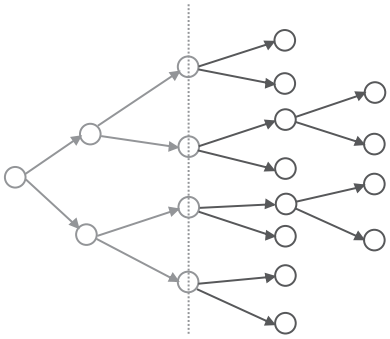
Final Program and Abstracts

2013 SIAM International Conference on DATA MINING

May 2–4, 2013



Sheraton Austin Hotel at the Capitol
Austin, Texas, USA

Sponsored by the SIAM Activity Group on Data Mining and Analytics (SIAG/DMA)

The purpose of the SIAM Activity Group on Data Mining and Analytics (SIAG/DMA) is to advance the mathematics of data mining, to highlight the importance and benefits of the application of data mining, and to identify and explore the connections between data mining and other applied sciences.

This conference is held in cooperation with the American Statistical Association.

2013 is designated as the year of Math of Planet Earth.
SIAM supports MPE2013.
www.mpe2013.org

Society for Industrial and Applied Mathematics
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA
Telephone: +1-215-382-9800 Fax: +1-215-386-7999
Conference E-mail: meetings@siam.org
Conference Web: www.siam.org/meetings/
Membership and Customer Service: (800) 447-7426 (US & Canada) or +1-215-382-9800 (worldwide)

www.siam.org/meetings/sdm13

Table of Contents

Program-at-a-Glance.....Fold out section
General Information.....2
Get-togethers.....8
Invited Plenary Presentations 10
Tutorials.....11
Workshops 13
Poster Session 24
Program Schedule.....19
Abstracts 32
Conference and Tutorial Speaker Index 53
Conference Budget .....Inside Back cover
Hotel Meeting Room Map.....Back cover

Organizing Committee

Steering Committee Co-Chairs

Chandrika Kamath
Lawrence Livermore National Laboratory, USA

Srinivasan Parthasarathy
The Ohio State University, USA

Conference Chairs

Joydeep Ghosh
University of Texas - Austin, USA

Zoran Obradovic
Temple University, USA

Program Chairs

Jennifer Dy
Northeastern University, USA

Zhi-Hua Zhou
Nanjing University, China

Workshop Chairs

Balaji Krishnapuram
Siemens, USA

Hui Xiong
Rutgers University, USA

Tutorial Chair

Jieping Ye
Arizona State University, USA

Publicity Chair

Yan Liu
University of Southern California, USA

Sponsorship Chairs

Ying Li
Concurix Corporation, USA

Chandan Reddy
Wayne State University, USA

Local Chair and Doctoral Forum Organizer

Latifur Khan
University of Texas, Dallas

Senior Program Committee Members

Arindam Banerjee
University of Minnesota, USA

Roberto Bayardo
Google Research, USA

Miguel Carreira-Perpinan
University of California, Merced, USA

Nitesh Chawla
University of Notre Dame, USA

Ian Davidson
University of California, Davis, USA

Tina Eliassi-Rad
Rutgers University, USA

Wei Fan
IBM T. J. Watson Research Center and IBM CRL, USA

Xiaoli Fern
Oregon State University, USA

Jiawei Han
University of Illinois at Urbana-Champaign, USA

Xiaofei He
Zhejiang University, China

George Karypi
University of Minnesota, USA

Tamara Kolda
Sandia National Laboratories, USA

Charles Ling
University of Western Ontario, Canada

Sofus A. Macskassy
University of Southern California, USA

Jennifer Neville
Purdue University, USA

Zaiqing Nie
Microsoft Research Asia, China

Dino Pedreschi
University of Pisa, Italy

Jian Pei
Simon Fraser University, Canada

Romer Rosales
LinkedIn, USA

Jimeng Sun
IBM T.J. Watson Research Center, USA

Pang-Ning Tan
Michigan State University, USA

Jianyong Wang
Tsinghua University, China

Wei Wang
University of California, Los Angeles, USA

Takashi Washio
Osaka University, Japan

Jeffrey Xu Yu
Chinese University of Hong Kong, Hong Kong

Philip Yu
University of Illinois at Chicago, USA

Min-Ling Zhang
Southeast University, China

Xiaojin (Jerry) Zhu
University of Wisconsin-Madison, USA

Program Committee Members

Naoki Abe
IBM T. J. Watson Research Center, USA

Evrim Acar
University of Copenhagen, Denmark

Leman Akoglu
SUNY Stony Brook, USA

David Andrzejewski
Sumo Logic, USA

Javad Azimi
Insightsone, USA

James Bailey
University of Melbourne, Australia

Natasha Balac
San Diego Supercomputer Center, USA

Justin Basilico
Netflix, USA

Paul Bennett
Microsoft Research, USA

Jonathan Berry
Sandia National Laboratories, USA

Liefeng Bo
University of Washington, USA

Daniel Boley
University of Minnesota, USA

Ulf Brefeld
Technical University of Darmstadt, Germany

Gavin Brown
University of Manchester, United Kingdom

Longbing Cao
University of Technology Sydney, Australia

Sharma Chakravarthy
The University of Texas at Arlington, USA

Varun Chandola
Oak Ridge National Laboratory, USA

Sanjay Chawla
University of Sydney, Australia

Shu-Ching Chen
Florida International University, USA

Phoebe Chen
La Trobe University, Australia

Enhong Chen
University of Science and Technology of China, China

Jian Chen
South China University of Technology, China

Jinchuan Chen
Renmin University of China, China

Lei Chen
Hong Kong University of Science and Technology, Hong Kong

Ming-Syan Chen
Academia Sinica, Taiwan

Songcan Chen
Nanjing University of Aeronautics and Astronautics, China

Yixin Chen
Washington University in St. Louis, USA

Hong Cheng
Chinese University of Hong Kong, Hong Kong

William K. Cheung
Hong Kong Baptist University, Hong Kong

Gao Cong
Nanyang Technological University, Singapore

Diane Cook
Washington State University, USA

Alfredo Cuzzocrea
Italian National Research Council, Italy

Brian Dalessandro
New York University, USA

Christian Desrosiers
Ecole de Technologie Superieure, Canada

Chris Ding
University of Texas at Arlington, USA

Wei Ding
University of Massachusetts Boston, USA

Carlotta Domeniconi
George Mason University, USA

Zhicheng Dou
Microsoft Research Asia, China

Murat Dundar
Indiana University-Purdue University Indianapolis, USA

Danny Dunlavy
Sandia National Laboratories in Albuquerque, USA

Saso Dzeroski
Jozef Stefan Institute, Slovenia

Hoda Eldardiry
Palo Alto Research Center, USA

Yun Fu
Northeastern University, USA

Xiaoming Fu
Georg-August-University of Goettingen, Germany

Glenn Fung
Siemens, USA

Amita Gajewar
Microsoft, USA

Brian Gallagher
Lawrence Livermore National Laboratory, USA

Jing Gao
University of Buffalo, USA

Yang Gao
Nanjing University, China

Rayid Ghani
Obama for America, USA

Amol Ghoting
IBM TJ Watson, USA

Fosca Giannotti
ISTI-CNR Pisa, Italy

Aristides Gionis
Aalto University, Finland

Aris Gkoulalas-Divanis
IBM Research, Ireland

David Gleich
Purdue University, USA

Bart Goethals
University of Antwerp, Belgium

Yuhong Guo
Temple University, USA

Lawrence Hall
University of South Florida, USA

Yi Han
National University of Defense Technology, China

Haibo He
University of Rhode Island, USA

Jingrui He
IBM Research, USA

Tu Bao Ho
Japan Adavanced Institute of Science and Technology, Japan

Steven Hoi
Nanyang Technological University, Singapore

Junling Hu
PayPal, USA

Xian-Sheng Hua
Microsoft, USA

Ming Hua
Facebook, USA

Jun (Luke) Huan
University of Kansas, USA

Joshua Huang
Shenzhen Institute of Advanced Technology, CAS, China

Akihiro Inokuchi
Osaka University, Japan

Nathalie Japkowicz
University of Ottawa, Canada

Christopher Jermaine
Rice University, USA

Shuiwang Ji
Old Dominion University, USA

Jing Jiang
Singapore Management University, Singapore

Ruoming Jin
Kent State University, USA

Vladimir Jojic
University of North Carolina, Chapel Hill, USA

Yoshinobu Kawahara
Osaka University, Japan

Yiping Ke
A*Star, Singapore

Philip Kegelmeyer
Sandia National Laboratories, USA

Eamonn Keogh
University of California, USA

Kristian Kersting
University of Bonn, Germany

Irwin King
The Chinese University of Hong Kong, Hong Kong

James Kwok
The Hong Kong University of Science and Technology, Hong Kong

Dongwon Lee
Pennsylvania State University, USA

Zhenhui Li
Pennsylvania State University, USA

Yuefeng Li
Queensland University of Technology, Australia

Xiaoli Li
Institute for Infocomm Research, Singapore

Qing Li
City University of Hong Kong, Hong Kong

Jiuyong Li
University of South Australia, Australia

Xue Li
The University of Queensland, Australia

Lei Li
University of California, Berkeley, USA

Chun-Hung Li
Hong Kong Baptist University, Hong Kong

Tao Li
Florida International University, USA

Ming Li
Nanjing University, China

Ying Li
Concurix Corporation, USA

Ee Peng Lim
Singapore Management University, Singapore

Lek-Heng Lim
University of Chicago, USA

Shou-de Lin
National Taiwan University, Taiwan

Chih-Jen Lin
National Taiwan University, Taiwan

Fei Tony Liu
Australian Taxation Office, Australia

Yan Liu
University of Southern California, USA

Tie-Yan Liu
Microsoft Research Asia, China

Jun Liu
Siemens Corporate Research, USA

Huan Liu
Arizona State University, USA

Xu-Ying Liu
Southeast University, China

David Lo
Singapore Management University, Singapore

Bo Long
LinkedIn Inc., USA

Chang-Tien Lu
Virginia Tech, USA

Ping Luo
HP Labs, China

Nikos Mamoulis
University of Hong Kong, Hong Kong

Gideon Mann
Google, USA

Fabian Moerchen
Amazon.com, USA

Abdullah Mueen
Microsoft, USA

Emmanuel Müller
Karlsruhe Institute of Technology, Germany

Art Munson
Sandia National Laboratories, USA

Hung Son Nguyen
Warsaw University, Poland

Xia Ning
NEC Labs America, USA

Nikunj Oza
NASA Ames Research Center, USA

Jia-Yu Pan
Google, USA

Weike Pan
Hong Kong Baptist University, Hong Kong

Spiros Papadimitriou
Rutgers University, USA

Ali Pinar
Sandia National Laboratories, USA

B. Aditya Prakash
Virginia Tech, USA

Yuan Qi
Purdue University, USA

Tao Qin
Microsoft, China

Huzefa Rangwala
George Mason University, USA

Chandan Reddy
Wayne State University, USA

Steffen Rendle
University of Konstanz, Germany

Jorg Sander
University of Alberta, Canada

Berkant Savas
Linkoping University, Sweden

Hanhuai Shan
Microsoft, USA

Dou Shen
CityGrid Media, USA

Victor S. Sheng
University of Central Arkansas, USA

Yong Shi
Kennesaw State University, USA

Vikas Sindhwani
IBM T.J.Watson Research Center, USA

Zhao Song
Microsoft, USA

Jaideep Srivastava
University of Minnesota, USA

Karsten Steinhaeuser
University of Minnesota, USA

Masashi Sugiyama
Tokyo Institute of Technology, Japan

Liang Sun
Opera Solutions, USA

Yizhou Sun
Northeastern University, USA

Jian-Tao Sun
Microsoft, China

Einoshin Suzuki
Kyushu University, Japan

Andrea Tagarelli
University of Calabria, Italy

Ichigaku Takigawa
Hokkaido University, Japan

Ah Hwee Tan
Nanyang Technological University, Singapore

Jie Tang
Tsinghua University, China

Dacheng Tao
University of Technology Sydney, Australia

Alexandre Termier
Universite Joseph Fourier, France

Evimaria Terzi
Boston University, USA

Hanghang Tong
IBM Research, USA

Vincent S. Tseng
National Cheng Kung University, Taiwan

Grigorios Tsoumakas
Aristotle University of Thessaloniki, Greece

Koji Tsuda
National Institute of Advanced Industrial Science and Technology, Japan

Shusaku Tsumoto
Shimane University, Japan

Hamed Valizadegan
University of Pittsburgh, USA

Ranga Raju Vatsavai
Oak Ridge National Laboratory, USA

Suresh Venkatasubramanian
University of Utah, USA

Haifeng Wang
Baidu Inc., China

Haixun Wang
Microsoft Resarch Asia, China

Fei Wang
IBM TJ Watson, USA

Ke Wang
Simon Fraser University, Canada

Weng-Keen Wong
Oregon State University, USA

Raymond C Wong
Hong Kong University of Science and Technology, Hong Kong

Junjie Wu
Beihang University, China

Xintao Wu
University of North Carolina at Charlotte, USA

Rongjing Xiang
Foursquare, USA

Zhengzheng Xing
Amazon, USA

Hui Xiong
Rutgers University, USA

Guirong Xue
Shanghai Jiao Tong University, China

Hui Xue
Southeast University, China

Rong Yan
Facebook, USA

Christopher Yang
Drexel University, USA

Mi-Yen Yeh
Academia Sinica, Taiwan

Yang Yu
Nanjing University, China

Shipeng Yu
Siemens Medical Solutions, USA

De-Chuan Zhan
Nanjing University, China

Yu Zhang
Hong Kong Baptist University, Hong Kong

Daoqiang Zhang
Nanjing University of Aeronautics and Astronautics, China

Kun Zhang
MPI for Intelligent Systems, Germany

Zhongfei (Mark) Zhang
SUNY Binghamton, USA

Chengqi Zhang
The University of Technology, Sydney, Australia

Dell Zhang
Birkbeck, University of London, United Kingdom

Wenjie Zhang
University of New South Wales, Australia

Ying Zhao
Tsinghua University, China

Yu Zheng
Microsoft Research Asia, China

Zijian Zheng
Microsoft, USA

Bin Zhou
University of Maryland Baltimore County, USA

De-Yu Zhou
Southeast University, China

Xueyuan Zhou
Quantitative Risk Management, USA

Feida Zhu
Singapore Management University, Singapore

Jun Zhu
Tsinghua University, China

Xingquan Zhu
Florida Atlantic University, USA

SIAM Registration Desk

The SIAM registration desk is located in the Ballroom Foyer – 3rd Floor. It is open during the following times:

Wednesday, May 1
5:00 PM – 7:00 PM

Thursday, May 2
7:00 AM – 7:30 PM

Friday, May 3
7:30 AM – 3:30 PM

Saturday, May 4
7:30 AM – 4:00 PM

Hotel Address

Sheraton Austin Hotel at the Capitol

701 East 11th Street

Austin, Texas 78701

Direct Telephone: +1-512-478-1111

Toll Free (Central Reservations): +1-800-325-3535

Hotel Fax: +1-512-478-3700

Hotel Website:
http://www.sheratonaustin.com

Hotel Telephone Number

To reach an attendee or to leave a message, call +1-512-478-1111. The hotel operator can either connect you with the SIAM registration desk or to the attendee's room. Messages taken at the SIAM registration desk will be posted to the message board located in the registration area.

Hotel Check-in and Check-out Times

Check-in time is 3:00 PM and check-out time is 12:00 PM.

Child Care

The Sheraton Austin Hotel at the Capitol does not offer childcare. If you contact the concierge directly, they will be happy to arrange for local child care services to come to the hotel for you.

Corporate Members and Affiliates

SIAM corporate members provide their employees with knowledge about, access to, and contacts in the applied mathematics and computational sciences community through their membership benefits. Corporate membership is more than just a bundle of tangible products and services; it is an expression of support for SIAM and its programs. SIAM is pleased to acknowledge its corporate members and sponsors. In recognition of their support, non-member attendees who are employed by the following organizations are entitled to the SIAM member registration rate.

Corporate Institutional Members

The Aerospace Corporation

Air Force Office of Scientific Research

AT&T Laboratories - Research

Bechtel Marine Propulsion Laboratory

The Boeing Company

CEA/DAM

Department of National Defence (DND/CSEC)

DSTO- Defence Science and Technology Organisation

Hewlett-Packard

IBM Corporation

IDA Center for Communications Research, La Jolla

IDA Center for Communications Research, Princeton

Institute for Computational and Experimental Research in Mathematics (ICERM)

Institute for Defense Analyses, Center for Computing Sciences

Lawrence Berkeley National Laboratory

Lockheed Martin

Los Alamos National Laboratory

Mathematical Sciences Research Institute

Max-Planck-Institute for Dynamics of Complex Technical Systems

Mentor Graphics

National Institute of Standards and Technology (NIST)

National Security Agency (DIRNSA)

Oak Ridge National Laboratory, managed by UT-Battelle for the Department of Energy

Sandia National Laboratories

Schlumberger-Doll Research

Tech X Corporation

U.S. Army Corps of Engineers, Engineer Research and Development Center

United States Department of Energy

**List current March 2013*

Funding Agency

SIAM and the conference organizing committee wish to extend their thanks and appreciation to the National Science Foundation for its support of this conference.

Leading the applied mathematics community . . .

Join SIAM and save!

SIAM members save up to $130 on full SIAM members save up to $130 on full registration for the 2013 SIAM International Conference on Data Mining! Join your peers in supporting the premier professional society for applied mathematicians and computational scientists. SIAM members receive subscriptions to SIAM Review and SIAM News and enjoy substantial discounts on SIAM books, journal subscriptions, and conference registrations.

If you are not a SIAM member and paid the Non-Member rate to attend the conference, you can apply the difference between what you paid and what a member would have paid ($130 for a Non-Member) towards a SIAM membership. Contact SIAM Customer Service for details or join at the conference registration desk.

If you are a SIAM member, it only costs $10 to join the SIAM Activity Group on Data Mining and Analytics (SIAG/DMA). As a SIAG/DMA member, you are eligible for an additional $10 discount on this conference, so if you paid the SIAM member rate to attend the conference, you might be eligible for a free SIAG/DMA membership. Check at the registration desk.

Free Student Memberships are available to students who attend an institution that is an Academic Member of SIAM, are members of Student Chapters of SIAM, or are nominated by a Regular Member of SIAM.

Join onsite at the registration desk, go to *www.siam.org/joinsiam* to join online or download an application form, or contact SIAM Customer Service

Telephone: +1-215-382-9800 (worldwide); or 800-447-7426 (U.S. and Canada only)

Fax: +1-215-386-7999

E-mail: membership@siam.org

Postal mail: Society for Industrial and Applied Mathematics, 3600 Market Street, 6th floor, Philadelphia, PA 19104-2688 USA

Standard Audio/Visual Set-Up in Meeting Rooms

SIAM does not provide computers for any speaker. When giving an electronic presentation, speakers must provide their own computers. SIAM is not responsible for the safety and security of speakers' computers.

The Plenary Session Room will have one (1) overhead projector, two (2) screens and one (1) data projector. Cables or adaptors for Apple computers are not supplied, as they vary for each model. Please bring your own cable/adaptor if using a Mac computer.

All other concurrent/breakout rooms will have one (1) screen and one (1) data projector. Cables or adaptors for Apple computers are not supplied, as they vary for each model. Please bring your own cable/adaptor if using a Mac computer. Overhead projectors will be provided only when requested.

If you have questions regarding availability of equipment in the meeting room of your presentation, or to request an overhead projector for your session, please see a SIAM staff member at the registration desk.

E-mail Access

SIAM attendees booked within the SIAM room block will have complimentary wired or wireless Internet access in their guest rooms. SIAM will also provide a limited number of email stations. Wireless email access is available in the conference meeting space to all participants.

Registration Fee Includes

- Admission to all technical sessions
- Admission to tutorial sessions
- Admission to a workshop
- CD of conference proceedings, workshop and tutorial notes
- Coffee breaks daily
- Continental Breakfast daily
- Room set-ups and audio/visual equipment
- Welcome Reception and Poster Session

Job Postings

Please check with the SIAM registration desk regarding the availability of job postings or visit *http://jobs.siam.org.*

Important Notice to Poster Presenters

The poster session is scheduled for Thursday, May 2 at 6:30 PM. Poster presenters are requested to set up their poster material on the provided 4' x 8' poster boards in the Capitol View Terrace beginning Thursday, May 2 at 7:00 AM. All materials must be posted by Thursday, May 2 by 6:30 PM, the official start time of the session. Poster displays must be removed by Friday, May 3 by 10:00 AM. Posters remaining after this time will be discarded. SIAM is not responsible for discarded posters.

SIAM Books and Journals

Display copies of books and complimentary copies of journals are available on site. SIAM books are available at a discounted price during the conference. If a SIAM books representative is not available, completed order forms and payment (credit cards are preferred) may be taken to the SIAM registration desk. The books table will close at 3:30 PM on Saturday, May 4.

Table Top Displays

Oxford University Press

SIAM

Springer

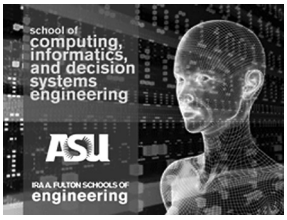
Taylor & Francis

Conference Sponsors

Google™

IBM Research

IBM Research is pleased to sponsor the 2nd Workshop on Data Mining for Medicine and Healthcare and is a Student Travel Sponsor.



Data Mining and Knowledge Discovery

Data Mining and Knowledge Discovery is pleased to sponsor the SIAM SDM Best Student Paper Award. This premier technical journal in data mining is published by Springer Science & Business Media, a leading publisher of scholarly books and journals.

Name Badges

A space for emergency contact information is provided on the back of your name badge. Help us help you in the event of an emergency!

Comments?

Comments about SIAM meetings are encouraged! Please send to:

Sven Leyffer, SIAM Vice President for Programs (*vpp@siam.org*)

Get-togethers

- Welcome Reception
  and Poster Session,
  Thursday, May 2
  6:30 PM – 9:00 PM
- Business Meeting
  (open to SIAG/DMA members)
  Complimentary wine and beer will be served.
  Friday, May 3
  6:30 PM – 7:00 PM

Please Note

SIAM is not responsible for the safety and security of attendees' computers. Do not leave your laptop computers unattended. Please remember to turn off your cell phones, pagers, etc. during sessions.

Recording of Presentations

Audio and video recording of presentations at SIAM meetings is prohibited without the written permission of the presenter and SIAM.

Social Media

SIAM is promoting the use of social media, such as Facebook and Twitter, in order to enhance scientific discussion at its meetings and enable attendees to connect with each other prior to, during and after conferences. If you are tweeting about a conference, please use the designated hashtag to enable other attendees to keep up with the Twitter conversation and to allow better archiving of our conference discussions. The hashtag for this meeting is #SIAMSDM13

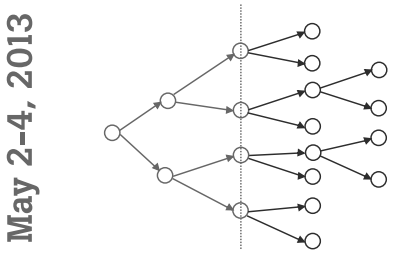
SIAM Activity Group on Data Mining (SIAG/DMA)

www.siam.org/activity/dma

A GREAT WAY TO GET INVOLVED!

Collaborate and interact with mathematicians and applied scientists whose work involves data mining.

2013 SIAM International Conference on DATA MINING



May 2–4, 2013

Sheraton Austin Hotel at the Capitol
Austin, Texas, USA

ACTIVITES INCLUDE:

- Annual conference and proceedings
- Special sessions at SIAM Annual Meetings
- Website

BENEFITS OF SIAG/DMA MEMBERSHIP:

- Listing in the SIAG's online membership directory
- Additional $10 discount on registration at the SIAM International Conference on Data Mining (excludes student)
- Electronic communications about recent developments in your specialty
- Eligibility for candidacy for SIAG/DMA office
- Participation in the selection of SIAG/DMA officers

ELIGIBILITY:

- Be a current SIAM member

COST:

- $10 per year
- Student SIAM members can join 2 activity groups for free!

2012-2013 SIAG/DMA OFFICERS:

- Chair: Chandrika Kamath, Lawrence Livermore National Laboratory
- Vice Chair: Charu Aggarwal, IBM T. J. Watson Research Center
- Program Director: Huan Liu, Arizona State University
- Secretary: Michael Mahoney, Stanford University

TO JOIN:

SIAG/DMA: ***my.siam.org/forms/join_siag.htm***

SIAM: ***www.siam.org/joinsiam***

Invited Plenary Speakers

* *All Invited Plenary Presentations will take place in Capitol Ballroom DE – 3rd Floor**

Thursday, May 2

8:15 AM – 9:30 AM

IP1 Regional Climate Informatics: A Statistical Perspective

Doug Nychka, *National Center for Atmospheric Research, USA*

1:30 PM – 2:45 PM

IP2 Mining Clinical Data to Build Predictive Models

Peter Szolovits, *Massachusetts Institute of Technology, USA*

Friday, May 3

8:15 AM – 9:30 AM

IP3 Modeling Individual-Level Data in the 21st Century

Padhraic Smyth, *University of California, Irvine, USA*

1:30 PM – 2:45 PM

IP4 Social Networks as Information Filters

Lada Adamic, *University of Michigan, Ann Arbor, USA*

Tutorials

Thursday, May 2

10:00 AM – 12:05 PM

TS1: Tutorial Session: Sampling and Summarization for Social Networks

Shou-De Lin, *National Taiwan University, Taiwan*

Mi-Yen Yeh, *Institute of Information Science, Academia Sinica, Taiwan*

Cheng-Te Li, *National Taiwan University, Taiwa*n

Capitol Ballroom ABC - 3rd Floor

3:00 PM – 5:05 PM

TS2: Tutorial Session: Big Data Analytics for Healthcare

Jimeng Sun, *IBM T. J. Watson Research Center, USA*

Chandan Reddy, *Wayne State University, USA*

Capitol Ballroom ABC - 3rd Floor

Friday, May 3

10:00 AM – 12:05 PM

TS3: Tutorial Session: Crowdsourcing & Human Computation:
Data Labeling & Building Hybrid Systems

Matthew Lease, *University of Texas at Austin, USA*

Capitol Ballroom ABC - 3rd Floor

3:00 PM – 5:05 PM

TS4: Tutorial Session:
Recent Advances in Applied Matrix Technologies

Fei Wang, *IBM T. J. Watson Research Center, USA*

Hanghang Tong, *IBM T. J. Watson Research Center, USA*

Capitol Ballroom ABC - 3rd Floor

Tutorials

Saturday, May 4

8:30 AM – 12:00 PM

TS5: Tutorial Session: Outlier Detection for Temporal Data

Manish Gupta, *University of Illinois at Urbana-Champaign, USA*

Jing Gao, *State University of New York, Buffalo, USA*

Charu Aggarwal, *IBM T. J. Watson Research Center, USA*

Jiawei Han, *University of Illinois at Urbana-Champaign, USA*

Capitol Ballroom ABC - 3rd Floor

1:30 PM – 5:00 PM

TS6: Tutorial Session: Online Learning for Big Data Mining: Methods and Applications

Steven C.H. Hoi, *Nanyang Technological University, Singapore*

Capitol Ballroom ABC - 3rd Floor

Full Day Workshops

Saturday, May 4

Data Mining for Medicine and Healthcare Workshop
8:30 AM – 5:00 PM
Co-Chairs:
David Gotz, *IBM T.J. Watson Research Center, USA*
Nigam Shah, *Stanford University, USA*
Gregor Stiglic, *University of Maribor, Slovenia*
Fei Wang, *IBM T.J. Watson Research Center, USA*
Capitol Ballroom D - 3rd Floor

See page 15 for schedule

Data Mining for Service and Maintenance Workshop
9:00 AM – 5:30 PM
Co-Chairs:
Nikunj Oza, *NASA AMES Research Center, USA*
Dragos Margineantu, *Boeing Research & Technology, USA*
Shipeng Yu, *Siemens Healthcare, USA*
Matan Ninio, *IBM Research Haifa, Israel*
Capitol Ballroom E– 3rd Floor

Half Day Workshops

Saturday, May 4

Analytics for Cyber-Physical Systems Workshop
8:05 AM – 12:30 PM
Co-Chairs:
Chetan Gupta, *Hewlett Packard Laboratories, USA*
Varun Chandola, *Oak Ridge National Laboratory, USA*
Ranga Raju Vatsavai, *Oak Ridge National Laboratory, USA*
Capitol Ballroom FGH - 3rd Floor

See page 17 for schedule

Reliability Aware Data Fusion in Participatory Networks Workshop
1:30 PM – 5:00 PM
Co-Chairs:
Sameep Mehta, *IBM Research India, India*
L. Venkata Subramaniam, *IBM Research India, India*
Capitol Ballroom FGH - 3rd Floor

See page 18 for schedule

Data Mining for Medicine and Healthcare

Full Day Workshop Schedule
Saturday, May 4
Capitol Ballroom D - 3rd Floor

8:30 AM – 8:35 AM

Workshop Opening

8:35 AM – 9:25 AM

Invited Talk I
Joydeep Ghosh, University of Texas - Austin, USA

9:25 AM – 10:20 AM

Classification of Clinical Data using Sequential Patterns: A case study in Amyotrophic Lateral Sclerosis
André V. Carreiro, Susana Pinto, Mamede de Carvalho, Sara C. Madeira and Cláudia Antunes

Incremental SampleBoost for Efficient Learning from Multi-Class Data Sets
Mohamed Abouelenien and Xiaohui Yuan

Discover Temporal Dynamics of Biomarkers in Predictive Modeling with Longitudinal Data
Jiayu Zhou, Jimeng Sun, Fei Wang, Jianying Hu, Shahram Ebadollahi and Jieping Ye

10:20 AM – 10:30 AM

Coffee Break

10:30 AM – 11:20 AM

Invited Talk II
Marc Suchard, University of California, Los Angeles, USA

11:20 AM – 12:00 PM

A Temporal Association Mining Method for Signaling Drug-Drug Interactions
Yanqing Ji, Hao Ying, John Tran, Peter Dews, Ayman Mansour, See Yan and R Michael Massanari

Temporal Mining of Integrated Healthcare Data: Methods, Revealings and Implications
Rui Henriques, Sílvia Moura Pina and Cláudia Antunes

12:00 PM – 1:30 PM

Lunch Break (attendees on your own)

1:30 PM – 2:20 PM

Invited Talk III
Speaker: Jimeng Sun, IBM T.J. Watson Research Center, USA

2:20 PM – 3:00 PM

Data Mining for Integrated Sensing of Multiple Wearable Body Sensors: A Healthcare Application
Gaurav Pradhan and B Prabhakaran

Multivariate Methods for Classifying Physiological Data
Patricia Ordóñez, Tom Armstrong, Tim Oates, Jim Fackler and Christoph U. Lehmann

3:00 PM – 3:15 PM

Coffee Break

3:15 PM – 4:05 PM

Invited Talk IV
Speaker: Nitesh Chawla, University of Notre Dame, USA

4:05 PM – 4:55 PM

Real-Time Digital Flu Surveillance using Twitter Data
Kathy Lee, Ankit Agrawal and Alok Choudhary

Classification and Diagnosis of Myopathy from EMG Signals
Brian Bue, Erzsébet Merényi and James Killian

Mining Hospital Admission-discharge Data to Discover the Chance of Readmission
Arian Hosseinzadeh, Aman Verma, Masoumeh Izadi, Doina Precup and David Buckeridge

4:55 PM – 5:00 PM

Best Paper and Travel Awards
Sponsored by IBM

Analytics for Cyber Physical Systems

Half Day Workshop Schedule

Saturday, May 4

Capitol Ballroom FGH - 3rd Floor

8:05 AM – 8:10 AM

Opening Remarks
Dr. Chetan Gupta, Hewlett Packard Laboratories, USA

8:10 AM – 9:00 AM

Keynote I: Fundamental Challenges in Modeling and Design Space Exploration for CPS
Professor John Baras, University of Maryland, USA

9:00 AM – 9:20 AM

Knowledge Discovery from Nuclear Reactor Simulation Data
Neeti Pokhriyal, Oak Ridge National Laboratory, USA

9:20 AM – 9:30 AM

Coffee Break

9:30 AM – 10:15 AM

Invited Talk II: NIST on CPS
Dr. Vijay Srinivasan, NIST, USA

10:15 AM – 10:35 AM

Traffic Analytics using Probabilistic Graphical Models Enhanced with Knowledge Bases
Pramod Anantharam, Wright State University, USA

10:35 AM – 10:45 AM

Break

10:45 AM – 11:30 AM

Keynote II: Title to be Announced
Alok Batra, CTO & Chief Architect, Software COE GE Global Research, USA

11:30 AM – 11:50 AM

Detection of Classes of Heart Arrhythmias Based on Heatbeat Morphology Patterns
Tony Basil, Indian Institute of Technology, Hyderabad, India

11:50 AM – 12:25 PM

Panel: Analytics Challenges in CPS
Panelists: Professor John Baras, University of Maryland, USA;
Dr. Vijay Srinivasan, NIST, USA;
Dr. Fei Wang, IBM T. J. Watson Research Center, USA;
Lin Nease, Hewlett Packard Laboratories, USA;
Hosted by Dr. R. Vatsavai, Oak Ridge National Laboratory, USA

12:25 PM – 12:30 PM

Closing Remarks
Dr. Varun Chandola, Oak Ridge National Laboratory, USA

Reliability Aware Data Fusion in Participatory Networks

Half day Workshop Schedule

Saturday, May 4

Capitol Ballroom FGH - 3rd Floor

Research Talks

Information Virtualization in the Real World

Dan Wolfson, IBM Software Group, USA

Quality-aware Large-scale Analysis of Uncertain Data

Pontus Svenson, Defence Research Agency, Sweden

QODI: An Automatic Mediator for SPARQL End-points

Daniel Miranker, University of Texas at Austin, USA

Data Stream Mining and its Applications

Latifur Khan, University of Texas at Dallas, USA

Correlation Between Turkish Stock Market and Economy News

Sadi Evren Seker, Istanbul University, Turkey

Information Fusion Based Learning for Frugal Traffic State Sensing

Raghuram Krishnapuram, IBM Research India, India

Information Quality in Information Fusion

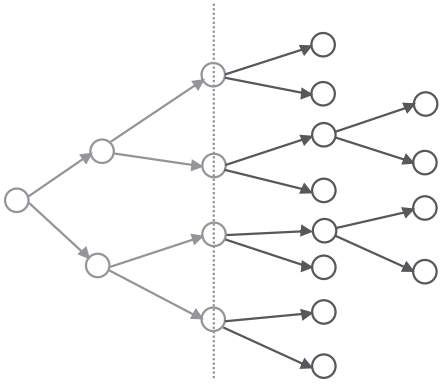
Galina Rogova, State University of New York at Buffalo, USA

Notes

SDM13 Program

2013 SIAM International Conference on DATA MINING

May 2–4, 2013



Sheraton Austin Hotel at the Capitol
Austin, Texas, USA

Wednesday, May 1

Registration

5:00 PM-7:00 PM

Room:Ballroom Foyer - 3rd Floor

Thursday, May 2

Registration

7:00 AM-7:30 PM

Room:Ballroom Foyer - 3rd Floor

Continental Breakfast

7:30 AM

Room:Ballroom Foyer - 3rd Floor

Announcements

8:00 AM-8:15 AM

Room:Capitol Ballroom DE - 3rd Floor

Thursday, May 2

IP1

Regional Climate Informatics: A Statistical Perspective

8:15 AM-9:30 AM

Room:Capitol Ballroom DE - 3rd Floor

Chair: Zhi-Hua Zhou, Nanjing University, China

As attention shifts from broad global summaries of climate change to more specific regional impacts there is a need for the data sciences to quantify the uncertainty in regional predictions. This talk will provide an overview on regional climate experiments with an emphasis on the data science problems for interpreting these large and complex simulations. Here a flexible spatial model based on fixed rank Kriging is implemented to handle a large number of spatial locations (LatticeKrig) and also include nonstationary spatial dependence.

Doug Nychka
National Center for Atmospheric Research, USA

Coffee Break

9:30 AM-10:00 AM

Room:Ballroom Foyer - 3rd Floor

Thursday, May 2

CP1

Mining with Big Data

10:00 AM-12:05 PM

Room:Capitol Ballroom D - 3rd Floor

Chair: B. Aditya Praka, Virginia Tech, USA

10:00-10:20 Fast Exact Max-Kernel Search
Parikshit Ram, Ryan Curtin and Alexander Gray, Georgia Institute of Technology, USA

10:25-10:45 Triadic Measures on Graphs: The Power of Wedge Sampling
C. Seshadhri, Ali Pinar, and Tamara G. Kolda, Sandia National Laboratories, USA

10:50-11:10 Maximal Deviations of Incomplete U-Statistics with Applications to Empirical Risk Sampling
Stephan Clémençon and Sylvain Robbiano, Télécom ParisTech, France; Jessica Tressou, INRA, France

11:15-11:35 NetSpot: Spotting Significant Anomalous Regions on Dynamic Networks
Misael Mongiovi, University of Catania, Italy; *Petko Bogdanov* and Razvan Ranca, University of California, Santa Barbara, USA; Evangelos Papalexakis, and Christos Faloutsos, Carnegie Mellon University, USA; Ambuj Singh, University of California, Santa Barbara, USA

11:40-12:00 Mining Connection Pathways for Marked Nodes in Large Graphs
Leman Akoglu, Carnegie Mellon University, USA; Jilles Vreeken, Universiteit Antwerpen, Belgium; Hanghang Tong, IBM T.J. Watson Research Center, USA; Polo Chau, Georgia Institute of Technology, USA; Nikolaj Tatti, Katholieke Universiteit Leuven, Belgium; Christos Faloutsos, Carnegie Mellon University, USA

Thursday, May 2

CP2

Mining with Uncertain and Noisy Data

10:00 AM-12:05 PM

Room:Capitol Ballroom E - 3rd Floor

Chair: Philip Kegelmeyer, Sandia National Laboratories, USA

10:00-10:20 Missing Or Inapplicable: Treatment of Incomplete Continuous-Valued Features in Supervised Learning
Pang-Ning Tan, Prakash Mandayam Comar, and Lei Liu, Michigan State University, USA; Sabyasachi Saha and Antonio Nucci, Narus Inc., USA

10:25-10:45 Patient Risk Prediction Model via Top-*k* Stability Selection
Jiayu Zhou, Arizona State University, USA; Jimeng Sun, IBM T.J. Watson Research Center, USA; Yashu Liu, Arizona State University, USA; Jianying Hu, IBM T.J. Watson Research Center, USA; Jieping Ye, Arizona State University, USA

10:50-11:10 Collective Kernel Construction in Noisy Environment
Miao Zhang, Chris Ding, and Deguang Kong, University of Texas at Arlington, USA

11:15-11:35 Mining Probabilistic Representative Frequent Patterns From Uncertain Data
Chunyang Liu, Ling Chen, and Chengqi Zhang, University of Technology, Sydney, Australia

11:40-12:00 Discriminative Feature Selection for Uncertain Graph Classification
Xiangnan Kong and Philip Yu, University of Ilinois at Chicago, USA; Xue Wang and Ann Ragin, Northwestern University, USA

Thursday, May 2

CP3

Clustering

10:00 AM-12:05 PM

Room:Capitol Ballroom FGH - 3rd Floor

Chair: Chandan Reddy, Wayne State University, USA

10:00-10:20 Determining the Number of Clusters Via Iterative Consensus Clustering
Shaina L. Race, North Carolina State University, USA

10:25-10:45 Constrained Spectral Clustering Using L1 Regularization
Jaya Kawale and *Daniel L. Boley*, University of Minnesota, USA

10:50-11:10 Efficient Anytime Density-Based Clustering
Son T. Mai, Xiao He, Jing Feng, and Christian Boehm, University of Munich, Germany

11:15-11:35 Evolutionary Soft Co-Clustering
Shuiwang Ji and *Wenlu Zhang*, Old Dominion University, USA; Rui Zhang, City College of New York, USA

11:40-12:00 Sparse Subspace Clustering Via Group Sparse Coding
Budhaditya Saha, Deakin University, Australia; Duc Son Pham, Curtin University, Australia; Dinh Phung and Svetha Venkatesh, Deakin University, Australia

Thursday, May 2

TS1

Tutorial Session:

Sampling and Summarization for Social Networks

10:00 AM-12:05 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, we will introduce the state-of-the-art solutions for social network sampling and summarization, and highlight the research challenges and unsolved issues.

Shou-De Lin, National Taiwan University, Taiwan

Mi-Yen Yeh, Academia Sinica, Taiwan

Cheng-Te Li, National Taiwan University, Taiwan

Lunch Break

12:05 PM-1:30 PM

Attendees on their own

Thursday, May 2

IP2

Mining Clinical Data to Build Predictive Models

1:30 PM-2:45 PM

Room:Capitol Ballroom DE - 3rd Floor

Chair: Jennifer Dy, Northeastern University, USA

The IOM's envisioned "learning health care system," uses data from each patient to teach about prevention, diagnosis, prognosis, and treatment. Several trends will enable the realization of this goal: (1) universal electronic health records, (2) recording of clinically important details to support "meaningful use" for quality of care, (3) wearable sensors that provide real-time data on every individual, (4) the "$1000 genome" to study relations between clinical and genomic factors, and (5) "big data" techniques for analysis. With enough data, even simple machine learning techniques find strong predictive relationships. I describe our experiences predicting mortality and other important clinical risks and therapeutic opportunities for intensive care patients. Our current efforts abstract more informative features from both coded (tabular) data and narrative descriptions of the patient, biasing those unsupervised processes toward representing existing medical knowledge. I will outline what seem like further promising approaches. (Joint work with R. Joshi, A. Rumshisky, C. Hug, K. Kshetri, M. Ghassemi and T. Naumann.)

Peter Szolovits

Massachusetts Institute of Technology, USA

Coffee Break

2:45 PM-3:00 PM

Room:Ballroom Foyer - 3rd Floor

Thursday, May 2

CP4

Graph Stream Data Mining

3:00 PM-5:05 PM

Room:Capitol Ballroom D - 3rd Floor

Chair: Feida Zhu, Singapore Management University, Singapore

3:00-3:20 On Graph Stream Clustering with Side Information
Yuchen Zhao and Philip Yu, University of Ilinois at Chicago, USA

3:25-3:45 Dynamic Community Detection in Weighted Graph Streams
Chang-Dong Wang and Jian-Huang Lai, Sun Yat-Sen University, China; Philip Yu, University of Ilinois at Chicago, USA

3:50-4:10 DeltaCon: A Principled Massive-Graph Similarity Function
Danai Koutra, Carnegie Mellon University, USA; Joshua Vogelstein, Duke University, USA; Christos Faloutsos, Carnegie Mellon University, USA

4:15-4:35 What's Your Next Move: User Activity Prediction in Location-Based Social Networks
Jihang Ye, Zhe Zhu, and *Hong Cheng*, The Chinese University of Hong Kong, Hong Kong

4:40-5:00 CoFiSet: Collaborative Filtering Via Learning Pairwise Preferences over Item-Sets
Weike Pan and Li Chen, Hong Kong Baptist University, Hong Kong

Thursday, May 2

CP5

Anomaly & Outlier Detection

3:00 PM-5:05 PM

Room:Capitol Ballroom E - 3rd Floor

Chair: Nikunj Oza, NASA Ames Research Center, USA

3:00-3:20 *k*-Means{-}{-}: A Unified Approach to Clustering and Outlier Detection

Aristides Gionis, Yahoo! Research Barcelona, Spain; Sanjay Chawla, University of Sydney, Australia

3:25-3:45 CMI: An Information-Theoretic Contrast Measure for Enhancing Subspace Cluster and Outlier Detection

Hoang Vu Nguyen and Emmanuel Müller, Karlsruhe Institute of Technology, Germany; Jilles Vreeken, Universiteit Antwerpen, Belgium; Fabian Keller and Klemens Böhm, Karlsruhe Institute of Technology, Germany

3:50-4:10 Cost-Sensitive Double Updating Online Learning and Its Application to Online Anomaly Detection

Peilin Zhao and *Steven C.H. Hoi*, Nanyang Technological University, Singapore

4:15-4:35 Efficient Selection of Globally Optimal Rules on Large Imbalanced Data Based on Rule Coverage Relationship Analysis

Jinjiu Li, *Can Wang*, and Longbing Cao, University of Technology, Sydney, Australia; Philip S. Yu, University of Ilinois at Chicago, USA

4:40-5:00 Outlier Detection with Space Transformation and Spectral Analysis

Xuan-Hong Dang, Barbora Micenkova, and Ira Assent, Aarhus University, Denmark; Raymond T. Ng, University of British Columbia, Canada

Thursday, May 2

CP6

Multi-View and Multi-Source Data Mining

3:00 PM-5:05 PM

Room:Capitol Ballroom FGH - 3rd Floor

Chair: Shuiwang Ji, Old Dominion University, USA

3:00-3:20 Multi-Objective Multi-View Spectral Clustering Via Pareto Optimization

Xiang Wang, Buyue Qian, and *Ian Davidson*, University of California, Davis, USA; Jieping Ye, Arizona State University, USA

3:25-3:45 Multi-Transfer: Transfer Learning with Multiple Views and Multiple Sources

Ben Tan, Erheng Zhong, Evan Wei Xiang, and Qiang Yang, Hong Kong University of Science and Technology, Hong Kong

3:50-4:10 Multi-View Clustering Via Joint Nonnegative Matrix Factorization

Jialu Liu and Chi Wang, University of Illinois at Urbana-Champaign, USA; Jing Gao, State University of New York at Buffalo, USA; Jiawei Han, University of Illinois at Urbana-Champaign, USA

4:15-4:35 On Handling Negative Transfer and Imbalanced Distributions in Multiple Source Transfer Learning

Liang Ge, Jing Gao, Hung Ngo, Kang Li, and Aidong Zhang, State University of New York at Buffalo, USA

4:40-5:00 Unsupervised Feature Selection for Multi-View Data in Social Media

Jiliang Tang, Xia Hu, Huiji Gao, and Huan Liu, Arizona State University, USA

Thursday, May 2

TS2

Tutorial Session: Big Data Analytics for Healthcare

3:00 PM-5:05 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, we introduce the characteristics and related mining challenges on dealing with big medical data. Many of those insights come from medical informatics community, which is highly related to data mining but focuses on biomedical specifics. We survey various related papers from data mining venues as well as medical informatics venues to share with the audiences key problems and trends in healthcare analytics research, with different applications ranging from clinical text mining, predictive modeling, patient similarity, genetic data analysis, privacy on medical data and medical images.

Jimeng Sun
IBM T.J. Watson Research Center, USA

Chandan Reddy
Wayne State University, USA

Organizational Break

5:05 PM-5:15 PM

Thursday, May 2

PP1

Poster Spotlights

5:15 PM-6:30 PM

Room:Capitol Ballroom DE - 3rd Floor

Time-Sensitive Classification of Behavioral Data
Shin Ando, Gunma University, Japan; Einoshin Suzuki, Kyushu University, Japan

An Examination of Practical Granger Causality Inference
Mohammad Taha Bahadori and Yan Liu, University of Southern California, USA

Bregman Divergence and Triangle Inequality
Sreangsu Acharyya, University of Texas at Austin, USA; *Arindam Banerjee* and Daniel L. Boley, University of Minnesota, USA

Joint Segmentation and Clustering in Text Corpuses
Samuel J. Blasiak, George Mason University, USA; Sithu Sudarsan, Food and Drug Administration, USA; Huzefa Rangwala, George Mason University, USA

Automatic Detection and Correction of Multi-Class Classification Errors Using System Whole-Part Relationships
Zhengzhang Chen, Northwestern University, USA; John Jenkins, North Carolina State University, USA; Alok Choudhary, Northwestern University, USA; Jinfeng Rao, Zhejiang University, China and North Carolina State University, USA; Fredrick Semazzi and Anatoli Melechko, North Carolina State University, USA; Vipin Kumar, University of Minnesota, USA; Nagiza Samatova, North Carolina State University and Oak Ridge National Laboratory, USA

Contextual Time Series Change Detection
XI Chen, Karsten Steinhaeuser, and Shyam Boriah, University of Minnesota, USA; Snigdhansu Chatterjee, University of Minnesota, Twin Cities, USA; Vipin Kumar, University of Minnesota, USA

continued in next column

Very Fast Similarity Queries on Semi-Structured Data from the Web
Bhavana Dalvi and William Cohen, Carnegie Mellon University, USA

Topic Models For Feature Selection in Document Clustering
Anna Drummond, Rice University, USA; Zografoula Vagena, LogicBlox Inc., USA; Chris Jermaine, Rice University, USA

A Nonparametric Mixture Model for Topic Modeling over Time
Ahmed Hefny, *Avinava Dubey*, Sinead Williamson, and Eric Xing, Carnegie Mellon University, USA

Discriminative Transfer Learning on Manifold
Zheng Fang and Zhongfei Zhang, Zhejiang University, China

SemInf: A Burst-Based Semantic Influence Model for Biomedical Topic Influence
Dan He, IBM T.J. Watson Research Center, USA; Douglas Parker, University of California, Los Angeles, USA

Pinch Ratio Clustering from a Topologically Intrinsic Lexicographic Ordering
Douglas R. Heisterkamp and Jesse Johnson, Oklahoma State University, USA

Retweeting: An Act of Viral Users, Susceptible Users, Or Viral Topics?
Tuan-Anh Hoang and Ee-Peng Lim, Singapore Management University, Singapore

Time Series Classification under More Realistic Assumptions
Bing Hu, Yanping Chen, and Eamonn Keogh, University of California, Riverside, USA

Finding Affordable and Collaborative Teams from a Network of Experts
Aijun An, *Mehdi Kargar*, and Morteza Zihayat, York University, Canada

IBSM: Interval-Based Sequence Matching
Alexios Kotsifakos, University of Texas at Arlington, USA; Panagiotis Papapetrou, Birkbeck, University of London, United Kingdom; Vassilis Athitsos, University of Texas at Arlington, USA

continued in next column

Modeling the Diffusion of Preferences on Social Networks
J Lou, Fu-Min Wang, Chin-Hua Tsai, San-Chuan Hung, Perng-Hwa Kung, and *Shou-De Lin*, National Taiwan University, Taiwan

Mining Labelled Tensors by Discovering Both Their Common and Discriminative Subspaces
Wei Liu, Jeffrey Chan, James Bailey, Christopher Leckie, and Ramamohanarao Kotagiri, The University of Melbourne, Australia

Modeling Clinical Time Series Using Gaussian Process Sequences
Zitao Liu and Milos Hauskrecht, University of Pittsburgh, USA

Integrity Verification of K-Means Clustering Outsourced to Infrastructure As a Service (IaaS) Providers
Wendy Hui Wang, *Ruilin Liu*, and Philippos Mordohai, Stevens Institute of Technology, USA; Hui Xiong, Rutgers University, USA

Selective Transfer Learning for Cross Domain Recommendation
Zhongqi Lu, Erheng Zhong, Lili Zhao, Wei Xiang, Weike Pan, and Qiang Yang, Hong Kong University of Science and Technology, Hong Kong

Change Detection from Temporal Sequences of Class Labels: Application to Land Cover Change Mapping
Varun Mithal and Ankush Khandelwal, University of Minnesota, Twin Cities, USA; Shyam Boriah, University of Minnesota, USA; Karsten Steinhaeuser, University of Minnesota, Twin Cities, USA; Vipin Kumar, University of Minnesota, USA

Fractional Immunization in Networks
B. Aditya Prakash, Virginia Tech, USA; Lada Adamic, and Theodore Iwashyna, University of Michigan, Ann Arbor, USA; Hanghang Tong, City College of CUNY, USA; Christos Faloutsos, Carnegie Mellon University, USA

continued on next page

Fast Shapelets: A Scalable Algorithm for Discovering Time Series Shapelets
Thanawin Rakthanamanon and Eamonn Keogh, University of California, Riverside, USA

Mc-MinH: Metagenome Clustering Using Minwise Based Hashing
Zeehasham Rasheed and Huzefa Rangwala, George Mason University, USA

Shattering and Compressing Networks for Betweenness Centrality
A. Erdem Sariyuce, Erik Saule, Kamer Kaya, and Umit V. Catalyurek, The Ohio State University, USA

CoSelect: Feature Selection with Instance Selection for Social Media Data
Jiliang Tang and Huan Liu, Arizona State University, USA

A Hierarchical Probabilistic Model for Low Sample Rate Home-Use Energy Disaggregation
Bingsheng Wang and Haili Dong, Virginia Tech, USA; Arnold Boedihardjo, US Army Corps of Engineers, USA; Feng Chen and Chang-Tien Lu, Virginia Tech, USA

On the Detectability of Node Grouping in Networks
Chi Wang, Hongning Wang, Jialu Liu, Ming Ji, Lu Su, Yuguo Chen, and Jiawei Han, University of Illinois at Urbana-Champaign, USA

Mods: Multiple One-Class Data Streams Learning from Homogeneous Data
Yanshan Xiao and Bo Liu, University of Technology, Sydney, Australia; Yanshan Xiao, Guangdong University of Technology, China; Philip Yu, University of Ilinois at Chicago, USA; Zhifeng Hao, Guangdong University of Technology, China

continued in next column

Robust Textual Data Streams Mining Based on Continuous Transfer Learning
Yanshan Xiao and Bo Liu, University of Technology, Sydney, Australia; Yanshan Xiao, Guangdong University of Technology, China; Philip Yu, University of Ilinois at Chicago, USA; Longbing Cao, University of Technology, Sydney, Australia; Zhifeng Hao, Guangdong University of Technology, China

Graphical Modeling of Macro Behavioral Targeting in Social Networks
Yusheng Xie, Zhengzhang Chen, Kunpeng Zhang, Md. Mostofa Ali Patwary, Yu Cheng, Haotian Liu, Ankit Agrawal, and Alok Choudhary, Northwestern University, USA

Learning Topics in Short Texts by Non-Negative Matrix Factorization on Term Correlation Matrix
Xiaohui Yan, Jiafeng Guo, Shenghua Liu, and Xueqi Cheng, Chinese Academy of Sciences, China; Yanfeng Wang, Sogou Inc., China

Set Coverage Problems in a One-Pass Data Stream
Huiwen Yu and Dayu Yuan, Pennsylvania State University, USA

Sentiment Topic Model with Decomposed Prior
Chengtao Li, Tsinghua University, P. R. China; *Jianwen Zhang*, Jian-Tao Sun, and Zheng Chen, Microsoft Research Asia

Topic-Level Expert Modeling in Community Question Answering
Tong Zhao, Naiwen Bian, Chunping Li, and Mengya Li, Tsinghua University, P. R. China

Butterfly Mixing: Accelerating Incremental-Update Algorithms on Clusters
Huasha Zhao and John Canny, University of California, Berkeley, USA

Feature Selection by Joint Graph Sparse Coding
Xiaofeng Zhu, University of Queensland, Australia; Xindong Wu, University of Vermont, USA; Wei Ding, University of Massachusetts, Boston, USA; Shichao Zhang, Guangxi Normal University, China

continued in next column

It Is Not Just What We Say, But How We Say Them: Lda-Based Behavior-Topic Model
Minghui Qiu, *Feida Zhu*, and Jing Jiang, Singapore Management University, Singapore

Welcome Reception and Poster Session

6:30 PM-9:00 PM

Room:Capitol View Terrace - 3rd Floor

Friday, May 3

Registration
7:30 AM-3:30 PM

Room:Ballroom Foyer - 3rd Floor

Continental Breakfast

7:30 AM

Room:Ballroom Foyer - 3rd Floor

Announcements
8:00 AM-8:15 AM

Room:Capitol Ballroom DE - 3rd Floor

Friday, May 3

IP3
Modeling Individual-Level Data in the 21st Century
8:15 AM-9:30 AM

Room:Capitol Ballroom DE - 3rd Floor

Chair: Joydeep Ghosh, University of Texas at Austin, USA

The collection and analysis of data related to human behavior has changed dramatically over the past 40 years, from the collection of small amounts of relatively static demographic data (such as a person's zipcode and education level), to much more detailed and dynamic transaction data (such as credit card and telephone records). More recently we have seen the rapid advent of individual-level "micro-data," including Web search, email, microblogs, online social media, geolocation data, and more. In this talk we will discuss some of the new research challenges and opportunities presented by such data. We will look at common themes across the variety of data sets and research projects in this general area, focusing both on what new types of data analysis techniques are likely to be needed, and what new scientific questions and applications are emerging in areas such as computational social science and public health.

Padhraic Smyth
University of California, Irvine, USA

Coffee Break

9:30 AM-10:00 AM

Room:Ballroom Foyer - 3rd Floor

Friday, May 3

CP7
Semi-Supervised and Active Learning
10:00 AM-12:05 PM

Room:Capitol Ballroom D - 3rd Floor

Chair: Romer Rosales, LinkedIn, USA

10:00-10:20 Smart: Semi-Supervised Music Emotion Recognition with Social Tagging
Bin Wu, Erheng Zhong, Derek Hao Hu, Andrew Horner, and Qiang Yang, Hong Kong University of Science and Technology, Hong Kong

10:25-10:45 Probabilistic Combination of Classifier and Cluster Ensembles for Non-Transductive Learning
Ayan Acharya, University of Texas at Austin, USA; Eduardo Hruschka, University of São Paulo at São Carlos, Brazil; Joydeep Ghosh, University of Texas at Austin, USA; Badrul Sarwar and Jean-David Ruvini, eBay Research Labs, USA

10:50-11:10 Active Learning to Rank Using Pairwise Supervision
Buyue Qian, University of California, Davis, USA; Hongfei Li and Jun Wang, IBM T.J. Watson Research Center, USA; Xiang Wang and Ian Davidson, University of California, Davis, USA

11:15-11:35 ActNeT: Active Learning for Networked Texts in Microblogging
Xia Hu, Jiliang Tang, Huiji Gao, and Huan Liu, Arizona State University, USA

11:40-12:00 Active Class Discovery and Learning for Networked Data
Meng Fang, University of Technology, Sydney, Australia; Jie Yin, CSIRO, Australia; Xingquan Zhu and Chengqi Zhang, University of Technology, Sydney, Australia

Friday, May 3

CP8

Applications

10:00 AM-12:05 PM

Room:Capitol Ballroom E - 3rd Floor

Chair: Hui Xiong, Rutgers University, USA

10:00-10:20 Climate Multi-Model Regression Using Spatial Smoothing
Karthik Subbian and Arindam Banerjee, University of Minnesota, USA

10:25-10:45 A Distribution Regularized Regression Framework for Climate Modeling
Zubin Abraham, Michigan State University, USA; Pang-Ning Tan, Perdinan Perdinan, Julie Winkler, and Shiyuan Zhong, Michigan State University, USA; Malgozata Liszewska, University of Warsaw, Poland

10:50-11:10 Sparse Representation for Hiv-1 Protease Drug Resistance Prediction
Xiaxia Yu, Irene Weber, and Robert Harrison, Georgia State University, USA

11:15-11:35 Dynamic Shaker Detection from Evolving Entities
Xiaoxiao Shi, University of Ilinois at Chicago, USA; Wei Fan, IBM T.J. Watson Research Center, USA; Philip Yu, University of Ilinois at Chicago, USA

11:40-12:00 Monitoring and Mining Gps Traces in Transit Space
Leon O. Stenneth and Philip Yu, University of Ilinois at Chicago, USA

Friday, May 3

CP9

Social Network Analysis

10:00 AM-12:05 PM

Room:Capitol Ballroom FGH - 3rd Floor

Chair: Shou-De Lin, National Taiwan University, Taiwan

10:00-10:20 Exploiting Synchronicity Networks for Finding Valuables in Heterogeneous Networks
Zhen Wen, IBM T.J. Watson Research Center, USA; Ching-Yung Lin, IBM Research, USA

10:25-10:45 Exploring and Inferring User-User Pseudo-Friendship for Sentiment Analysis with Heterogeneous Networks
Hongbo Deng and Jiawei Han, University of Illinois at Urbana-Champaign, USA; Hao Li and Heng Ji, City University of New York, USA; Hongning Wang, University of Illinois at Urbana-Champaign, USA; Yue Lu, Twitter Inc., USA; Chi Wang, University of Illinois at Urbana-Champaign, USA

10:50-11:10 Opinion Maximization in Social Networks
Aristides Gionis, Yahoo! Research Barcelona, Spain; Evamaria Terzi, Boston University, USA; Panayiotis Tsaparas, University of Ioannina, Greece

11:15-11:35 Point-of-Interest Recommendation in Location Based Social Networks with Topic and Location Awareness
Bin Liu and Hui Xiong, Rutgers University, USA

11:40-12:00 Community Detection with Prior Knowledge
Karthik Subbian, University of Minnesota, USA; Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Jaideep Srivastava, University of Minnesota, USA; Philip Yu, University of Ilinois at Chicago, USA

Friday, May 3

TS3

Tutorial Session: Crowdsourcing & Human Computation: Data Labeling & Building Hybrid Systems

10:00 AM-12:05 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

This tutorial is aimed at those with little to intermediate experience with crowdsourcing and human computation, who are interested in learning more about the capabilities and limitations of crowdsourcing for: 1) collecting labeled data; and 2) integrating human computation with automation to build more effective, hybrid intelligent systems.

Matthew Lease
University of Texas at Austin, USA

Lunch Break

12:05 PM-1:30 PM

Attendees on their own

Friday, May 3

IP4

Social Networks As Information Filters

1:30 PM-2:45 PM

Room:Capitol Ballroom DE - 3rd Floor

Chair: Zoran Obradovic, Temple University, USA

Social networks, especially online social networks, are driven by information sharing. But just how much information sharing is influenced by social networks? A large-scale experiment measured the effect of the social network on the quantity and diversity of information being shared within Facebook. While strong ties were found to be individually more influential, collectively it is the strong ties that wield more influence and provide more diverse information exposure. Furthermore, the network not only transmits information, but also often modifies it, allowing it to evolve.

Lada Adamic

University of Michigan, Ann Arbor, USA

Coffee Break

2:45 PM-3:00 PM

Room:Ballroom Foyer - 3rd Floor

Friday, May 3

CP10

Classification and Sparse Methods

3:00 PM-5:05 PM

Room:Capitol Ballroom FGH - 3rd Floor

Chair: Steven Hoi, Nanyang Technological University, Singapore

3:00-3:20 Regularization of Latent Variable Models to Obtain Sparsity
Ramnath Balasubramanyan and William Cohen, Carnegie Mellon University, USA

3:25-3:45 Sparse Max-Margin Multiclass and Multi-Label Classifier Design for Fast Inference
Tanuja Ganu, IBM Research, India; Shirish Shevade, Indian Institute of Science, Bangalore, India; S Sudararajan, Microsoft Research, India

3:50-4:10 An Empirical Study of the Suitability of Class Decomposition for Linear Models: When Does It Work Well?
Francisco Ocegueda-Hernandez and Ricardo Vilalta, University of Houston, USA

4:15-4:35 Reduced Set Kpca for Improving the Training and Execution Speed of Kernel Machines
Hassan A. Kingravi, Georgia Institute of Technology, USA

4:40-5:00 A New Perspective on Convex Relaxations of Sparse Svm
Noam Goldberg, Carnegie Mellon University, USA; Sven Leyffer and Todd Munson, Argonne National Laboratory, USA

Friday, May 3

TS4

Tutorial Session: Recent Advances in Applied Matrix Technologies

3:00 PM-5:05 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

Matrix is a natural representation for many real world data, such as an image, a collection of documents and an adjacency graph, and matrix related technologies has been very popular in data analytics research because of its nice interpretability, effectiveness and efficiency. This tutorial will overview the recent advances in applied matrix technologies, include nonnegative matrix factorization, rank minimization, sparse learning and online/distributed learning strategies. We will also present how these technologies are applied in social and healthcare informatics.

Fei Wang
Cornell University, USA

Hanghang Tong
IBM T.J. Watson Research Center, USA

Organizational Break

5:05 PM-5:15 PM

NSF Panel

5:15 PM-6:30 PM

Room:Capitol Ballroom DE - 3rd Floor

SIAG/DMA Business Meeting

6:30 PM-7:00 PM

Room:Capitol Ballroom DE - 3rd Floor

Complimentary wine and beer will be served.

Doctoral Forum and Student Posters

7:00 PM-9:00 PM

Room:Capitol View Terrace - 3rd Floor

Saturday, May 4

Registration
7:30 AM-4:00 PM

Room:Ballroom Foyer - 3rd Floor

Continental Breakfast

7:30 AM

Room:Ballroom Foyer - 3rd Floor

Analytics for Cyber-Physical Systems Workshop
8:05 AM-9:20 AM

Room:Capitol Ballroom FGH - 3rd Floor

For schedule, see page 17.

Data Mining for Medicine and Healthcare Workshop
8:30 AM-10:00 AM

Room:Capitol Ballroom D - 3rd Floor

For schedule, see pages 15-16.

Data Mining for Service and Maintenance Workshop
9:00 AM-10:00 AM

Room:Capitol Ballroom E - 3rd Floor

Saturday, May 4

TS5

Tutorial Session: Outlier Detection for Temporal Data
8:30 AM-10:00 AM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, we will present an organized picture of recent research in temporal outlier detection. We will begin by motivating the importance of temporal outlier detection and briefing the challenges beyond usual outlier detection. Then, we will list down a taxonomy of proposed techniques for temporal outlier detection and cover some of the techniques in detail. We will summarize by presenting a collection of applications where temporal outlier detection techniques have been applied to discover interesting outliers.

Manish Gupta
University of Illinois at Urbana-Champaign, USA

Jing Gao
State University of New York at Buffalo, USA

Charu C. Aggarwal
IBM T.J. Watson Research Center, USA

Jiawei Han
University of Illinois at Urbana-Champaign, USA

Saturday, May 4

Analytics for Cyber-Physical Systems Workshop, continued
9:30 AM-10:35 AM

Room:Capitol Ballroom FGH - 3rd Floor

For schedule, see page 17.

Coffee Break

10:00 AM-10:30 AM

Room:Ballroom Foyer - 3rd Floor

Data Mining for Medicine and Healthcare Workshop, continued
10:30 AM-12:00 PM

Room:Capitol Ballroom D - 3rd Floor

For schedule, see pages 15-16.

Data Mining for Service and Maintenance Workshop, continued
10:30 AM-12:00 PM

Room:Capitol Ballroom E - 3rd Floor

Saturday, May 4

TS5

Tutorial Session: Outlier Detection for Temporal Data, continued

10:30 AM-12:00 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, we will present an organized picture of recent research in temporal outlier detection. We will begin by motivating the importance of temporal outlier detection and briefing the challenges beyond usual outlier detection. Then, we will list down a taxonomy of proposed techniques for temporal outlier detection and cover some of the techniques in detail. We will summarize by presenting a collection of applications where temporal outlier detection techniques have been applied to discover interesting outliers.

Manish Gupta
University of Illinois at Urbana-Champaign, USA

Jing Gao
State University of New York at Buffalo, USA

Charu C. Aggarwal
IBM T.J. Watson Research Center, USA

Jiawei Han
University of Illinois at Urbana-Champaign, USA

Saturday, May 4

Analytics for Cyber-Physical Systems Workshop, continued

10:45 AM-12:30 PM

Room:Capitol Ballroom FGH - 3rd Floor

For schedule, see page 17.

Lunch Break

12:00 PM-1:30 PM

Attendees on their own

Data Mining for Service and Maintenance Workshop, continued

1:30 PM-3:00 PM

Room:Capitol Ballroom E - 3rd Floor

Reliability Aware Data Fusion in Participatory Networks Workshop

1:30 PM-3:00 PM

Room:Capitol Ballroom FGH - 3rd Floor

For schedule, see page 18.

Data Mining for Medicine and Healthcare Workshop, continued

1:30 PM-3:00 PM

Room:Capitol Ballroom D - 3rd Floor

For schedule, see pages 15-16.

Saturday, May 4

TS6

Tutorial Session: Online Learning for Big Data Mining: Methods and Applications

1:30 PM-3:00 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, the main content consists of two parts:

(i) online learning methods for linear classification/prediction, and

(ii) kernel-based online learning methods for nonlinear classification/prediction.

Steven C.H. Hoi
Nanyang Technological University, Singapore

Coffee Break

3:00 PM-3:30 PM

Room:Ballroom Foyer - 3rd Floor

Data Mining for Medicine and Healthcare Workshop, continued

3:15 PM-5:00 PM

Room:Capitol Ballroom D - 3rd Floor

For schedule, see pages 15-16.

Data Mining for Service and Maintenance Workshop, continued

3:30 PM-5:30 PM

Room:Capitol Ballroom E - 3rd Floor

Reliability Aware Data Fusion in Participatory Networks Workshop, continued

3:30 PM-5:00 PM

Room:Capitol Ballroom FGH - 3rd Floor

For schedule, see pages 18.

Saturday, May 4

TS6

Tutorial Session: Online Learning for Big Data Mining: Methods and Applications, continued

3:30 PM-5:00 PM

Room:Capitol Ballroom ABC - 3rd Floor

Chair: Jieping Ye, Arizona State University, USA

In this tutorial, the main content consists of two parts:

(i) online learning methods for linear classification/prediction, and

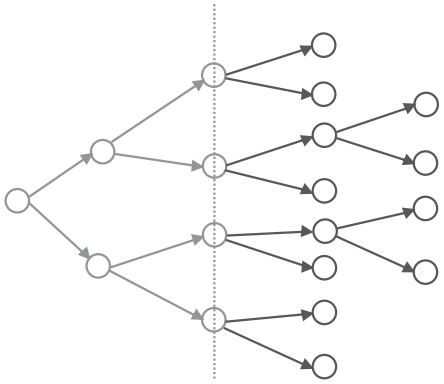
(ii) kernel-based online learning methods for nonlinear classification/prediction.

Steven C.H. Hoi
Nanyang Technological University, Singapore

SDM13 Abstracts

2013 SIAM International Conference on DATA MINING

May 2–4, 2013



Sheraton Austin Hotel at the Capitol
Austin, Texas, USA

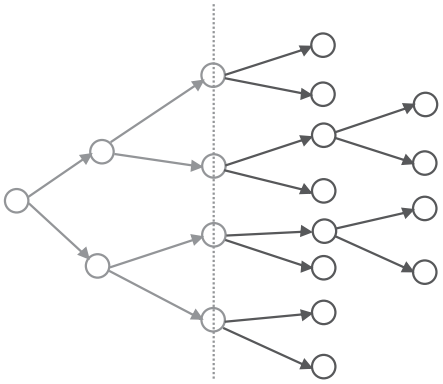
Abstracts are printed as submitted by the authors.

last page of abstracts

Organizer and Speaker Index

2013 SIAM International Conference on DATA MINING

May 2–4, 2013



Sheraton Austin Hotel at the Capitol
Austin, Texas, USA

A

Abraham, Zubin, CP8, 10:25 Fri

Acharya, Ayan, CP7, 10:25 Fri

Adamic, Lada, IP4, 1:30 Fri

Aggarwal, Charu C., TS5, 8:30 Sat

Akoglu, Leman, CP1, 11:40 Thu

Ando, Shin, PP1, 5:15 Thu

B

Bahadori, Mohammad Taha, PP1, 5:15 Thu

Balasubramanyan, Ramnath, CP10, 3:00 Fri

Banerjee, Arindam, PP1, 5:15 Thu

Blasiak, Samuel J., PP1, 5:15 Thu

Bogdanov, Petko, CP1, 11:15 Thu

Boley, Daniel L., CP3, 10:25 Thu

C

Chen, XI, PP1, 5:15 Thu

Chen, Zhengzhang, PP1, 5:15 Thu

Cheng, Hong, CP4, 4:15 Thu

Clémençon, Stephan, CP1, 10:50 Thu

D

Dalvi, Bhavana, PP1, 5:15 Thu

Dang, Xuan-Hong, CP5, 4:40 Thu

Davidson, Ian, CP6, 3:00 Thu

Drummond, Anna, PP1, 5:15 Thu

Dubey, Avinava, PP1, 5:15 Thu

F

Fang, Meng, CP7, 11:40 Fri

Fang, Zheng, PP1, 5:15 Thu

G

Ganu, Tanuja, CP10, 3:25 Fri

Gao, Jing, TS5, 8:30 Sat

Ge, Liang, CP6, 4:15 Thu

Gionis, Aristides, CP5, 3:00 Thu

Gionis, Aristides, CP9, 10:50 Fri

Goldberg, Noam, CP10, 4:40 Fri

Gupta, Manish, TS5, 8:30 Sat

H

Han, Jiawei, TS5, 8:30 Sat

He, Dan, PP1, 5:15 Thu

Heisterkamp, Douglas R., PP1, 5:15 Thu

Hoang, Tuan-Anh, PP1, 5:15 Thu

Hoi, Steven C.H., CP5, 3:50 Thu

Hoi, Steven C.H., TS6, 1:30 Sat

Hu, Bing, PP1, 5:15 Thu

Hu, Xia, CP7, 11:15 Fri

K

Kargar, Mehdi, PP1, 5:15 Thu

Kingravi, Hassan A., CP10, 4:15 Fri

Kong, Xiangnan, CP2, 11:40 Thu

Kotsifakos, Alexios, PP1, 5:15 Thu

Koutra, Danai, CP4, 3:50 Thu

L

Lease, Matthew, TS3, 10:00 Fri

Li, Cheng-Te, TS1, 10:00 Thu

Lin, Shou-De,TS1, 10:00 Thu

Lin, Shou-De, PP1, 5:15 Thu

Liu, Bin, CP9, 11:15 Fri

Liu, Chunyang, CP2, 11:15 Thu

Liu, Jialu, CP6, 3:50 Thu

Liu, Ruilin, PP1, 5:15 Thu

Liu, Wei, PP1, 5:15 Thu

Liu, Zitao, PP1, 5:15 Thu

Lu, Zhongqi, PP1, 5:15 Thu

M

Mai, Son T., CP3, 10:50 Thu

Mithal, Varun, PP1, 5:15 Thu

N

Nguyen, Hoang Vu, CP5, 3:25 Thu

Nychka, Doug, IP1, 8:15 Thu

O

Ocegueda-Hernandez, Francisco, CP10, 3:50 Fri

P

Pan, Weike, CP4, 4:40 Thu

Prakash, B. Aditya, PP1, 5:15 Thu

Q

Qian, Buyue, CP7, 10:50 Fri

R

Race, Shaina L., CP3, 10:00 Thu

Rakthanamanon, Thanawin, PP1, 5:15 Thu

Ram, Parikshit, CP1, 10:00 Thu

Rasheed, Zeehasham, PP1, 5:15 Thu

Reddy, Chandan K., TS2, 3:00 Thu

S

Saha, Budhaditya, CP3, 11:40 Thu

Sariyuce, A. Erdem, PP1, 5:15 Thu

Seshadhri, C., CP1, 10:25 Thu

Shi, Xiaoxiao, CP8, 11:15 Fri

Smyth, Padhraic, IP3, 8:15 Fri

Stenneth, Leon O., CP8, 11:40 Fri

Subbian, Karthik, CP8, 10:00 Fri

Subbian, Karthik, CP9, 11:40 Fri

Sun, Jimeng, TS2, 3:00 Thu

Szolovits, Peter, IP2, 1:30 Thu

T

Tan, Ben, CP6, 3:25 Thu

Tan, Pang-Ning, CP2, 10:00 Thu

Tang, Jiliang, CP6, 4:40 Thu

Tang, Jiliang, PP1, 5:15 Thu

Tong, Hanghang, TS4, 3:00 Fri

Italicized names indicate session organizers.

W

Wang, Bingsheng, PP1, 5:15 Thu
Wang, Can, CP5, 4:15 Thu
Wang, Chang-Dong, CP4, 3:25 Thu
Wang, Chi, PP1, 5:15 Thu
Wang, Chi, CP9, 10:25 Fri
Wang, Fei, TS4, 3:00 Fri
Wen, Zhen, CP9, 10:00 Fri
Wu, Bin, CP7, 10:00 Fri

X

Xiao, Yanshan, PP1, 5:15 Thu
Xiao, Yanshan, PP1, 5:15 Thu
Xie, Yusheng, PP1, 5:15 Thu

Y

Yan, Xiaohui, PP1, 5:15 Thu
Yeh, Mi-Yen, TS1, 10:00 Thu
Yu, Huiwen, PP1, 5:15 Thu
Yu, Xiaxia, CP8, 10:50 Fri

Z

Zhang, Jianwen, PP1, 5:15 Thu
Zhang, Miao, CP2, 10:50 Thu
Zhang, Wenlu, CP3, 11:15 Thu
Zhao, Huasha, PP1, 5:15 Thu
Zhao, Tong, PP1, 5:15 Thu
Zhao, Yuchen, CP4, 3:00 Thu
Zhou, Jiayu, CP2, 10:25 Thu
Zhu, Feida, PP1, 5:15 Thu
Zhu, Xiaofeng, PP1, 5:15 Thu

Italicized names indicate session organizers.

Notes

SDM13 Budget

Conference Budget
SIAM Conference on Data Mining
May 2-4, 2013
Austin, TX

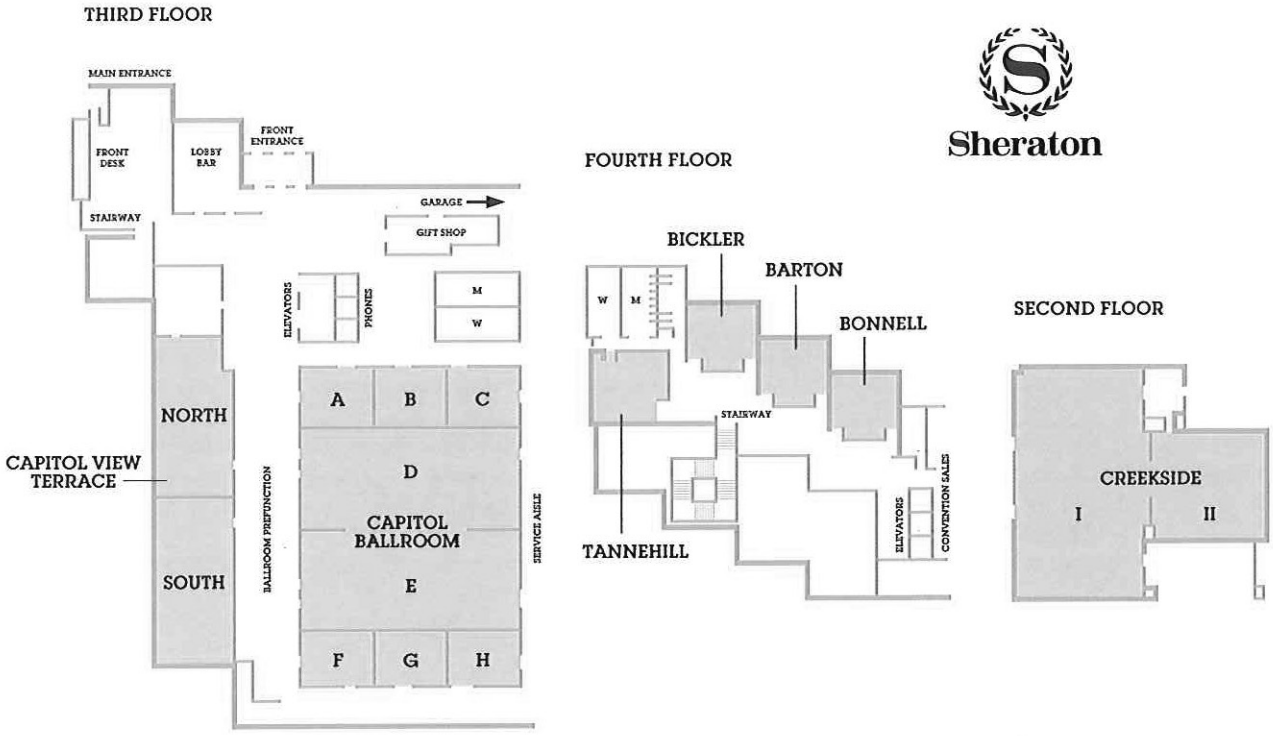
Expected Paid Attendance 250

Revenue		
Registration Income		$99,130
	Total	$99,130
Expenses		
Printing		$1,300
Organizing Committee		$1,700
Invited Speakers		$10,400
Food and Beverage		$22,600
AV Equipment and Telecommunication		$18,300
Advertising		$4,000
Proceedings		$14,200
Conference Labor (including benefits)		$30,580
Other (supplies, staff travel, freight, misc.)		$4,800
Administrative		$10,926
Accounting/Distribution & Shipping		$7,667
Information Systems		$10,443
Customer Service		$3,933
Marketing		$6,391
Office Space (Building)		$4,289
Other SIAM Services		$4,783
	Total	$156,312
Net Conference Expense		($57,182)
Support Provided by SIAM		$57,182
		$0

Estimated Support for Travel Awards not included above:

Post Docs	2	$1,600

Sheraton Austin Hotel at the Capitol Map



FSC logo text box indicating size & layout of logo. Conlins to insert logo.