neera Agarwal (<i>Startup</i>),	
• Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining	2125
Sara Hajian (<i>Eurecat</i>), Francesco Bonchi (<i>ISI Foundation</i>), Carlos Castillo (<i>Eurecat</i>)	
• Collective Sensemaking via Social Sensors: Extracting, Profiling, Analyzing, and Predicting Real-world Events	2127
Yuheng Hu (<i>University of Illinois at Chicago</i>), Yu-Ru Lin (<i>University of Pittsburgh</i>), Jiebo Luo (<i>University of Rochester</i>)	
• Extracting Optimal Performance from Dynamic Time Warping	2129
Abdullah Mueen (<i>University of New Mexico</i>), Eamonn Keogh (<i>University of California, Riverside</i>)	
• Scalable Learning of Graphical Models	2131
Francois Petitjean, Geoffrey I. Webb (<i>Monash University</i>)	

• Leveraging Propagation for Data Mining: Models, Algorithms and Applications	2133
B. Aditya Prakash, Naren Ramakrishnan (<i>Virginia Tech.</i>)	
• CNTK: Microsoft's Open-Source Deep-Learning Toolkit	2135
Frank Seide, Amit Agarwal (<i>Microsoft Research</i>)	
• Healthcare Data Mining with Matrix Models	2137
Fei Wang (<i>Cornell University</i>), Ping Zhang (<i>IBM T. J. Watson Research Center</i>), Joel Dudley (<i>Icahn School of Medicine at Mount Sinai</i>)	
• Business Applications of Predictive Modeling at Scale	2139
Qiang Zhu, Songtao Guo, Paul Ogilvie, Yan Liu (<i>LinkedIn</i>)	
Author Index	2141

KDD 2016 Conference Organization

General Chairs: Balaji Krishnapuram (*IBM*)
Mohak Shah (*Bosch*)

Research Track Program Chairs: Alex Smola (*Amazon*)
Charu Aggarwal (*IBM*)

Applied Data Science Track Program Dou Shen (*Baidu*)
Chairs: Rajeev Rastogi (*Amazon*)

Associate General Chair Shipeng Yu (*LinkedIn*)

Workshop Chairs: Quoc Le (*Google*)
Zhi-Hua Zhou (*Nanjing University*)

Tutorial Chairs: Hanghang Tong (*Arizona State University*)
S V N Vishwanathan (*UCSC*)

KDD Cup Chairs: Markus Weimer (*Microsoft*)
Shou-De Lin (*National Taiwan University*)
Vijay Narayanan (*Microsoft*)

Panel Chair: Andrei Broder (*Google*)

Applied Data Science Invited Talks Chairs: Evangelos Simoudis (*Synapse Partners*)
Usama Fayyad (*Barclays Bank and Oasis500*)
Rajesh Parekh (*Facebook*)

Poster Chair: Jiliang Tang (*Yahoo Labs*)

Best Paper Chairs: Jiliang Tang (*Yahoo Labs*)

Doctoral Dissertation Award Chair: Haixun Wang (*Facebook*)

Innovation and Service Award Chair: Jian Pei (*Simon Fraser University*)

Test-of-Time Award Chair: Wei Wang (*UCLA*)

Local Arrangements Chairs: Joaquin Quiñonero Candela (*Facebook*)
Olivier Chapelle (*Google*)

Student Travel Award Chairs: Jure Leskovec (*Stanford University*)
Leman Akoglu (*Stony Brook University*)

Media and Publicity Chairs: Ankur Teredesai (*University of Washington*)
Gabor Melli (*Opengov*)

Proceedings Chairs: Rómer Rosales (*LinkedIn*)
Stefanie Jegelka (*MIT*)

Social Networking Chairs: Faisal Farooq (*IBM*)
Ron Bekkerman (*University of Haifa*)

Sponsorship Chair: Claudia Perlich (*Dstillery*)
Yu Zheng (*Microsoft Research*)

Treasurer: Ying Li (*EV Analysis*)

Video and Streaming Chairs: Jie Tang (*Tsinghua University*)
Marko Grobelnik (*Jozef Stefan Institute*)

Web Chairs: David Hazel (*University of Washington*)
Derek Young (*Seasonal View*)

Job Match Chair: Unmesh Kurup (*Bosch*)

Social Impact Chair: Rayid Ghani (*University of Chicago*)

Research Track Senior Program Committee: Alice Oh (*KAIST*)
Ambuj Singh (*UCSB*)
Amr Ahmed (*Google*)
Andrew Tomkins (*Google Research*)
Anil Vullikanti (*Virginia Tech*)
Aristides Gionis (*Aalto University*)
Arno Siebes
Bart Goethals (*University of Antwerp*)
Chih-Jen Lin (*National Taiwan University*)
Christos Faloutsos (*Carnegie Mellon University*)
Corinna Cortes (*Google*)
Dacheng Tao (*The University of Technology, Sydney*)
Deepak Agarwal (*LinkedIn*)
Diane Cook (*Washington State University*)
Dimitrios Gunopulos
Eamonn Keogh (*UC Riverside*)
Edo Liberty (*Yahoo Labs*)
Evgeniy Gabrilovich (*Google*)
Evimaria Terzi (*Boston University*)
Fei Wang (*University of Connecticut*)
Francesco Bonchi (*Yahoo Barcelona*)
George Karypis
Graham Cormode (*Warwick*)
Guojun Qi (*University of Central Florida*)
Hong Cheng (*Chinese University of Hong Kong*)
Huan Liu (*ASU*)
Hui Xiong (*Rutgers*)
Ian Davidson (*University of California-Davis*)
James Bailey

Research Track Senior Program Jeffrey Yu (*Chinese University of Hong Kong*)
Committee (continued): Jennifer Neville (*Purdue*)
Jennifer Dy (*Northeastern University*)
Jian Pei (*Simon Fraser University*)
Jiawei Han (*University of Illinois at Urbana-Champaign*)
Jie Tang (*Tsinghua University*)
Jieping Ye (*University of Michigan at Ann Arbor*)
Jimeng Sun (*Georgia Tech*)
Jing Gao
Johannes Fuernkranz (*TU Darmstadt*)
Johannes Gehrke (*Microsoft Research*)
Jure Leskovec (*Stanford University*)
Karsten Borgwardt (*ETH Zurich*)
Katharina Morik (*TU Dortmund*)
Kristian Kersting (*TU Dortmund*)
Martin Ester (*Simon Fraser University*)
Mohammed Zaki (*RPI*)
Naoki Abe (*IBM T J Watson Research Center*)
Osmar Zaiane (*University of Alberta*)
Paul Bradley (*MethodCare*)
Paul Bennett (*Microsoft Research*)
Peter Flach (*University of Bristol*)
Philip S. Yu (*UI Chicago*)
Qiang Yang (*HKUST*)
Ravi Kumar (*Google*)
Saharon Rosset (*Tel Aviv University*)
Sanjay Chawla
Siegfried Nijssen (*Universiteit Leiden*)
Stefan Wrobel (*Universitaet Bonn*)
Suchi Saria (*Johns Hopkins University*)
Tanya Berger-Wolf (*University of Illinois*)
Thomas Seidl (*LMU Munich*)
Thomas Gaertner (*University of Nottingham*)
Thorsten Joachims (*Cornell University*)
Tie-Yan Liu (*Microsoft Research*)
Tijl De Bie (*Ghent University*)
Vanja Josifovski (*Pinterest*)
Wei Wang (*UC Los Angeles*)
Wray Buntine (*Monash University*)
Xavier Amatriain (*Quora*)
Yan Liu
Yehuda Koren (*Google*)
Yisong Yue (*Caltech*)
Yizhou Sun (*Northeastern University*)
Zhi-Hua Zhou

Applied Data Science Track Senior Anitha Kannan (*Facebook*)
Program Committee: Balaraman Ravindran (*Indian Institute of Technology Madras*)
Brian D’Alessandro (*Facebook*)
Chao Liu (*Jindi Data Technology Ltd Co.*)
Chiranjib Bhattacharya (*Indian Institute of Science Bangalore*)
Divesh Srivastava (*AT&T Research*)
Gautam Das (*University of Texas Arlington*)
Gautam Shroff (*Tata Consultancy Services Research*)
Graham Williams (*Australian Taxation Office*)
Haixun Wang (*Google*)
Hang Li (*Huawei Technologies*)
Hema Raghavan (*LinkedIn*)
Hui Xiong (*Rutgers University*)
Jan Pedersen (*Microsoft*)
Jeffrey Pan (*Facebook*)
Jun Yan (*Microsoft Research Asia*)
Minos Garofalakis (*University of Chania*)
Ping Luo (*Chinese Academy of Sciences*)
Ralf Herbrich (*Amazon*)
Rayid Ghani (*University of Chicago*)
Roberto Bayardo (*Google*)
S. Muthukrishnan (*Rutgers Univ*)
Sinno Pan (*Nanyang Technological University*)
Tina Eliassi-Rad (*Rutgers Univ*)
Vincent Tseng (*National Chiao Tung University*)
Vipin Kumar (*University of Minnesota*)
Wayne Thompson (*SAS*)
Wei Fan (*Baidu*)
Xiaoming Jin (*Tsinghua University*)
Ying Li (*EV Analysis Corp*)
Yoelle Maarek (*Yahoo! Labs*)
Yu Zheng (*Microsoft Research*)

KDD 2016 Sponsors & Supporters

Sponsor: ACM SIGKDD



Supporters: Diamond



Supporters: Platinum



Supporters: Gold



BOSCH

Invented for life



facebook



IBM **Research**



Orchestrating a brighter world

NEC

Supporters: Silver

ebay

Tencent 腾讯

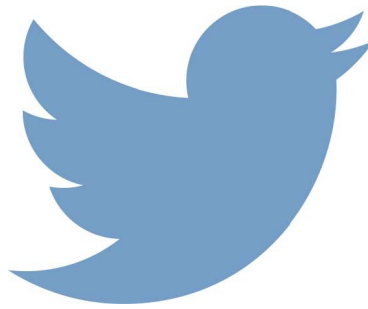
SIEMENS

Ingenuity for life

UBER

sas

criteo.
labs



Supporters: Bronze



NETFLIX

Google™
Booking.com



Supporters: Lanyard Sponsor



Supporters: KDD CUP Sponsor



Supporters: Student Travel Awards



Supporters: Program Ad Sponsor



Supporters: Media Sponsor

BIG DATA

Mary Ann Liebert, Inc.  publishers

Exhibitors:



INDIANA UNIVERSITY
DATA SCIENCE PROGRAM



MORGAN & CLAYPOOL
PUBLISHERS



ELSEVIER





spare 5



Georgia Institute of Technology

