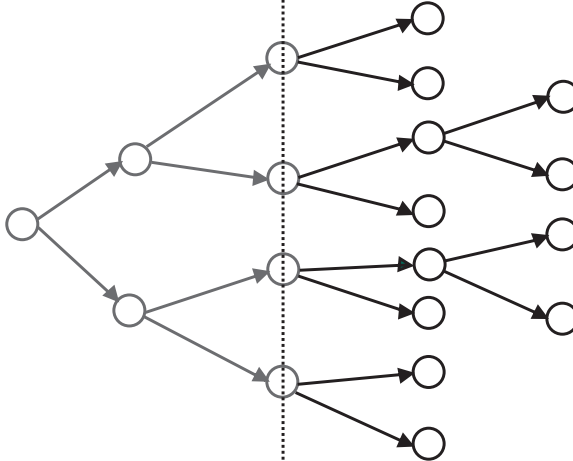


2017 SIAM  
International Conference  
on **DATA MINING**

April 27 - 29, 2017



The Westin Galleria Houston  
Houston, Texas, USA

Sponsored by the SIAM Activity Group on  
Data Mining and Analytics  
(SDM17)

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.



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](mailto:meetings@siam.org)

Conference Web: [www.siam.org/meetings/](http://www.siam.org/meetings/)

Membership and Customer Service:

(800) 447-7426 (USA & Canada) or +1-215-382-9800 (worldwide)

[www.siam.org/meetings/sdm17](http://www.siam.org/meetings/sdm17)

## Table of Contents

Program-At-A-Glance...	Fold out section
General Information.....	2
Get-togethers.....	7
Invited Plenary Presentations.....	10
Tutorials .....	11
Workshops.....	12
Program Schedule .....	15
Poster Sessions.....	21 & 25
Abstracts .....	29
Speaker and Organizer Index.....	57
Conference Budget...Inside Back Cover	
Hotel Meeting Room Map...Back Cover	

## Organizing Committee

### Steering Committee Chair

**Srinivasan Parthasarathy**  
Ohio State University, USA

### Steering Committee

**Chid Apte**  
IBM T.J. Watson Research Center, USA

**Christos Faloutsos**  
Carnegie Mellon University, USA

**Joydeep Ghosh**  
The University of Texas at Austin, USA

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

**Chandrika Kamath**  
Lawrence Livermore National Laboratory, USA

**Vipin Kumar**  
University of Minnesota, USA

**Haesun Park**  
Georgia Institute of Technology, USA

**Srinivasan Parthasarathy**  
Ohio State University, USA

**Qiang Yang**  
Hong Kong University of Science and Technology, USA

**Philip Yu**  
University of Illinois at Chicago, USA

### Conference Co-Chairs

**Carlotta Domeniconi**  
George Mason University, USA

**Dimitrios Gunopulos**  
University of Athens, Greece

### Program Co-Chairs

**Nitesh Chawla**  
University of Notre Dame, USA

**Wei Wang**  
University of California, Los Angeles, USA

### Workshops Co-Chairs

**Leman Akoglu**  
Carnegie Mellon University, USA

**Arthur Zimek**  
University of Southern Denmark, Denmark

### Tutorials Chair

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

### Doctoral Forum Chair

**Shuiwang Ji**  
Washington State University, USA

### Sponsorship Co-Chairs

**James Faghmous**  
Cahn School of Medicine, Mount Sinai, USA

**Jiliang Tang**  
Yahoo Research, USA

### Panels Chair

**Chris Jermaine**  
Rice University, USA

### Publicity Co-Chairs

**Danai Koutra**  
University of Michigan, USA

**Matthijs van Leeuwen**  
Leiden University, The Netherlands

**Panagiotis Papapetrou**  
Stockholm University, Sweden

### Awards Chair

**Zoran Obradovic**  
Temple University, USA

### Senior Program Committee

**Dino Pedreschi**  
University of Pisa, Italy

**Bianca Zadrozny**  
IBM-Research, Brazil

### Charu Aggarwal

IBM T.J. Watson Research Center, USA

### Leman Akoglu

Carnegie Mellon University, USA

### Arindam Banerjee

University of Minnesota, USA

### Ron Bekkerman

University of Haifa, Israel

### Tanya Berger-Wolf

University of Illinois, Chicago, USA

### Bettina Berned

KU Leuven, Belgium

### Francesco Bonchi

Yahoo! Barcelona, Spain

### Sanjay Chawla

University of Sydney, Australia

### Carlotta Domeniconi

George Mason University, USA

### Christos Faloutsos

Carnegie Mellon University, USA

### Joao Gama

University of Porto, Portugal

### Aristides Gionis

Alto University, Finland

### David Gleich

Purdue University, USA

### Jiawei Han

University of Illinois Urbana Champaign, USA

### Pallika Kanani

Oracle Labs, USA

### Eamonn Keogh

University of California Riverside, USA

### Chih-Jen Lin

National Taiwan University, Taiwan

### Wagner Meira

Universidade Federal de Minas Gerais, Brazil

### Ernestina Mensalvas

Universidad Politecnica de Madrid, Spain

### Zoran Obradovic

Temple University, USA

### Srini Parthasarathy

Ohio State University, USA

### Jian Pei

Simon Fraser University, Canada

### B. Aditya Prakash

Virginia Tech, USA

### Rajeev Rastogi

Amazon, India

**Celine Robarde**

National Institute of Applied Science,  
France

**Mohak Shah**

Bosch, USA

**Kyuseok Shim**

Seoul National University, Korea

**Arno Siebes**

Universiteit Utrecht, Netherlands

**Masashi Sugiyama**

RIKEN/University of Tokyo, Japan

**Yizhou Sun**

University of California, Los Angeles

**Hui Xiong**

Rutgers University, USA

**Jeffrey Yu**

Chinese University of Hong Kong, Hong  
Kong

**Zhi-Hua Zhou**

Nanjing University, China

**Program Committee****Aris Anagnostopoulos**

Sapienza University of Rome, Italy

**Annalisa Appice**

University of Bari, Italy

**Gowtham Atluri**

University of Cincinnati, USA

**Tony Bagnall**

University of East Anglia, United Kingdom

**Alex Beutel**

Google, USA

**K. Selcuk Candan**

Arizona State University, USA

**Nan Cao**

BM T. J. Watson Research Center, USA

**Andre Carvalho**

ICMC-USP, Brazil

**Loic Cerf**

Universidade Federal de Minas Gerais,  
Brazil

**Shiyu Chang**

UIUC, USA

**Xiangyu Chang**

Xi'an Jiaotong University, China

**Yi Chang**

Yahoo! labs

**Duen Horng Chau**

Georgia Tech, USA

**Feng Chen**

University at Albany, SUNY

**Haifeng Chen**

NEC Research Lab, USA

**Jianhui Chen**

Yahoo Labs

**Wei Chen**

Microsoft Research, Asia

**Hong Cheng**

Hong Kong Chinese University, Hong  
Kong

**Jian Cheng**

Chinese Academy of Science, China

**Wei Cheng**

NEC Lab, USA

**Yu Cheng**

IBM T.J. Watson Research Center, USA

**Malolan Chettur**

IBM Research, India

**Diane Cook**

Washington State University, USA

**Robson Cordeiro**

Universidade de Sao Paulo, Brazil

**Peng Cui**

Tsinghua University, China

**Alfred Cuzzocrea**

University of Calabria, Italy

**Xuan-Hong Dang**

University of California Santa Barbara,  
USA

**Mahashweta Das**

HP Labs, USA

**Hasan Davulcu**

Arizona State University, USA

**Yuxiao Dong**

University of Notre Dame, USA

**Haimonti Dutta**

Columbia University, USA

**William Eberle**

Tennessee Technological University, USA

**Dora Erdos**

Boston University, USA

**Lisa Friedland**

Northeastern University, USA

**Sorelle Friedler**

Haverford College, USA

**Yanjie Fu**

Missouri University of Science and  
Technology, USA

**Yun Fu**

Northeastern University, USA

**Amita Gajewar**

Microsoft Corp, USA

**Jing Gao**

University of Buffalo, USA

**Mohamed Ghalwash**

Temple University, USA

**Manuel Gomez-Rodriguez**

Max Planck Institute, Germany

**Quanquan Gu**

University of Virginia, USA

**Ziyu Guan**

Northwest University, China

**Jie Gui**

Chinese Academy of Science, China

**Francesco Gullo**

Yahoo! Research, Spain

**Stephan Günnemann**

Technical University of Munich, Germany

**Yuan Hao**

Google, USA

**Xia Hu**

Arizona State University, USA

**Jun (Luke) Huan**

University of Kansas, USA

**Heng Huang**

University of Texas at Arlington, USA

**Sheng-Jun Huang**

Nanjing University of Aeronautics and  
Astronautics, China

**Xin Huang**

University of British Columbia, Canada

**Bo Jin**

Dalian University of Technology, China

**Reid Johnson**

University of Notre Dame, USA

**Chandrika Kamath**

LLNL, USA

**Murat Kantarcioglu**

University of Texas, Dallas, USA

**Masayuki Karasuyama**

Nagoya Institute of Technology, Japan

**George Karypis**

University of Minnesota, USA

**Philip Kegelmeyer**

Sandia National Labs, USA

**Daniel Keren**

Haifa University, Israel

**Deguang Kong**

Samsung, USA

**Xiangnan Kong**

Worcester Polytechnic Institute, USA

**Danai Koutra**

University of Michigan, USA

**Georg Krempf**Otto-von-Guericke Univ. Magdeburg,  
Germany**Cheng-Te Li**

Academic Sinica, Taiwan

**Geng Li**

Yahoo!, USA

**Jiuyong Li**

University of South Australia, Australia

**Rong-Hua Li**

Shenzhen University, China

**Shuai Li**

University of Cambridge, United Kingdom

**Yuefeng Li**Queensland University of Technology,  
Australia**Zhenhui Li**

Penn State University, USA

**Jessica Lin**

George Mason University, USA

**Shou-de Lin**

National Taiwan University, Taiwan

**Xuemin Lin**

New South Wales University, Australia

**Bin Liu**

IBM Watson, USA

**Chuanren Liu**

Drexel University, USA

**Huan Liu**

Arizona State University, USA

**Tongliang Liu**

University of Technology Sydney, Australia

**Yan Liu**

University of Southern California, USA

**David Lo**Singapore Management University,  
Singapore**Aurelie Lozano**

IBM Research, USA

**Chang-Tien Lu**

Virginia Tech, USA

**Shuai Ma**

Beihang University, China

**Arun Maiya**

Institute for Defense Analysis, USA

**Julian McAuley**

University of California, San Diego, USA

**Luiz Merschmann**Universidade Federal de Ouro Preto  
(UFOP), Brazil**Pauli Miettinen**Max-Planck-Institute for Informatics,  
Germany**Benjamin Miller**

MIT Lincoln Laboratory, USA

**Abdullah Mueen**

University of New Mexico, USA

**Vinh Nguyen**

University of Melbourne, Australia

**Xia Ning**Indiana University - Purdue University,  
USA**Eduardo Ogasawara**Centro Federal de Educação Tecnológica  
(CEFET-RJ), Brazil**Nikunj Oza**

NASA Ames Research Center, USA

**Aline Paes**Universidade Federal Fluminense (UFF),  
Brazil**Evangelos Papalexakis**

University of California, Riverside, USA

**Douglass Parker**

University of California Los Angeles, USA

**Ruggero Pensa**

University of Torino, Italy

**Bryan Perozzi**

Google, USA

**Marc Plantevit**

University de Lyon, France

**Buyue Qian**

IBM T. J. Watson Research Center, USA

**Chao Qian**University of Science and Technology of  
China**Mu Qiao**

IBM, USA

**Zhiwei Qin**

Walmart Labs, USA

**Chedy Raissi**

LORIA INRIA, France

**Arti Ramesh**

SUNY Binghamton, USA

**Balaraman Ravindran**

IIT-Madras, India

**Alejandro Rodriguez**

University of Madrid, Spain

**Mahsa Saleh**

IBM, Australia

**Saket Sathe**

IBM T. J. Watson Research Center, USA

**Lisa Singh**

Georgetown University, USA

**Dongjin Song**

NEC Labs America, USA

**Le Song**

Georgia Tech, USA

**Yangqiu Song**

UIUC, USA

**Arnaud Soulet**University François Rabelais of Tours,  
France**Sucheta Soundarajan**

Syracuse University, USA

**Jerzy Stefanowski**

Technical University Poznan, Poland

**Andrea Tagarelli**

University of Calabria, Italy

**Hana Tai**

National Taipei University, Taiwan

**Pang-Ning Tan**

Michigan State University, USA

**Nan Tang**

QCRI, Qatar

**Dacheng Tao**The University of Technology, Sydney,  
Australia**Hanghang Tong**

Arizona State University, USA

**Caetano Traina**

ICMC-USP, Brazil

**Panayiotis Tsaparas**

University of Ioannina, Greece

**Grigorios Tsoumakas**

Aristotle University Thessaloniki, Greece

**Matthijs van Leeuwen**

Leiden University, Netherlands

**Adriano Veloso**Universidade Federal de Minas Gerais,  
Brazil**Suresh, Venkatasubramiam**

University of Utah, USA

**Jilles Vreeken**

Max Planck Institute, Germany

**Fei Wang**

University of Connecticut, USA

**Jianyang Wang**

Tsinghua University, China

**Ke Wang**

Simon Fraser University, Canada

**Ting Wang**

Lehigh University, USA

**Xiang Wang**

Google, USA

**Zheng Wang**

University of Michigan, USA

**Takashi Washio**

Osaka University, Japan

**Ingmar Weber**

QCRI, Qatar

**Raymond Chi-Wing Wong**

Hong Kong University of Science and Technology, Hong Kong

**Jia Wu**

University of Technology, Australia

**Xintao Wu**

University of Arkansas, USA

**Ren Xiang**

University of Illinois Urbana-Champaign, USA

**Keli Xiao**

Stony Brook University, USA

**Yanghua Xiao**

Fudan University, China

**Jian Xu**

University of Notre Dame, USA

**Tong Xu**

University of Science and Technology of China, China

**Makoto Yamada**

Kyoto University, Japan

**Yang Yang**

Northwestern University, USA

**Dayu Yuan**

Google, USA

**Jing Yuan**

Microsoft Research, Asia

**Carlo Zaniolo**

University of California Los Angeles, USA

**De-Chuan Zhan**

Nanjing University, China

**Chao Zhang**

Dalian University of Technology, China

**Kun Zhang**

Xavier University of Louisiana, USA

**Lumin Zhang**

Facebook, USA

**Min-Ling Zhang**

Southeast University, China

**Ping Zhang**

IBM Thomas J. Watson Research Center, USA

**Xiang Zhang**

Penn State University, USA

**Xiangliang Zhang**

KAUST, Saudi Arabia

**Xiaoming Zhang**

Beihang University, China

**Yao Zhang**

Virginia Tech, USA

**Zhao Zhang**

Soochow University, China

**Yu Zheng**

Microsoft, China

**Jiayu Zhou**

Michigan State University, USA

**Yuanchun Zhou**

Chinese Academy of Sciences, China

**Hengshu Zhu**

Baidu, China

**Xingquan Zhu**

Florida Atlantic University, USA

**Arthur Zimek**

University of Southern Denmark, Denmark

**Albrecht Zimmermann**

University de Caen, France

**SIAM Registration Desk**

The SIAM registration desk is located in the Galleria Foyer except Wednesday night, when it will be in the Chevy Chase Room. It is open during the following hours:

Wednesday, April 26

5:00 PM – 7:00 PM

Thursday, April 27

7:00 AM – 7:30 PM

Friday, April 28

7:15 AM – 3:30 PM

Saturday, April 29

7:15 AM – 4:00 PM

**Hotel Address**

Westin Galleria Houston

5060 West Alabama Street

Houston, Texas 77056 USA

Phone Number: +1-713-960-8100

Toll Free Reservations (USA and Canada):  
+1-888-627-8514

Fax: +1-713-960-6549

Hotel web address:

<http://westingalleriahoustonhotel.com>**Hotel Telephone Number**

To reach an attendee or leave a message, call +1-713-960-8100. 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**

The Westin Galleria Houston recommends Mom's Best Friend (<http://momsbestfriend.com/houston/>) for attendees interested in child care services. Attendees are responsible for making their own child care arrangements. Their phone number is +1-713-776-2669 ext. 204.

**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  
 Amazon  
 Aramco Services Company  
 Bechtel Marine Propulsion Laboratory  
 The Boeing Company  
 CEA/DAM  
 Department of National Defence (DND/  
 CSEC)  
 DSTO- Defence Science and  
 Technology Organisation  
 Hewlett-Packard  
 Huawei FRC French R&D Center  
 IBM Corporation  
 IDA Center for Communications  
 Research, La Jolla  
 IDA Center for Communications  
 Research, Princeton  
 Institute for Defense Analyses, Center  
 for Computing Sciences  
 Lawrence Berkeley National Laboratory  
 Lawrence Livermore National Labs  
 Lockheed Martin  
 Los Alamos National Laboratory  
 Max-Planck-Institute for Dynamics of  
 Complex Technical Systems  
 Mentor Graphics  
 National Institute of Standards and  
 Technology (NIST)  
 National Security Agency (DIRNSA)  
 Naval PostGrad  
 Oak Ridge National Laboratory,  
 managed by UT-Battelle for the  
 Department of Energy  
 Sandia National Laboratories  
 Schlumberger-Doll Research  
 United States Department of Energy  
 U.S. Army Corps of Engineers,  
 Engineer Research and Development  
 Center  
 US Naval Research Labs

*List current March 2017.*

**Funding Agency**

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

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

SIAM members save up to \$130 on full registration for the 2017 SIAM Conference on 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 \$15 to join the SIAM Activity Group on Data Mining and Analytics (SIAG/DMA). As a SIAG/DMA member, you are eligible for an additional \$15 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](http://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 (USA  
 and Canada only)

Fax: +1-215-386-7999

E-mail: [membership@siam.org](mailto:membership@siam.org)

Postal mail: Society for Industrial and Applied Mathematics, 3600 Market Street, 6<sup>th</sup> 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. The data projectors support VGA connections only. Presenters requiring an HDMI or alternate connection must provide their own adaptor.

All other concurrent/breakout rooms will have one (1) screen and one (1) data projector. The data projectors support VGA connections only. Presenters requiring an HDMI or alternate connection must provide their own adaptor.

If you have questions regarding availability of equipment in the meeting room of your presentation, please see a SIAM staff member at the registration desk.

## Internet Access

Complimentary wireless Internet access will be available for SIAM attendees in the meeting space and guestrooms in the SIAM block.

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 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 sessions are scheduled for Thursday, April 27, 7:00 PM – 9:00 PM and Friday, April 28, 7:00 PM – 9:00 PM. Presenters are requested to put up their posters no later than 7:00 PM, the official start time of both sessions.

Papers presented on Thursday and Saturday will have their poster slots during the Welcome Reception and Poster Session on Thursday, April 27. Papers presented on Friday will have their poster slots during the Doctoral Forum and Poster Session on Friday, April 28. Boards and push pins will be available to Thursday's presenters at 12:00 PM on Thursday, April 27 and at 9:00 AM on Friday, April 28 for Friday's presenters.

For information about preparing a poster, please visit <http://www.siam.org/meetings/guidelines/presenters.php>.

## 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. Titles on display forms are available with instructions on how to place a book order.

## 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!

## Comments?

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

## Get-togethers

Welcome Reception and Poster Session

Thursday, April 27

7:00 PM – 9:00 PM



Business Meeting  
(open to SIAG/DMA members)

Friday, April 28

6:15 PM – 7:00 PM



*Complimentary beer and wine will be served.*

## Statement on Inclusiveness

As a professional society, SIAM is committed to providing an inclusive climate that encourages the open expression and exchange of ideas, that is free from all forms of discrimination, harassment, and retaliation, and that is welcoming and comfortable to all members and to those who participate in its activities. In pursuit of that commitment, SIAM is dedicated to the philosophy of equality of opportunity and treatment for all participants regardless of gender, gender identity or expression, sexual orientation, race, color, national or ethnic origin, religion or religious belief, age, marital status, disabilities, veteran status, field of expertise, or any other reason not related to scientific merit. This philosophy extends from SIAM conferences, to its publications, and to its governing structures and bodies. We expect all members of SIAM and participants in SIAM activities to work towards this commitment.

## Please Note

SIAM is not responsible for the safety and security of attendees' computers. Do not leave your personal electronic devices unattended. Please remember to turn off your cell phones and other devices 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 #SIAMSDM17.

SIAM's Twitter handle is @TheSIAMNews.



# SIAM Activity Group on Data Mining and Analytics (SIAG/DMA)

[www.siam.org/activity/dma/index.php](http://www.siam.org/activity/dma/index.php)



## A GREAT WAY TO GET INVOLVED!

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

### ACTIVITIES INCLUDE:

- Special sessions at SIAM Annual Meetings
- Annual conference

### BENEFITS OF SIAG/DMA MEMBERSHIP:

- Listing in the SIAG's online membership directory
- Additional \$15 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:

- \$15 per year
- Student members can join two activity groups for free!

### 2016-17 SIAG/DMA OFFICERS

- Chair: Jeremy Kepner, Massachusetts Institute of Technology
- Vice Chair: Jaideep Vaidya, Rutgers University
- Program Director: Wagner Meira, Universidade Federal de Minas Gerais
- Secretary: Danai Koutra, University of Michigan

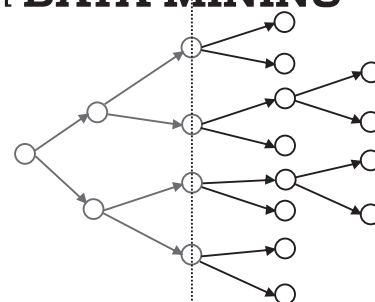
### TO JOIN:

SIAG/DMA: [my.siam.org/forms/join\\_siag.htm](http://my.siam.org/forms/join_siag.htm)

SIAM: [www.siam.org/joinsiam](http://www.siam.org/joinsiam)

### 2017 SIAM International Conference on **DATA MINING**

April 27 - 29, 2017



The Westin Galleria Houston  
Houston, Texas, USA

## Invited Plenary Speakers

**\*\* All Invited Plenary Presentations will take place in Galleria Ballroom I\*\***

**Thursday, April 27**

**8:15 AM - 9:30 AM**

**IP1** Title Not Available at Time of Publication

**William Craig Fugate**, *Federal Emergency Management Agency, USA*

**1:30 PM - 2:45 PM**

**IP2** Collective Learning in Society and the Economy

**Cesar Hidalgo**, *Massachusetts Institute of Technology, USA*

**Friday, April 28**

**8:15 AM - 9:30 AM**

**IP3** Machine Learning under Resource Constraints

**Katharina Morik**, *Technische Universität Dortmund, Germany*

**1:30 PM - 2:45 PM**

**IP4** Repeated Choice, Markov Models, and LAMP

**Andrew Tomkins**, *Google, Inc., USA*

## Tutorials

**Thursday, April 27**

**10:00 AM - 12:00 PM**

**TS1: Tutorial Session:** An Introduction to Redescription Mining  
*Post Oak*

**2:45 PM - 4:45 PM**

**TS2: Tutorial Session:** Leveraging Propagations for Data Mining  
*Post Oak*

**5:00 PM - 7:00 PM**

**TS3: Tutorial Session:** Opportunities, Challenges and Methods  
for Higher Education Data Mining  
*Post Oak*

**Friday, April 28**

**10:00 AM - 12:00 PM**

**TS4: Tutorial Session:** IoT Big Data Stream Mining  
*Post Oak*

**2:45 PM - 3:45 PM**

**TS4: Tutorial Session (Continued):** IoT Big Data Stream Mining  
*Post Oak*

**Saturday, April 29**

**3:45 PM - 5:45 PM**

**TS5: Tutorial Session:** Summarizing Large-Scale Graph Data  
*Galleria Ballroom I*

## Workshops

**Saturday, April 29**

**8:00 AM - 12:15 PM (Half Day Workshop)**

**Workshop 2:** Data Mining for Medicine and Healthcare

*Tanglewood*

**8:00 AM - 5:45 PM (Full Day Workshops)**

**Workshop 1:** Machine Learning Methods for Recommender Systems

*Galleria Ballroom II*

**Workshop 4:** Inferring Networks From Non-Network Data

*Post Oak*

**Workshop 5:** Mining Big Data in Climate and Environment

*Bellaire*

**Workshop 6:** Women in Data Science

*San Felipe*

**1:30 PM - 5:45 PM (Half Day Workshop)**

**Workshop 3:** Workshop on Data Mining for Oil & Gas

*Tanglewood*



## SIAM Presents is an audio-visual archive

comprised of more than 2,000 presentations posted in over 40 searchable topics, including:

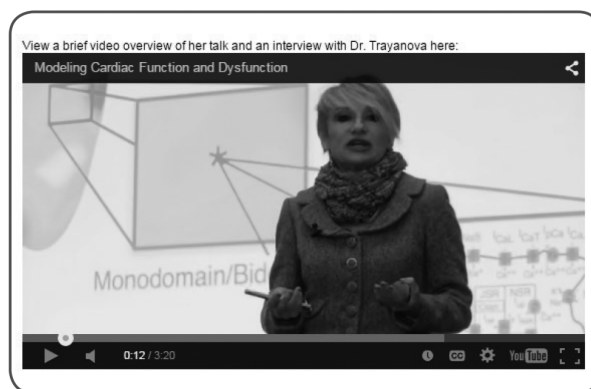
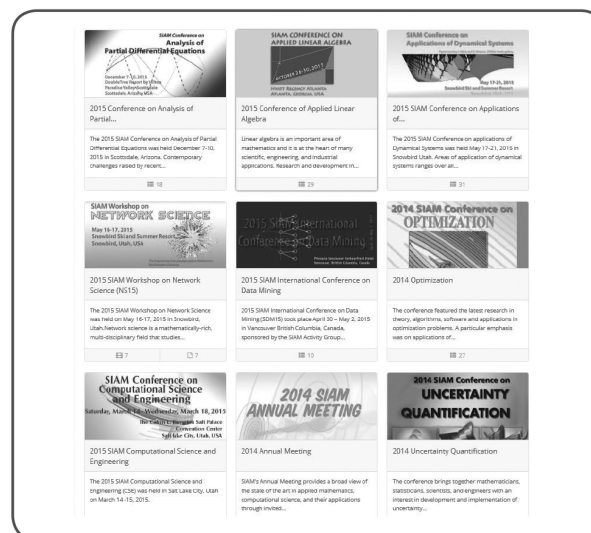
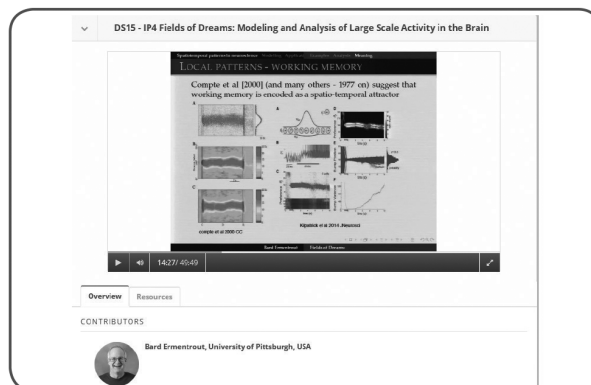
- algebraic geometry
- atmospheric and oceanographic science
- computational science
- data mining
- geophysical science
- optimization
- uncertainty quantification and more...

The collection, *Featured Lectures from our Archives*, includes audio and slides from more than 30 conferences since 2008, including talks by invited and prize speakers, select minisymposia, and minitutorials. Presentations from SIAM meetings are being added throughout the year.

In addition you can view short video clips of speaker interviews from sessions at Annual Meetings starting in 2010.

Plans for adding more content are on the horizon. Keep an eye out!

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 partially support this project.



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 key words.

[www.siam.org/meetings/presents.php](http://www.siam.org/meetings/presents.php)



Society for Industrial and Applied Mathematics • 3600 Market Street, 6th Floor • Philadelphia, PA 19104-2688 USA  
Phone: +1-215-382-9800 • Fax +1-215-386-7999 • [service@siam.org](mailto:service@siam.org) • [www.siam.org](http://www.siam.org)

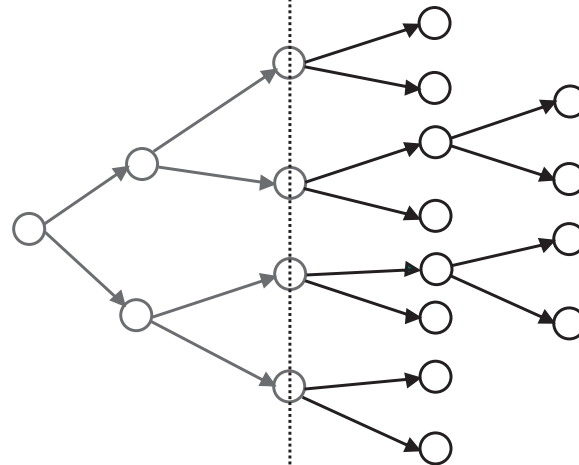
## Notes



## Program Schedule

2017 SIAM  
International Conference  
on **DATA MINING**

April 27 - 29, 2017



The Westin Galleria Houston  
Houston, Texas, USA

## Wednesday, April 26

### Registration

5:00 PM-7:00 PM

Room: Chevy Chase

## Thursday, April 27

### Registration

7:00 AM-7:30 PM

Room: Galleria Foyer

### Continental Breakfast

7:30 AM-8:00 AM

Room: Galleria Foyer



### Welcome Remarks

8:00 AM-8:15 AM

Room: Galleria Ballroom I

Thursday, April 27

## IP1

### Title Not Available at Time of Publication

8:15 AM-9:30 AM

Room: Galleria Ballroom I

Chair: Wei Wang, University of California,  
Los Angeles, USA

Abstract not available at time of  
publication.

William Craig Fugate

Federal Emergency Management Agency,  
USA

### Coffee Break

9:30 AM-10:00 AM

Room: Galleria Foyer



Thursday, April 27

**TS1****Tutorial Session:  
An Introduction to  
Redescription Mining**

10:00 AM-12:00 PM

Room: Post Oak

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

In this tutorial we will give an overview of redescription mining, from the intuition behind the concept and its links to existing data analysis techniques to more recent developments in algorithms and interactive mining techniques. We will also cover five areas where applications for redescription mining have been proposed. The tutorial will give the attendants knowledge of the state-of-the-art techniques in redescription mining and open problems in method development, as well as examples and information on how to apply redescription mining to real-world data analysis problems.

**Pauli Miettinen**Max Planck Institute for Informatics,  
Germany**Esther Galbrun**

Inria, France

Thursday, April 27

**CP1****Classification I**

10:00 AM-11:40 AM

Room: Galleria Ballroom I

Chair: Xin Huang, Hong Kong Baptist  
University, Hong Kong**10:00-10:15 Using a Random Forest  
to Inspire a Neural Network and  
Improving on It**Suhang Wang, Arizona State University,  
USA; Charu C. Aggarwal, IBM T.J.  
Watson Research Center, USA; Huan Liu,  
Arizona State University, USA**10:20-10:35 Pruning Decision Trees Via  
Max-Heap Projection**Zhi Nie, Arizona State University, USA;  
Binbin Lin, University of Michigan, Ann  
Arbor, USA; Shuai Huang, University of  
Washington, USA; Naren Ramakrishnan,  
Virginia Tech, USA; Wei Fan, Baidu  
Research Big Data Lab, China; Jieping Ye,  
University of Michigan, Ann Arbor, USA**10:40-10:55 Margin Distribution  
Logistic Machine**Yi Ding, Sheng-Jun Huang, Chen Zu, and  
Daoqiang Zhang, Nanjing University of  
Aeronautics and Astronautics, China**11:00-11:15 Active Learning of  
Classification Models with Likert-Scale  
Feedback**Milos Hauskrecht and Yanbing Xue,  
University of Pittsburgh, USA**11:20-11:35 Risk Clearance with  
Guaranteed Precision**Ryan McBride, Ke Wang, and Viswanadh  
Nekkanti, Simon Fraser University,  
Canada; Wenyan Li, Chongqing  
University, P.R. China

Thursday, April 27

**CP2****Bioinformatics**

10:00 AM-11:40 AM

Room: Galleria Ballroom II

Chair: Yao Zhang, Virginia Tech, USA

**10:00-10:15 Computational Drug  
Discovery with Dyadic Positive-  
Unlabeled Learning**Yashu Liu, Arizona State University, USA;  
Shuang Qiu, University of Michigan, Ann  
Arbor, USA; Ping Zhang, IBM T.J. Watson  
Research Center, USA; Pinghua Gong,  
University of Michigan, Ann Arbor, USA;  
Fei Wang, IBM T.J. Watson Research  
Center, USA; Guoliang Xue, Arizona State  
University, USA; Jieping Ye, University of  
Michigan, Ann Arbor, USA**10:20-10:35 Multi-Region Neural  
Representation: A Novel Model for  
Decoding Visual Stimuli in Human  
Brains**Muhammad Yousefnezhad and Daoqiang  
Zhang, Nanjing University of Aeronautics  
and Astronautics, China**10:40-10:55 Alpine: Progressive Itemset  
Mining with Definite Guarantees**Qiong Hu and Tomasz Imielinski, Rutgers  
University, USA**11:00-11:15 Polyadic Regression and  
Its Application to Chemogenomics**Ioakeim Perros, Georgia Institute of  
Technology, USA; Fei Wang, Weill Cornell  
Medicine, USA; Ping Zhang, IBM T.J.  
Watson Research Center, USA; Peter  
Walker, Naval Medical Research Center,  
USA; Richard Vuduc, Georgia Institute of  
Technology, USA; Jyotishman Pathak, Weill  
Cornell Medicine, USA; Jimeng Sun, Georgia  
Institute of Technology, USA**11:20-11:35 Active Learning of  
Functional Networks from Spike Trains**Honglei Liu, University of California, Santa  
Barbara, USA; Bian Wu, Washington State  
University, USA

Thursday, April 27

## CP3

### Anomaly and Outlier Detection

10:00 AM-11:40 AM

Room: Tanglewood

Chair: Xia Ben Hu, Texas A&M University, USA

#### 10:00-10:15 Outlier Detection with Autoencoder Ensembles

Jinghui Chen, University of Virginia, USA;  
Saket Sathe, Charu C. Aggarwal, and  
Deepak Turaga, IBM T.J. Watson Research  
Center, USA

#### 10:20-10:35 Gleaning Wisdom from the Past: Early Detection of Emerging Rumors in Social Media

Liang Wu and Jundong Li, Arizona State  
University, USA; Xia Hu, Texas A&M  
University, USA; Huan Liu, Arizona State  
University, USA

#### 10:40-10:55 VolTime: Unsupervised Anomaly Detection on Users' Online Activity Volume

Daniel Chino, Alceu Costa, and Agma J.  
Traina, University of Sao Paulo, Brazil;  
Christos Faloutsos, Carnegie Mellon  
University, USA

#### 11:00-11:15 Detecting Malicious Behavior in Computer Networks Via Cost-Sensitive and Connectivity Constrained Classification

Houping Xiao and Jing Gao, State University  
of New York at Buffalo, USA; Long Vu and  
Deepak Turaga, IBM T.J. Watson Research  
Center, USA

#### 11:20-11:35 Efficiently Discovering Unexpected Pattern-Co-Occurrences

Arno Siebes and Roel Bertens, Universiteit  
Utrecht, The Netherlands; Jilles Vreeken,  
Max Planck Institute for Informatics,  
Germany

## Lunch Break

12:00 PM-1:15 PM

Attendees on their own

Thursday, April 27

## IP2

### Collective Learning in Society and the Economy

1:15 PM-2:30 PM

Room: Galleria Ballroom I

Chair: Nitesh Chawla, University of Notre  
Dame, USA

How do economies learn to develop new products, industries, skills, and occupations? Where do they get the knowledge they need? And how is that learning affected by technologies, institutions, and culture? In this presentation I will present work describing the key questions defining the field of collective learning, and also, will present novel data visualization tools designed to facilitate collective learning in teams and nations.

Cesar Hidalgo

Massachusetts Institute of Technology, USA

## Coffee Break

2:30 PM-2:45 PM

Room: Galleria Foyer



Thursday, April 27

## TS2

### Tutorial Session: Leveraging Propagations for Data Mining

2:45 PM-4:45 PM

Room: Post Oak

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

Can we guess if a user is sick from her tweet? How do opinions get formed in online forums? Which people should we immunize, to prevent an epidemic as fast as possible? How to quickly zoom out of a graph? Graphs---also known as networks---are powerful tools for modeling processes and situations of interest in real-life like social-systems, cyber-security, epidemiology and biology. They are ubiquitous, from online social networks, gene-regulatory networks, to router graphs.

This tutorial will cover recent and state-of-the-art research on how propagation-like processes can help big-data mining specifically large networks and time-series, algorithms behind network problems, and their practical applications in various diverse settings. Topics include diffusion and virus propagation in networks, anomaly and outbreak detection, event prediction and connections with work in public health, the web and online media, social sciences/humanities and cyber security.

**B. Aditya Prakash**

Virginia Tech, USA

**Naren Ramakrishnan**

Virginia Tech, USA

Thursday, April 27

**CP4****Classification II**

2:45 PM-4:25 PM

Room: Galleria Ballroom I

*Chair: Polo Chau, Georgia Institute of Technology, USA***2:45-3:00 Time-Aware Subscription Prediction Model for User Acquisition in Digital News Media***Heidar Davoudi, Morteza Zihayat, and Aijun An, York University, Canada***3:05-3:20 The Power of Certainty: A Dirichlet-Multinomial Model for Belief Propagation***Dhivya Eswaran, Carnegie Mellon University, USA; Stephan Guennemann, Technical University of Munich, Germany; Christos Faloutsos, Carnegie Mellon University, USA***3:25-3:40 From Theory to Practice: Efficient Active Cost-Sensitive Classification with Expected Error Reduction***Yexun Zhang and Yanfeng Wang, Shanghai Jiao Tong University, China; Wenbin Cai, Microsoft Research Asia; Siyuan Zhou and Ya Zhang, Shanghai Jiao Tong University, China***3:45-4:00 Generalized Inverse Classification***Michael T. Lash, Qihang Lin, Nick Street, Jennifer Robinson, and Jeffrey Ohlmann, University of Iowa, USA***4:05-4:20 Predict Land Covers with Transition Modeling and Incremental Learning***Xiaowei Jia, University of Minnesota, Twin Cities, USA; Ankush Khandelw, Guruprasad Nayak, and James Gerber, University of Minnesota, USA; Kimberly Carlson, University of Hawaii, Manoa, USA; Paul West and Vipin Kumar, University of Minnesota, USA*

Thursday, April 27

**CP5****Health Informatics**

2:45 PM-4:25 PM

Room: Galleria Ballroom II

*Chair: Danai Koutra, University of Michigan, USA***2:45-3:00 Unified and Contrasting Graphical Lasso for Brain Network Discovery***Xinyue Liu and Xiangnan Kong, Worcester Polytechnic Institute, USA; Ann Ragin, Northwestern University, USA***3:05-3:20 T-Bne: Tensor-Based Brain Network Embedding***Bokai Cao, University of Illinois, Chicago, USA; Lifang He, South China University of Technology, China; Xiaokai Wei, University of Illinois at Chicago, USA; Mengqi Xing, Philip Yu, Heide Klumpp, and Alex Leow, University of Illinois, Chicago, USA***3:25-3:40 An Rnn Architecture with Dynamic Temporal Matching for Personalized Predictions of Parkinson's Disease***Chao Che, Dalian University of Technology, China; Cao Xiao, University of Washington and IBM T.J. Watson Research Center, USA; Jian Liang, Tsinghua University, P. R. China; Bo Jin, Dalian University of Technology, China; Jiayu Zhou, Michigan State University, USA; Fei Wang, IBM T.J. Watson Research Center, USA***3:45-4:00 Clustering with Domain-Specific Usefulness Scores***Yale Chang and Junxiang Chen, Northeastern University, USA; Michael Cho, Peter Castaldi, and Edwin Silverman, Harvard Medical School, USA; Jennifer Dy, Northwestern University, USA***4:05-4:20 Brainzoom: High Resolution Reconstruction from Multi-Modal Brain Signals***Xiao Fu, University of Minnesota, Twin Cities, USA; Kejun Huang, University of Minnesota, USA; Otilia Stretcu and Hyun Ah Song, Carnegie Mellon University, USA; Evangelos Papalexakis, University of California, Riverside, USA; Partha Talukdar, Indian Institute of Science, Bangalore, India; Tom Mitchell, Carnegie Mellon University, USA; Nicholas Sidiropoulos, University of Minnesota, USA; Christos Faloutsos, Carnegie Mellon University, USA*

Thursday, April 27

**CP6****Clustering and Recommender Systems**

2:45 PM-4:25 PM

Room: Tanglewood

*Chair: Leman Akoglu, Carnegie Mellon University, USA***2:45-3:00 Uncovering Group Level Insights with Accordant Clustering***Amit Dhurandhar, IBM Research, USA; Margareta Ackerman, San Jose State University, USA; Xiang Wang, Google, Inc., USA***3:05-3:20 A Method to Accelerate Human in the Loop Clustering***Anni Coden, Marina Danilevsky, Daniel Gruhl, Linda Kato, and Meenakshi Nagarajan, IBM Research, USA***3:25-3:40 Model-Based Von Mises-Fisher Co-Clustering with a Conscience***Aghiles Salah and Mohamed Nadif, Paris Descartes, France***3:45-4:00 Targeted Matrix Completion***Natali Ruchansky, University of Southern California, USA; Evimaria Terzi and Mark Crovella, Boston University, USA***4:05-4:20 Active Positive-Definite Matrix Completion***Charalampos Mavroforakis, Dora Erdos, Mark Crovella, and Evimaria Terzi, Boston University, USA***Organizational Break**

4:45 PM-5:00 PM

Thursday, April 27

## TS3

### Tutorial Session: Opportunities, Challenges and Methods for Higher Education Data Mining

5:00 PM-7:00 PM

Room: Post Oak

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

The application of big data approaches, specifically methods inspired from recommender system domain to predict student performance is largely a new area of research. The types of solutions available depend largely on the type of available data, and problem definition. For instance, for the purposes of degree planning, one task is to predict grades for a student in a class in the future (or in the next term). To predict student performance within individual class assessments, either in a typical classroom, an online, or a MOOC setting, students interaction with a learning management system (LMS) provides different ways to address this problem.

In this tutorial we will provide a formal definition of higher educational mining and discuss different opportunities for inter-disciplinary research in this field. Challenges as they relate to data acquisition and ethics will be discussed in detail followed with current practices surrounding the development of degree planning tools and course analytics to assist students and teachers.

**Huzefa Rangwala**

George Mason University, USA

**Aditya Johri**

George Mason University, USA

**Asmaa El Badrawy**

University of Minnesota, USA

**George Karypis**

University of Minnesota, USA

Thursday, April 27

## CP7

### Scalable Data Mining

5:00 PM-6:40 PM

Room: Galleria Ballroom I

Chair: George Karypis, University of  
Minnesota, USA

#### 5:00-5:15 Multi-Core K-Means

Christian Böhm, Ludwig-Maximilians-  
Universität München, Germany; Martin  
Perdacher and Claudia Plant, University of  
Vienna, Austria

#### 5:20-5:35 Indexing and Classifying Gigabytes of Time Series under Time Warping

Chang Wei Tan, Geoffrey I Webb, and  
Francois Petitjean, Faculty of Information  
Technology, Monash University, Australia

#### 5:40-5:55 A Sparse Nonlinear Classifier Design Using Auc Optimization

Vishal Kakkar and Shirish Shevade, Indian  
Institute of Science, Bangalore, India;  
S Sundararajan, Microsoft Research,  
India; Dinesh Garg, Indian Institute of  
Technology, Gandhinagar, India

#### 6:00-6:15 A Dual-Tree Algorithm for Fast $k$ -Means Clustering With Large $k$

Ryan Curtin, Symantec, USA

#### 6:20-6:35 Specious Rules: An Efficient and Effective Unifying Method for Removing Misleading and Uninformative Patterns in Association Rule Mining

Wilhelmiina Hämmäläinen, Aalto University,  
Finland; Geoff Webb, Monash University,  
Australia

Thursday, April 27

## CP8

### Mining Graphs I

5:00 PM-6:40 PM

Room: Galleria Ballroom II

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

#### 5:00-5:15 MeiKe: Influence-Based Communities in Networks

Yao Zhang, Bijaya Adhikari, Steve Jan, and  
B. Aditya Prakash, Virginia Tech, USA

#### 5:20-5:35 Signed Network Embedding in Social Media

Suhang Wang, Arizona State University,  
USA; Jiliang Tang, Michigan State  
University, USA; Charu Aggarwal, IBM  
T.J. Watson Research Center, USA; Yi  
Chang, Huawei Research, USA; Huan Liu,  
Arizona State University, USA

#### 5:40-5:55 Finding Low-Tension Communities

Esther Galbrun, Inria, France; Behzad  
Golshan, Recruit Institute of Technology,  
USA; Aristides Gionis, Aalto University,  
Finland; Evimaria Terzi, Boston University,  
USA

#### 6:00-6:15 Sensitivity of Community Structure to Network Uncertainty

Marc Mitri, Ecole Polytechnique, France;  
Fragkiskos D. Malliaros, University of  
Patras, Greece; Michalis Vazirgiannis,  
Ecole Polytechnique, France

#### 6:20-6:35 Subnetworks Mining with Spatial and Temporal Smoothness

Xuan-Hong Dang, Hongyuan You, Ambuj  
Singh, and Scott Grafton, University of  
California, Santa Barbara, USA



Thursday, April 27

**CP9****Recommender Systems**

5:00 PM-6:40 PM

Room: Tanglewood

Chair: Xia Ning, Purdue University, USA

**5:00-5:15 Selection of Negative Samples for One-Class Matrix Factorization**

Hsiang-Fu Yu, Amazon.com, USA; Mikhail Bilenko, Microsoft Corporation, USA; Chih-Jen Lin, National Taiwan University, Taiwan

**5:20-5:35 HBGG: a Hierarchical Bayesian Geographical Model for Group Recommendation**

Ziyu Lu, Hui Li, Nikos Mamoulis, and David Cheung, University of Hong Kong, Hong Kong

**5:40-5:55 Collaborative User Network Embedding for Social Recommender Systems**

Chuxu Zhang, Rutgers University, USA; Lu Yu, King Abdullah University of Science & Technology (KAUST), Saudi Arabia; Yan Wang and Chirag Shah, Rutgers University, USA; Xiangliang Zhang, King Abdullah University of Science & Technology (KAUST), Saudi Arabia

**6:00-6:15 Redundancies in Data and Their Effect on the Evaluation of Recommendation Systems: A Case Study on the Amazon Reviews Datasets**

Daniel Basaran, Ludwig-Maximilians-Universität München, Germany; Eirini C. Ntoutsi, Leibniz University Hannover, Germany; Arthur Zimek, University of Southern Denmark, Denmark

**6:20-6:35 Price Recommendation on Vacation Rental Websites**

Yang Li, Northwestern Polytechnical University, China; Suhan Wang, Arizona State University, USA; Tao Yang and Quan Pan, Northwestern Polytechnical University, China; Jiliang Tang, Michigan State University, USA

**Welcome Reception and Poster Session**

7:00 PM-9:00 PM



Room: Galleria Ballroom III and Galleria Foyer

Papers presented on Thursday and Saturday will have their poster slots during this session.

**Friday, April 28****Registration**

7:15 AM-3:30 PM

Room: Galleria Foyer

**Continental Breakfast**

7:30 AM-8:00 AM

Room: Galleria Foyer

**Announcements**

8:00 AM-8:15 AM

Room: Galleria Ballroom I

Friday, April 28

**IP3****Machine Learning under Resource Constraints**

8:15 AM-9:30 AM

Room: Galleria Ballroom I

Chair: Dimitrios Gunopulos, University of Athens, Greece

Big data are produced by various sources. Most often, they are distributedly stored at computing farms or clouds. Analytics on the Hadoop Distributed File System (HDFS) then follows the MapReduce programming model. According to the Lambda architecture of Nathan Marz and James Warren, this is the batch layer. It is complemented by the speed layer, which aggregates and integrates incoming data streams in real time. When considering big data and small devices, obviously, we imagine the small devices being hosts of the speed layer, only. Analytics on the small devices is restricted by memory and computation resources. The interplay of streaming and batch analytics offers a multitude of configurations. In this talk, we discuss opportunities for using sophisticated models for learning spatio-temporal models. In particular, we investigate graphical models, which generate the probabilities for connected (sensor) nodes. We even approximate likelihood estimates such that they can be computed on very restricted devices.

Katharina Morik

Technische Universität Dortmund, Germany

**Coffee Break**

9:30 AM-10:00 AM

Room: Galleria Foyer



Friday, April 28

## TS4

### Tutorial Session: IoT Big Data Stream Mining

10:00 AM-12:00 PM

Room: Post Oak

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

The challenge of deriving insights from the Internet of Things (IoT) has been recognized as one of the most exciting and key opportunities for both academia and industry. Advanced analysis of big data streams from sensors and devices is bound to become a key area of data mining research as the number of applications requiring such processing increases. Dealing with the evolution over time of such data streams, i.e., with concepts that drift or change completely, is one of the core issues in IoT stream mining. This tutorial is a gentle introduction to mining IoT big data streams. The first part introduces data stream learners for classification, regression, clustering, and frequent pattern mining. The second part deals with scalability issues inherent in IoT applications, and discusses how to mine data streams on distributed engines such as Spark, Flink, Storm, and Samza.

**Gianmarco De Francisci Morales**

Qatar Computing Research Institute, Qatar

**Albert Bifet**

Télécom ParisTech, France

**Abdul R. Khan**

King Fahd University of Petroleum and Minerals, Saudi Arabia

**Joao Gama**

University of Porto, Portugal

**Wei Fan**

Baidu Research Big Data Lab, China

Friday, April 28

## CP10

### Social Network Analysis

10:00 AM-11:40 AM

Room: Galleria Ballroom I

Chair: B. Aditya Prakash, Virginia Tech, USA

#### 10:00-10:15 Ranking in Heterogeneous Networks with Geo-Location Information

Abhinav Mishra, Stony Brook University, USA; Leman Akoglu, Carnegie Mellon University, USA

#### 10:20-10:35 Condensing Temporal Networks Using Propagation

Bijaya Adhikari, Yao Zhang, Aditya Bharadwaj, and B. Aditya Prakash, Virginia Tech, USA

#### 10:40-10:55 Community-Aware Network Sparsification

Aristides Gionis and Rozenshtein Polina, Aalto University, Finland; Nikolaj Tatti, Helsinki Institute for Information Technology HIIT, Finland; Evimaria Terzi, Boston University, USA

#### 11:00-11:15 Graph-Based Semi-Supervised Learning for Relational Networks

Leto Peel, Université Catholique de Louvain, Belgium

#### 11:20-11:35 Toward Personalized Relational Learning

Jundong Li and Liang Wu, Arizona State University, USA; Osmar Zaiane, University of Alberta, Canada; Huan Liu, Arizona State University, USA

Friday, April 28

## CP11

### Text and Web Mining

10:00 AM-11:40 AM

Room: Galleria Ballroom II

Chair: Bianca Zadrozny, IBM Research, USA

#### 10:00-10:15 Biclustering: An Application of Dual Topic Models

Daniel Rugeles, Kaiqi Zhao, and Cong Gao, Nanyang Technological University, Singapore; Manoranjan Dash and Shonali Krishnaswamy, Agency for Science Technology and Research, Singapore

#### 10:20-10:35 A Graduated Non-Convexity Relaxation for Large Scale Seriation

Xenophon Evangelopoulos and Austin Brockmeier, University of Liverpool, United Kingdom; Tingting Mu, University of Manchester, United Kingdom; John Goulernas, University of Liverpool, United Kingdom

#### 10:40-10:55 User-Guided Cross-Domain Sentiment Classification

Arun Reddy Nelakurthi, Hang Hang Tong, Ross Maciejewski, Nadya Bliss, and Jingrui He, Arizona State University, USA

#### 11:00-11:15 Exploring Latent Semantic Factors to Find Useful Product Reviews

Subhabrata Mukherjee, Max-Planck-Institut fuer Informatik, Germany; Kashyap Popat and Gerhard Weikum, Max Planck Institute for Informatics, Germany

#### 11:20-11:35 Outlier Detection for Text Data

Ramakrishnan Kannan, Oak Ridge National Laboratory, USA; Hyenkyun Woo, Korea University of Technology and Education, South Korea; Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Haesun Park, Georgia Institute of Technology, USA

Friday, April 28

**CP12****Feature, Extraction, Selection, and Dimension Reduction**

10:00 AM-11:40 AM

*Room: Tanglewood**Chair: Jiayu Zhou, Michigan State University, USA***10:00-10:15 Roflmao: Robust Oblique Forests with Linear Matrix Operations***Tyler M. Tomita, Johns Hopkins University, USA***10:20-10:35 Exploiting Hierarchical Structures for Unsupervised Feature Selection***Suhang Wang and Yilin Wang, Arizona State University, USA; Jiliang Tang, Michigan State University, USA; Charu C. Aggarwal, IBM T.J. Watson Research Center, USA; Suhas Ranganath and Huan Liu, Arizona State University, USA***10:40-10:55 Embedded Supervised Feature Selection for Multi-Class Data***Lin Chen, Arizona State University, USA; Jiliang Tang, Michigan State University, USA; Baoxin Li, Arizona State University, USA***11:00-11:15 Correlation by Compression***Kailash Budhathoki and Jilles Vreeken, Max Planck Institute for Informatics, Germany***11:20-11:35 A Deflation Method for Structured Probabilistic PCA***Rajiv Khanna and Joydeep Ghosh, University of Texas at Austin, USA; Russell Poldrack, Stanford University, USA; Oluwasanmi Koyejo, University of Illinois, Urbana-Champaign, USA***Lunch Break**

12:00 PM-1:15 PM

*Attendees on their own*

Friday, April 28

**IP4****Repeated Choice, Markov Models, and LAMP**

1:15 PM-2:30 PM

*Room: Galleria Ballroom I**Chair: Carlotta Domeniconi, George Mason University, USA*

In this talk I'll discuss modeling of user consumption data, using the framework of Discrete Choice. I'll start with some background on choice models, survey the key ideas, then give an example modeling repeated consumption of the same item (for instance a song, restaurant, or web page). I'll then show how choice theory can be integrated with first-order markov models to develop lightweight but powerful "LAMP" sequence models that efficiently learn how to take history into account. Finally, I'll show some comparisons between these models and standard deep network sequence models.

**Andrew Tomkins***Google, Inc., USA***Coffee Break**

2:30 PM-2:45 PM

*Room: Galleria Foyer*

Friday, April 28

**TS4****Tutorial Session: IoT Big Data Stream Mining , Continued**

2:45 PM-3:45 PM

*Room: Post Oak**Chair: Jilles Vreeken, Saarland University and Max Planck Institute for Informatics, Germany*

The challenge of deriving insights from the Internet of Things (IoT) has been recognized as one of the most exciting and key opportunities for both academia and industry. Advanced analysis of big data streams from sensors and devices is bound to become a key area of data mining research as the number of applications requiring such processing increases. Dealing with the evolution over time of such data streams, i.e., with concepts that drift or change completely, is one of the core issues in IoT stream mining. This tutorial is a gentle introduction to mining IoT big data streams. The first part introduces data stream learners for classification, regression, clustering, and frequent pattern mining. The second part deals with scalability issues inherent in IoT applications, and discusses how to mine data streams on distributed engines such as Spark, Flink, Storm, and Samza.

**Gianmarco De Francisci Morales***Qatar Computing Research Institute, Qatar***Albert Bifet***Télécom ParisTech, France***Abdul R. Khan***King Fahd University of Petroleum and Minerals, Saudi Arabia***Joao Gama***University of Porto, Portugal***Wei Fan***Baidu Research Big Data Lab, China*

Friday, April 28

## CP13

### Novel Applications

2:45 PM-4:45 PM

Room: Galleria Ballroom I

Chair: Xuan-Hong Dang, University of California, Santa Barbara, USA

#### 2:45-3:00 Identifying Deep Contrasting Networks from Time Series Data: Application to Brain Network Analysis

John Boaz Lee, Xiangnan Kong, and Yihan Bao, Worcester Polytechnic Institute, USA; Constance Moore, University of Massachusetts Medical School, USA

#### 3:05-3:20 Cumulative Knowledge-based Regression Models for Next-term Grade Prediction

Sara Morsy, University of Minnesota, USA; George Karypis, University of Minnesota, USA

#### 3:25-3:40 BreachRadar: Automatic Detection of Points-of-Compromise

Miguel Araujo, Carnegie Mellon University, USA; Miguel Almeida, Jaime Ferreira, Luis Silva, and Pedro Bizarro, Feedzai, Portugal

#### 3:45-4:00 Hidden: Hierarchical Dense Subgraph Detection with Application to Financial Fraud Detection

Si Zhang, Dawei Zhou, and Mehmet Yigit Yildirim, Arizona State University, USA; Scott Alcorn, Early Warnings LLC, USA; Jingrui He, Hasan Davulcu, and Hanghang Tong, Arizona State University, USA

#### 4:05-4:20 MultiC<sup>2</sup>: An Optimization Framework for Learning from Task and Worker Dual Heterogeneity

Yao Zhou, Lei Ying, and Jingrui He, Arizona State University, USA

#### 4:25-4:40 Uplift Modeling with Multiple Treatments and General Response Types

Yan Zhao, Xiao Fang, and David Simchi-Levi, Massachusetts Institute of Technology, USA

Friday, April 28

## CP14

### Mining Graphs II

2:45 PM-4:45 PM

Room: Galleria Ballroom II

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

#### 2:45-3:00 FACETS: Adaptive Local Exploration of Large Graphs

Robert Pienta and Minsuk Kahng, Georgia Institute of Technology, USA; Zhiyuan Lin, Stanford University, USA; Jilles Vreeken, Max Planck Institute for Informatics, Germany; Partha Talukdar, Indian Institute of Science, Bangalore, India; James Abello, Rutgers University, USA; Ganesh Parameswaran, Yahoo! Inc., USA; Duen Horng Chau, Georgia Institute of Technology, USA

#### 3:05-3:20 Absenteeism Detection in Social Media

Fang Jin, Virginia Tech, USA

#### 3:25-3:40 Multimodal Network Alignment

Huda Nassar and David F. Gleich, Purdue University, USA

#### 3:45-4:00 Near-Optimal and Practical Algorithms for Graph Scan Statistics

Jose Cadena and Anil Vullikanti, Virginia Tech, USA; Feng Chen, University of Albany - State University of New York, USA

#### 4:05-4:20 Accelerated Attributed Network Embedding

Xiao Huang, Texas A&M University, USA; Jundong Li, Arizona State University, USA; Xia Hu, Texas A&M University, USA

#### 4:25-4:40 Meta-Path Graphical Lasso for Learning Heterogeneous Connectivities

Yao Zhang and Yun Xiong, Fudan University, China; Xinyue Liu and Xiangnan Kong, Worcester Polytechnic Institute, USA; Yangyong Zhu, Fudan University, China

Friday, April 28

## CP15

### Novel Learning Methods

2:45 PM-4:45 PM

Room: Tanglewood

Chair: Huan Liu, Arizona State University, USA

#### 2:45-3:00 Statistical Learning Theory Approach for Data Classification with $\ell$ -Diversity

Koray Mancuhan and Chris Clifton, Purdue University, USA

#### 3:05-3:20 Private and Right-Protected Big Data Publication: An Analysis

Michail Vlachos, IBM Research, USA; Reinhard Heckel, University of California, Berkeley, USA

#### 3:25-3:40 Differentially Private Rank Aggregation

Michael Hay, Colgate University, USA; Liudmila Elagina and Gerome Miklau, University of Massachusetts, Amherst, USA

#### 3:45-4:00 Hash-Based Feature Learning for Incomplete Continuous-Valued Data

Shuai Yuan, Pang-Ning Tan, Kendra Cheruvilil, Emi Last Name, Nicholas Skaff, and Patricia Soranno, Michigan State University, USA

#### 4:05-4:20 Multi-Task Multiple Kernel Relationship Learning

Keerthiram Murugesan and Jaime Carbonell, Carnegie Mellon University, USA

#### 4:25-4:40 Multivariate Confidence Intervals

Jussi Korpela, Emilia Oikarinen, Kai Puolamäki, and Antti Ukkonen, Finnish Institute of Occupational Health, Finland

## Organizational Break

4:45 PM-5:00 PM

Friday, April 28

**PD1****Data Mining, Clinical Medicine, and Medical Research: The Current State and Future of the Nexus**

5:00 PM-6:15 PM

Room: Galleria Ballroom I

Chair: Chris Jermaine, Rice University, USA

**Award Ceremony and SIAG/DMA Business Meeting**

6:15 PM-7:00 PM

Room: Galleria Ballroom I

Complimentary beer and wine will be served.

**Doctoral Forum and Poster Session**

7:00 PM-9:00 PM

Room: Galleria Ballroom III and Galleria Foyer

Papers presented on Friday will have their poster slots during the Doctoral Forum Session.

**Saturday, April 29****Registration**

7:15 AM-4:00 PM

Room: Galleria Foyer

**Continental Breakfast**

7:30 AM-8:00 AM

Room: Galleria Foyer



Saturday, April 29

**CP16****Optimization Methods**

8:00 AM-9:40 AM

Room: Galleria Ballroom I

Chair: Xiangnan Kong, Worcester Polytechnic Institute, USA

**8:00-8:15 A Fast Trust-Region Newton Method for Softmax Logistic Regression**

Nayyar Zaidi and Geoff Webb, Monash University, Australia

**8:20-8:35 Learning from Multi-Modality Multi-Resolution Data: An Optimization Approach**

Yada Zhu, IBM T.J. Watson Research Center, USA; Jianbo Li, Three Bridges Capital, USA; Jingrui He, Arizona State University, USA

**8:40-8:55 Sparse Graphical Modeling Via Stochastic Complexity**

Kohei Miyaguchi, Shin Matsushima, and Kenji Yamanishi, University of Tokyo, Japan

**9:00-9:15 Limited-Memory Common-Directions Method for Distributed Optimization and Its Application on Empirical Risk Minimization**

Ching-Pei Lee, University of Wisconsin, Madison, USA; Po-Wei Wang, Carnegie Mellon University, USA; Weizhu Chen, Microsoft, USA; Chih-Jen Lin, National Taiwan University, Taiwan

**9:20-9:35 Automatic Frankensteining: Creating Complex Ensembles Autonomously**

Martin Wistuba, Nicolas Schilling, and Lars Schmidt-Thieme, University of Hildesheim, Germany



Saturday, April 29

**Workshop 2 (half day): Data Mining for Medicine and Healthcare**

8:00 AM-12:15 PM

Room: Tanglewood

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

8:00 AM-5:45 PM

Room: Galleria Ballroom II

**Workshop 4 (full day): Inferring Networks From Non-Network Data**

8:00 AM-5:45 PM

Room: Post Oak

**Workshop 5 (full day): Mining Big Data in Climate and Environment**

8:00 AM-5:45 PM

Room: Bellaire

**Workshop 6 (full day): Women in Data Science**

8:00 AM-5:45 PM

Room: San Felipe

**Coffee Break**

10:00 AM-10:15 AM

Room: Galleria Foyer



Saturday, April 29

**CP17**

**Data Stream Mining**

10:15 AM-11:55 AM

Room: Galleria Ballroom I

Chair: Kun Zhang, Xavier University, USA

**10:15-10:30 Error Metrics for Learning Reliable Manifolds from Streaming Data**

Frank Schoeneman, Suchismit Mahapatra, Varun Chandola, Nils Napp, and Jaroslaw Zola, University of Buffalo, USA

**10:35-10:50 CSTG: An Effective Framework for Cost-Sensitive Sparse Online Learning**

Zhong Chen, Xavier University, USA; Zhide Fang, LSU Health Sciences Center, USA; Wei Fan, Baidu Research Big Data Lab, China; Andrea Edwards and Kun Zhang, Xavier University, USA

**10:55-11:10 Concept Drift Detection with Hierarchical Hypothesis Testing**

Shujian Yu, University of Florida, USA; Zubin Abraham, Robert Bosch LLC, USA

**11:15-11:30 Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting**

Rose Yu, Yaguang Li, Cyrus Shahabi, Ugur Demiryurek, and Yan Liu, University of Southern California, USA

**11:35-11:50 H-Fuse: Efficient Fusion of Aggregated Historical Data**

Zongge Liu and Hyun Ah Song, Carnegie Mellon University, USA; Vladimir Zadorozhny, University of Pittsburgh, USA; Christos Faloutsos, Carnegie Mellon University, USA; Nicholas Sidiropoulos, University of Minnesota, USA

**Lunch Break**

12:15 PM-1:30 PM

Attendees on their own

Saturday, April 29

**Workshop 3 (half-day): Workshop on Data Mining for Oil & Gas**

1:30 PM-5:45 PM

Room: Tanglewood



Saturday, April 29

## CP18

### Spatial-Temporal Mining

1:30 PM-3:10 PM

Room: Galleria Ballroom I

Chair: Jilles Vreeken, Max Planck Institute for Informatics, Germany

#### 1:30-1:45 Efficiently Summarising Event Sequences with Rich Interleaving Patterns

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

#### 1:50-2:05 Discovery of Causal Time Intervals

Zhenhui Li, Guanjie Zheng, Amal Agarwal, Lingzhou Xue, and Thomas Lauvaux, Pennsylvania State University, USA

#### 2:10-2:25 Robust Map Matching for Heterogeneous Data Via Dominance Decompositions

Martin P. Seybold, University of Stuttgart, Germany

#### 2:30-2:45 Uncovering the Spatiotemporal Patterns of Collective Social Activity

Martin Jankowiak, Courant Institute of Mathematical Sciences, New York University, USA; Manuel Gomez-Rodriguez, Max Planck Institute for Software Systems, Germany

#### 2:50-3:05 Discovering Bursts Revisited: Guaranteed Optimization of the Model Parameters

Nikolaj Tatti, Helsinki Institute for Information Technology HIIT, Finland

## Coffee Break

3:30 PM-3:45 PM



Room: Galleria Foyer

Saturday, April 29

## TS5

### Tutorial Session: Summarizing Large-Scale Graph Data

3:45 PM-5:45 PM

Room: Post Oak

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

Recent advances in computing resources have made possible collecting enormous amounts of data, such as social media interactions, web browsing, product and service purchases, autonomous vehicle routing, activities via smart home sensors and health / wellness sensors, and more. Since summarization helps humans find structure and meaning in data, the data mining community has taken a strong interest in the task and accordingly has proposed many methods for a variety of data types. This tutorial aims to provide a comprehensive overview of summarization techniques for large-scale graphs, which are very prevalent due to their intrinsic ability to represent many natural phenomena and encode relationships between entities.

Graph summarization, compression and pattern discovery are tightly coupled: describing a graph succinctly leads to the discovery of interesting patterns, as well as to the detection of deviations or outliers. The objective of this tutorial is to give a systematic overview of graph summarization methods for static and dynamic networks with main emphasis on the important concepts, intuition, and main objectives, describe successful real-world applications, and present the open research problems in the field. Connections to compression and pattern/outlier discovery will be drawn throughout the tutorial.

#### Danai Koutra

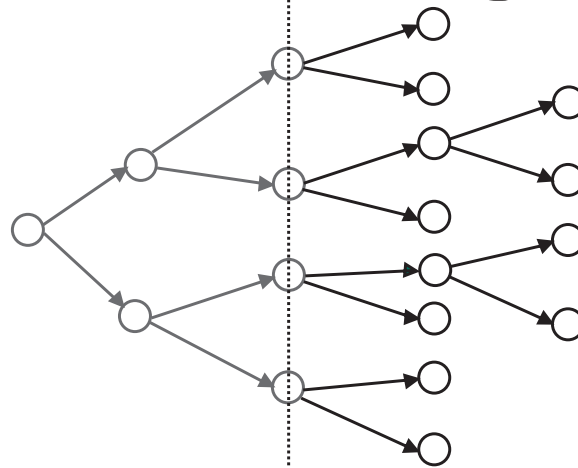
University of Michigan, USA

## Notes

# SDM17 Abstracts

2017 SIAM  
International Conference  
on **DATA MINING**

April 27 - 29, 2017



The Westin Galleria Houston  
Houston, Texas, USA

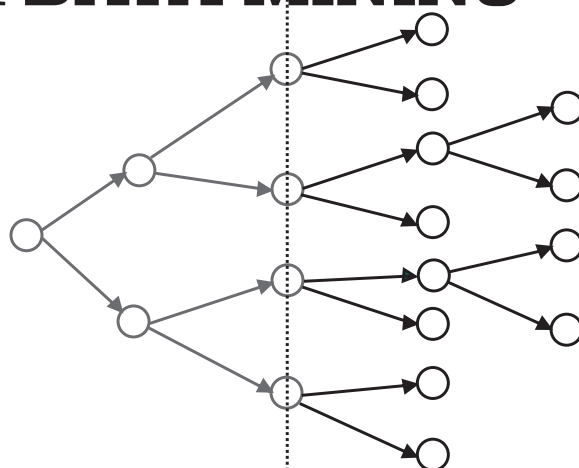
**Abstracts are printed as submitted by the author.**

## Notes

## Organizer and Speaker Index

### 2017 SIAM International Conference on **DATA MINING**

April 27 - 29, 2017



The Westin Galleria Houston  
Houston, Texas, USA

**A**

Abraham, Zubin, CP17, 10:55 Sat  
 Adhikari, Bijaya, CP10, 10:20 Fri  
 Akoglu, Leman, CP10, 10:00 Fri  
 Araujo, Miguel, CP13, 3:25 Fri

**B**

Bhattacharyya, Apratim, CP18, 1:30 Sat  
 Budhathoki, Kailash, CP12, 11:00 Fri

**C**

Cadena, Jose, CP14, 3:45 Fri  
 Cao, Bokai, CP5, 3:05 Thu  
 Chang, Yale, CP5, 3:45 Thu  
 Che, Chao, CP5, 3:25 Thu  
 Chen, Jinghui, CP3, 10:00 Thu  
 Chen, Lin, CP12, 10:40 Fri  
 Chen, Zhong, CP17, 10:35 Sat  
 Chino, Daniel, CP3, 10:40 Thu  
 Curtin, Ryan, CP7, 6:00 Thu

**D**

Dang, Xuan-Hong, CP8, 6:20 Thu  
 Danilevsky, Marina, CP6, 3:05 Thu  
 Davoudi, Heidar, CP4, 2:45 Thu  
 Dhurandhar, Amit, CP6, 2:45 Thu

**E**

Eswaran, Dhivya, CP4, 3:05 Thu  
 Evangelopoulos, Xenophon, CP11, 10:20 Fri

**F**

Fan, Wei, TS4, 10:00 Fri  
 Fan, Wei, TS4, 2:45 Fri  
 Fu, Xiao, CP5, 4:05 Thu  
 Fugate, William Craig, IP1, 8:15 Thu

**G**

Galbrun, Esther, CP8, 5:40 Thu  
 Gomez-Rodriguez, Manuel, CP18, 2:30 Sat

**H**

Hämäläinen, Wilhelmiina, CP7, 6:20 Thu  
 Hay, Michael, CP15, 3:25 Fri  
 He, Jingrui, CP16, 8:20 Sat  
 Heckel, Reinhard, CP15, 3:05 Fri  
 Hidalgo, Cesar, IP2, 1:15 Thu  
 Hu, Qiong, CP2, 10:40 Thu  
 Huang, Xiao, CP14, 4:05 Fri

**J**

*Jermaine, Chris, PD1, 5:00 Fri*  
 Jia, Xiaowei, CP4, 4:05 Thu  
 Jin, Fang, CP14, 3:05 Fri

**K**

Kakkar, Vishal, CP7, 5:40 Thu  
 Kannan, Ramakrishnan, CP11, 11:20 Fri  
 Khanna, Rajiv, CP12, 11:20 Fri  
 Korpela, Jussi, CP15, 4:25 Fri  
 Koutra, Danai, TS5, 3:45 Sat

**L**

Lash, Michael T., CP4, 3:45 Thu  
 Lee, Ching-Pei, CP16, 9:00 Sat  
 Lee, John Boaz, CP13, 2:45 Fri  
 Li, Jundong, CP10, 11:20 Fri  
 Li, Yaguang, CP17, 11:15 Sat  
 Li, Zhenhui, CP18, 1:50 Sat  
 Liu, Honglei, CP2, 11:20 Thu  
 Liu, Xinyue, CP5, 2:45 Thu  
 Liu, Yashu, CP2, 10:00 Thu  
 Liu, Zongge, CP17, 11:35 Sat  
 Lu, Ziyu, CP9, 5:20 Thu

**M**

Malliaros, Fragkiskos D., CP8, 6:00 Thu  
 Mancuhan, Koray, CP15, 2:45 Fri  
 Mavroforakis, Charalampos, CP6, 4:05 Thu

Mcbride, Ryan, CP1, 11:20 Thu  
 Miettinen, Pauli, TS1, 10:00 Thu  
 Miyaguchi, Kohei, CP16, 8:40 Sat  
 Morik, Katharina, IP3, 8:15 Fri  
 Morsy, Sara, CP13, 3:05 Fri  
 Mukherjee, Subhabrata, CP11, 11:00 Fri  
 Murugesan, Keerthiram, CP15, 4:05 Fri

**N**

Nassar, Huda, CP14, 3:25 Fri  
 Nelakurthi, Arun Reddy, CP11, 10:40 Fri  
 Nie, Zhi, CP1, 10:20 Thu  
 Ntoutsis, Eirini C., CP9, 6:00 Thu

**P**

Peel, Leto, CP10, 11:00 Fri  
 Perdacher, Martin, CP7, 5:00 Thu  
 Perros, Ioakeim, CP2, 11:00 Thu  
 Pienta, Robert, CP14, 2:45 Fri  
 Polina, Rozenshtein, CP10, 10:40 Fri  
 Prakash, B. Aditya, TS2, 2:45 Thu

**R**

Rangwala, Huzefa, TS3, 5:00 Thu  
 Ruchansky, Natali, CP6, 3:45 Thu  
 Rugeles, Daniel, CP11, 10:00 Fri

**S**

Salah, Aghiles, CP6, 3:25 Thu  
 Schoeneman, Frank, CP17, 10:15 Sat  
 Seybold, Martin P., CP18, 2:10 Sat  
 Siebes, Arno, CP3, 11:20 Thu

**T**

Tan, Chang Wei, CP7, 5:20 Thu  
 Tang, Jiliang, CP8, 5:20 Thu  
 Tatti, Nikolaj, CP18, 2:50 Sat  
 Tomita, Tyler M., CP12, 10:00 Fri  
 Tomkins, Andrew, IP4, 1:15 Fri



**W**

Wang, Suhang, CP1, 10:00 Thu  
Wang, Yilin, CP12, 10:20 Fri  
Wistuba, Martin, CP16, 9:20 Sat  
Wu, Liang, CP3, 10:20 Thu

**X**

Xiao, Houping, CP3, 11:00 Thu  
Xue, Yanbing, CP1, 11:00 Thu

**Y**

Yang, Tao, CP9, 6:20 Thu  
Yu, Hsiang-Fu, CP9, 5:00 Thu  
Yuan, Shuai, CP15, 3:45 Fri

**Z**

Zaidi, Nayyar, CP16, 8:00 Sat  
Zhang, Chuxu, CP9, 5:40 Thu  
Zhang, Daoqiang, CP2, 10:20 Thu  
Zhang, Daoqiang, CP1, 10:40 Thu  
Zhang, Si, CP13, 3:45 Fri  
Zhang, Yao, CP8, 5:00 Thu  
Zhang, Yao, CP14, 4:25 Fri  
Zhang, Yexun, CP4, 3:25 Thu  
Zhao, Yan, CP13, 4:25 Fri  
Zhou, Yao, CP13, 4:05 Fri

## Notes

## SDM17 Budget

### Conference Budget

### SIAM International Conference on Data Mining

**April 27 - 29, 2017**

**Houston, TX**

**Expected Paid Attendance** 225

#### Revenue

Registration Income		\$109,810.00
	Total	<u>\$109,810.00</u>

#### Expenses

Printing	\$1,100.00
Organizing Committee	\$2,900.00
Invited Speakers	\$8,350.00
Food and Beverage	\$37,000.00
AV Equipment and Telecommunication	\$15,500.00
Advertising	\$6,500.00
Proceedings	\$6,000.00
Conference Labor (including benefits)	\$37,095.00
Other (supplies, staff travel, freight, misc.)	\$5,050.00
Administrative	\$12,101.00
Accounting/Distribution & Shipping	\$6,614.00
Information Systems	\$12,161.00
Customer Service	\$4,557.00
Marketing	\$7,033.00
Office Space (Building)	\$4,649.00
Other SIAM Services	\$4,845.00
	<hr/>
Total	\$171,455.00

Net Conference Expense -\$61,645.00

Support Provided by SIAM	<u>\$61,645.00</u>
	<u>\$0.00</u>

#### Estimated Support for Travel Awards not included above:

Early Career and Students	18	\$13,500.00
---------------------------	----	-------------

# Westin Galleria Houston Floor Plan

## THE WESTIN GALLERIA & WESTIN OAKS HOUSTON

### THIRD LEVEL

