2015 SIAM International
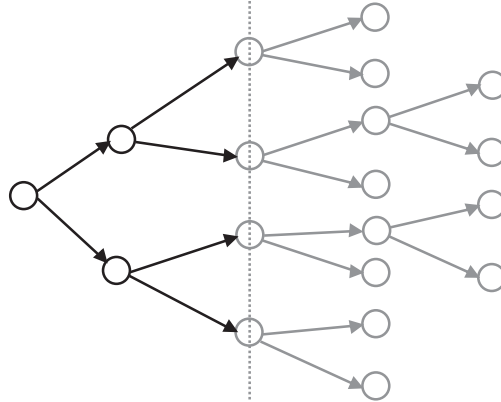## Conference on **DATA MINING**

April 30-May 2, 2015

Pinnacle Vancouver Harbourfront Hotel
Vancouver, British Columbia, Canada

Sponsored by the SIAM Activity Group on
Data Mining and Analytics

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. The activity group organizes the yearly SIAM International Conference on Data Mining (SDM),organizes minisymposia at the SIAM Annual Meeting, and maintains a membership directory and electronic mailing list.

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

siam.

Table of Contents

Organizing Committee

Steering Committee Chair

Srinivasan Parthasarathy
Ohio State University, USA

Conference Chairs

Ke Wang
Simon Fraser University, Canada

Mohammed Zaki
Qatar Computing Research Institute, Qatar and Rensselaer Polytechnic Institute, USA

Program Chairs

Suresh Venkatasubramanian
University of Utah, USA

Jieping Ye
Arizona State University, USA

Workshop Chairs

Xiaoli Fern
Oregon State University, USA

Xifeng Yan
University of California Santa Barbara, USA

Tutorial Chair

Tao Li
Florida International University, USA

Doctoral Forum Chair

Anoop Sarkar
Simon Fraser University, Canada

Publicity Chairs

Leman Akoglu
Stony Brook University, USA

Mohammad Al Hasan
Indiana University-Purdue University Indianapolis, USA

Panels Chair

Sanjay Chawla
University of Sydney, Australia

Sponsorship Chairs

Nitesh Chawla
University of Notre Dame, USA

Fred Popowich
Simon Fraser University, Canada

Senior Program Committee

Shivani Agarwal
Indian Institute of Science, Bangalore, India

James Bailey
University of Melbourne, Australia

Tanya Berger-Wolf
University of Illinois at Chicago, USA

Francesco Bonchi
Yahoo! Barcelona, Spain

Rajmonda Caceras
Massachusetts Institute of Technology: Lincoln Laboratory, USA

Diane Cook
Washington State University, USA

Carlotta Domeniconi
George Mason University, USA

Tina Eliassi-Rad
Rutgers University, USA

Wei Fan
Huawei Noah's Ark Lab, Hong Kong

Minos Garofalakis
Technical University of Crete, Greece

Fosca Gianotti
University of Pisa, Italy

Aristides Gionis
Aalto University, Finland

Jiawei Han
University of Illinois at Urbana–Champaign, USA

Jingrui He
Arizona State University, USA

Shuiwang Ji
Old Dominion University, USA

George Karypis
University of Minnesota, USA

Eammon Keogh
University of California, Riverside, USA

Laks Lakshmanan
University of British Columbia, Canada

Jure Leskovec
Stanford Unviersity, USA

Zhoujun Li
Beihang University, China

Huan Liu
Arizona State University, USA

Jun Liu
SAS Institute Inc., USA

Jennifer Neville
Purdue University, USA

Haesun Park
Georgia Institute of Technology, USA

Jian Pei
Simon Fraser University, Canada

Balaraman Ravindran
Indian Institute of Technology Chennai, India

Tamás Sarlós
Google, USA

Myra Spiliopoulou
Otto-von-Guericke-University Magdeburg, Germany

Evimaria Terzi
Boston University, USA

Hanghang Tong
Arizona State University, USA

Sergei Vassilvitskii
Google, USA

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

Wei Wang
University of California, Los Angeles, USA

Hui Xiong
Rutgers University, USA

Philip Yu
University of Illinois at Chicago, USA

Program Committee

Bruno Abrahao
Cornell University, USA

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

Leman Akoglu
SUNY Stony Brook, USA

Tony Bagnall
University of East Anglia, United Kingdom

Kanishka Bhaduri
Netflix, USA

Smriti Bhagat
Technicolor Labs, USA

Albert Bifet
Huawei Noah's Ark Lab, Hong Kong

Selçuk Candan
Arizona State University, USA

Huiping Cao
New Mexico State University, USA

Nan Cao
IBM T. J. Watson Research Center, USA

Shiyu Chang
University of Illinois at Urbana-Champaign, USA

Xiangyu Chang
Xi'an Jiaotong University, China

Polo Chau
Georgia Institute of Technology, USA

Sanjay Chawla
University of Sydney, Australia

Nitesh Chawla
University of Notre Dame, USA

Jianhui Chen
GE Global Research,USA

Wei Chen
Microsoft Research Asia

Mingsyan Chen
National Taiwan University, Taiwan

Jian Cheng
Chinese Academy of Science, China

Alfredo Cuzzocrea
Italian National Research Council, Italy

Xuan-Hong Dang
University of California Santa Barbara, USA

Kamalika Das
NASA/University Affliated Research Center, USA

Mahashweta Das
HP Labs, USA

Hasan Davulcu
Arizona State University, USA

Antonios Deligiannakis
Technical University of Crete, Greece

Wei Ding
University of Massachusetts Boston, USA

Haimonti Dutta
Columbia University, USA

Saso Dzeroski
Jozef Stefan Institute, Slovenia

William Eberle
Tennessee Technological University, USA

Dora Erdos
Boston University, USA

Chunsheng Victor Fang
EMC / Pivotal, USA

Sorelle Friedler
Haverford College, USA

Esther Galbrun
Boston University, USA

Brian Gallagher
Lawrence Livermore National Laboratory, USA

Jing Gao
University of Buffalo, USA

Jun Gao
Peking University, China

Yong Ge
University of North Carolina at Charlotte, USA

Amol Ghoting
American Express, USA

Joachim Giesen
Friedrich-Schiller-Universität Jena, Germany

Manuel Gomez-Rodriguez
Max Planck Institute for Intelligent Systems, Germany

Pinghua Gong
Arizona State University, USA

Ziyu Guan
Northwest University, USA

Jie Gui
Chinese Academy of Sciences, China

Francesco Gullo
Yahoo! Research, Spain

Stephan Günnemann
Carnegie Mellon University, USA

Manish Gupta
Microsoft Research, India

Yuan Hao
Google, USA

Larry Holder
Washington State University, USA

Xia Hu
Arizona State University, USA

Xiaohua Hu
Drexel University, USA

Bing Hu
Samsung Research, USA

Jun (Luke) Huan
University of Kansas, USA

Akihiro Inokuchi
Kwansei Gakuin University, USA

Ruoming Jin
Kent State University, USA

Chandrika Kamath,
Lawrence Livermore National Laboratory, USA

Purushottam Kar,
Microsoft Research, India

Jeremy Kepner
Massachusetts Institute of Technology, USA

Daniel Keren
Haifa University, Israel

Kristian Kersting
University of Bonn, Germany

Xiangnan Kong
Worcester Polytechnic Institute, USA

Georg Krempl
Otto-von-Guericke Universität Magdeburg, Germany

Vincent Lemaire
Université Pierre et Marie Curie - Paris VI, France

Jing Li
Arizona State University, USA

Zhenguo Li
Huawei Noah's Ark Lab, Hong Kong

Lei Li
Florida International University, USA

Jiuyong Li
University of South Australia, Australia

Qin Lijing
Tsinghua University, China

Shou-de Lin
National Taiwan University, Taiwan

Binbin Lin
Arizona State University, USA

Jessica Lin
George Mason University, USA

Qi Liu
University of Science and Technology of China, China

Ji Liu
University of Rochester, USA

Wei Liu
University of Technology, Australia

Shuai Ma
Beihang University, China

Arun Maiya
Institute for Defense Analysis, USA

Julian McAuley
University of California, San Diego

Rosa Meo
University of Torino, Italy

Pauli Miettinen
Max-Planck-Institute for Informatics, Germany

Benjamin Miller
Massachusetts Institute of Technology Lincoln Laboratory, USA

Abdullah Mueen
University of New Mexico, USA

Emmanuel Müller
Karlsruhe Institute of Technology (KIT), Germany

Sriraam Natarajan
Indiana University, USA

Vinh Nguyen
University of Melbourne, Australia

Xia Ning
Indiana University - Purdue University Indianapolis, USA

Nikunj Oza
NASA Ames Research Center, USA

Rasmus Pagh
IT University of Copenhagen, Denmark

Spiros Papadimitriou
Rutgers University, USA

Laurence Park
University of Western Sydney, Australia

B. Aditya Prakash
Virginia Tech, USA

Buyue Qian
IBM T. J. Watson Research Center, USA

Zhiwei Qin
Walmart Labs, USA

Xiaojun Quan
Institute for Infocomm Research, Singapore

Naren Ramakrishnan
Virginia Tech, USA

Parasaran Raman
Yahoo!, USA

Huzefa Rangwala
George Mason University, USA

Sayan Ranu
Indian Institute of Technology Madras, India

Chandan Reddy
Wayne State University, USA

Dafna Shahaf
Stanford University, USA

Yangqiu Song
University of Illinois at Urbana-Champaign, USA

Sucheta Soundarajan
Rutgers University, USA

Jerzy Stefanowski
Poznan University of Technology, Poland

Karthik Subbian
University of Minnesota, USA

Masashi Sugiyama
Tokyo Institute of Technology, Japan

Yizhou Sun
Northeastern University, USA

Jimeng Sun
Georgia Institute of Technology, USA

Andrea Tagarelli
University of Calabria, Italy

Hana Tai
National Taipei University, Taiwan

Pang-Ning Tan
Michigan State University, USA

Jie Tang
Tsinghua University, China

Dacheng Tao
University of Technology Sydney, Australia

Nikolaj Tatti
Aalto University, Finland

Justin Thaler
Yahoo! Labs, USA

Panayiotis Tsaparas
University of Ioannina, Greece

Grigorios Tsoumakas
Aristotle University of Thessaloniki, Greece

Johan Ugander
University of Washington and Microsoft Research, USA

Antti Ukkonen
Helsinki Institute for Information Technology, Finland

Matthijs van Leeuwen
Katholieke Universiteit Leuven, Belgium

Jilles Vreeken
Max Planck Institute for Informatics, Germany

Jianyong Wang
Tsinghua University, China

Jie Wang
Arizona State University, USA

Ke Wang
Simon Fraser University, Canada

Ting Wang
IBM T. J. Watson Research Center, USA

Wendy(Hui) Wang
Stevens Institute of Technology, USA

Xiang Wang
IBM T. J. Watson Research Center, USA

Takashi Washio
Osaka University, Japan

Tim Weninger
University of Notre Dame, USA

Raymond Wong
Hong Kong University of Science and Technology, China

Teresa Wu
Arizona State University, USA

Xintao Wu
University of Arkansas, USA

Shuo Xiang
Arizona State University, USA

Keli Xiao
Stony Brook University, USA

Jaewon Yang
LinkedIn, USA

Sen Yang
Arizona State University, USA

Hongxia Yang
IBM T.J. Watson Research Center, USA

Dawei Yin
Yahoo! Labs, USA

Lei Yu
Binghamton University, USA

Guoxian Yu
Southwest University, China

Jing Yuan
Microsoft Research Asia

Nicholas Yuan
Microsoft Research Asia

Ruiwen Zhang
SAS Institute, USA

Xiang Zhang
Case Western Reserve University, USA

Xiaoming Zhang
Beihang University, China

Kun Zhang
Xavier University of Louisiana, USA

Zheng Zhao
SAS Institute, USA

Jiayu Zhou
Samsung Research, USA

Wenjun Zhou
The University of Tennessee Knoxville, USA

Zhi-Hua Zhou
Nanjing University, China

Xingquan Zhu
Florida Atlantic University, USA

Zhenfeng Zhu
Beijing Jiaotong University, China

Shiai Zhu
University of Ottawa, Canada

Hengshu Zhu
Baidu.com, China

Arthur Zimek
Ludwig-Maximilians-Universität München, Germany

SIAM Registration Desk

The SIAM registration desk is located in the Port of Vancouver – 2nd Level. It is open during the following hours:

Wednesday, April 29
5:00 PM – 7:00 PM

Thursday, April 30
7:00 AM – 7:30 PM

Friday, May 1
7:30 AM – 3:30 PM

Saturday, May 2
7:30 AM – 4:00 PM

Hotel Address

Pinnacle Vancouver Harbourfront Hotel
1133 West Hastings Street
Vancouver, British Columbia,
Canada V6E3T3
Direct Telephone: +1-604-689-9211
Toll Free (Hotel Direct):
+1-844-337-3118
Hotel Fax: +1-604-691-2731
http://www.pinnacleharbourfronthotel.com/

Hotel Telephone Number

To reach an attendee or leave a message, call +1-604-689-9211.

If the attendee is a hotel guest, the hotel operator can connect you with the attendee's room.

Hotel Check-in and Check-out Times

Check-in time is 3:00 PM.

Check-out time is 12:00 PM.

Child Care

For local child care information, the hotel concierge recommends Nannies on Call. They can be reached online at *www.nanniesoncall.com*, or by telephone at

604-734-1776 or 877-214-2828.

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

Aramco Services Company

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 Almos National Laboratory

Mathematical Sciences Research Institute

Max-Planck-Institute for Dynamics of Complex Technical Systems

Mentor Graphics

The MITRE Corporation

National Institute of Standards and Technology (NIST)

National Security Agency (DIRNSA)

Naval Surface Warfare Center, Dahlgren Division

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 2015

Funding Agency

Leading the applied mathematics community . . . Join SIAM and save!

SIAM members save up to $130 on full registration for the 2015 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*, *SIAM News* and *SIAM Unwrapped*, 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. Lists are available at registration or online.

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 two (2) screens, one (1) data projector and one (1) overhead projector. Cables or adaptors for Apple computers are not supplied, as they vary for each model. Please bring your own cable/adaptor if using an Apple 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 an Apple computer. Overhead projectors will be provided only if 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.

Internet Access

The Pinnacle Vancouver Harbourfront Hotel offers wireless Internet access to guestrooms in the SIAM block and public areas of the hotel.

Complimentary wireless Internet access in the meeting space is also available to SIAM attendees.

In addition, a limited number of computers with Internet access will be available during registration hours.

Registration Fee Includes

- Admission to all technical sessions
- Admission to all tutorial sessions
- Admission to workshops
- Business Meeting (open to SIAG/DMA members)
- Coffee breaks daily
- Continental breakfast daily
- Room set-ups and audio/visual equipment
- USB of conference proceedings, workshop and tutorial notes
- Welcome Reception and Poster Session
- Doctoral Forum and Student Posters

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, April 30 at 7:00 PM. Poster presenters are requested to set up their poster material on the provided 4' x 8' poster boards in the Tuscany Room no later than 7:00 PM on Thursday, the official start time of the session. Boards and push pins will be available to presenters beginning Thursday, April 30 at 7:00 AM. For information about preparing a poster, please visit
http://www.siam.org/meetings/guidelines/presenters.php.

Table Top Displays

SIAM
Springer

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!

Conference Sponsors

The SIAM Data Mining Conference thanks the National Science Foundation for its generous support towards student travel awards to attend the conference and participate in the doctoral forum.

Best Paper Award and Doctoral Forum Sponsor

SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Vice President, Academic

FACULTY OF
APPLIED SCIENCES

Silver Sponsor

LIVING ANALYTICS
RESEARCH CENTRE

ExxonMobil
Research and Engineering

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 lunch time on Saturday, May 2.

Comments?

Comments about SIAM meetings are encouraged! Please send to:
Cynthia Phillips, SIAM Vice President for Programs (*vpp@siam.org*).

Get-togethers

- Welcome Reception and Poster Session
  Thursday, April 30
  7:00 PM – 9:00 PM

- Business Meeting
  (*open to SIAG/DMA members*)
  Friday, May 1
  6:30 PM – 7:00 PM
  Complimentary beer and wine will be served.

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 #SIAMSDM15.

Invited Plenary Speakers

*** All Invited Plenary Presentations will take place in Pinnacle Harbourfront Ballroom II & III - 2nd Level***

Thursday, April 30

8:15 AM - 9:30 AM
IP1 Efficient Personalization Algorithms for Social Networks
Ashish Goel, *Stanford University, USA*

1:30 PM - 2:45 PM
IP2 Analysis of Cancer Genomes to Aid in Therapeutic Choice
Steven Jones, *BC Cancer Agency,*
Simon Fraser University and University of British Columbia, Canada

Friday, May 1

8:15 AM - 9:30 AM
IP3 Information at Your Fingertips: Only a Dream
for Enterprises?
Surajit Chaudhuri, *Microsoft Research, USA*

1:30 PM - 2:45 PM
IP4 The Multi-facets of a Data Science Project to Answer:
How are Organs Formed?
Bin Yu, *University of California, Berkeley, USA*

Tutorials

Thursday, April 30

10:00 AM - 12:05 PM
TS1: Tutorial Session: Visual Text Analytics
Port of San Francisco – 3rd Level

3:00 PM - 6:15 PM
TS2: Tutorial Session: Finding Repeated Structure in Time Series:
Algorithms and Applications
Port of San Francisco – 3rd Level

Friday, May 1

10:00 AM - 12:05 PM
TS3: Tutorial Session: Methods and Applications of Net*work Sampling*
Port of San Francisco – 3rd Level

Saturday, May 2

8:30 AM - 12:00 PM
TS4: Tutorial Session: Information Theory in Data Mining
Port of Singapore – 3rd Level

SIAM Activity Group on Data Mining and Analytics(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.

2015 SIAM International Conference on **DATA MINING**

April 30-May 2, 2015

**Pinnacle Vancouver Harbourfront Hotel
Vancouver, British Columbia, Canada**

ACTIVITIES INCLUDE:

*   Special sessions at SIAM Annual Meetings
*   Annual conference
*   Website

BENEFITS OF SIAG/DMA MEMBERSHIP:

*   Listing in the SIAG's online membership directory
*   Additional $10 discount on registration at SIAM International Conference on Data Mining (excludes student)
*   Electronic communications from your peers 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.
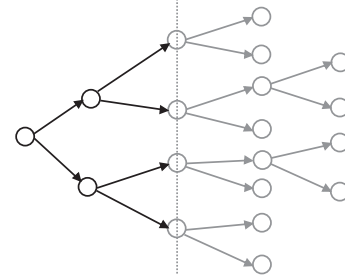
COST:

* $10 per year
*   Student members can join two activity groups for free!

2014-15 SIAG/DMA OFFICERS

*   Chair: Zoran Obradovic, Temple University
*   Vice Chair: Jeremy Kepner, Massachusetts Institute of Technology
*   Program Director: Takashi Washio, Osaka University
*   Secretary: Kirk Borne, George Mason University

TO JOIN:

SIAG/DMA: my.siam.org/forms/join_siag.htm
SIAM: www.siam.org/joinsiam

Workshops and Minisymposium

Friday, May 1

Minisymposium
3:00 PM – 5:05 PM

MS1: BigBrain: Mini-Symposium on Big Data for Brain Science
Port of San Francisco - 3rd Level

Saturday, May 2

Workshops
8:30 AM - 5:00 PM

Workshop 1: Machine Learning Methods for Recommender Systems
Port of New York - 3rd Level

Workshop 2: 4th Workshop on Data Mining for Medicine and Healthcare
Pinnacle Harbourfront Ballroom III - 2nd Level

1:30 PM - 5:00 PM

Workshop 3: Adaptive Learning On-a-chip: Hardware and Algorithms
Pinnacle Harbourfront Ballroom I - 2nd Level

Workshop 4: Mining Networks and Graphs: A Big Data Analytic Challenge
Pinnacle Harbourfront Ballroom II - 2nd Level

Workshop 5: Big Data and Stream Analytics
Port of Singapore - 3rd Level

Workshop 6: Heterogeneous Learning
Port of San Francisco - 3rd Level

For detailed schedules of each workshop, visit the conference workshop page http://www.siam.org/meetings/sdm15/workshops.php .

12 Reasons Your Library Needs SIAM E-Books

1 SIAM books are recognized everywhere as a leading source of advanced knowledge in applied mathematics and computational science, the gold standard for mathematics publishing.

2 SIAM's Atypon® Literatum platform serves subscribers with secure reliability and unlimited simultaneous users.

3 SIAM e-books and journals share the same platform, providing quick and easy cross-referencing between articles and book content.

4 SIAM offers two payment options—perpetual access and annual subscription—so you can budget according to your needs.

5 E-book files are in PDF format with no DRM and are widely compatible with most devices.

6 Books are viewable by chapter, searchable by keyword, and stored in manageable file sizes.

7 New titles are available immediately upon publication.

8 Wide-ranging content in 13 different series and dozens of subject categories.

9 Comprehensive MARC files for every title, assuring easy handling and cataloging.

10 Substantial discounts for SIAM academic and corporate members.

11 A primary resource for nearly all technical researchers, now offered at more than 100 institutions worldwide.

12 Access to SIAM e-books also means access to SIAM's individualized customer service, with secure billing and attention to your individual needs.

siam e-books

Visit the SIAM e-books website at www.siam.org/ebooks, featuring:

• MARC records • Sample licensing agreements • Link to the e-book platform • FAQ about SIAM e-books
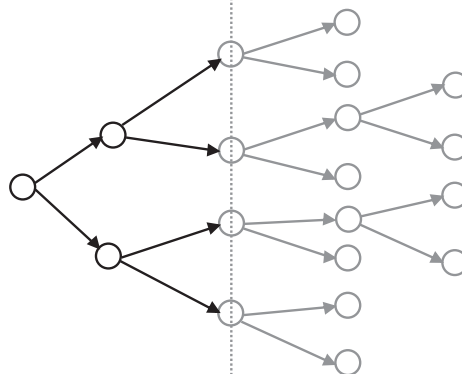• SIAM's newest e-book titles

siam

To buy, subscribe, or for more information contact:
Joanne Cassetti, SIAM, Phone: +1-215-382-9800 x-321 or 800-447-SIAM (7426) / ebooks@siam.org / www.siam.org/ebooks

SDM15 Program

2015 SIAM International
Conference on **DATA MINING**

April 30-May 2, 2015

Pinnacle Vancouver Harbourfront Hotel
Vancouver, British Columbia, Canada

Wednesday, April 29

Registration
5:00 PM-7:00 PM

Room:Port of Vancouver - 2nd Level

Thursday, April 30

Registration
7:00 AM-7:30 PM

Room:Port of Vancouver - 2nd Level

Continental Breakfast
7:30 AM-8:00 AM

Room:Ballroom Foyer - 2nd Level

Announcements
8:00 AM-8:15 AM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Thursday, April 30

IP1
Efficient Personalization Algorithms for Social Networks
8:15 AM-9:30 AM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Chair: Suresh Venkatasubramanian, University of Utah, USA

We will present efficient algorithms for many social search and recommendation problems. One particular emphasis will be algorithms that can be easily implemented on distributed platforms such as MapReduce/Hadoop, Spark, and Storm. The algorithms covered will
include incremental and personalized Page Rank, Cosine Similarity, and searching for the closest instance of a word in a social network. We will motivate these problems with
concrete examples of deployed systems.

Ashish Goel
Stanford University, USA

Coffee Break
9:30 AM-10:00 AM

Room:Ballroom Foyer - 2nd Level

Thursday, April 30

CP1

Networks, Graphs I

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

Chair: Hanghang Tong, Arizona State University, USA

10:00-10:20 Functional Node Detection on Linked Data

Kang Li, Jing Gao, Suxin Guo, Nan Du, and *Aidong Zhang*, State University of New York at Buffalo, USA

10:25-10:45 Where Graph Topology Matters: The Robust Subgraph Problem

Leman Akoglu, *Hau Chan*, and Shuchu Han, Stony Brook University, USA

10:50-11:10 Same Bang, Fewer Bucks: Efficient Discovery of the Cost-Influence Skyline

Antti Ukkonen, Finnish Institute of Occupational Health, Finland ; Matthijs Van Leeuwen, Katholieke Universiteit Leuven, Belgium

11:15-11:35 Selecting Shortcuts for a Smaller World

Nikos Parotsidis, Evaggelia Pitoura, and Panayiotis Tsaparas, University of Ioannina, Greece

11:40-12:00 Significant Subgraph Mining with Multiple Testing Correction

Mahito Sugiyama, Osaka University, Japan; Felipe Llinares Lopez, ETH Zürich, Switzerland; Niklas Kasenburg, University of Copenhagen, Denmark; Karsten Borgwardt, ETH Zürich, Switzerland

Thursday, April 30

CP2

Metric Learning, Feature Selection/Extraction

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

Chair: Bo Jin, Dalian University of Technology, China

10:00-10:20 From Categorical to Numerical: Multiple Transitive Distance Learning and Embedding

Kai Zhang, NEC Laboratories America, USA

10:25-10:45 Spectral Embedding of Signed Networks

David Skillicorn and Quan Zheng, Queen's University, Canada

10:50-11:10 An Lle Based Heterogeneous Metric Learning for Cross-Media Retrieval

Yi-Dong Shen, *Peng Zhou*, and Liang Du, Chinese Academy of Sciences, China; Mingyu Fan, Wenzhou University, China

11:15-11:35 Feature Selection for Nonlinear Regression and Its Application to Cancer Research

Yijun Sun, State University of New York at Buffalo, USA

11:40-12:00 Efficient Partial Order Preserving Unsupervised Feature Selection on Networks

Xiaokai Wei, Sihong Xie, and Philip S. Yu, University of Ilinois at Chicago, USA

Thursday, April 30

CP3

Clustering

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom III - 2nd Level

Chair: Xia Ning, University of Minnesota, Twin Cities, USA

10:00-10:20 NetCodec: Community Detection from Individual Activities

Long Q. Tran, University of Engineering and Technology, Pakistan and Vietnam National University at Hanoi, Vietnam; *Mehrdad Farajtabar*, Le Song, and Honguyan Zha, Georgia Institute of Technology, USA

10:25-10:45 Efficient Algorithms for a Robust Modularity-Driven Clustering of Attributed Graphs

Patricia Iglesias Sanchez, *Emmanuel Müller*, Uwe Leo Korn, Klemens Böhm, Andrea Kappes, Tanja Hartmann, and Dorothea Wagner, Karlsruhe Institute of Technology, Germany

10:50-11:10 Vertex Clustering of Augmented Graph Streams

Ryan Mcconville, Weiru Liu, and Paul Miller, Queen's University, Belfast, United Kingdom

11:15-11:35 Tensor Spectral Clustering for Partitioning Higher-Order Network Structures

Austin Benson, Stanford University, USA; David F. Gleich, Purdue University, USA; Jure Leskovec, Stanford University, USA

11:40-12:00 Community Detection for Emerging Networks

Jiawei Zhang and Philip S. Yu, University of Ilinois at Chicago, USA

Thursday, April 30

TS1

Tutorial Session: Visual Text Analytics

10:00 AM-12:05 PM

Room:Port of San Francisco - 3rd Level

Chair: Tao Li, Florida International University, USA

Text mining methods have been extensively used in the last 30 years. However, in scenarios where the path from data to decisions is unclear, or where different users may be interested in different solutions, the involvement of the user or analyst in the text mining process becomes crucial. Visual Text Analytics aims at addressing these problems by incorporating concepts from Visual Analytics to text mining and natural language processing techniques. In this tutorial we will introduce Visual Text Analytics as a multi-disciplinary field of research. We will cover conceptual and practical tools, review state-of-the-art systems as well as identify interesting research avenues in the area. The course will be of benefit to any participant interested in improving the impact of their text mining methods.

Axel Soto
Dalhousie University, Canada

Evangelos Milios
Dalhousie University, Canada

Lunch Break

12:05 PM-1:30 PM

Attendees on their own

Thursday, April 30

IP2

Analysis of Cancer Genomes to Aid in Therapeutic Choice

1:30 PM-2:45 PM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Chair: Ke Wang, Simon Fraser University, Canada

We are conducting an approach to conduct complete genomic and transcriptomic analysis of patient tumours to aid in clinical decision making. Determining the genetic changes that have accrued in patient tumours will provide an understanding of the molecular biology underlying their oncogenesis, the biological drivers and potential drug sensitivities. Using this information for therapeutic decision-making could result in more effective cancer treatment. Overall, the generation of genomic data for cancer therapy has great potential for the incorporation of machine learning methods and the development of algorithms that can perform analyses in a clinically relevant timeframe.

Steven Jones
BC Cancer Agency, Simon Fraser University and University of British Columbia, Canada

Coffee Break

2:45 PM-3:00 PM

Room:Ballroom Foyer - 2nd Level

Thursday, April 30

CP4

Applications

3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

Chair: Jing Gao, State University of New York at Buffalo, USA

3:00-3:20 Labeling Educational Content with Academic Learning Standards

Danish Contractor, Kashyap Popat, Shajith Ikbal, Sumit Negi, Bikram Sengupta, and Mukesh Mohania, IBM Research, India

3:25-3:45 Data Mining for Real Mining: A Robust Algorithm for Prospectivity Mapping with Uncertainties

Justin Granek and Eldad Haber, University of British Columbia, Canada

3:50-4:10 Product Adoption Rate Prediction: A Multi-Factor View

Le Wu, University of Science and Technology of China, China

4:15-4:35 PatentCom: A Comparative View of Patent Document Retrieval

Longhui Zhang, Lei Li, Chao Shen, and Tao Li, Florida International University, USA

4:40-5:00 Combating Product Review Spam Campaigns via Multiple Heterogeneous Pairwise Features

Chang Xu and Jie Zhang, Nanyang Technological University, Singapore

Thursday, April 30

CP5

Recommendation, Classification

3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

Chair: George Karypis, University of Minnesota, USA

3:00-3:20 A Bayesian Framework for Modeling Human Evaluations

Himabindu Lakkaraju and Jure Leskovec, Stanford University, USA; Jon M. Kleinberg, Cornell University, USA; Sendhil Mullainathan, Harvard University, USA

3:25-3:45 Feature-Based Factorized Bilinear Similarity Model for Cold-Start Top-N Item Recommendation

Mohit Sharma, University of Minnesota, USA; Jiayu Zhou and Junling Hu, Samsung Research America, USA; George Karypis, University of Minnesota, USA

3:50-4:10 Cross-Modal Retrieval: A Pairwise Classification Approach

Aditya K. Menon, The Australian National University, Australia; Didi Surian and Sanjay Chawla, University of Sydney, Australia

4:15-4:35 Binary Classifier Calibration Using a Bayesian Non-Parametric Approach

Mahdi Pakdaman Naeini, Gregory Cooper, and Milos Hauskrecht, University of Pittsburgh, USA

4:40-5:00 Semi-Supervised Learning for Structured Regression on Partially Observed Attributed Graphs

Jelena Z. Stojanovic, Temple University, USA; Milos Jovanovic, University of Belgrade, Serbia; Djordje Gligorijevic and Zoran Obradovic, Temple University, USA

Thursday, April 30

CP6

Security/Privacy, Social Media

3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom III- 2nd Level

Chair: B. Aditya Prakash, Virginia Tech, USA

3:00-3:20 Health Insurance Market Risk Assessment: Covariate Shift and K-Anonymity

Dennis Wei, *Karthikeyan Natesan Ramamurthy*, and Kush R. Varshney, IBM T.J. Watson Research Center, USA

3:25-3:45 Attacking Dbscan for Fun and Profit

Jonathan Crussell and Philip Kegelmeyer, Sandia National Laboratories, USA

3:50-4:10 Result Integrity Verification of Outsourced Privacy-Preserving Frequent Itemset Mining

Wendy Hui Wang and Ruilin Liu, Stevens Institute of Technology, USA

4:15-4:35 Modeling Users' Adoption Behaviors with Social Selection and Influence

Ziqi Liu, Xi'an Jiaotong University, P.R. China; Fei Wang, University of Connecticut, USA; Qinghua Zheng, Xi'an Jiaotong University, P.R. China

4:40-5:00 Exploring the Impact of Dynamic Mutual Influence on Social Event Participation

Tong Xu, University of Science and Technology of China, China; Hao Zhong, Rutgers University, USA; Hengshu Zhu, Baidu, USA; Hui Xiong, Rutgers University, USA; Enhong Chen, University of Science and Technology of China, China; Guannan Liu, Tsinghua University, P. R. China

Thursday, April 30

TS2

Tutorial Session: Finding Repeated Structure in Time Series: Algorithms and Applications

3:00 PM-6:15 PM

Room:Port of San Francisco - 3rd Level

Chair: Tao Li, Florida International University, USA

Repeated patterns in time series data are indicative to identical dynamics in the origin. Such patterns can be used to summarize, classify, compress, cluster and classify time series data. In this tutorial, we will present several algorithms for repeated pattern discovery in univariate and multivariate time series data. The algorithms cover a wide range of settings from in-memory to online data, from approximate to exact algorithms and, from one length to all lengths. We will present applications of repeated patterns in several domains and in various data types. We will cover Exact and approximate algorithms for finding repeated patterns in time series Clustering, classification and rule discovery algorithms using repeated patterns Applications to Entomology, Data center management, Activity recognition The tutorial will conclude with a list of open problems and research directions.

Abdullah Mueen

University of New Mexico, USA

Eamonn Keogh

University of California, Riverside, USA

Organizational Break

5:05 PM-5:15 PM

Poster Spotlights

5:15 PM-6:45 PM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Thursday, April 30

PP1

Welcome Reception and Poster Session

7:00 PM-9:00 PM

Room:Tuscany - Lobby Level

Towards Classification of Social Streams

Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Min-Hsuan Last Tsai, Google, Inc., USA; Thomas Huang, University of Illinois at Urbana-Champaign, USA

Mobile App Security Risk Assessment: A Crowdsourcing Ranking Approach from User Comments

Lei Cen, Purdue University, USA; Deguang Kong, University of Texas at Arlington, USA; Hongxia Jin, Samsung Research America, USA; Luo Si, Purdue University, USA

Learning Stroke Treatment Progression Models for An MDP Clinical Decision Support System

Dan C. Coroian, Indiana University, USA; Kris Hauser, Duke University, USA

OnlineCM: Real-Time Consensus Classification with Missing Values

Bowen Dong and Sihong Xie, University of Ilinois at Chicago, USA; Jing Gao, State University of New York at Buffalo, USA; Wei Fan, IBM T.J. Watson Research Center, USA; Philip Yu, University of Ilinois at Chicago, USA

What Shall I Share and with Whom? - A Multi-Task Learning Formulation using Multi-Faceted Task Relationships

Sunil K. Gupta, Deakin University, Australia

A Generalized Mixture Framework for Multi-Label Classification

Charmgil Hong, University of Pittsburgh, USA; Iyad Batal, GE Global Research, USA; Milos Hauskrecht, University of Pittsburgh, USA

Domain-Knowledge Driven Cognitive Degradation Modeling for Alzheimer's Disease

Shuai Huang, University of Washington, USA

Optimizing Hashing Functions for Similarity Indexing in Arbitrary Metric and Nonmetric Spaces

Pat Jangyodsuk, University of Texas at Arlington, USA; Panagiotis Papapetrou, Birkbeck, University of London, United Kingdom; *Vassilis Athitsos*, University of Texas at Arlington, USA

Ensemble Learning Methods for Binary Classification with Multi-Modality Within the Classes

Anuj Karpatne, University of Minnesota, USA; Ankush Khandelwal, University of Minnesota, Twin Cities, USA; Vipin Kumar, University of Minnesota, USA

A Framework for Simplifying Trip Data into Networks Via Coupled Matrix Factorization

Chia-Tung Kuo, University of California, Davis, USA; James Bailey, The University of Melbourne, Australia; Ian Davidson, University of California, Davis, USA

MET: A Fast Algorithm for Minimizing Propagation in Large Graphs with Small Eigen-Gaps

Long Le and Tina Eliassi-Rad, Rutgers University, USA; Hanghang Tong, Arizona State University, USA

Learning Compressive Sensing Models for Big Spatio-Temporal Data

Dongeun Lee and Jaesik Choi, Ulsan National Institute of Science and Technology, South Korea

Multi-View Low-Rank Analysis for Outlier Detection

Sheng Li, Ming Shao, and Yun Fu, Northeastern University, USA

Reafum: Representative Approximate Frequent Subgraph Mining

Ruirui Li and Wei Wang, University of California, Los Angeles, USA

Dias: A Disassemble-Assemble Framework for Highly Sparse Text Clustering

Hongfu Liu, Northeastern University, USA; Junjie Wu, BeiHang University, China; Dacheng Tao, University of Technology, Sydney, Australia; Yuchao Zhang, Beijing Institute of System Engineering, China; Yun Fu, Northeastern University, USA

Optimal Event Sequence Sanitization

Grigorios Loukides, Cardiff University, United Kingdom; Robert Gwadera, École Polytechnique Fédérale de Lausanne, Switzerland

Predicting Neighbor Distribution in Heterogeneous Information Networks

Yuchi Ma, Ning Yang, Chuan Li, and Lei Zhang, Sichuan University, China; Philip S. Yu, University of Ilinois at Chicago, USA

Temporally Coherent CRP: A Bayesian Non-Parametric Approach for Clustering Tracklets with Applications to Person Discovery in Videos

Adway Mitra, Soma Biswas, and Chiranjib Bhattacharyya, Indian Institute of Science, Bangalore, India

Correlating Surgical Vital Sign Quality with 30-Day Outcomes Using Regression on Time Series Segment Features

Risa Myers, Rice University, USA; John Frenzel and Joseph Ruiz, University of Texas MD Anderson Cancer Center, USA; Christopher Jermaine, Rice University, USA

Multi-Layered Framework for Modeling Relationships Between Biased Objects

Iku Ohama, Panasonic Co., Ltd., USA; Takuya Kida and Hiroki Arimura, Hokkaido University, Japan

Speclda: Modeling Product Reviews and Specifications to Generate Augmented Specifications

Dae Hoon Park and ChengXiang Zhai, University of Illinois at Urbana-Champaign, USA; Lifan Guo, TCL Research America, USA

Mining Multi-Relational Gradual Patterns

Nhathai Phan, University of Oregon, USA

Modeling User Arguments, Interactions, and Attributes for Stance Prediction in Online Debate Forums

Minghui Qiu, Singapore Management University, Singapore; Yanchuan Sim and Noah Smith, Carnegie Mellon University, USA; Jing Jiang, Singapore Management University, Singapore

Predicting Preference Tags to Improve Item Recommendation

Huzefa Rangwala, Tanwistha Saha, and Carlotta Domeniconi, George Mason University, USA

Data Stream Classification Guided by Clustering on Nonstationary Environments and Extreme Verification Latency

Vinicius Souza and Diego Silva, University of Sao Paulo, Brazil; Joao Gama, University of Porto, Portugal; Gustavo Batista, University of Sao Paulo, Brazil

Mining Block I/O Traces for Cache Preloading with Sparse Temporal Non-Parametric Mixture of Multivariate Poisson

Lavanya S. Tekumalla and Chiranjib Bhattacharyya, Indian Institute of Science, Bangalore, India

Taming the Empirical Hubness Risk in Many Dimensions

Nenad Tomašev, Jozef Stefan Institute, Slovenia

Scalable Clustering of Time Series with U-Shapelets

Liudmila Ulanova, University of California, Riverside, USA

Causal Inference by Direction of Information

Jilles Vreeken, Max Planck Institute for Informatics, Germany and Saarland University, Germany

Graph Regularized Meta-path Based Transductive Regression in Heterogeneous Information Network

Mengting Wan and Yunbo Ouyang, University of Illinois at Urbana-Champaign, USA; Lance Kaplan, U.S. Army Research Laboratory, USA; Jiawei Han, University of Illinois at Urbana-Champaign, USA

Localizing Temporal Anomalies in Large Evolving Graphs

Teng Wang, University of California, Davis, USA; *Chunsheng Fang* and Derek Lin, Pivotal Software, Inc., USA; S. Felix Wu, University of California, Davis, USA

Non-Exhaustive, Overlapping k-Means

Joyce J. Whang, and Inderjit S. Dhillon, University of Texas at Austin, USA; David F. Gleich, Purdue University, USA

Festival, Date and Limit Line: Predicting Vehicle Accident Rate in Beijing

Xinyu Wu, University of Chinese Academy of Sciences, China; Ping Luo and Qing He, Chinese Academy of Sciences, China; Tianshu Feng, University of Science and Technology of China, China; Fuzhen Zhuang, Chinese Academy of Sciences, China

A Multi-Label Least-Squares Hashing For Scalable Image Search

Xin-Shun Xu and Shengsheng Wang, Shandong University, China; Zi Huang, University of Queensland, Australia

Simpleppt: A Simple Principal Tree Algorithm

Qi Mao and *Le Yang*, State University of New York at Buffalo, USA; Li Wang, Brown University, USA; Steve Goodison, Mayo Clinic, USA; Yijun Sun, State University of New York at Buffalo, USA

Spatiotemporal Event Forecasting in Social Media

Liang Zhao, Virginia Tech, USA; Feng Chen, University of Albany - State University of New York, USA; Chang-Tien Lu and Naren Ramakrishnan, Virginia Tech, USA

Friday, May 1

Continental Breakfast
7:30 AM-8:00 AM

Room:Ballroom Foyer - 2nd Level

Registration
7:30 AM-3:30 PM

Room:Port of Vancouver - 2nd Level

Announcements
8:00 AM-8:15 AM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Friday, May 1

IP3

Information at Your Fingertips: Only a Dream for Enterprises?

8:15 AM-9:30 AM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Chair: Mohammed Zaki, Qatar Computing Research Institute, Qatar and Rensselaer Polytechnic Institute, USA

In the last decade, the information worker at enterprises has benefited from the evolution of technology in scalable data platforms, ad-hoc data analysis as well as data visualization. While such accomplishments have advanced the state of the art of enterprise analytics, much less progress has been made in the critical area of enterprise search and data discovery, especially for structured data. We will discuss why the promise of enterprise search has not been realized yet and discuss a few research opportunities to advance the state-of-the-art.

Surajit Chaudhuri is a Distinguished Scientist at Microsoft Research and leads the Data Management, Exploration and Mining group. In addition, as a Deputy Managing Director of Microsoft Research Lab at Redmond, he also has oversight of Distributed Systems, Networking, Security, Programming languages and Software Engineering groups of Microsoft Research, Redmond. He serves on the Senior Leadership Team of Microsoft's Cloud and Enterprises division. His current areas of interest are data discovery, self-manageability, and cloud database services. Working with his colleagues in Microsoft Research, he helped incorporate the Index Tuning Wizard (and subsequently Database Engine Tuning Advisor) and data cleaning technology into Microsoft SQL Server. Surajit is an ACM Fellow, a recipient of the ACM SIGMOD Edgar F. Codd Innovations Award, ACM SIGMOD Contributions Award, a VLDB 10 year Best Paper Award, and an IEEE Data Engineering Influential Paper Award. Surajit received his Ph.D. from Stanford University in 1992.

Surajit Chaudhuri

Microsoft Research, USA

Coffee Break

9:30 AM-10:00 AM

Room:Ballroom Foyer - 2nd Level

Friday, May 1

CP7

Time Series, Online Learning

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

Chair: Leman Akoglu, Stony Brook University, USA

10:00-10:20 Efficient Online Relative Comparison Kernel Learning

Eric Heim, University of Pittsburgh, USA; Matthew Berger and Lee Seversky, Air Force Research Laboratory, USA; Milos Hauskrecht, University of Pittsburgh, USA

10:25-10:45 *Cheetah* Fast Graph Kernel Tracking on Dynamic Graphs

Liangyue Li and Hanghang Tong, Arizona State University, USA; Yanghua Xiao, Fundan University, China; Wei Fan, Baidu, USA

10:50-11:10 On the Non-Trivial Generalization of Dynamic Time Warping to the Multi-Dimensional Case

Mohammad Shokoohi-Yekta, University of California, Riverside, USA; Jun Wang, University of Texas, Dallas, USA; Eamonn Keogh, University of California, Riverside, USA

11:15-11:35 Fast Mining of a Network of Coevolving Time Series

Yongjie Cai, City University of New York, USA; Hanghang Tong, Arizona State University, USA; Wei Fan, Baidu, USA; Ping Ji, City University of New York, USA

11:40-12:00 Shapelet Ensemble for Multi-Dimensional Time Series

Mustafa S. Cetin, University of New Mexico, USA

Friday, May 1

CP8

Matrix/Tensor

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

Chair: Shuiwang Ji, Arizona State University, USA

10:00-10:20 Low Rank Representation on Riemannian Manifold of Symmetric Positive Definite Matrices

Yifan Fu and Junbin Gao, Charles Sturt University, Australia; Xia Hong, University of Reading, United Kingdom; David Tien, Charles Sturt University, Australia

10:25-10:45 Getting to Know the Unknown Unknowns: Destructive-Noise Resistant Boolean Matrix Factorization

Sanjar Karaev, and Pauli Miettinen, Max-Planck Institute for Informatics, Germany; Jilles Vreeken, Max Planck Institute for Informatics, Germany and Saarland University, Germany

10:50-11:10 Convex Matrix Completion: A Trace-Ball Optimization Perspective

Guangxiang Zeng, University of Science and Technology of China, China; Ping Luo, Chinese Academy of Sciences, China; Enhong Chen, University of Science and Technology of China, China; Hui Xiong, Rutgers University, USA; Hengshu Zhu, Baidu, USA; Qi Liu, University of Science and Technology of China, China

11:15-11:35 Near-Separable Non-Negative Matrix Factorization with ℓ_1 and Bregman Loss Functions

Abhishek Kumar, IBM Research, USA; Vikas Sindhwani, Google Research, USA

11:40-12:00 Personalized TV Recommendation with Mixture Probabilistic Matrix Factorization

Huayu Li, University of North Carolina, Charlotte, USA; Hengshu Zhu, Baidu, USA; Yong Ge, University of North Carolina, Charlotte, USA; Yanjie Fu, Rutgers University, USA; Yuan Ge, Anhui Polytechnic University, China

Friday, May 1

CP9

Multi-source and Heterogeneous Learning

10:00 AM-12:05 PM

Room:Pinnacle Harbourfront Ballroom III - 2nd Level

Chair: Feng Chen, University of Albany - State University of New York, USA

10:00-10:20 Legislative Prediction with Dual Uncertainty Minimization from Heterogeneous Information

Yu Cheng and Ankit Agrawal, Northwestern University, USA; Huan Liu, Arizona State University, USA; Alok Choudhary, Northwestern University, USA

10:25-10:45 Plums: Predicting Links Using Multiple Sources

Karthik Subbian and Arindam Banerjee, University of Minnesota, USA; Sugato Basu, Google, Inc., USA

10:50-11:10 SourceSeer: Forecasting Rare Disease Outbreaks Using Multiple Data Sources

Theodoros Rekatsinas, University of Maryland, USA; Saurav Ghosh, Virginia Tech, USA; Sumiko Mekaru, Elaine Nsoesie and John Brownstein, Boston Children's Hospital, USA; Lise Getoor, University of California, Santa Cruz, USA; Naren Ramakrishnan, Virginia Tech, USA

11:15-11:35 Gin: A Clustering Model for Capturing Dual Heterogeneity in Networked Data

Jialu Liu, University of Illinois at Urbana-Champaign, USA; Chi Wang, Microsoft Research, USA; Jing Gao, State University of New York at Buffalo, USA; Quanquan Gu, University of Virginia, USA; Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Lance Kaplan, U.S. Army Research Laboratory, USA; Jiawei Han, University of Illinois at Urbana-Champaign, USA

11:40-12:00 Believe It Today Or Tomorrow? Detecting Untrustworthy Information from Dynamic Multi-Source Data

Houping Xiao, Yaliang Li, and Jing Gao, State University of New York at Buffalo, USA; Fei Wang, University of Connecticut, USA; Liang Ge, Google, Inc., USA; Wei Fan, Baidu, USA; Long Vu and Deepak Turaga, IBM T.J. Watson Research Center, USA

Friday, May 1

TS3

Tutorial Session: Methods and Applications of Network Sampling

10:00 AM-12:05 PM

Room:Port of San Francisco - 3rd Level

Chair: Tao Li, Florida International University, USA

In this tutorial, we aim to cover a diverse collection of methodologies and applications of network sampling. We will begin with a discussion of the problem setting in terms of objectives (such as, sampling a representative subgraph, sampling graphlets, etc.), population of interest (vertices, edges, motifs), and sampling methodologies (such as Metropolis-Hastings, random walk, and snowball sampling). We will then present a number of applications of these methods, and will outline both the resulting opportunities and possible biases of different methods in each application.

Mohammad Hasan

Indianana University–Purdue University, Indianapolis, USA

Nesreen Ahmed

Purdue University, Indianapolis, USA

Jennifer Neville

Purdue University, USA

Lunch Break

12:05 PM-1:30 PM

Attendees on their own

Friday, May 1

IP4

The Multi-facets of a Data Science Project to Answer: How Are Organs Formed?

1:30 PM-2:45 PM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Chair: Jieping Ye, Arizona State University, USA

Understanding local gene networks is key for advancing systems biology and developing treatments for human diseases. It requires the integration of biology, statistics and computer science (or data science) to turn into knowledge the recently available large and complex systematic spatial data.

In this talk, I present results from a data science project co-led by biologist Frise from LBNL to answer the question in the talk title.

Our team consists of Wu, Joseph, Kumbier from my group, Dr. Frise other biologists (Hommands) in Celniker's Lab at LBNL that generate the Drosophila embryonic image data of spatial expression, and Dr. Xu from Tsinghua Univ to devise a scalable open software package to manage the acquisition and computation of imaged data.

Bin Yu

University of California, Berkeley, USA

Coffee Break

2:45 PM-3:00 PM

Room:Ballroom Foyer - 2nd Level

Friday, May 1
MS1
BigBrain: Mini-Symposium on Big Data for Brain Science
3:00 PM-5:05 PM

Room:Port of San Francisco - 3rd Level

This mini-symposium will focus on exploring the forefront between data mining and brain science and inspiring fundamentally new ways of mining and knowledge discovery from a variety of brain data. The presentations and discussions will lead to novel insights and knowledge on the function and dysfunction of brain at various levels, ranging from molecular, cellular, circuitry to systems levels by mining, integrating, and interpreting large-scale, multi-modality brain data.

Organizer: Shuai Huang
University of Washington, USA

Organizer: Shuiwang Ji
Arizona State University, USA

Detecting Genetic Risk Factors for Alzheimer's Disease in Whole Genome Sequence Data via Lasso Screening
Jieping Ye, University of Michigan, USA

Neuroimage-based Diagnosis of Brain Disorders
Dinggang Shen, University of North Carolina at Chapel Hill, USA

Friday, May 1
CP10
Networks, Graphs II
3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

Chair: Rajmonda Caceres, Massachusetts Institute of Technology, USA

3:00-3:20 Rare Class Detection in Networks
Karthik Subbian, University of Minnesota, USA; Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Jaideep Srivastava, Qatar Computing Research Institute, Qatar; Vipin Kumar, University of Minnesota, USA

3:25-3:45 Hidden Hazards: Finding Missing Nodes in Large Graph Epidemics
Aditya Prakash, Virginia Tech, USA

3:50-4:10 Clustering and Ranking in Heterogeneous Information Networks via Gamma-Poisson Model
Junxiang Chen, Wei Dai, Yizhou Sun, and Jennifer Dy, Northeastern University, USA

4:15-4:35 A Devide-and-Conquer Algorithm for Betweenness Centrality
Dora Erdos, Boston University, USA; Vatche Ishakian, IBM T.J. Watson Research Center, USA; Azer Bestavros and Evimaria Terzi, Boston University, USA

4:40-5:00 Frameworks to Encode User Preferences for Inferring Topic-Sensitive Information Networks
Qingbo Hu, Sihong Xie, and Shuyang Lin, University of Ilinois at Chicago, USA; Wei Fan, Baidu, USA; Philip Yu, University of Ilinois at Chicago, USA

Friday, May 1
CP11
Optimization
3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

Chair: Fei Wang, IBM T.J. Watson Research Center, USA

3:00-3:20 Dropout Training of Matrix Factorization and Autoencoder for Link Prediction in Sparse Graphs
Shuangfei Zhai, State University of New York, Binghamton, USA

3:25-3:45 An ADMM Algorithm for Clustering Partially Observed Networks
Necdet S. Aybat, Sahar Zarmehri, and Soundar Kumara, Pennsylvania State University, USA

3:50-4:10 Scaling Log-Linear Analysis to Datasets with Thousands of Variables
Francois Petitjean and Geoffrey Webb, Monash University, Australia

4:15-4:35 A Distributed Frank-Wolfe Algorithm for Communication-Efficient Sparse Learning
Aurélien Bellet, Télécom ParisTech, France; Yingyu Liang, Princeton University, USA; Alireza Bagheri Garakani, University of Southern California, USA; Maria-Florina Balcan, Carnegie Mellon University, USA; Fei Sha, University of Southern California, USA

4:40-5:00 Exceptional Model Mining with Tree-Constrained Gradient Ascent
Ad Feelders and *Thomas Krak*, Universiteit Utrecht, The Netherlands

Friday, May 1

CP12

Multi-task/Transfer Learning

3:00 PM-5:05 PM

Room:Pinnacle Harbourfront Ballroom III - 2nd Level

Chair: Jiayu Zhou, Samsung Research America, USA

3:00-3:20 FORMULA: FactORized MUlti-task LeArning for Task Discovery in Personalized Medical Models

Jianpeng Xu, Michigan State University, USA; Jiayu Zhou, Arizona State University, USA; Pang-Ning Tan, Michigan State University, USA

3:25-3:45 Active Multi-Task Learning Via Bandits

Meng Fang and Dacheng Tao, University of Technology, Sydney, Australia

3:50-4:10 Hierarchical Active Transfer Learning

David Kale, Marjan Ghazvininejad, and Anil Ramakrishna, University of Southern California, USA; Jingrui He, Arizona State University, USA; Yan Liu, University of Southern California, USA

4:15-4:35 Learning Complex Rare Categories with Dual Heterogeneity

Pei Yang, Arizona State University, USA; Jingrui He, Stevens Institute of Technology, USA; Jia-Yu Pan, Google, Inc., USA

4:40-5:00 Faster Jobs in Distributed Data Processing Using Multi-Task Learning

Neeraja J. Yadwadkar, Bharath Hariharan, Joseph Gonzalez, and Randy Katz, University of California, Berkeley, USA

Organizational Break

5:05 PM-5:15 PM

Friday, May 1

NSF Panel:

Data Science: A Multi-Disciplinary Viewpoint and Shootout

5:15 PM-6:30 PM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Chair: Sanjay Chawla, Qatar Computing Research Institute, Qatar and University of Sydney, Australia

A panel consisting of leading players from related but distinct disciplines including Statistics, Data Mining, Machine Learning, Data Management, Computer Science theory and algorithms, will discuss the Data Science phenomenon. To keep the discussion grounded the panel members will be a presented with a concrete "problem" and will be asked to briefly present how the problem will be addressed using tools and formalisms from within their discipline. The other panel members and the audience will then get a chance to comment and start a discussion.

Panelists

Ashish Goel
Stanford University, USA

Arno Siebes
Universiteit Utrecht, The Netherlands

David Skillicorn
Queen's University, Canada

Bin Yu
University of California, Berkeley, USA

SIAG/DMA Business Meeting

6:30 PM-7:00 PM

Room:Pinnacle Harbourfront Ballroom II & III - 2nd Level

Complimentary beer and wine will be served.

Doctoral Forum and Student Posters

7:00 PM-9:00 PM

Room:Tuscany - Lobby Level

Saturday, May 2

Registration

7:30 AM-4:00 PM

Room:Port of Vancouver - 2nd Level

Continental Breakfast

8:00 AM-8:30 AM

Room:Ballroom Foyer - 2nd Level

Saturday, May 2

CP13

Text Mining, Applications - Part I of II

8:30 AM-9:20 AM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

For Part 2 see CP15

Chair: Lijing Qin, Tsinghua University, P. R. China

8:30-8:50 Selecting Social Media Responses to News: A Convex Framework Based On Data Reconstruction

Linli Xu, University of Science and Technology of China, China

8:55-9:15 Tracking Events Using Time-Dependent Hierarchical Dirichlet Tree Model

Rumeng Li, Peking University, China; Tao Wang, Wuhan University, China; Xun Wang, Peking University, China

Saturday, May 2

CP14

Networks, Graphs, Applications - Part I of II

8:30 AM-9:20 AM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

For Part 2 see CP16

Chair: Manuel Gomez-Rodriguez, Max Planck Institute for Intelligent Systems, Germany

8:30-8:50 Fast Eigen-Functions Tracking on Dynamic Graphs

Chen Chen and Hanghang Tong, Arizona State University, USA

8:55-9:15 Approximation Algorithms for Reducing the Spectral Radius to Control Epidemic Spread

Sudip Saha, Abhijin Adiga, B. Aditya Prakash, and Anil Vullikanti, Virginia Tech, USA

Saturday, May 2

Workshop 1 (full day): Machine Learning Methods for Recommender Systems

8:30 AM-5:00 PM

Room:Port of New York - 3rd Level

Workshop 2 (full day): 4th Workshop on Data Mining for Medicine and Healthcare

8:30 AM-5:00 PM

Room:Pinnacle Harbourfront Ballroom III - 2nd Level

Saturday, May 2
TS4
Tutorial Session: Information Theory in Data Mining
8:30 AM-9:30 AM

Room:Port of Singapore - 3rd Level

Chair: Tao Li, Florida International University, USA

In the last decade Information Theoretic methods slowly but surely became popular in the data mining community for selecting the best model for the data at hand. In this tutorial we present an overview of these methods. Starting from the basics, to how one defines and finds good solutions using Information Theory, to how such models provide highly competitive solutions to many data mining tasks.

Matthijs Van Leeuwen
Katholieke Universiteit Leuven, Belgium

Arno Siebes
Universiteit Utrecht, The Netherlands

Jilles Vreeken
Max Planck Institute for Informatics, Germany and Saarland University, Germany

Coffee Break
9:30 AM-10:00 AM

Room:Ballroom Foyer - 2nd Level

Saturday, May 2
CP15
Text Mining, Applications - Part II of II
10:00 AM-11:40 AM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

For Part 1 see CP13

Chair: Lijing Qin, Tsinghua University, P. R. China

10:00-10:20 Propagation-Based Sentiment Analysis for Microblogging Data

Jiliang Tang, Arizona State University, USA; *Chikashi Nobata*, Anlei Dong, and Yi Chang, Yahoo! Labs, USA; Huan Liu, Arizona State University, USA

10:25-10:45 Polyglot-Ner Massive Multilingual Named Entity Recognition

Rami Al-Rfou and *Bryan Perozzi,* Stony Brook University, USA

10:50-11:10 Online Resource Allocation with Structured Diversification

Nicholas A. Johnson and Arindam Banerjee, University of Minnesota, USA

11:15-11:35 Towards Permission Request Prediction on Mobile Apps Via Structure Feature Learning

Deguang Kong, University of Texas at Arlington, USA; Hongxia Jin, Samsung Research America, USA

Saturday, May 2
CP16
Networks, Graphs, Applications - Part II of II
10:00 AM-11:40 AM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

For Part 1 see CP14

Chair: Manuel Gomez-Rodriguez, Max Planck Institute for Software Systems, Germany

10:00-10:20 On Influential Nodes Tracking in Dynamic Social Networks

Xiaodong Chen and Guojie Song, Peking University, China; Xinran He, University of South Carolina, USA; Kunqing Xie, Peking University, China

10:25-10:45 Less Is More: Building Selective Anomaly Ensembles with Application to Event Detection in Temporal Graphs

Leman Akoglu and *Shebuti Rayana*, Stony Brook University, USA

10:50-11:10 Principled Neuro-Functional Connectivity Discovery

Kejun Huang and Nicholas Sidiropoulos, University of Minnesota, USA; Evangelos Papalexakis and Christos Faloutsos, Carnegie Mellon University, USA; Partha Talukdar, Indian Institute of Science, Bangalore, India; Tom Mitchell, Carnegie Mellon University, USA

11:15-11:35 Estimating Ad Impact on Clicker Conversions for Causal Attribution: A Potential Outcomes Approach

Joel Barajas, University of California, Santa Cruz, USA; Ram Akella, University of California, Berkeley, USA; Aaron Flores and Marius Holtan, AOL Research, USA

Saturday, May 2

TS4

Tutorial Session: Information Theory in Data Mining, continued

10:00 AM-12:00 PM

Room:Port of Singapore - 3rd Level

Chair: Tao Li, Florida International University, USA

In the last decade Information Theoretic methods slowly but surely became popular in the data mining community for selecting the best model for the data at hand. In this tutorial we present an overview of these methods. Starting from the basics, to how one defines and finds good solutions using Information Theory, to how such models provide highly competitive solutions to many data mining tasks.

Matthijs Van Leeuwen
Katholieke Universiteit Leuven, Belgium

Arno Siebes
Universiteit Utrecht, The Netherlands

Jilles Vreeken
Max Planck Institute for Informatics, Germany and Saarland University, Germany

Lunch Break

12:00 PM-1:00 PM

Attendees on their own

Saturday, May 2

Workshop 3 (half day): Adaptive Learning On-a-chip: Hardware and Algorithms

1:00 PM-5:00 PM

Room:Pinnacle Harbourfront Ballroom I - 2nd Level

Workshop 4 (half day): Mining Networks and Graphs: A Big Data Analytic Challenge

1:00 PM-5:00 PM

Room:Pinnacle Harbourfront Ballroom II - 2nd Level

Workshop 5 (half day): Big Data and Stream Analytics

1:00 PM-5:00 PM

Room:Port of Singapore - 3rd Level

Workshop 6 (half day): Heterogeneous Learning

1:00 PM-5:00 PM

Room:Port of San Francisco - 3rd Level
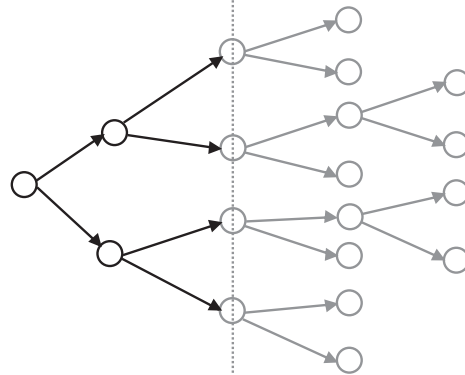
Coffee Break

3:00 PM-3:30 PM

Room:Ballroom Foyer - 2nd Level

SDM15 Abstracts

2015 SIAM International
Conference on **DATA MINING**

April 30-May 2, 2015

**Pinnacle Vancouver Harbourfront Hotel
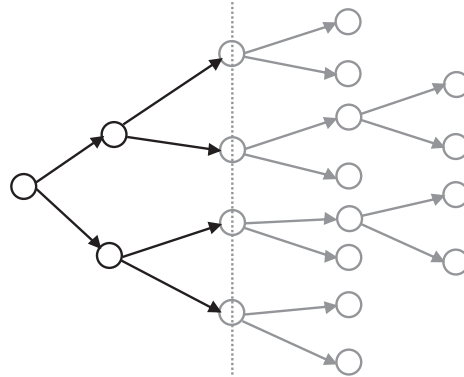Vancouver, British Columbia, Canada**

Abstracts are printed as submitted by the authors.

abstracts 28-52

Organizer and Speaker Index

2015 SIAM International
Conference on **DATA MINING**

April 30-May 2, 2015

**Pinnacle Vancouver Harbourfront Hotel
Vancouver, British Columbia, Canada**

A

Aggarwal, Charu C., PP1, 7:00 Thu

Ahmed, Nesreen, TS3, 10:00 Fri

Athitsos, Vassilis, PP1, 7:00 Thu

Aybat, Necdet S., CP11, 3:25 Fri

B

Barajas, Joel, CP16, 11:15 Sat

Bellet, Aurélien, CP11, 4:15 Fri

Benson, Austin, CP3, 11:15 Thu

C

Cai, Yongjie, CP7, 11:15 Fri

Cen, Lei, PP1, 7:00 Thu

Cetin, Mustafa S., CP7, 11:40 Fri

Chan, Hau, CP1, 10:25 Thu

Chaovalitwongse, Wanpracha Art, MS1, 10:25 Thu

Chaovalitwongse, Wanpracha Art, MS1, 3:00 Fri

Chaudhuri, Surajit, IP3, 8:15 Fri

Chen, Chen, CP14, 8:30 Sat

Chen, Junxiang, CP10, 3:50 Fri

Chen, Xiaodong, CP16, 10:00 Sat

Cheng, Yu, CP9, 10:00 Fri

Contractor, Danish, CP4, 3:00 Thu

Coroian, Dan C., PP1, 7:00 Thu

Crussell, Jonathan, CP6, 3:25 Thu

D

Dong, Bowen, PP1, 7:00 Thu

E

Erdos, Dora, CP10, 4:15 Fri

F

Fang, Chunsheng, PP1, 7:00 Thu

Fang, Meng, CP12, 3:25 Fri

Farajtabar, Mehrdad, CP3, 10:00 Thu

Fu, Yifan, CP8, 10:00 Fri

G

Goel, Ashish, IP1, 8:15 Thu

Granek, Justin, CP4, 3:25 Thu

Gupta, Sunil K., PP1, 7:00 Thu

H

Hasan, Mohammad, TS3, 10:00 Fri

Hauskrecht, Milos, CP5, 4:15 Thu

Heim, Eric, CP7, 10:00 Fri

Hong, Charmgil, PP1, 7:00 Thu

Hu, Qingbo, CP10, 4:40 Fri

Huang, Kejun, CP16, 10:50 Sat

Huang, Shuai, PP1, 7:00 Thu

Huang, Shuai, MS1, 3:00 Fri

J

Ji, Shuiwang, MS1, 3:00 Fri

Johnson, Nicholas A., CP15, 10:50 Sat

Jones, Steven, IP2, 1:30 Thu

K

Kale, David, CP12, 3:50 Fri

Karaev, Sanjar, CP8, 10:25 Fri

Karpatne, Anuj, PP1, 7:00 Thu

Keogh, Eamonn, TS2, 3:00 Thu

Kong, Deguang, CP15, 11:15 Sat

Krak, Thomas, CP11, 4:40 Fri

Kumar, Abhishek, CP8, 11:15 Fri

Kuo, Chia-Tung, PP1, 7:00 Thu

L

Lakkaraju, Himabindu, CP5, 3:00 Thu

Le, Long, PP1, 7:00 Thu

Lee, Dongeun, PP1, 7:00 Thu

Li, Huayu, CP8, 11:40 Fri

Li, Liangyue, CP7, 10:25 Fri

Li, Ruirui, PP1, 7:00 Thu

Li, Rumeng, CP13, 8:55 Sat

Li, Sheng, PP1, 7:00 Thu

Li, Tao, TS1, 10:00 Thu

Li, Tao, TS2, 3:00 Thu

Li, Tao, TS3, 10:00 Fri

Li, Tao, TS4, 8:30 Sat

Liu, Hongfu, PP1, 7:00 Thu

Liu, Jialu, CP9, 11:15 Fri

Liu, Ziqi, CP6, 4:15 Thu

Loukides, Grigorios, PP1, 7:00 Thu

M

Ma, Yuchi, PP1, 7:00 Thu

Mcconville, Ryan, CP3, 10:50 Thu

Menon, Aditya K., CP5, 3:50 Thu

Milios, Evangelos, TS1, 10:00 Thu

Mitra, Adway, PP1, 7:00 Thu

Mueen, Abdullah, TS2, 3:00 Thu

Müller, Emmanuel, CP3, 10:25 Thu

Myers, Risa, PP1, 7:00 Thu

N

Natesan Ramamurthy, Karthikeyan, CP6, 3:00 Thu

Neville, Jennifer, TS3, 10:00 Fri

Nobata, Chikashi, CP15, 10:00 Sat

O

Ohama, Iku, PP1, 7:00 Thu

P

Park, Dae Hoon, PP1, 7:00 Thu

Parotsidis, Nikos, CP1, 11:15 Thu

Perozzi, Bryon, CP15, 10:25 Sat

Petitjean, Francois, CP11, 3:50 Fri

Phan, Nhathai, PP1, 7:00 Thu

Prakash, Aditya, CP10, 3:25 Fri

Q

Qiu, Minghui, PP1, 7:00 Thu

R

Rangwala, Huzefa, PP1, 7:00 Thu

Rayana, Shebuti, CP16, 10:25 Sat

Rekatsinas, Theodoros, CP9, 10:50 Fri

S

Saha, Sudip, CP14, 8:55 Sat

Sharma, Mohit, CP5, 3:25 Thu

Shen, Dinggang, MS1, 3:25 Thu

Shokoohi-Yekta, Mohammad, CP7, 10:50 Fri

Siebes, Arno, TS4, 8:30 Sat

Skillicorn, David, CP2, 10:25 Thu

Soto, Axel, TS1, 10:00 Thu

Souza, Vinicius, PP1, 7:00 Thu

Stojanovic, Jelena Z., CP5, 4:40 Thu

Subbian, Karthik, CP9, 10:25 Fri

Subbian, Karthik, CP10, 3:00 Fri

Sugiyama, Mahito, CP1, 11:40 Thu

Sun, Yijun, CP2, 11:15 Thu

T

Tekumalla, Lavanya S., PP1, 7:00 Thu

Tomašev, Nenad, PP1, 7:00 Thu

U

Ukkonen, Antti, CP1, 10:50 Thu

Ulanova, Liudmila, PP1, 7:00 Thu

V

Van Leeuwen, Matthijs, TS4, 8:30 Sat

Vreeken, Jilles, PP1, 7:00 Thu

Vreeken, Jilles, TS4, 8:30 Sat

W

Wan, Mengting, PP1, 7:00 Thu

Wang, Wendy Hui, CP6, 3:50 Thu

Wei, Xiaokai, CP2, 11:40 Thu

Whang, Joyce J., PP1, 7:00 Thu

Wu, Le, CP4, 3:50 Thu

Wu, Xinyu, PP1, 7:00 Thu

X

Xiao, Houping, CP9, 11:40 Fri

Xu, Chang, CP4, 4:40 Thu

Xu, Jianpeng, CP12, 3:00 Fri

Xu, Linli, CP13, 8:30 Sat

Xu, Tong, CP6, 4:40 Thu

Xu, Xin-Shun, PP1, 7:00 Thu

Y

Yadwadkar, Neeraja J., CP12, 4:40 Fri

Yang, Le, PP1, 7:00 Thu

Yang, Pei, CP12, 4:15 Fri

Ye, Jieping, MS1, 4:15 Fri

Yu, Bin, IP4, 1:30 Fri

Z

Zeng, Guangxiang, CP8, 10:50 Fri

Zhai, Shuangfei, CP11, 3:00 Fri

Zhang, Aidong, CP1, 10:00 Thu

Zhang, Jiawei, CP3, 11:40 Thu

Zhang, Kai, CP2, 10:00 Thu

Zhang, Longhui, CP4, 4:15 Thu

Zhao, Liang, PP1, 7:00 Thu

Zhou, Peng, CP2, 10:50 Thu
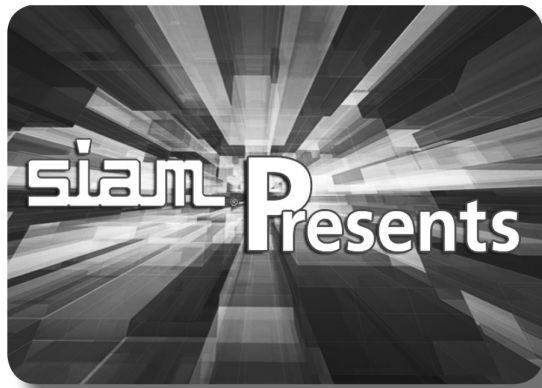
SIAM Presents

An audio-visual archive, comprised of more than 1800 presentations posted in 28 searchable topics, including:

- algebraic geometry
- atmospheric and oceanographic science
- computational science
- data mining
- geophysical science
- optimization
- uncertainty quantification and more…

The collection, *Featured Lectures from our Archives,* includes audio and slides from 25 conferences since 2008, including talks by invited and prize speakers, select minisymposia, and minitutorials from the 2014 Annual Meeting and four 2014 SIAG meetings.

In addition, you can view brief video clips of speaker interviews and topic overviews from sessions at Annual Meetings starting in 2010, as well as the 2013 SIAM Conference on Computational Science and Engineering and the 2014 SIAM Conference on the Life Sciences.

Plans for adding more content from SIAM meetings abound, including presentations from six meetings in 2015.

New presentations are posted every few months as the program expands with sessions from additional SIAM meetings. Users can search for presentations by category, speaker name, and/or keywords.
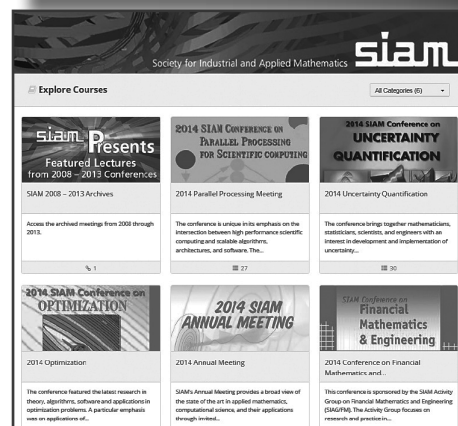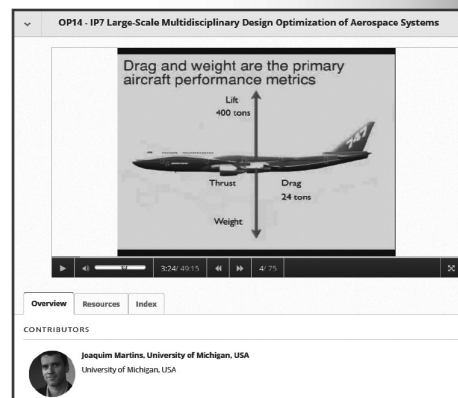
The audio, slide, and video presentations are part of SIAM's outreach activities to increase the public's awareness of mathematics and computational science in the real world, and to bring attention to exciting and valuable work being done in the field. Funding from SIAM, the National Science Foundation, and the Department of Energy was used to support this project.

www.siam.org/meetings/presents.php

SDM15 Budget

Conference Budget
SIAM Conference on Data Mining
April 30 - May 2, 2015
Vancouver, British Columbia

Expected Paid Attendance 250

Revenue

Registration Income		$102,185
	Total	$102,185

Expenses

Printing		$1,100
Organizing Committee		$2,500
Invited Speakers		$8,350
Food and Beverage		$32,300
AV Equipment and Telecommunication		$11,200
Advertising		$5,000
Proceedings		$9,400
Conference Labor (including benefits)		$46,580
Other (supplies, staff travel, freight, misc.)		$7,100
Administrative		$13,570
Accounting/Distribution & Shipping		$7,236
Information Systems		$13,047
Customer Service		$4,928
Marketing		$7,740
Office Space (Building)		$4,896
Other SIAM Services		$5,171
	Total	$180,118

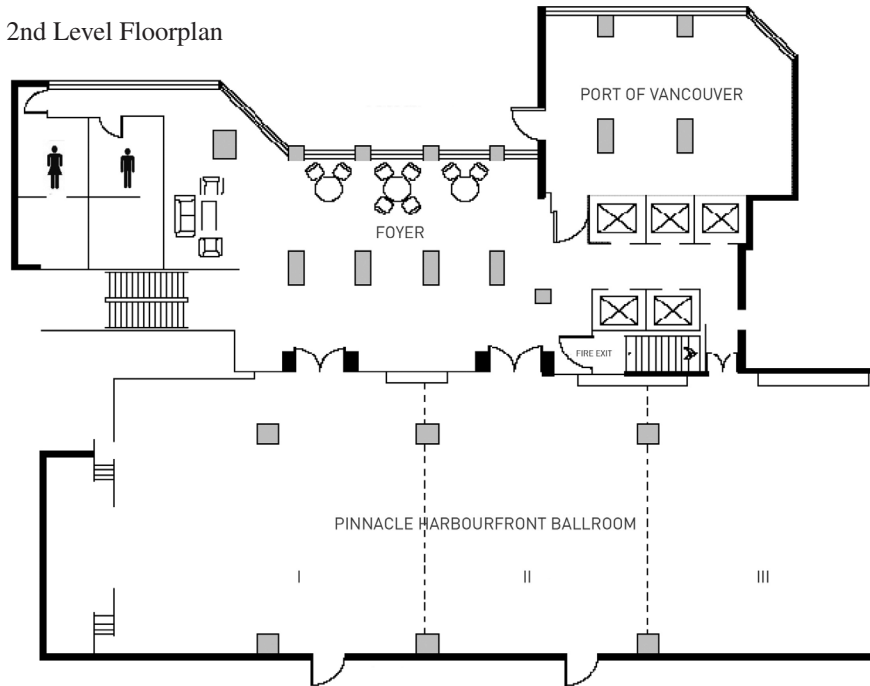|		
Net Conference Expense		($77,933)
Support Provided by SIAM		$77,933
		$0

Estimated Support for Travel Awards not included above:

Early Career / Students	27	$20,450

Pinnacle Vancouver Harbourfront Hotel
Vancouver, British Columbia, Canada

2nd Level Floorplan

PORT OF VANCOUVER

FOYER

FIRE EXIT

PINNACLE HARBOURFRONT BALLROOM

I   II   III

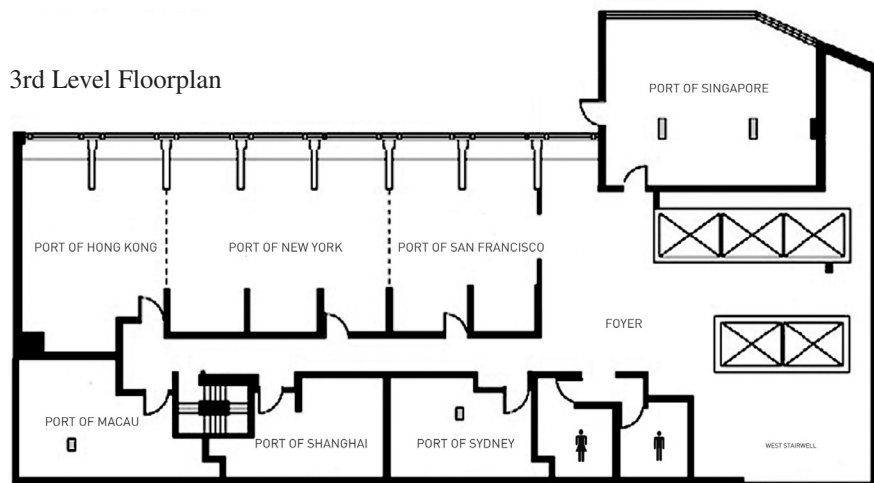**PINNACLE HOTEL**
VANCOUVER HARBOURFRONT

1133 West Hastings Street, Vancouver, BC, V6E 3T3
t: 604.689.2720 f: 604.691.2791 tf: 1.800.905.8582 w: pinnacleharbourfronthotel.com

3rd Level Floorplan

PORT OF SINGAPORE

PORT OF HONG KONG   PORT OF NEW YORK   PORT OF SAN FRANCISCO

FOYER

PORT OF MACAU   PORT OF SHANGHAI   PORT OF SYDNEY   WEST STAIRWELL

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

PINNACLE HOTEL
VANCOUVER HARBOURFRONT

1133 West Hastings Street, Vancouver, BC, V6E 3T3
t: 604.689.2720 f: 604.691.2791 tf: 1.800.905.8582 w: pinnacleharbourfronthotel.com