ICDM 2010 Program

IEEE International Conference on Data Mining

10th Anniversary
2001 – 2010

Sydney, Australia
December 14 - 17

Conference Sponsors

Platinum Sponsorship:

National Science Foundation

TOMTOM

Gold Sponsorship:

UNIVERSITY OF TECHNOLOGY SYDNEY

UTS:QCIS — QUANTUM COMPUTATION & INTELLIGENT SYSTEMS

IEEE COMPUTER SOCIETY

Silver Sponsorship:

RULEQUEST RESEARCH — data mining tools

UNSW COMPUTING — THE UNIVERSITY OF NEW SOUTH WALES

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH — UNITED STATES AIR FORCE

ASIAN OFFICE OF AEROSPACE R&D — BRIDGING SCIENCE & TECHNOLOGY

Bronze Sponsorship:

THE UNIVERSITY OF SYDNEY

ei² — ARC Research Network in Enterprise Information Infrastructure

NASA

IBM Research

Capital Markets CRC Limited

HONORARY PATRONAGE OF THE MAYOR OF WARSAW

KNIME — Data Analytics Made Easy

CSIRO

sas — THE POWER TO KNOW.

Table of Contents

Message from the Conference and Program Chairs

ICDM 2010

Welcome to the 10th IEEE International Conference on Data Mining which will be held in Sydney, Australia from 14 to 17 December 2010. While this is the tenth instalment of the iconic ICDM conference series, it is the first one held in the southern hemisphere and consequently, it is the first to be held in the summer, even though it is in December. You will know Sydney by its iconic landmarks, the Sydney Opera House, the Sydney Harbour Bridge, and the 2000 Summer Olympic Games. December in Sydney is also the time to enjoy a warm and sunny Christmas.

ICDM 2010 continues the series' tradition of stimulating technical, workshop, and contest paper presentations, as well as tutorial presentations and demonstrations. To celebrate the 10th anniversary of ICDM, this year's conference also includes events that look at the state of Data Mining research from different perspectives – specifically, a panel which focuses on Top-10 Data Mining Case Studies, a keynote that focuses on the past 10 years, a forum devoted to Early Career Researchers, and a forum which will explore interactions between academic researchers, industry practitioners, and government end-users in data mining. We hope these specialised academic activities will create a platform for data mining researchers to share their problems, ideas, and solutions in data mining research and its applications.

As well as academic activities, the conference will also offer delegates various social activities. Morning and afternoon teas and lunch will be provided daily, as well as a reception, a boat cruise with banquet and live show, and the choice of either a City, or an Opera House tour. In addition, the conference offers every delegate an individual Internet account with unlimited bandwidth so that all participants will be able to maintain connections with their network as if they were in their own office.

The success of any big conference depends on the hard work of a dedicated team. ICDM 2010 is no exception, and the team members deserve acknowledgment. First, we would like to thank the Honorary Chair, **Ramamohanarao Kotagiri,** for bringing this conference to Sydney. Our thanks and gratitude are also due to the Program Committee Vice Chairs and members, who undertook the enormous task of advising Geoff and Bing in choosing the 155 best papers from nearly 800 submissions. We would like to thank the Workshop Co-Chairs, **Wynne Hsu** and **Wei Fan**, the Tutorial Co-Chairs, **Sanjay Chawla** and **Myra Spiliopoulou**, the Contest Chair, **Demetris Zeinalipour**, the Exhibits and Demos Co-Chairs, **Bart Goethals** and **Yucel Saygin**, the Panels Chair, **Hillol Kargupta**, and the Award Chair, **Wei Wang**, for their hard work and contribution in creating an excellent program for the delegates.

We would also like to acknowledge the behind-the-scenes contributions, including the Publicity Co-Chairs **Jian Yang**, **Taghi Khoshgoftaar**, and **Wenjie Zhang**, who assisted ICDM 2010 to attract a record of nearly 800 submissions; Sponsorship Co-Chairs **Gabor Melli** and **Paul Beinat** who attracted substantial conference sponsors; the webmaster, **Xinhua Zhu** for timely updates of the webpage, the secretary **Li Liu** who efficiently handled registrations; and Student Travel Award Chair **Wei Ding** who selected those students most deserving of travel awards.

We would like to especially acknowledge the Local Arrangements Chair, **Longbing Cao**, and Finance Chair, **Xingquan Zhu**, who have managed all logistical arrangements including venue booking, hotels, lunches, banquets, tours, student volunteer coordination, etc. They have spent considerable time ensuring the conference will run smoothly. We would like to thank the ICDM Steering Committee for their guidance and help. Special thanks also go to the ICDM Steering Committee Chair, **Xindong Wu**, who has generously offered help, guidance and solutions at all levels to this conference series since it was initiated in 2001.

Special acknowledgement also goes to the financial sponsors, without whose support we could not have organised this conference. They are: the US **National Science Foundation (NSF)**, **TomTom International**, **IEEE Computer Society**, **QCIS Centre (UTS)**, **Advanced Analytics Institute (UTS)**, **Rulequest Research** (Data Mining Tools), **UNSW School of Computer Science and Engineering**, **Asian Office of Aerospace Research & Development** (US Air Force Office of Scientific Research), **University of Sydney**, **ARC Research Network in Enterprise Information Infrastructure**, **NASA**, **IBM Research**, **Capital Markets CRC**, **KNIME** (Data Analytics Made Easy), **CSIRO**, **SAS**, and the **Honorary Patronage of the Mayor of Warsaw**. The direct impact of sponsors from NSF, the IEEE Computer Society, and

IBM Research has enabled partial financial support of 40 students to present their papers at the conference.

Finally, our thanks to the Centre for Quantum Computation & Intelligent Systems, and the Advanced Analytics Institute at UTS for hosting this conference, as well as to the IEEE Computer Society for their sponsorship. We appreciate the contributions of all the student volunteers who have offered their time to help make the conference a successful and memorable event.

Chengqi Zhang
Dimitrios Gunopulos
ICDM 2010 General Co-Chairs

Geoff Webb
Bing Liu
ICDM 2010 Program Committee Co-Chairs

December 2010

General Information

Your Badge

Each badge carries the name and affiliation of the badge holder. Admission to the conference and workshop sessions is by badge only. If you lose your badge, please go to the Registration Desk for a replacement.

Lunches

Lunches, in general from 12:20-13:30 (except 17 December), are included in the conference registration. You need to carry your badge for lunches. The location is at the Marigold Restaurant, Level 5, 683-689 George Street.

Conference Registration

December 13, 4:00pm – 8:00pm, Lobby Block C, Building 5, UTS Haymarket Campus
December 14, 8:00 – 6:00pm, Lobby Block C, Building 5, UTS Haymarket Campus
December 15, 8:00 – 6:00pm, Lobby Block B, Building 5, UTS Haymarket Campus
December 16, 8:00 – 12:00pm, Lobby Block B, Building 5, UTS Haymarket Campus
December 17, 8:00 – 12:00pm, Lobby Block B, Building 5, UTS Haymarket Campus

Conference Reception

December 15, 6:30pm – 8:30pm, Lobby Block D, Building 5, UTS Haymarket Campus

Conference Banquet/Award Ceremony

December 16, 6:30pm – 10:00pm, Sydney Harbour Dinner Cruise and Live Show.

Please meet at 6.30pm at Campbell Cove Wharf, Sydney Harbour foreshore. Campbell Cove is opposite the Opera House – walk past Circular Quay and follow the foreshore (waterfront) around past the historic Rocks to Campbell Cove (see map, page 56). Your Showboat Two* banquet dinner (look for the red and white paddle-wheeler) will board at 7.00pm for a 7.15pm departure, the boat won't wait for anyone after 7.15pm sharp (please bring your banquet ticket).* **Note**: 'Showboat Two' is the official name of the company who runs the boat and show. There is only one boat – a red and white paddle-wheeler.

Morning and Afternoon Teas

Workshop morning and afternoon teas are served at Lobby Block D, Building 5, UTS Haymarket Campus. Conference morning and afternoon teas are served at Lobby Block D, Building 5.

Volunteers

You may ask Volunteers (see page 51) for help for any questions. They will be happy to help you.

Language

The conference and all its activities will be conducted in English.

Smoking Policy

No Smoking is permitted in the ICDM'10 rooms.

General Inquiries
If you have any queries regarding the conference organization, please contact the conference Local Arrangements Chair: Prof Longbing Cao at longbing.cao-1@uts.edu.au, or 0423 255 257.

Information for Session Chairs and Presenters

Facilities at the Presentation Room

All rooms are equipped with a desktop and digital projector connected to the desktop.

Presentation Time

The presentation time allocated to each regular paper is 20 minutes, and 10 minutes for each short paper, including questions and answers.

Session Chairs

If you cannot fulfill your duties as a session chair, please ensure that someone else will take your place as the session chair, or contact the Program Chairs to arrange a back-up.

Session chairs are kindly requested to help with the followings:

1. Note the time allocated for each paper in your session. Each regular paper is allocated 20 minutes (18 minutes for the presentation plus 2 minutes for discussion). Each short paper is allocated 10 minutes (8 minutes for the presentation plus 2 minutes for discussion).
2. Arrive at the room of the session 5 minutes before the session starts and identify each of the speakers for the session.
3. Suggest each speaker to keep corresponding time for discussions (questions and answers), and for transition to the next presentation. If a presentation extends into the time for discussions, please shorten the discussions accordingly or postpone the discussions until after the session.
4. Do not allow presentations or the subsequent discussions to spill beyond the starting time of the next presentation.
5. If the presenter of a paper is absent ("no-show"), please continue to the next presentation. Please check again at the end of the last presentation whether the "no-show" shows up. Best efforts have been made to reduce the number of no-shows; however, they may not be eliminated.
6. Each presentation room is equipped with a digital projector. If something is not working properly, please contact a volunteer for help.

Presenters

Please check your presentation time and room. Please go to the room 5 minutes before the session starts and identify yourself to the session chairs.

1. Note the time allocated for each regular paper is 20 minutes (18 minutes for your presentation plus 2 minutes for discussion) and 10 minutes for each short paper (2 minutes for discussions).
2. When it is your turn to present, please leave corresponding time for discussion (questions and answers), and for transition to the next presentation. If your presentation extends into the time for discussions, discussions on your paper will be shortened by the session chair accordingly or postpone until after the session.
3. Please do not exceed your allocated time. Please follow the instructions of the Session Chairs.

If you cannot find your name in sessions or your information is incorrect in the Program Booklet, please contact the Program Chairs.

Internet Access Guide

(Your account is valid ONLY between December 14 to 17 for the ICDM'10 conference)

An Account Example (use your own account please):
CAO, Longbing
Login ID: m0027276
Password: lcao6840

How can you access UTS Internet during the conference of ICDM'10?

PC Desktop Access:

There are several PCs in the foyer of Block C (CM05C) which is the area next to the conference venue in the same building (see map, Conference venue is CM05B.). Please use these PCs to check your emails.

Wireless Access:

(Note: Signal only covers Block C, CM05C, See map.)

- **Step 1**: You will need to ensure your laptop has the wireless connection enabled. To do this, go to your Control Panel, select Network and Internet Connections and then Network Connections. Click on the Wireless Network Connection and select Enable.

- **Step 2**: Then select "UTS-Wireless" as your wireless connection. Depending on your location within the campus, there may be more than one wireless networks available. You should always select UTS-Wireless as it appears on the wireless network profiles.

- **Step 3**: Then open your web browser. You will be redirected to the UTS Wireless Network login page. If not, attempting to browse to any site on the Internet will take you to the login page. Connecting to the UTS Wireless Network requires the Java Runtime Environment (JRE) to be installed on your laptop. JRE comes pre-installed on Mac OS X but may or may not be installed on Windows or Linux. If not installed, download JRE by clicking on the Windows or Linux JRE links and then run the downloaded executable.

- **Step 4**: The first time you connect to the wireless network, you may need to install an ActiveX control if you see the message below appear below your navigation bar in your browser. Click on the "Click here to install" message if it appears under your task bar.

- **Step 5**: When authenticated and successfully connected to the UTS Wireless Network, you will find the Network Connect icon located in the system tray.

- **Step 6**: To disconnect from the Wireless Network, double click on the icon to display the Network Connect application screen and then click on "Sign Out", or the "Lock" icon on the taskbar.

5
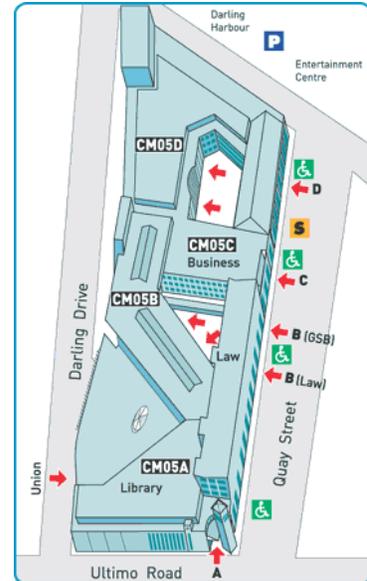
ICDM 2010 Conference Program at-a-glance

Workshop, Contest, Colloquium Program at-a-glance

Tuesday December 14, 2010			
08:30-12:20 Including **Tea/Coffee Break** (10:00-10:20) [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
OEDM10 [B112]	**BioDM10** [B113]	**SSTDM10** [B102]	**DMAGI10** [C119]
Optimization based Methods for Emerging Data Mining Problems Chairs: Chris Ding, Tao Li, Fei Wang	*Biological Data Mining and its Applications in Healthcare* Chairs: Xiao-li Li, See-Kiong Ng	*Spatial and Spatio-temporal Data Mining* Chairs: Shashi Shekhar, Peggy Agouris, Raju Vatsavai, Sanjay Chawla, Matt Duckh	*Applications of Data Mining and Modeling in Government and Industry* Chairs: Graham Williams, Wei Lin, Warwick Graco, Peter O'Hanlon
DDDM10 [C121]	**MMIS10** [C240]	**CLIMKD10** [C241]	**LACIS10** [C242]
Domain Driven Data Mining Chairs: Philip Yu, Graham Williams, Ling Chen, Huaifeng Zhang, Gabor Melli	*Mining Multiple Information Sources* Chairs: Ruoming Jin, Xingquan Zhu, Haixun Wang, Zoran Obradovic	*Knowledge Discovery from Climate DataL Prediction, Extremes, and Impacts* Chairs: Nitesh V. Chawla, Vipin Kumar, Karsten Steinhaeuser, Auroop R. Ganguly, Mechael Steinbach	*Large-scale Analytics for Complex Instrumented Systems* Chairs: Wray Buntime, Chid Apte, Yan Liu, Jimeng Sun, and Jie Tang
DMS10 [C243]	**SIASP10** [C244]	**PADM10** [C245]	**TFDOM10** [C246]
Data Mining for Service Chairs: Katsutoshi Yada, Yukio Ohsawa, Russ Winer	*Social Interactions Analysis and Services Providers* Chair: Hakim Hacid, Cecile Favre, Tetsuya Yoshida	*Privacy Aspects of Data Mining* Chairs: Aris Gkoulalas-Divanis, Kun Liu, Vassilios S. Verykios, Ran Wolff	*Topic Feature Discovery and Opinion Mining* Chairs: Yuefeng Li, Ning Zhong, Raymond Lau
IWCD10 [C247]	**VAKD10** [C254]	**DSKDBI Colloquium** [C120]	
Chance Discovery Chairs: Akinori Abe, Yukio Ohsawa	*Visual Analytics and Knowledge Discovery* Chairs: Simeon Simoff, Pak Chung Wong, Mike Sips, Arturas Mazeika	*Data Science, Knowledge Discovery and Business Intelligence*	
12:20-13:30 **Lunch** Marigold Restaurant, Level 5, 683-689 George Street, Sydney			
13:30-17:40 Including **Tea/Coffee Break** (15:30-16:00) [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
OEDM10 [B112]	**BioDM10** [B113]	**SSTDM10** [B102]	**DMAGI10** [C119]
Optimization based Techniques for Emerging Data Mining Problems Chairs: Chris Ding, Tao Li, Fei Wang	*Biological Data Mining and its Applications in Healthcare* Chairs: Xiao-li Li, See-Kiong Ng	*Spatial and Spatio-temporal Data Mining* Chairs: Shashi Shekhar, Peggy Agouris, Raju Vatsavai, Sanjay Chawla, Matt Duckh	*Applications of Data Mining and Modeling in Government and Industry* Chairs: Graham Williams, Wei Lin, Warwick Graco, Peter O'Hanlon
DDDM10 [C121]	**MMIS10** [C240]	**KDCloud10** [C241]	**SADM10** [C242]
Domain Driven Data Mining Chairs: Philip Yu, Graham Williams, Ling Chen, Huaifeng Zhang, Gabor Melli	*Mining Multiple Information Sources* Chairs: Ruoming Jin, Xingquan Zhu, Haixun Wang, Zoran Obradovic	*Knowledge Discovery Using Cloud and Distributed Computing Platforms* Chairs: Ranga Raju Vatsavai	*Semantic Aspects in Data Mining* Chairs: Hui Xiong, Vania Bogorny Chiara Renso
RIKD10 [C243]	**SIASP10** [C244]	**PADM10** [C245]	**TFDOM10** [C246]
Reliability Issues in Knowledge Discovery Chairs: Honghua Dai, James Liu, Evgueni Smirnov	*Social Interactions Analysis and Services Providers* Chairs: Hakim Hacid, Cecile Favre, Tetsuya Yoshida	*Privacy Aspects of Data Mining* Chairs: Aris Gkoulalas-Divanis, Kun Liu, Vassilios S. Verykios, Ran Wolff	*Topic Feature Discovery and Opinion Mining* Chairs: Yuefeng Li, Ning Zhong, Raymond Lau
IWCD10 [C247]	**VAKD10** [C254]	**DSKDBI Colloquium** [C120]	**Contest** [C256]
Chance Discovery Chairs: Akinori Abe, Yukio Ohsawa	*Visual Analytics and Knowledge Discovery* Chairs: Simeon Simoff, Pak Chung Wong, Mike Sips, Arturas Mazeika	*Data Science, Knowledge Discovery and Business Intelligence*	*ICDM'10 Data Mining Contest* Chairs: Demetris Zeinalipour
18:15 -19:15 **Business Analytics Forum** [B102]			

ICDM'10 Technical Program at-a-glance

Wednesday December 15, 2010	08:30-09:00	**Conference Opening** by Prof. Attila Brungs from UTS DVCR [B111]			
	09.00-10.00	**KEYNOTE** [B111] Chair: Geoff Webb Mining Billion-node Graphs: Patterns, Generators and Tools by Christos Faloutsos			
	10:00-10:20	**Tea/Coffee Break** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
	10:20 - 12:20	**Session 1A** [B111] *Clustering* Chair: Chris Ding	**Session 1B** [B112] *Co-learning, semi-supervised learning, active learning, instance-based learning & evaluation* Chair: Jimeng Sun	**Session 1C** [B102] *Applications* Chair: Shusaku Tsumoto	**Tutorial 1** [C119] *Spatial and Spatio-temporal Data Mining*	**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	12:20-13:30	**Lunch** [Marigold Restaurant, Level 5, 683-689 George Street, Sydney]			
	13:30 - 15:30	**Session 2A** [B111] *Clustering* Chair: Kai Ming Ting	**Session 2B** [B112] *SVM, kernel, vector & matrix techniques* Chair: Hisashi Kashima	**Session 2C** [B102] *Privacy preserving data mining, ensemble learning & regression* Chair: Aris Gkoulalas-Divanis	**ICDM Contest Overview** [C119] Chair: Marcin Wojnarski Note: Start at 14:00	**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	15:30-16:00	**Tea/Coffee Break** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
	16:00 - 18:00	**Session 3A** [B111] *Sequence mining and clustering* Chair: Ke Wang	**Session 3B** [B112] *Social networks, social media and probabilistic techniques* Chair: Jie Tang	**Session 3C** [B102] *Stream mining, multimodal data and inference in factorial spaces* Chair: Jeffrey Yu	**Tutorial 2** [C119] *Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data*	**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	18:00 - 18:30	**ICDM 10 Year Impact Award Address** [B111] Chair: Wei Wang			
	18:30-20:30	**Reception** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
Thursday December 16, 2010	09:00-10:00	**KEYNOTE** [B111] Chair: Chengqi Zhang 10 Years of Data Mining Research: Retrospect and Prospect By Xindong Wu			
	10:00-10:20	**Tea/Coffee Break** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
	10:20 - 12:20	**Session 4A** [B111] *Clustering* Chair: George Karypis	**Session 4B** [B112] *Association discovery & uplift modeling* Chair: Laitfur Khan	**Session 4C** [B102] *Classification* Chair: Yan Liu	**Tutorial 3** [C119] *Knowledge Discovery in Academic Drug Discovery Programs: Opportunities and Challenges*	**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	12:20-13:30	**Lunch** [Marigold Restaurant, Level 5, 683-689 George Street, Sydney]			
	13:30	**Excursion -** 13.40. Foyer, Building 5 for bus pickup for your choice of either a City Tour or Opera House Tour			
	18:30-22:00	**Banquet -** 18:30 - 22:00, 'Showboat Two' for Sydney Harbour dinner cruise, Campbell Cover Wharf, Sydney Harbour Foreshore (see page 3 for details, and page 57 for map).			
Friday December 17, 2010	09:00 - 10:00	**Session 5A** [B111] *Text mining* Chair: Arindam Banerjee	**Session 5B** [B112] *Graph & network mining* Chair: Ee-Peng Lim	**Session 5C** [B102] *Feature selection & construction* Chair: katharina morik	**Tutorial 4** [C119] *How to do Good Data Mining Research and get it Published in Top Venues*	**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	10:00-10:20	**Tea/Coffee Break** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
	10:20 - 12:20	**Session 6A** [B111] *Text mining, transfer learning, time series & spatio-temporal mining* Chair: Toon Calders	**Session 6B** [B112] *Graph & network mining* Chair: Xiaoli Li	**Session 6C** [B102] *Outlier detection, distance learning & minhash computation* Chair: Jianyong Wang		**DSKDBI Colloquium** [C120] *Data Science, Knowledge Discovery and Business Intelligence*
	12:20-14:00	**Lunch and ICDM 2010 Community Meeting** [Marigold Restaurant, Level 5, 683-689 George Street, Sydney] (everyone is welcome)			
	14:00-14:30	**Session 7A** [B111] Image retrieval, Chair: Hui Xiong			
	14:00-16:00	**Session 7B** [B112] Text Mining, Chair: Kyuseok Shim			
	14:00-16:00	**Session 7C** [B102] Collaborative filtering, link prediction and recommender systems, Chair: Dejing Dou			
	14:30-16:00	**ICDM Panel:** *ICDM'10 Top 10 Case Studies* [B111] Chairs: Gabor Melli, Xindong Wu			
	16:00-16:15	**Tea/Coffee Break** [Lobby, Block D, Building 5, Haymarket Campus, UTS]			
	16:15-17:15	**KEYNOTE** [B111] Chair: Bing Liu Assessing the Significance of Groups in High-Dimensional Data by Geoff McLachlan			
	17:15	**CLOSING SESSION**			

ICDM 2010 Conference Program

Workshop and Contest Program

Tuesday, December 14, 2010 (Building 5, Haymarket Campus, UTS)

General Information for All Workshops

Morning Tea/coffee break (10:00-10:20)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Conference Lunch (12:20-13:30)
Location: Marigold Restaurant, Level 5, 683-689 George Street, Sydney

Afternoon tea/coffee break (15:30-16:00)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Optimization based Methods for Emerging Data Mining Problems (OEDM 2010)

Workshop Co-Chairs: Chris Ding, Tao Li, Fei Wang
Location: B112

9:30-9:45 *ALPOS -- Assisted Learning for Partial Observation: A Machine Learning Approach for Classifying Tweets*, Dan Zhang, Yan Liu, Rick Lawrence, and Vijil Chenthamarakshan

9:45-10:00 *Semi-supervised PLSA for Document Clustering*, Lingfeng Niu and Yong Shi

10:00-10:30 Tea/coffee break

10:30-11:00 Invited talk

11:00-11:15 *A Novel Ranking Algorithm for Service Matching Based on Agent Association Graphs*, Haolan Zhang, Clement H. C. Leung, Gitesh K. Raikundalia, Jing He

11:15-11:30 *Distributed Flow Algorithms for Scalable Similarity Visualisation*, Novi Quadrianto, Dale Schuurmans, and Alex Smola

11:30-11:45 *XML Documents Clustering using Tensor Space Model,* Sangeetha Kutty, Richi Nayak, and Yuefeng Li

11:45-12:00 *Find Intelligent Knowledge by Second Time Mining: Three Cases from China*, Guangli Nie, Lingling Zhang, Yuejin Zhang, and Yong Shi

12:00-12:15 *Integer Programming for Multi-Class Active Learning*, Dragomir Yankov, Suju Rajan, and Adwait Ratnaparkhi

12:15-1:30 Lunch break

1:30-2:10 *Invited talk: Sub-space Distance-Based Sampling Method for SVM*, Xiaofei Zhou and Yong Shi

2:10-2:30 *Integrating Symmetric Nonnegative Matrix Factorization and Normalized Cut Spectral Clustering*, Zhichen Xia and Chris Ding

2:30-2:45 *From Convex to Non-Convex: a Loss Function Analysis for Binary Classification*, Lei Zhao, Musa Mammadov, and John Yearwood

2:45-3:00 *Cluster Cores and Modularity Maximization*, Michael Ovelgönne and Andreas Geyer-Schulz

3:00-3:15 *Reviewer Profiling Using Sparse Matrix Regression*, Evaggelos Papalexakis, Nikolaos Sidiropoulos, and Minos Garofalakis

3:15-3:30 *A Comparison of Objective Functions in Network Community Detection*, Chuan Shi, Yanan Cai, Philip Yu, ZhenYu Yan, and Bin Wu

3:30-4:00   Tea/coffee break

4:00-4:15 *Large-Scale Matrix Factorization using MapReduce*, Zhengguo Sun, Tao Li, and Naphtali Rishe

4:15-4:30 *Efficient Dimensionality Reduction on Undersampled Problems through Incremental Discriminative Common Vectors*, Francesc J. Ferri, Katerine Diaz-Chito, and Wladimiro Diaz-Villanueva

4:30-4:45 *Robust Low-Rank Subspace Segmentation with Semidefinite Guarantees*, Yuzhao Ni, Ju Sun, Xiaotong Yuan, Simon yan, and Loong-Fah Cheong

4:45-5:00 *ProCommend: A Flexible Profit-Oriented Recommender System*, Leman Akoglu and Christos Faloutsos

5:00-5:15 *Efficient additive models via higher order ℓ_1 regularization*, Dimitri Semenovich and Nobuyuki Morioka

5:15-5:30 *Ant Colony Optimization Algorithm Based on Immunity Vaccine and Dynamic Pheromone Updating*, Wanjun Liu, Juan Zhang, and Junli liu

Biological Data Mining and its Applications in Healthcare (BioDM 2010)

Workshop Co-Chairs: Xiao-li Li, See-Kiong Ng
Location: B113

8:30 - 8:35 Opening: Xiao-Li Li & See-Kiong Ng

8:35 - 9:20 Keynote 1: *Using Clinical Terminologies to Enhance Data Mining of Genomic Data*, David Hansen

9:20 - 9:40 *Is Topotecan Effective at Killing Cancer Cells?*, Ricardo Santiago Mozos, Imtiaz A. Khan and Michael Madden

9:40 - 10:00 *An integrative scoring approach to identify transcriptional regulations controlling lung surfactant homeostasis*, Minlu Zhang, Chunsheng Fang, Yan Xu, Raj Bhatnagar and Long Lu

10:00 - 10:30 Tea/coffee break

10:30 - 10:50 *Combining Time Series Similarity with Density-based Clustering to Identify Fiber Bundles in the Human Brain*, Junming Shao, Klaus Hahn, Qinli Yang, Christian Boehm, Afra Wohlschlaeger, Nicholas Myers and Claudia Plant

10:50 - 11:10 *Efficient Alignments of Metabolic Networks with Bounded Treewidth*, Qiong Cheng, Piotr Berman, Rob Harrison, Ion Mandoiu, and Alex Zelikovsky

11:10 - 11:30 *System Biology Approach for Relation of Indonesian Herbal Plants and Efficacy of Jamu*, Farit Afendi, Aki Hirai, Md. Altaf-Ul-in, Hiroki Takahashi, Kensuke Nakamura and Shigehiko Kanaya

11:30 - 11:50 *Fixed-Parameter Tractable Combinatorial Algorithms for Metabolic Networks Alignments*, Qiong Cheng, Jinpeng Wei, Alexander Zelikovsky and Mitsunori Ogihara

11:50 - 12:10 *The Multiple Alignment Algorithm for Metabolic Pathways without Abstraction*, Wenbin Chen, William Hendrix, Andrea Rocha, Matt Schmidt and Nagiza Samatova

12:10 - 1:25 Lunch Break

1:25 - 2:10 Keynote 2: *Influence of genomic and other biological data sets in the understanding of protein structures, functions and interactions,* N. Srinivasan

2:10 - 2:30 *Evaluation of Protein Backbone Alphabets: Using Predicted Local Structure for Fold Recognition*, Kyong Jin Shim

2:30 - 2:50 *Parametric Templates: A New Enzyme Active-Site Prediction Algorithm*, Tsuyoshi Kato, Kazuhiro Suwa and Nozomi Nagano

2:50 - 3:10 *Overcoming the Curse of Dimensionality in a Statistical Geometry Based Computational Protein Mutagenesis*, Majid Masso

3:10 - 3:30 *An Empirical Comparison of Platts Calibration and Inductive Confidence Machines for Predictions in Drug Discovery*, Nikil Wale

3:30 - 4:00 Tea/coffee break

4:00 - 4:20 *Cost-sensitive feature selection based on the Set Covering Machine*, Raul Santos-Rodriguez and Darío García-García

4:20 - 4:40 *FD Cluster: Mining Frequent Closed Discriminative Bi-cluster without Candidate Maintenance in Multiple Microarray Datasets*, Miao wang, Xuequn Shang, Shaohua Zhang, and Zhanhuai Li

4:40 - 5:00 *Curve Profiling Feature: Novel Compact Representation for Drosophila Embroynic Gene Expression Pattern Mining*, Chunsheng Fang, Anca Ralescu and Long Lu

5:00 - 5: 20 *Discovering Significant Patterns of Interaction*, Martin McGrane and Simon Poon

5:20 - 5:40 *Fast Anomaly Detection in Dynamic Clinical Datasets Using Near-Optimal Hashing with Concentric Expansions*, Zeeshan Syed and Ilan Rubinfeld

Spatial and Spatio-temporal Data Mining (SSTDM 2010)

Workshop Co-Chairs: Shashi Shekhar, Peggy Agouris, Raju Vatsavai, Sanjay Chawla, Matt Duckh
Location: B102

8:30 – 8:45 Opening Comments

8:45 – 9:30 Invited Talk

9:30 – 9:50 *Curvature Maxima based Trajectories Mining*, Shoji Hirano and Shusaku Tsumoto

9:50 – 10:10 *Sample Bias due to Missing Data in Mobility Surveys*, Dirk Hecker, Hendrik Stange, Christine Körner, Michael May

10:10 – 10:30 Tea/coffee break

10:30 – 10:50 *Comparing Vessel Trajectories using Geographical Domain Knowledge and Alignments*, Gerben de Vries and Maarten van Someren

10:50 – 11:10 *Measuring similarity for multidimensional sequences*, Hui Wang, Zhiwei Lin, Sally McClean, and Jun Liu

11:10 – 11:30 *Spatio-Temporal Symbolization of Multidimensional Time Series*, Shohei Hidaka and Chen Yu,

11:30 – 11:50 *Unsupervised Semantic Labeling Framework for Identification of Complex Facilities in High-resolution Remote Sensing Images*, Ranga Raju Vatsavai, Anil Cheriyadat, and Shaun Gleason

12:00 – 1:45 Lunch

1:45 – 2:30 Keynote: Prof. Vipin Kumar, Minnesota

2:30 – 2:50 *Computing Popular Places using Graphics Processors*, Marta Fort, J. Antoni Sellarès, and Nacho Valladares

2:50 – 3:10 *Using SelfOrganizing Map and Heuristics to Identify Small Statistical Areas Based on Household Socio-Economic Indicators in Turkey's Address Based Population Register System*, Şebnem Düzgün and Şeyma Özcan Yavuzoğlu

3:10 – 3:30 *Putting Pixels in Place: A Storage Layout Language for Scientific Data*, Peter Baumann, Shams Fezyabadi, and Constantin Jucovschi

3:30 – 4:00 Tea/coffee break

4:00 – 4:20 *A K-Main Routes Approach to Spatial Network Activity Summarization: A Summary of Results,* Dev Oliver, Abdussalam Bannur, James Kang, Shashi Shekhar, and Renee Bousselaire

4:20 – 5:20 Panel

5:20 – 5:30 Closing Remarks

Applications of Data Mining and Modeling in Government and Industry (DMAGI 2010)

Workshop Co-Chairs: Graham Williams, Wei Lin, Warwick Graco, Peter O'Hanlon
Location: C119

09.00-10.00 Retail Administration and Keynote Presentation: *The Craft of Credit Scoring: Data Mining Applied to Retail Finance*, Ross Gayler

10.00-10.20 Tea/coffee Break

10.20-10.45 *Unsupervised DRG Upcoding Detection in Healthcare Databases*, Wei Luo and Marcus Gallagher

10.45-11.10 *Anomaly Detection in Spatio-Temporal Health Data: Experience with a Real Life Application*, Kee Siong Ng, Yin Shan, D. Wayne Murray, Alison Sutinen, Bernhard Schwarz, David Jeacocke, and Josephine Farrugia

11.10-11.35 *Apply data mining and natural computing in detecting suspicious cases of money laundering in an investment bank: A case study*, Nhien An Le Khac and M-Tahar Kechadi

11.35-12.00 *Empirical Analysis: News Impact on Stock Prices based on News Density*, Xiaodong LI, Xiaotie DENG, Feng WANG, Keren DONG, and Qing LI

12.00-12.25 *Visual Predictions of Currency Crises using Self-Organizing Maps*, Peter Sarlin and Dorina Marghescu

12.25-1.30 Lunch break

1.30-1.55 *Less Effort, More Outcome: Optimising Debt Recovery with Decision Tree*, Yanchang Zhao, Hans Bohlscheid, and Shanshan Wu

1.55-2.20 *Using SOM-Ward clustering and predictive analytics for conducting customer segmentation*, Zhiyuan Yao, Tomas Eklund, and Barbro Back

2.20-2.45 *Sales Forecasting of IT Products using A Hybrid MARS and SVR Model*, Chi-Jie Lu, Tian-Shyug Lee, and Chia-Mei Lian

2.45-3.10 Retail *Quantum Path Integral Inspired Query Sequence Suggestion for User Search Task Simplification*, Baojun Yue

3.10-3.35 *Targeting Television Audiences using Demographic Similarity*, Brendan Kitts

3.35-3.55 Tea/coffee break

3.55-4.20 *Mining Heterogeneous ADS-B Data Sets for Probabilistic Models of Pilot Behavior*, Ronald Marsh and Kirk Ogaard

4.20-4.45 *Gaussian Processes for Dispatching Rule Selection in Production Scheduling*, Bernd Scholz-Reiter, Jens Heger, and Torsten Hildebrandt

4.45-5.10 *Variable Productivity Adjustment Estimation for Function Point Project Delivery*, Saeed Bagheri, Krishna Ratakonda, and Rakesh Mohan

5.10-5.35 *Graphics Classification for Enterprise Knowledge Management,* Divna Djordjevic and Rayid Ghani

5.35-6.00 *Efficient Rule Generation for Dominant Class Problems on LARM*, JuiHsi Fu and SingLing Lee

6.00 Welcome Drinks

6:15 - 7:15 Business Analytics Forum

Domain Driven Data Mining (DDDM 2010)

Workshop Co-Chairs: Philip Yu, Graham Williams, Ling Chen, Huaifeng Zhang, Gabor Melli
Location: C121

8:45 - 9:00 Opening Address Workshop Co-Chairs

9:00 - 9:30 Keynote Speech: *Highly Efficient Predictive Feature Vector Mining from Raw Data,* Wei Fan

9:30 - 10:00 Feature Talk: *Actionable Knowledge Discovery and Transfer,* Hans Bohlscheid

10:00 - 10:30 Tea/coffee break

10:30 - 12:30 Session I - Involving Ubiquitous Intelligence

Using Taxonomies to Perform Aggregated Querying over Imprecise Data, Atanu Roy, Chandrima Sarkar and Rafal Angryk

Sequence Alignment Based Analysis of Player Behavior in Massively Multiplayer Online Role-Playing Games (MMORPGs), Kyong Jin Shim and Jaideep Srivastava

Network anomaly detection using a commute distance based approach, Khoa Nguyen, Tahereh Babaie, Sanjay Chawla and Zainab Zaidi

Link Prediction across Multiple Social Networks, Muhammad Ahmad and Jaideep Srivastava

Controlling consistency in Top-N Recommender Systems, Paolo Cremonesi and Roberto Turrin

Agent Assignment for Process Management: Agent Performance Evaluation Framework, Muhammad Ramzan Talib, Bernhard Volz, and Stefan Jablonski

12:30-1:30:00 Lunch break

1:30-2:00 Keynote Speech: *Interestingness in the Real World: Two Case Studies*, Bing Liu

2:00 - 3:30 Session II: Mining Complex Data and Applications

Learning with Actionable Attributes: Attention – Boundary Cases! Indre Zliobaite and Mykola Pechenizkiy

Contextual Sequential Pattern Mining, Julien Rabatel, Sandra Bringay, and Pascal Poncelet

Mining Textual Data Streams with Intra-Chunk Concept Drifting, Bo Liu, Yanshan Xiao, Longbing Cao and Philip Yu

Evaluating Association Rules by Quantitative Pairwise Property Comparisons, Elnaz Delpisheh and John Zhang

3:30 -4:00 Tea/coffee break

4:00 -5:40 Session III: Domain-specific Applications

Medically Driven Data Mining Application: Recognition of Health Problems from Gait Patterns of the Elderly, Bogdan Pogorelc and Matjaz Gams

Insights from Applying Sequential Pattern Mining to E-Commerce Click Strame Data, Arthur Pitman and Markus Zanker

Domain-driven Data Mining for IT Infrastructure Support, Girish Palshikar, Harrick Vin, Mohammed Mudassar, and Maitreya Natu

Trading Tests of Long-Term Market Forecast by Text Mining, Kiyoshi Izumi, Takashi Goto, and Tohgoroh Matsui

5:40- 6:00 Closing comments, Workshop Co-chairs
6:15 - 7:15 Business Analytics Forum

Mining Multiple Information Sources (MMIS 2010)

Workshop Co-Chairs: Ruoming Jin, Xingquan Zhu, Haixun Wang, Zoran Obradovic
Location: C240

8: 50 - 9:00 Welcome & Opening Comments

9:00 - 10:00 Keynote Speech: *Computational and Algorithmic Challenges for the Square Kilometre Array in Imaging Astronomy*, Tim Cornwell, Astronomy and Space Science Division,

CSIRO, Australia

10:00 – 10: 30 Tea/coffee break

10:30 – 12:10pm Section I: [20 minutes/each]

Differential Analysis on Deep Web Data Sources
    Tantan Liu, Jiedan Zhu, Fan Wang, and Gagan Agrawal, Ohio State University, USA

Learning Document Labels from Enriched Click Graphs
    Lan Nie Yahoo! Labs, USA

Mining research topics evolving over time using a diachronic multi-source approach
    Jean-Charles Lamirel, INRIA TALARIS team LORIA, France

Multiple Feature-based Classifier and Its Application to Image Classification
    Dong-Chul Park, Myong Ji University, Republic of Korea

Learning a Combination of Dissimilarities from a Set of Equivalence Constraints
    Manuel Martin-Merino, University Pontificia of Salamanca, Spain and Javier De Las Rivas, Cancer Research Center of Salamanca, Spain

12:20pm-1:30pm Lunch Break [Marigold Restaurant: http://www.marigold.com.au]

1:30pm – 2:30 pm Keynote Speech: *Learning from Heterogeneous Data Sources*, Prof. Philip S. Yu, University of Illinois at Chicago, USA

2:30pm – 3:30 pm Section II [20 minutes /each]

W-LEACH: Weighted Low Energy Adaptive Clustering Hierarchy Aggregation Algorithm for Data Streams in Wireless Sensor Networks
    Hanady Abdulsalam and Layla Kamel, Kuwait University, Kuwait

PTCR-Miner: Programessive Temporal Class Rule Mining for Multivariate Temporal Data Classification
    Chao-Hui Lee and Vincent, S. Tseng, National Cheng Kung University, Tainan, Taiwan

Event Data Mining and Classification from Multiple Streaming Sources
    Ashit Talukder, Jet Propulsion Laboratory, California Institute of Technology, USA

3:30pm – 4:00 pm Tea/coffee break

4:00pm – 5:50 pm Section III [20 minutes/each]

A Triple-Random Ensemble Classification Method for Mining Multi-label Data
    Gulisong Nasierding, Deakin University, Australia

CUBS: Multivariate Sequence Classification Using Bounded Z-score with Sampling
    Ariella Richardson, Gal Kaminka, and Sarit Kraus, Bar Ilan University, Israel

Distributed classification on peers with variable data spaces and distributions
    Vivekanand Gopalkrishnan, Nanyang Technological University, Singapore

Diffusion Maps: A Superior Semantic Method to Improve Similarity Join Performance
    Bilal Hawashin, Farshad Fotouhi, Wayne State University, USA and William Grosky, University of Michigan - Dearborn, USA

5: 50 pm – 6:00 pm Discussion and Concluding Comments

Knowledge Discovery from Climate Data Prediction, Extremes, and Impacts (CLIMKD 2010)

Workshop Co-Chairs: Nitesh V. Chawla, Vipin Kumar, Karsten Steinhaeuser, Auroop R. Ganguly, Mechael Steinbach
Location: C241

8:30 Welcome & Opening Comments

Session I (Chair: TBA)

8:40 S2205: *Mining Spatial Co-occurence of Drought Events from Climate Data of India*. S. Kollukuduru, K. S. Rajan.

9:00 S2206: D*ata Analysis in Los Angeles Long Beach with Seasonal Time Series Model*. W. Wang, Z. Niu.

9:20 S2202: *Detecting Climate Change in Multivariate Time Series Data by Novel Clustering and Cluster Tracing Techniques*. H. Kremer, S. Günnemann, T. Seidl.

9:40 S5214: *Using Time Series Segmentation for Deriving Vegetation Phenology Indices from MODIS NDVI Data*. V. Chandola, D. Hui, L. Gu, R. R. Vatsavai.

10:00 Tea/coffee break

10:30 Invited Talk: TBA

Session II (Chair: TBA)

11:20 S2204: *A New Data Mining Model for Hurricane Intensity Prediction*. Y. Su, S. Chelluboina, M. Hahsler, M. H. Dunham.

11:40 DM272: *Learning Maximum Lag for Grouped Graphical Granger Models*. A. Dhurandhar.

12:00 Closing Comments

12:20pm-1:30pm Lunch Break [Marigold Restaurant: http://www.marigold.com.au]

Knowledge Discovery Using Cloud and Distributed Computing Platforms (KDCloud 2010)

Workshop Co-Chairs: Ranga Raju Vatsavai
Location: C241

2:00 – 2:05 Opening Comments

2:05 – 2:45 Keynote Talk: Prof. Nitesh Chawla, University of Notre Dame

2:45 – 3:05 *Parallel EMClustering: Fast Convergence by Asynchronous Model Updates*, Claudia Plant and Christian Böhm

3:05 – 3:25 *Parallelizing Information Theoretic Coclustering Algorithm Using a Cloud Middleware*, Venkatram Ramanathan, Wenjing Ma, Vignesh Ravi, and Gagan Agrawal

3:25 – 3:45 *dSimpleGraph: a Novel Distributed Clustering Algorithm for Exploring Very Large*

Scale Unknown Data Sets, Li Lu, Yunhong Gu, and Robert Grossman

3:45 - 4:00 Tea/coffee break

4:00 - 4:20 *Distributed, Scalable Clustering for Detecting Halos in Terascale Astronomy*, Srivatsava Daruru, Sankari Dhandapani, Gunjan Gupta, Ilian Iliev, Weijia Xu, Paul Navratil, Nena Marin, and Joydeep Ghosh

4:20 - 4:40 *Parallel User profiling based on folksonomy for Large Scaled Recommender Systems-An implementation of Cascading MapReduce*, Huizhi Liang,

4:40 - 5:00 *Challenges in Scheduling Aggregation in CyberPhysical Information Processing Systems*, James Horey

5:00 – 5:20 *Distributed Stream Computing Platform,* Leonardo Neumeyer, Bruce Robbins, Anish Nair, and Anandsudhakar Kesari

5:20 - 5:40 *On Finding Similar Items in a Stream of Transactions*, Andrea Campagna and Rasmus Pagh,

5:40 – 6:00 *Minimum spanning tree based classification model for massive data with MapReduce implementation*, Jin Chang, Jun Luo, and Joshua Zhexue Huang

6:00 Closing Comments

Large-scale Analytics for Complex Instrumented Systems (LACIS 2010)

Workshop Co-Chairs: Wray Buntime, Chid Apte, Yan Liu, Jimeng Sun, and Jie Tang
Location: C242

8:45-9:00 Opening

9:00-10:00 Keynote 1

10:00-10:30 Tea/coffee break

10:30-12:30 Technical Presentation (each presentation has 15 minutes)
Enhancing Ubiquitous Systems through System Call Mining, Katharina Morik, Nico Piatkowski, Michael Engel, and Felix Jungermann

Parallelized Boosting with Map-Reduce, Indranil Palit and Chandan Reddy

Multi-step Time Series Prediction in Complex Instrumented Domains, I T Dhurandhar

EigenDiagnostics: Spotting Connection Patterns and Outliers in Large Graphs, Koji Maruhashi and Christos Faloutsos

Using Adaptive Downsampling to Compare Time Series with Warping, Chih-Chun Chia and Zeeshan Syed

Sliding HyperLogLog: Estimating cardinality in a data stream over a sliding window, Yousra Chabchoub and Yousra Chabchoub

dMaximalCliques: A Distributed Algorithm for Enumerating All Maximal Cliques and Maximal Clique Distribution, Li Lu, Yunhong Gu, and Robert Grossman

12:20pm-1:30pm Lunch Break [Marigold Restaurant: http://www.marigold.com.au]

Semantic Aspects in Data Mining (SADM 2010)

Workshop Co-Chairs: Hui Xiong, Vania Bogorny Chiara Renso
Location: C242

1:30 – 2:30 Invited talk: *Geospatial Schema Matching With High-Quality Cluster Assurance and Location Mining from Social Network*, Lathifur Khan

2:30 – 3:00 Session: Semantic Trajectory Mining: *From Pattern Discovery to Pattern Interpretation in Movement Data,* Rebecca Ong, Chiara Renso, Mirco Nanni, and Monica Wachowicz

3:00 – 3:30 Session: Semantic Text Mining: *Analysis of Collaborative Writing Process Using Hidden Markov Model*, Vilaythong Southavilay, Kalina Yacef, and Refael Calvo

3:30 – 4:00 Tea/Coffee Break

Session: Exploiting taxonomies and ontologies

4:00 – 4:30 *Semantic content filtering with Wikipedia and Ontologies*, Pekka Malo, Pyry-Antti Siitari, Oskar Ahlgren, Jyrki Wallenius, and Pekka Korhonen

4:30 – 5:00 *Using Ontology and Sequence Information for Extracting Behavior Patterns from Web Navigation Logs*, HAKAN YILMAZ and PINAR SENKUL

5:00 – 5:30 *Incorporating Multi-Partite Networks and Expertise to Construct Related-Term Graph,* Jyh-Ren Shieh, Ching-Yung Lin, Shun-Xuan Wang, Yung-Huan Hsieh, and Ja-Ling Wu

Data Mining for Service (DMS 2010)

Workshop Co-Chairs: Katsutoshi Yada, Yukio Ohsawa, Russ Winer
Location: C243

Session 1

Discovering Temporal Features and Relations of Activity Patterns, Ehsan Nazerfard, Parisa Rashidi, and Diane Cook

Knowledge File System - A principled approach to personal information management, Kuiyu Chang, I Wayan Tresna Perdana, Bramandia Ramadhana, Kailash Sethuraman, Truc Viet Le, and Neha Chachra

Towards Data-Oriented Hospital Services: Data Mining-based Hospital Management, Shusaku Tsumoto, Shoji Hirano, and Yuko Tsumoto,

Tea/coffee break (10:00 – 10:30)

Invited talk: *Not Enough Data? A Paradox of a Large Customer Database*, Makoto Abe

Session 2

Interesting Subset Discovery and its Application on Service Processes, Maitreya Natu and Girish Palshikar

Large-Scale Customized Models for Advertisers, Abraham Bagherjeiran, Andrew Hatch, Adwait Ratnaparkhi, and Rajesh Parekh

Automated Prompting in a Smart Home Environment, Barnan Das, Chao Chen, Nairanjana Dasgupta, and Diane J. Cook

12:20pm-1:30pm Lunch Break [Marigold Restaurant: http://www.marigold.com.au]

Reliability Issues in Knowledge Discovery (RIKD 2010)

Workshop Co-Chairs: Honghua Dai, James Liu, Evgueni Smirnov
Location: C243

Session 1 James Liu (Chair)

1:30-1:40 Chair Introduction

1:35-2:00pm *Co-occurence Based Predictors for Estimating Query Difficulty*, Hazra Imran and Aditi Sharan

2:00-2:25pm *Towards the Detection of Potential Contradictions in Fuzzy Ontology Using a High Level Net Approach Integrated with*, Ke Wang, James Liu and Wei-min Ma

2:25- 2:50pm *Semantic Image Retrieval Based on Multiple-Instance Learning*, Min Yao, Wenshen Yi, Rong Zhu and Ran Chen

2:50-3:15pm *Learning TAN Structure with Mixed Non-i.i.d. Spamling*, Zhongfeng Wang, Zhihai Wang and Bin Fu

3:30-4:00pm Tea/coffee break

Session 2 Evgueni Smirnov (Chair)

4:00-4:25pm *Towards a Reliable Framework of Uncertainty-based Group Decision Support System*, Junyi CHAI and James Liu

4:25-4:50pm *A Study on the Accuracy of Frequency Measures and Its Impact on Knowledge Discovery in Single Sequence*, Min Gan and Honghua Dai

4:50-5:15pm *Learning Robust Bayesian Network Classifiers in the Space of Markov Equivalent Classes*, Zhongfeng Wang, Zhihai Wang and Bin Fu,

5:15-5:40pm *Toward Robust Probabilistic Dependency Filtering*, Dingcheng Feng, Feng Chen and Wenli Xu

5:40-6:05pm *Batch Mode Sparse Active Learning*, Lixin Shi and Yuhang Zhao

Social Interactions Analysis and Services Providers (SIASP 2010)

Workshop Co-Chairs: Hakim Hacid, Cecile Favre, Tetsuya Yoshida
Location: C244

9:30: Opening session

9:45 – 12:00: Information and Networking Dynics

S6202 *Catching a Viral Video*, Mirjam Wattenhofer (Google, Switzerland),
Yannet Interian (Google, USA), Jon Vaver (Google, USA), Tom Broxton (Google,
Switzerland)

S6205 *The Meme Ranking Problem: Maximizing Microblogging Virality*,
Dino Ienco (University of Torino, Italy), Francesco Bonchi (Yahoo! Research, Spain), Carlos
Castillo (Yahoo! Research, Spain)

DM281 *Detecting and Tracking Community Dynamics in Evolutionary Networks*,
Zhengzhang Chen & Ye Jin (North Carolina State University and Oak Ridge National Laboratory,
USA), Kevin Wilson (North Carolina State University and RTI International, USA), William Hendrix
(North Carolina State University, USA), Nagiza Samatova (Oak Ridge National Lab, North
Carolina State University, USA)

12:20pm- 1:30pm: Lunch break

1:30pm- 3:30pm: Social Content

S6214 *RnR: A System for Extracting Rationale from Online Reviews and Ratings*, Dwi Rahayu
(Monash University, Australia), Shonali Krishnaswamy (Monash University, Australia), Cyril
Labbe (University of Grenoble, France), Oshadi Alahakoon (Monash University, Australia)

S6215 *Extracting Representative Tags for Flickr Users*, Xian Chen (Konkuk University, Seoul,
Korea), Hyoseop Shin (Konkuk University, Seoul, Korea)

DM612 *A Personalized Social Matching System using Past Relations and User Similarities*, Richi
Nayak (Queensland University of Technology, Australia) Meng Zhang (Queensland
University of Technology, Australia) Lin Chen (Queensland University of Technology,
Australia)

S6203 *A Social Recommendation System for Online Dating Networks*, Lin Chen (Queensland
University of Technology, Australia), Richi Nayak (Queensland University of Technology,
Australia), Yue Xu (Queensland University of Technology, Australia)

3:30-4:00pm Tea/coffee break

4:00pm – 6:00pm: COMMUNITIES, RECOMMANDATION, AND INFLUENCE

S6213 *Bridging folksonomies and domain ontologies: Getting out non-taxonomic relations*, Chiraz
Trabelsi (Faculty of Sciences of Tunis, Tunisia), Aicha Ben Jerad (Faculty of Sciences of
Tunis, Tunisia), Sadok Ben Yahia (Faculty of Sciences, Tunisia)

S6206 *Modeling and Comparing the Influence of Neighbors on the Behavior of Users in Social
and Similarity Networks*, Mohsen Jamali (Simon Fraser University, Canada), Martin Ester
(Simon Fraser University, Canada)

S6207 *Toward Finding Hidden Communities based on User Profile*, Tetsuya Yoshida (Hokkaido
University, Japan)

DM925 *Coalition and Opposition: Community Discovery in Signed Collaborative Interaction
Networks*, Petko Bogdanov (University of California, Santa Barbara, USA), Nicholas Larusso
(University of California, Santa Barbara, USA), buj K. Singh (University of California, Santa
Barbara, USA)

6:00pm: Closing Comments

Privacy Aspects of Data Mining (PADM 2010)

Workshop Co-Chairs: Aris Gkoulalas-Divanis, Kun Liu, Vassilios S. Verykios, Ran Wolff
Location: C245

9:00-9:10am Opening Session

9:10-10:00am Keynote Talk: *Analyzing Private Network Data*, Gerome Miklau

10:00-10:30am Tea/coffee break

10:30-12:00pm Session 1: Privacy in Graphs and Social Networks

Anonymizing Graphs Against Weight-based Attacks, Y. Li and H. Shen

Identifying Similar Neighborhood Structures in Private Social Networks, L. Singh and C. Schramm
Privacy Violations Using Microtargeted Ads: A Case Study, A. Korolova

12:20-1:30pm Lunch Break

1:45-3:15pm Session 2: Privacy Preserving Data Publishing

On Attribute Disclosure in Randomization based Privacy Preserving Data Publishing, L. Guo, X. Ying and X. Wu

Identifying and Preventing Data Leakage in Multi-relational Classification, G. Hongyu, H. Viktor and E. Paquet

Language-Based Enforcement of Access Constraints on Geoscientific Image Data, P. Baumann

3:30-4:00pm Tea/coffee break

4:00-5:00pm Session 3: Privacy Preservation in Distributed Data Mining

Frequent Closed Itemsets for Privacy Preserving Mining in Distributed Databases, S. Kuno, K. Doi and A. Yamamoto

A Privacy-Preserving Framework for Gaussian Mixture Models, M. Shashanka

5:00-5:15pm Closing Remarks

Topic Feature Discovery and Opinion Mining (TFDOM 2010)

Workshop Co-Chairs: Yuefeng Li, Ning Zhong, Raymond Lau
Location: C246

09:00-10:00 Session 1 (Chair: Yuefeng Li)

09:00-09:15 Workshop Opening, Yuefeng Li

09:15-10:00 Keynote: *Opinion Mining: What is there for Data Miners*, Bing Liu

10:00-10:30 Tea/coffee break

10:30-12:00 Session 2 (Chair: Raymond Y.K. Lau)

10:30-10:50 *Sentence-Level and Document-Level Sentiment Mining for Arabic Texts*, Noura Farra, Elie Challita, Rawad Abou Assi, and Hazem Hajj

10:50-11:10 *Mining Users' Opinions Based on Item Folksonomy and Taxonomy for Personalized Recommender Systems*, Huizhi Liang, Yue Xu, and Yuefeng Li

11:10-11:30 *A Framework for Emotion Mining from Text in Online Social Networks*, Mohamed Yassine and Hazem Hajj

11:30-11:50 *Augmenting Chinese Online Video Recommendations by Using Virtual Ratings Predicted by Review Sentiment Classification*, Weishi Zhang, Guiguang Ding, Li Chen, and Chunping Li

12:20-13:30 Lunch break

13:30-15:30 Session 3 (Chair: Yuefeng Li)

13:30-13:50 *The PARIS Algorithm for Determining Latent Topics*, Michal Aharon, Ira Cohen, Arik Itskovitch, Inbal Marhaim, and Ron Banner

13:50-14:10 *Infrequent Purchased Product Recommendation Making Based on User Behaviour and Opinions in E-commerce Sites*, Noraswaliza Abdullah, Yue Xu, Shlomo Geva, and Jinghong Chen

14:10-14:30 *A Block Mixture Model for Pattern Discovery in Preference Data*, Nicola Barbieri, Massimo Guarascio, and Giuseppe Manco

14:30-14:50 *Mining Arabic Business Reviews*, Mohamed Elhawary and Mohamed Elfeky

14:50-15:10 *High-Order Concept Associations Mining and Inferential Language Modeling for Online Review Spam Detection*, C.L. Lai, K.Q. Xu, Raymond Y.K. Lau, Yuefeng Li, and Dawei Song

15:30-16:00 Tea/coffee break

16:00-17:00 Session 4 (*Chair:* Yuefeng Li)

16:00-16:50 Panel Discussion Session

16:50-17:00 *Workshop Closing Comments,* Raymond Y K Lau

Chance Discovery (IWCD 2010)

Workshop Co-Chairs: Akinori Abe, Yukio Ohsawa
Location: C247

09:35: Opening remarks

09:40--10:20: *Innovators Marketplace as Game on the Table versus Board on the Web*, Yukio Ohsawa, Kensuke Okamoto, Yuma Takahashi, and Yoko Nishihara

10:20--10:30: Tea/coffee break

10:30--11:10: *Paired Evaluators Method to Track Concept Drift: An Application for Hedge Funds Operations*, Masabumi Furuhata, Takanobu Mizuta, and Jihei So

11:10--11:50: *Temporal Logic TL_DU Modeling Local and Global Discovery with Logical Uncertainty*, Vladimir Rybakov

12:20--13:30: Lunch break

13:30--14:10: *Automated Empirical Selection of Rule Induction Methods based on Recursive Iteration of Resampling Methods and Multiple testing*, Shusaku Tsumoto and Shoji Hirano

14:10--14:50: *Finding Rare Patterns with Weak Correlation Constraints*, Yoshiaki Okubo, Makoto Haraguchi, and Takeshi Nakajima

14:50--15:30: *Change with Delayed Labeling: when is it detectable?*, Indre Zliobaite

15:30--15:50: Tea/coffee break

15:50--16:30: *Text Categorization with Considering Temporal Patterns of Term Usages*, Hidenao Abe and Shusaku Tsumoto

16:30--16:30: *Bullshitting as Chance-Faking*, Lorenzo Magnani and Emanuele Bardone (not attending)

16:30--17:10: *Curation in Chance Discovery*, Akinori Abe

17:10--17:30: Discussion and closing remarks

Visual Analytics and Knowledge Discovery (VAKD 2010)

Workshop Co-Chairs: Simeon Simoff, Pak Chung Wong, Mike Sips, Arturas Mazeika
Location: C254

Session 1: Welcome and Invited Talk

9:00 Welcome

9:10 Invited Talk: *Its hard to draw a graph*, Peter Eades

Session 2: Oral Presentations

10:30 *User-based active learning*, Christin Seifert and Michael Granitzer

10:55 *Visually Controllable Data Mining Methods*, Kai Puolamaki, Panagiotis Papapetrou, and Jefrey Lijffijt

11:20 *SIMPLE: Interactive Analytics on Patent Data*, Scott Spangler, Ying Chen, Jeffrey Kreulen, Stephen Boyer, Thomas Griffin, Alfredo Alba, Linda Kato, Ana Lelescu, and Su Yan

11:45 *Evolving Ensemble-Clustering to a Feedback-Driven Process*, Martin Hahmann, Dirk Habich, and Wolfgang Lehner

12:20--1:30pm: Lunch break

Session 3: Invited Talk and Challenge Panel

1:30pm Invited Talk: *Visual Pattern Mining*, Bart Goethals

2:20pm *BODY - Buckets Of Disease sYmptoms for Disease Outbreak Analysis*, Hanisha

Veeramachaneni, Soujanya Vadapalli, and Kamalakar Karlapalem,

2:45pm *Analysing large-scale, time-oriented datasets with PRISMA*, Rafael Guimarães, Nikolas Carneiro, Bianchi Meiguins, and Aruanda Meiguins

3:10pm Challenge Panel Preferably folks who has seen the real world

3:30pm--4:00pm: Tea/coffee break

Session 4: Oral Presentation and Collaboration Panel

4:00pm *A Visual Analytics Tool for Analysing Microarray Data*, Martin èlie, Monique Noirhomme-Fraiture, Quang Vinh Nguyen, and Simeon Simoff

4:25pm *High-dimensional multimodal distribution embedding*, Eniko Szekely, Eric Bruno, and Stephane Marchand-Maillet,

4:50pm Panel Keynote: *Research Directions and Collaborations in Visual nalytics*, Joseph Kielman

Session 5: Collaboration Panel Research Directions in the interplay of Knowledge Discovery and Interactive Visualization

5:15pm Session 6: Closing Session

6.00pm End of Workshop

7.30pm Dinner: *Future of VAKD*

Business Analytics Forum

Tuesday, December 14, 2010 (Building 5, Haymarket Campus, UTS)

Location: B102
Time: 6:15pm – 7:15pm
Topic: A group of senior representatives from academic, industrial/government and vendor institutions will share their opinion on challenges and prospects in the real-world data mining and analytics, next-generation of data mining and analytics, and on bridging the gap between academic and business.

Note: Drinks and wine will be provided. Everyone is welcome.

DSKDBI Colloquium Program

Tuesday, December 14, 2010 (C120, Building 5, Haymarket Campus, UTS)

Session I: Business Solutions

9:00 -10:00 *Invited Talk: Analysing Extremely Large Volumes of Business Data: Practical Considerations*, Paul Segal, Teradata, Australia

10:00 – 10:20 Tea/coffee break

10:20 -12:20 *Invited Talk: Why Analytics isn't Enough,* Evan Stubbs, SAS, Australia

12:20 - 1:30 Lunch at Marigold Restaurant, Level 5, 683-689 George Street

Session II: Knowledge Advancement

1:30 – 3:30 *Invited talk: Association Discovery*, Prof Geoff Webb, Monash University, Australia

3:30 - 4:00 Tea/coffee break

Session III: Enterprise Applications

4:00 – 5:00 *Invited talk*. *Behavior Modeling, Analysis and Mining,* Prof Longbing Cao, University of Technology, Sydney

5:00 – 6:00 *Invited talk: Integrating risk analytics with processing systems*, Ben Slater, Accenture Information Management Services, Australia

6:15 – 7:15 Business Analytics Forum joint with ICDM workshops (Drinks) [B102]

Wednesday, December 15, 2010 (C120, Building 5, Haymarket Campus, UTS)

Session IV: ICDM2010 Keynote Speech

9:00 -10:00 ICDM2010 Keynote Speech, *Mining Billion-node Graphs: Patterns, Generators and Tools,* Prof Christos Faloutsos, Carnegie Mellon University (Venue: B111, Building 5, Haymarket, UTS)

10:00 – 10:20 Tea/coffee break

Session II: Knowledge Advancement (Continued)

10:20 -12:20 *Invited talk: On Mining Graph Patterns*, Prof Philip S Yu, University of Illinois at Chicago, USA

12:20 - 1:30 Lunch at Marigold Restaurant, Level 5, 683-689 George Street

1.30 – 3:30 *Invited talk: Discovery of Patterns in Global Earth Science Data using Data Mining*, Prof Vipin Kumar, University of Minnesota, USA

3:30 – 4:00 Tea/coffee break

4:00 – 6:00 *Invited talk: On Applications of the EM Algorithm*, Prof Geoff McLachlan, University of Queensland, Australia

6:30 – 8:30 Reception at Building 5, UTS

Thursday, December 16, 2010 (C120, Building 5, Haymarket Campus, UTS)

Session IV: ICDM2010 Keynote Speech

9:00 -10:00 ICDM2010 Keynote Speech, *10 Years of Data Mining Research: Retrospect and Prospect,* Prof Xindong Wu, University of Vermont
(Venue: B111, Building 5, Haymarket, UTS)

10:00 – 10:20 Tea/coffee break

Session II: Knowledge Advancement (Continued)

10:20 -12:20 *Invited talk:* *From Feature Construction, to Simple but Effective Modeling, to Domain Transfer*, Dr Wei Fan, IBM T.J.Watson Research, USA

12:20 – 1:30 Lunch at Marigold Restaurant, Level 5, 683-689 George Street

Excursion 1:40

Friday, December 17, 2010 (C120, Building 5, Haymarket Campus, UTS)

Session III: Enterprise Applications (Continued)

9:00 – 10:00 *Invited talk:* *Some Challenges with Analytics*, Dr Warwick Graco, Australian Taxation Office, Australia

10:00 – 10:20 Tea/coffee break

11:20 – 12:20 *Invited talk:* *Discriminative Subgraph Mining,* Prof Wei Wang, University of North Carolina at Chapel Hill, USA

12:20 – 1:30 Lunch at Marigold Restaurant, Level 5, 683-689 George Street

ICDM'10 Demos and Exhibitions

General Information for Demos

Morning tea/coffee (10:00-10:20)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Conference Lunch (12:20-13:30)
Location: Marigold Restaurant, Level 5, 683-689 George Street, Sydney

Afternoon tea/coffee (15:30-16:00)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Demo Time: 13:30 - 15:30, Wednesday, December 15 and 10:20 - 12:20, Friday, December 17
Location: Lobby, Block B, Building 5, Haymarket Campus, UTS

Accepted Demos

SK209   Philipp Kranen, Hardy Kremer, Timm Jansen, Thomas Seidl, Albert Bifet, Geoff Holmes, and Bernhard Pfahringer, *Clustering Performance on Evolving Data Streams: Assessing Algorithms and Evaluation Measures within MOA*

SK208 Jiyang Chen, Justin Fagnan, Randy Goebel, Reihaneh Rabbany, Farzad Sangi, Mansoureh Takaffoli, Eric Verbeek, and Osmar Zaiane, *Meerkat: Community Mining in Social Networks*

SK207 Stephan Gunnemann, Hardy Kremer, Ines Farber, and Thomas Seidl, *MCExplorer: Interactive Exploration of Multiple (Subspace) Clustering Solutions*

SK206 Chayapol Moemeng, Xinhua Zhu, Longbing Cao, and Jiahang Chen, *i-Analyst: An Agent-Based Distributed Data Mining Platform*

SK205   Anthony Goldbloom, *Data Mining Competitions - a potent tool that facilitates real-time science*

SK204 Brett Gillick, Hasnain AlTaiar, Jonathan Liono, Shonali Krishnaswamy, Nicholas Nicoloudis, Mohamed Gaber, Abhijat Sinha, and Arkady Zaslavsky, *Clutter-Adaptive Visualisation for Mobile Data Mining*

SK201   Brendad Kitts, Liang Wei, Dyang Au, John Sobieski, Jon Borchardt, and Brian Burdick, *Television Contextual Ad Targeting Heatmap*

Exhibitions

SAS

Springer

Salford Systems

ICDM'10 Technical Program

Wednesday, December 15, 2010 (Building 5, Haymarket Campus, UTS)

Conference Opening (08:30--09:00)
Location: B111
UTS Welcome: Professor Attila Brungs (UTS DVCR)

ICDM: Keynote (09:00-10:00)
Location: *B111*
Title: *Mining Billion-node Graphs: Patterns, Generators and Tools*
Speaker: Christos Falutsos
Chair: Geoff Webb

Tea/coffee Break (10:00-10:20)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Session 1A: Clustering (Room: B111)
Session Chair: Chris Ding
Session Time: 10:20-12:20
Regular Papers:

- *A Conscience On-line Learning Approach for Kernel-Based Clustering,* Chang-Dong Wang, Jian-Huang Lai, and Jun-Yong Zhu

- *Active Spectral Clustering,* Xiang Wang and Ian Davidson

- *Learning a Bi-Stochastic Data Similarity Matrix,* Fei Wang, Ping Li, and Arnd Christian Konig

Short Papers:

- *Efficient Constrained Co-clustering with Background Knowledge*, Xiaoxiao Shi, Wei Fan, and Philip Yu

- *Minimizing the Variance of Cluster Mixture Models for Clustering Uncertain Objects,* Francesco Gullo, Giovanni Ponti, and Andrea Tagarelli

- *Resilient K-d Trees: K-Means in Space Revisited,* Fabian Gieseke, Gabriel Moruz, and Jan Vahrenhold

- *K-AP: Generating Specified K Clusters by Efficient Affinity Propagation,* Xiangliang Zhang, Wei Wang, Kjetil Norvag, and Michele Sebag

- *Understanding and Enhancement of Internal Clustering Validation Measures,* Yanchi Liu, Zhongmou Li, Hui Xiong, Xuedong Gao, Junjie Wu

- *On normalizing fuzzy coincidence matrices to compare fuzzy and/or possibilistic partitions with the Rand index,* Romain Quere, Hoel Le Capitaine, Carl Frelicot, and Noel Fraisseix

Session 1B: Co-learning, semi-supervised learning, active learning, instance-based learning & evaluation (Room: B112)
Session Chair: Jimeng Sun
Session Time: 10:20-12:20
Regular Papers:

- *A Novel Contrast Co-Learning Framework for Generating High Quality Training Data,* Zeyu Zheng, Jun Yan, Shuicheng Yan, Ning Liu, Zheng Chen, Ming Zhang

- *An Extensive Empirical Study on Semi-supervised Learning,* Yuanyuan Guo, Xiaoda Niu and Harry Zhang

- *Active Learning from Multiple Noisy Labelers with Varied Costs,* Yaling Zheng, Stephen Scott, and Kun Deng

- *Consequences of Variability in Classifier Performance Estimates,* Troy Raeder, T. Ryan Hoens, and Nitesh Chawla

Short Papers:

- *Training Data Extraction from Completely Different Feature Spaces via Spectral Transformation,* Xiaoxiao Shi, Qi Liu, Wei Fan, Philip Yu, and Ruixin Zhu

- *Graph-Based Semi-supervised Learning with Adaptive Similarity Estimation,* Xianchao Zhang, Yansheng Jiang, Wenxin Liang, and Xin Han

- *Active Learning with Human-Like Noisy Oracle,* Jun Du and Charles Ling

- *SONNET: Efficient Approximate Nearest Neighbor using Multi-Core,* Mohammad Hasan, Hilmi Yildirim, and Ahbirup Chakraborty

Session 1C: Applications (Room: B102)
Session Chair: Shusaku Tsumoto
Session Time: 10:20-12:20
Regular Papers:

- *Separation of Interleaved Web Sessions with Heuristic Search,* Marko Pozenel, Viljan Mahnic, and Matjaž Kukar

- *An Unsupervised Approach to Modeling Personalized Contexts of Mobile Users,* Tengfei Bao, Happia Cao, Enhong Chen, Jilei Tian, and Hui Xiong

- *Viral Marketing for Multiple Products, Anirban Majumder,* Samik Datta, and Nisheeth Shrivastava

Short Papers:

- *Category Mining by Heterogeneous Data Fusion Using PdLSI Model in a Retail Service,* Tsukasa Ishigaki, Takeshi Takenaka, and Yoichi Motomura

- *Advertising Campaigns Management: Should We Be Greedy?,* Sertan Girgin, Jeremie Mary, Philippe Preux, and Olivier Nicol

- *Content-Based Methods for Predicting Web-Site Demographic Attributes,* Santosh Kabbur, Eui-Hong Han, and George Karypis

- *Attribution of Conversion Events to Multi-Channel Media,* Brendan Kitts, Liang Wei, Dyng Au, anda Powter, Brian Burdick

- *An Approach for Automatic Sleep Stage Scoring and Apnea-Hypopnea Detection,* Tim Schlüter and Stefan Conrad

- *Modeling Experts and Novices in Citizen Science data for Species Distribution Modeling,* Jun Yu, Weng-Keen Wong, and Rebecca Hutchinson

LUNCH (12:20-13:30)
Location: Marigold Restaurant, Level 5, 683-689 George Street

Session 2A: Clustering (Room: B111)
Session Chair: Kai Ming Ting
Session Time: 13:30-15:30
Regular Papers:

- *Clustering Large Attributed Graphs: An Efficient Incremental Approach,* Yang Zhou, Hong Cheng, and Jeffrey Xu Yu

- *minCEntropy: a Novel Information Theoretic Approach for the Generation of Alternative Clusterings,* Xuan Vinh Nguyen and Julien Epps

- *Discovering Correlated Subspace Clusters in 3D Continuous-Valued Data,* Kelvin Sim, Vivekanand Gopalkrishnan, and Zeyar Aung

- *Co-clustering of Lagged Data, Eran Shaham*, David Sarne, and Boaz Ben-Moshe

- *Bayesian Maximum Margin Clustering,* Bo Dai, Baogang Hu, and Gang Niu

- *Multi-agent random walks for local clustering on graphs,* Morteza Alamgir and Ulrike von Luxburg

Session 2B: SVM, kernel, vector & matrix techniques & inference in factorial spaces (Room: B112)
Session Chair: Hisashi Kashima
Session Time: 13:30-15:30
Regular Papers:

- *Permutations as Angular Data: Efficient Inference in Factorial Spaces,* Sergey Plis, Terran Lane and Vince Calhoun

- *Improving Kernel Methods through Complex Data Mapping,* Hang Zhou, Fabio Ramos, and Eric Nettleton

- *NESVM: a Fast Gradient Method for Support Vector Machines,* Tianyi Zhou, Dacheng Tao, and Xindong Wu

Short Papers:

- *Factorization Machines,* Steffen Rendle

- *Interval-valued Matrix Factorization with Applications,* Zhiyong Shen, Liang Du, Xukun Shen, and Yi-Dong Shen

- *Accelerating Radius-Margin Model Selection for SVMs Using Geometric Bounds,* Ben Goodrich, David Albrecht, and Peter Tischer

- *Sparse Boolean Matrix Factorizations,* Pauli Miettinen

- *Compressed Non-Negative Sparse Coding,* Fei Wang and Ping Li

- *Assessing Data Mining Results on Matrices with Randomization,* Markus Ojala

Session 2C: Privacy preserving data mining, ensemble learning & regression (Room: B102)
Session Chair: Aris Gkoulalas-Divanis
Session Time: 13:30-15:30
Regular Papers:

- *Bayesian Aggregation of Binary Classifiers,* Sunho Park and Seungjin Choi

- *Exploiting Unlabeled Data to Enhance Ensemble Diversity,* Min-Ling Zhang and Zhi-Hua Zhou

- *A Pairwise-Systematic Microaggregation for Statistical Disclosure Control,* Md Enamul Kabir, Hua Wang, and Yanchun Zhang

Short Papers:

- *Homotopy Regularization for Boosting,* Zheng Wang, Yangqiu Song, and Changshui Zhang

- *Anomaly Detection Using an Ensemble of Feature Models,* Keith Noto, Carla Brodley, and Donna Slonim

- *Enforcing Vocabulary k-Anonymity by Semantic Similarity Based Clustering,* Junqiang Liu and Ke Wang

- *Probabilistic Inference Protection on Uncertain Data,* Raymond Wong, Ada Fu, Ke Wang, Yabo Xu, Jian Pei, and Philip Yu

- *Anonymizing Temporal Data,* Ke Wang, Yabo Xu, Raymond Chi-Wing Wong, Ada Wai-Chee Fu

- *Passive Sampling for Regression,* Hwanjo Yu and Sungchul Kim

ICDM Contest Overview Session (14:00 – 15:30)
Location: Room C119
Session Chair: Marcin Wojnarski

Tea/coffee Break (15:30-16:00)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Session 3A: Sequence mining & clustering (Room: B111)
Session Chair: Ke Wang
Session Time: 16:00-18:00
Regular Papers:

- *Constraint Based Dimension Correlation and Distance Divergence for Clustering High-Dimensional Data,* Xianchao Zhang, Yao Wu, and Yang Qiu

- *Abstraction Augmented Markov Models, Cornelia Caragea,* Adrian Silvescu, Doina Caragea, and Vasant Honavar

- *Training Conditional Random Fields Using Transfer Learning for Gesture Recognition,* Jie Liu, Kai Yu, Yi Zhang, and Yalou Huang

- *Mining Closed Strict Episodes,* Nikolaj Tatti and Boris Cule

Short Papers:

- *Efficient Episode Mining with Minimal and Non-overlapping Occurrences,* Huisheng Zhu, Peng Wang, Xianmang He, Yujia Li, Wei Wang, and Baile Shi

- *Accelerating Dynamic Time Warping Subsequence Search with GPUs and FPGAs,* Doruk Sart, Abdullah Mueen, Walid Najjar, Vit Niennattrakul, and Eamonn Keogh

- *Frequent Instruction Sequential Pattern Mining in Hardware Sample Data,* Jia Zou, Jing Xiao, Rui Hou,

and Yanqi Wang

- *A Pseudo-CRF Approach to Sequence Labeling Problems,* Shing-Kit Chan and Wai Lam

Session 3B: Social networks, social media and probabilistic techniques (Room: B112)
Session Chair: Jie Tang
Session Time: 16:00-18:00
Regular Papers:

- *Learning Markov Network Structure with Decision Trees,* Daniel Lowd and Jesse Davis

- *Modeling Information Diffusion in Implicit Networks,* Jaewon Yang and Jure Leskovec

- *Discovering Overlapping Groups in Social Media,* Xufei Wang, Lei Tang, Huiji Gao, and Huan Liu

- *Scalable Influence Maximization in Social Networks under the Linear Threshold Model,* Wei Chen, Yifei Yuan, and Li Zhang

Short Papers:

- *Supervised Link Prediction Using Multiple Sources,* Berkant Savas, Wei Tang, Zhengdong Lu, and Inderjit Dhillon

- *A Generalized Linear threshold Model for Multiple Cascades,* Nishith Pathak, Arindam Banerjee, and Jaideep Srivastava

- *MoodCast: Emotion Prediction via Dyamnic Continuous Factor Graph Model,* Yuan Zhang, Jie Tang, Jimeng Sun, Yiran Chen, and Jinghai Rao

- *Leveraging D-Separation for Relational Data Sets,* Matthew Rattigan and David Jensen

Session 3C: Stream mining, multimodal data and inference in factorial spaces (Room: B102)
Session Chair: Jeffrey Yu
Session Time: 16:00-18:00
Regular Papers:

- *Polishing the Right Apple: Anytime Classification Also Benefits Data Streams with Constant Arrival Times,* Jin Shieh and Eamonn Keogh

Short Papers:

- *Multi-Stream Join Answering for Mining Significant Cross-Stream Correlations,* Robert Gwadera

- *A System for Mining Temporal Physiological Data Streams for Advanced Prognostic Decision Support,* Jimeng Sun, Daby Sow, Jianying Hu, and Shahram Ebadollahi

- *Causal Discovery from Streaming Features*, Kui Yu, Xindong Wu, Hao Wang, and Wei Ding

- *ABS: the Anti Bouncing Model for Usage Data Streams,* Chongsheng Zhang, Florent Masseglia, and Yves Lechevallier

- *Data Editing Techniques to Allow the Application of Distance-Based Outlier Detection to Strames,* Vit Niennattrakul, Chotirat Ann Ratanamahatana, and Eamonn Keogh

- *Addressing Concept-Evolution in Concept-Drifting Data Streams,* Mohammad Masud, Qing Chen, Latifur Khan, Charu Aggarwal, Jing Gao, Jiawei Han, and Bhavani Thuraisingham

- *Classifiers and Clusters Ensembling for Mining Concept Drifting Data Streams,* Peng Zhang, Xingquan Zhu, Jianlong Tan, and Li Guo

- *Block-GP: Scalable Gaussian Process Regression for Multimodal Data,* Kamalika Das and Ashok Srivastava

- *Tru-Alarm: Trustworthiness Analysis of Sensor Networks in Cyber-Physical Systems,* Lu-An Tang, Xiao Yu, Sangkyum Kim, Jiawei Han, Chih-Chieh Hung, and Wen-Chih Peng

- *QMAS: Querying, Mining And Summarization of Multi-modal Databases,* Robson L. F. Cordeiro, Fan Guo, Donna S. Haverkamp, James H. Horne, Ellen K. Hughes, Gunhee Kim, Agma J. M. Traina, Caetano Traina Jr., and Christos Faloutsos

ICDM 10 Year Impact Award Address (Room B111)
Session Chair: Wei Wang
Session Time: 18:00-18:30

Reception (18:30-20:30)

Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Thursday, December 16, 2010 (Building 5, Haymarket Campus, UTS)

ICDM: Keynote (09:00-10:00)
Location: B111
Title: *10 Years of Data Mining Research: Retrospect and Prospect*
Speaker: Xindong Wu
Chair: Chengqi Zhang

Tea/coffee Break (10:00-10:20)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Session 4A: Clustering (Room: B111)
Session Chair: George Karypis
Session Time: 10:20-12:20

Regular Papers:

- *gSkeletonClu: Density-based Network Clustering via Structure-Connected Tree Division or Agglomeration,* Heli Sun, Jianbin Huang, Jiawei Han, Hongbo Deng, Peixiang Zhao, and Boqin Feng
- *Multi-Dimensional Mass Estimation and Mass-based Clustering,* Kai Ming Ting and Jonathan Wells
- *CLUSMASTER: A Clustering Approach for Sampling Data Streams in Sensor Networks,* Alzennyr Da Silva, Raja Chiky, and Georges Hebrail
- *LogTree: A Framework for Generating System Events from Raw Textual Logs,* Liang Tang and Tao Li
- *Algorithm for Discovering Low-Variance 3-Clusters From Real-Valued Datasets,* Zhen Hu and Raj Bhatnagar

Short Papers:

- *Enhancing Single-Objective Projective Clustering Ensembles,* Carlotta Domeniconi, Francesco Gullo, and Andrea Tagarelli
- *Hierarchical Ensemble Clustering,* Li Zheng, Tao Li, and Chris Ding

Session 4B: Association discovery & uplift modeling (Room: B112)
Session Chair: Laitfur Khan
Session Time: 10:20-12:20
Regular Papers:

- *Subgroup Discovery meets Bayesian networks - an Exceptional Model Mining approach,* Wouter Duivesteijn, Arno Knobbe, Ad Feelders, and Matthijs van Leeuwen
- *Efficient discovery of the top-K optimal dependency rules with the Fisher's exact test of significance,* Wilhelmiina Hämäläinen
- *Decision trees for uplift modeling,* Szymon Jaroszewicz and Piotr Rzepakowski
- *PGLCM: Efficient Parallel Mining of Closed Frequent Gradual Itemsets,* Trong Dinh Thac Do, Anne Laurent, and Alexandre Termier
- *Stratified Sampling for Data Mining on the Deep Web,* Tantan Liu, Fan Wang, and Gagan Agrawal

Short Papers:

- *Testing the Significance of Patterns in Data with Cluster Structure,* Niko Vuokko and Petteri Kaski
- *Approximating Frequentness Probability of Itemsets in Uncertain Data,* Toon Calders, Calin Garboni, and Bart Goethals

Session 4C: Classification (Room: B102)
Session Chair: Yan Liu
Session Time: 10:20-12:20
Regular Papers:

- *A Graph-Based Approach for Multi-Folder Email Classification,* Sharma Chakravarthy, Aravind Venkatachalam, and Aditya Telang

- *A Binary Decision Diagram-Based One-Class Classifier,* Takuro Kutsuna

- *SMILE: A Similarity-Based Approach for Multiple Instance Learning,* Yanshan Xiao, Bo Liu, Longbing Cao, Jie Yin, and Xindong Wu

- *Towards Structural Sparsity: An Explicit l_2/l_0 Approach,* Dijun Luo, Chris Ding, and Heng huang

Short Papers:

- *A Robust, Accurate, and Fast Stochastic Gradient Training Method for Modeling Latent-Information in Data,* Xu Sun, Hisashi Kashima, Takuya Matsuzaki, and Naonori Ueda

- *Quantification via Probability Estimators,* Antonio Bella, Cèsar Ferri, José Hernández-Orallo, and María José Ramírez -Quintana

- *Monotone Relabeling in Ordinal Classification*, Ad Feelders

- *Discrimination Aware Decision Tree Learning,* Faisal Kamiran, Toon Calders, and Mykola Pechenizkiy

LUNCH (12:20-13:30)
Location: Marigold Restaurant, Level 5, 683-689 George Street

Excursion (1:40 - 6.30)

* see page 43 for detailed information

BANQUET: 7:00pm – 10:00pm, Showboat Two Sydney Harbour Dinner Cruise and Live Show
Locaton: Please meet at 6.30pm at Campbell Cove Wharf, Sydney Harbour foreshore (waterfront). Campbell Cove is opposite the Opera House – walk past Circular Quay and follow the foreshore around past the historic Rocks to Campbell Cove (see map, page 56). Your Showboat Two banquet dinner (look for the red and white paddle-wheeler) will board at 7.00pm for a 7.15pm departure (please bring your banquet ticket).

* See map, page 56

Friday, December 17, 2010 (Building 5, Haymarket Campus, UTS)

Session 5A: Text mining (Room: B111)
Session Chair: Arindam Banerjee
Session Time: 09:00-10:00
Regular Papers:

- *Multi-Document Summarization Using Minimum Distortion,* Tengfei Ma and Xiaojun Wan

- *Weighted Feature Subset Non-Negative Matrix Factorization and its Applications to Document Understanding,* Dingding Wang, Chris Ding, and Tao Li

Short Papers:

- *Document Similarity Self-Join with MapReduce,* Claudio Lucchese, Gianmarco De Francisci Morales, and Ranieri Baraglia

- *Learning Preferences using Millions of Features by Enforcing Sparsity,* Xi Chen, Bing Bai, Yanjun Qi, Qihang Lin, and Jaime Carbonell

Session 5B: Graph & network mining (Room: B112)
Session Chair: Ee-Peng Lim
Session Time: 09:00-10:00
Short Papers:

- *Patterns on the Connected Components of Terabyte-Scale Graphs,* U Kang, Mary McGlohon, Leman Akoglu, and Christos Faloutsos

- *Max-Clique: A Top-down Graph based Approach to Frequent Pattern Mining,* Yan Xie and Philip Yu

- *Subspace Clustering Meets Dense Subgraph Mining: A Synthesis of Two Paradigms,* Stephan Günnemann, Ines Färber, Brigitte Boden, and Thomas Seidl

- *On the Vulnerability of Large Graphs,* Hanghang Tong, B. Aditya Prakash, Charalampos Tsourakakis, Tina Eliassi-Rad, Christos Faloutsos, and Duen Horng Chau

- *Node Similarities from Spreading Activation,* Kilian Thiel and Michael Berthold

- *Mixed-Membership Stochastic Block-Models for Transactional Networks,* Mahdi Shafiei and Hugh Chipman

Session 5C: Feature selection & construction (Room: B102)
Session Chair: Katharina Morik
Session Time: 09:00-10:00
Regular Papers:

- *Feature Selection for Unsupervised Learning using Random Cluster Ensembles,* Haytham Elghazel and Alex Aussem

- *A Variance Reduction Framework for Stable Feature Selection,* Yue Han and Lei Yu

Short Papers:

- *Fast Stochastic Search Variable Selection inside the DBMS,* Mario Navas, Carlos Ordonez, and Veerabhadran Baladandayuthapani

- *Active Improvement of Hierarchical Object Features under Budget Constraints,* Nicolas Cebron

Tea/coffee Break (10:00-10:20)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

Session 6A: Text mining, transfer learning, time series & spatio-temporal mining (Room: B111)
Session Chair: Toon Calders
Session Time: 10:20-12:20
Regular Papers:

- *iSAX 2.0: Indexing and Mining One Billion Time Series,* Alessandro Camerra, Jin Shieh, Themis Palpanas, and Eamonn Keogh

- *Mining Sensor Streams for Discovering Human Activity Patterns Over Time,* Parisa Rashidi and Diane Cook

- *Spatiotemporal Event Detection in Mobility Network,* Tom Au, Rong Duan, Heeyoung Kim, and Guang Qin Ma

- *Fast and Flexible Multivariate Time Series Subsequence Search,* Kanishka Bhaduri, Qiang Zhu, Nikunj Oza, and Ashok Srivastava

Short Papers:

- *On Finding Frequent Patterns in Spatial Event Sequences,* Andrea Campagna and Rasmus Pagh

- *Financial Forecasting with Gompertz Multiple Kernel Learning,* Han Qin, Dejing Dou, and Yue Fang

- *Location and Scatter Matching for Dataset Shift in Text Mining,* Bo Chen, Wai Lam, Ivor W. Tsang, and Tak-Lam Wong

- *Transfer Learning via Cluster Correspondence Induction,* Mingsheng Long, Wei Cheng, Xiaoming Jin, Jianmin Wang, Dou Shen

Session 6B: Graph & network mining (Room: B112)
Session Chair: Xiaoli Li
Session Time: 10:20-12:20
Regular Papers:

- *Detecting Blackholes and Volcano Patterns in Directed Networks,* Zhongmou Li, Hui Xiong, Yanchi Liu, Aoying Zhou

- *Network Simplification with Minimal Loss of Connectivity,* Fang Zhou, Sébastien Mahler, and Hannu Toivonen

- *Multi-Label Feature Selection for Graph Classification,* Xiangnan Kong and Philip Yu

- *Edge Weight Regularization Over Multiple Graphs For Similarity Learning,* Pradeep Muthukrishnan, Dragomir Radev, and Qiaozhu Mei

- *Detecting Novel Discrepancies in Communication Networks,* James Abello, Tina Eliassi-Rad, and Nishchal Devanur

Short Papers:

- *Visualizing Graphs using Minimum Spanning Dendrograms,* Daniel Svonava and Michail Vlachos

- *Bonsai: Growing Interesting Small Trees,* Stephan Seufert, Srikanta Bedathur, Julián Mestre, and Gerhard Weikum

Session 6C: Outlier detection, distance learning & minhash computation (Room: B102)
Session Chair: Jianyong Wang
Session Time: 10:20-12:20
Regular Papers:

- *Exponential Family Tensor Factorization for Missing-Values Prediction and Anomaly Detection,* Kohei Hayashi, Takashi Takenouchi, Tomohiro Shibata, Yuki Kamiya, Daishi Kato, Kazuo Kunieda, Keiji Yamada, and Kazushi Ikeda

- *Exploiting Local Data Uncertainty to Boost Global Outlier Detection,* Bo Liu, Jie Yin, Yanshan Xiao, Longbing Cao, and Philip Yu

- *Finding Local Anomalies in Very High Dimensional Space,* Timothy de Vries, Sanjay Chawla, and Michael Houle

- *Rare Category Characterization,* Jingrui He, Hanghang Tong, and Jaime Carbonell

- *Adaptive Distances on Sets of Vectors,* Adam Woznica and Alexandros Kalousis

- *Improved Consistent Sampling, Weighted Minhash and L1 Sketching,* Sergey Ioffe

LUNCH & ICDM Community Meeting (12:20-14:00)
Location: Marigold Restaurant, Level 5, 683-689 George Street

Session 7A: Image retrieval (Room: B111)
Session Chair: Hui Xiong
Session Time: 14:00-14:30
Regular Papers:

- *Mother Fugger: Mining Historical Manuscripts with Local Color Patches,* Qiang Zhu and Eamonn Keogh

Short Papers:

- *Collaborative Learning between Visual Content and Hidden Semantic for Image Retrieval,* Jun Wu, Ming-Yu Lu, and Chun-Li Wang

Session 7B: Text mining (Room: B112)
Session Chair: Kyuseok Shim
Session Time: 14:00-16:00
Regular Papers:

- *An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews,* Peng Jiang, Chunxia Zhang, Hongping Fu, Zhendong Niu, and Qing Yang

- *D-LDA: A Topic Modeling Approach without Constraint Generation for Semi-Defined Classification,* Fuzhen Zhuang, Ping Luo, Zhiyong Shen, Qing He, Yuhong Xiong, and Zhongzhi Shi

- *Sequential Latent Dirichlet Allocation: Discover Underlying Topic Structures within a Document,* Lan Du, Wray Buntine, and Huidong Jin

- *Term Filtering with Bounded Error,* Zi Yang, Wei Li, Jie Tang, and Juanzi Li

Short Papers:

- *Efficient Probabilistic Latent Semantic Analysis with Sparsity Control,* Sen Liu, Chaolun Xia, Xiaohong Jiang

- *Topic Modeling Ensembles,* Zhiyong Shen, Ping Luo, Shengwen Yang, and Xukun Shen

- *What Do People Want in Microblogs? Measuring Interestingness of Hashtags in Twitter,* Jianshu Weng, Ee-peng Lim, Qi He, and Cane Wing-Ki Leung

- *Micro-blogging Sentiment Detection by Collaborative Online Learning,* Guangxia Li, Steven C.H. Hoi, Kuiyu Chang, and Ramesh Jain

Session 7C: Collaborative filtering, link prediction and recommender systems (Room: B102)
Session Chair: Dejing Dou
Session Time: 14:00-16:00
Regular Papers:

- *Learning Attribute-to-Feature Mappings for Cold-start Recommendations,* Zeno Gantner, Lucas Drumond, Christoph Freudenthaler, Steffen Rendle, and Lars Schmidt-Thieme

- *A Log-Linear Model with Latent Features for Dyadic Prediction,* Aditya Menon and Charles Elkan

Short Papers:

- *Recommending Social Events from Mobile Phone Location Data,* Daniele Quercia, Neal Lathia, Francesco Calabrese, Giusy Di Lorenzo, and Jon Crowcroft

- *Personalizing Web Page Recommendation via Collaborative Filtering and Topic-Aware Markov Model,* Qingyan Yang, Ju Fan, Jianyong Wang, and Lizhu Zhou

- *One Class Matrix Completion with Low Density Factorizations,* Vikas Sindhwani, Serhat Bucak, Jianying Hu, and Aleksandra Mojsilovic

- *Mining Public Transport Usage For Personalised Intelligent Transport Systems,* Neal Lathia, Jon Froehlich, and Licia Capra

- *The Effect of History on Modeling Systems' Performance: The Problem of the Demanding Lord,* George Giannakopoulos and Themis Palpanas

- *Generalized Probabilistic Matrix Factorizations for Collaborative Filtering,* Hanhuai Shan and Arindam Banerjee

- *Two of a Kind or the Ratings Game? Adaptive Pairwise Preferences and Latent Factor Models,* Suhrid Balakrishnan and Sumit Chopra

- *Learning Collaborative Filtering and Its Application on People-to-People Recommendation in Social Networks,* Xiongcai Cai, Michael Bain, Alfred Krzywicki, Wayne Wobcke, Yang-Sok Kim, Paul Compton, and Ashesh Mahidadia

ICDM'10 Top 10 Case Studies (14:30 – 16:00)
Chairs: Gabor Melli, Xindong Wu
Location: B111

Tea/coffee (16:00 – 16:15)
Location: Lobby, Block D, Building 5, Haymarket Campus, UTS

ICDM: Keynote (16:15-17:15)

Location: B111
Title: *Assessing the Significance of Groups in High-Dimensional Data*
Speaker: Geoff McLachlan
Chair: Bing Liu

CLOSING SESSION

ICDM'10 Top 10 Case Studies

Chairs: Gabor Melli, Xindong Wu
Panelists: Paul Beinat, Longbing Cao, Rong Duan, Christos Faloutsos, Rayid Ghani, Bart Goethals, Geoff McLachlan, and Ashok Srivasta
Location: B111
Title: *ICDM'10: Top 10 Case Studies*
Time: 14:30 – 16:00
Date: 17 Dec 2010

- *Analyzing System Logs for Online Failure Prediction on Supercomputers,* Zhiling Lan, Ziming Zheng, Jiexing Gu, Susan Coghlan, and Rajeev Thakur

- *Chemical and Biological Entity Extraction Techniques for Scientific Literature* , Su Yan, Stephen Boyer, Ying Chen, Alfredo Alba, Thomas D. Griffin, W. Scott Spangler, Ana Lelescu, and Jeffrey T. Kreulen

- *Click Fraud Attack Detection at Massive Scale over 5 Years at Microsoft* , Brendan Kitts, Jingying Zhang, Gang Wu, et al

- *Discovery of Precursors to Aviation Safety Incidents with Data Mining* , Ashok Srivastava, and Irving Statler

- *Enhancing the Chemical Safety of Agricultural Products* , Walter C P Magalhaes Junior, and Marilde Terezinha Prado Santos

- *Forecasting Skewed Biased Stochastic Ozone Days: analyses, solutions and beyond* , Wei Fan, Kun Zhang, and Xiaojing Yuan

- *MineFleet®: A Distributed Vehicle Performance Data Stream Mining System* , Hillol Kargupta, Kakali Sarkar, Michael Gilligan, Parag Namjoshi, Sai Subhash Paruchuru, Thiraphat Pongsudhiraks, and Robert Gilligan

- *Mining Medical Images* , Balaji Krishnapuram, and R. Bharat Rao

- *Social Security Data Mining for Public Services* , Longbing cao, Hans M Bohlscheid, Yanchang Zhao, Huaifeng zhang, Peter newbigin, Brett Clark, Yuming Ou, Jinjiu Li, Yong Yang, Chengqi Zhang, and Yanshan Xiao

- *US Department of Agriculture Risk Management Agency's crop insurance program* , Dr. Bert Little, Michael Schucking, and the CAE Research Center Team

ICDM'10 Tutorials

Tutorial 1: Spatial and Spatio-temporal Data Mining

Room: C119

Session Time: 10:20-12:20 Wednesday, December 15, 2010

Lecturers: Vania Bogorny (Univ. Federal de Santa Catarina, BRAZIL), Shashi Shekhar (Univ. Minnesota, USA)

Tutorial 2: Discovering Multiple Clustering Solutions: Grouping Objects in Different Views of the Data

Room: C119

Session Time: 16:00-18:00 Wednesday, December 15, 2010

Lecturer: Jun Huan (Univ. Kansas, USA)

Tutorial 3: Knowledge Discovery in Academic Drug Discovery Programs: Opportunities and Challenges

Room: C119

Session Time: 10:20-12:20 Thursday, December 16, 2010

Lecturer: Eamonn Keogh (Univ. California - Riverside, USA)

Tutorial 4: How to do Good Data Mining Research and get it Published in Top Venues

Room: C119

Session Time: 9:00-10:00 Friday, December 17, 2010

Lecturers: Emmanuel Müller (Karlsruhe Institute of Technology, Germany) & Stephan Günnemann & Ines Färber & Thomas Seidl (RWTH Aachen University, Germany)

ICDM'10 Keynote Speeches

Title: *Mining Billion-node Graphs: Patterns, Generators and Tools*

PROFESSOR Christos Faloutsos
Department of Computer Science, Carnegie Mellon University, USA

Abstract

What do graphs look like? How do they evolve over time?
How to handle a graph with a billion nodes?

We present a comprehensive list of static and temporal laws, and some recent observations on real graphs (eg., "eigenSpokes"). We describe some recent generators, which naturally match all of the known properties of real graphs. Finally, we present the "oddBall" tool for discovering anomalies and patterns, as well as an overview of the PEGASUS system which is designed to handle Billion-node graphs, running on top of the "hadoop" system.

Biography

Christos Faloutsos is a Professor at Carnegie Mellon University. He has received the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, fifteen "best paper" awards, and four teaching awards. He has served as a member of the executive committee of SIGKDD; and has published over 200 refereed articles, 11 book chapters and one monograph. He holds five patents and has given over 30 tutorials and 10 invited distinguished lectures. His research interests include data mining for graphs and streams, fractals, database performance, and indexing for multimedia and bio-informatics data.

Title: *Ten Years of Data Mining Research: Retrospect and Prospect*

PROFESSOR Xindong Wu
Department of Computer Science, University of Vermont, USA

Abstract

This talk reviews ten research activities with ICDM in the past ten years, discusses ten achievements in the data mining community, and presents ten research topics for the next ten years.

Biography

Xindong Wu is a Professor in the Department of Computer Science at the University of Vermont (USA) and a Yangtze River Scholar in the School of Computer Science and Information Engineering at the Hefei University of Technology (China). He holds a PhD in Artificial Intelligence from the University of Edinburgh, Britain. His research interests include data mining, knowledge-based systems, and Web information exploration. He has published over 200 refereed papers in these areas in various journals and conferences, including IEEE TKDE, TPI, ACM TOIS, DMKD, KAIS, IJCAI, AAAI, ICML, KDD, ICDM, and WWW, as well as 25 books and conference proceedings.

Dr. Wu is the founder and current Steering Committee Chair of the IEEE International Conference on Data Mining (ICDM), the founder and current Editor-in-Chief of Knowledge and Information Systems (KAIS, by Springer), the Founding Chair (2002-2006) of the IEEE Computer Society Technical Committee on Intelligent Informatics (TCII), and a Series Editor of the Springer Book Series on Advanced Information and Knowledge Processing (AI&KP). He was the Editor-in-Chief of the IEEE Transactions on Knowledge and Data Engineering (TKDE, by the IEEE Computer Society) between January 1, 2005 and December 31, 2008. He served as Program Committee Chair for ICDM '03 (the 2003 IEEE International Conference on Data Mining) and as Program Committee Co-Chair for KDD-07 (the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining), and is currently serving as Program Committee Chair (Knowledge Management Track) for CIKM 2010 (the 19th ACM Conference on Information and Knowledge Management).

Professor Wu is the 2004 ACM SIGKDD Service Award winner and the 2006 IEEE ICDM Outstanding Service Award winner. He has been an invited/keynote speaker at numerous international conferences including IEEE IRI 2010, IEEE GrC 2009, IDEAL 2009, JCKBSE 2008, HAIS 2008, NSF-NGDM'07, PAKDD-07, IEEE EDOC'06, IEEE ICTAI'04, IEEE/WIC/ACM WI'04/IAT'04, SEKE 2002, and PADD-97

Title: *Assessing the Significance of Groups in High-Dimensional Data*

PROFESSOR Geoff McLachlan
Department of Mathematics, University of Queensland, Australia

Abstract

We consider the problem of assessing the significance of groups in high-dimensional data. In the case of supervised classification where there is data of known origin with respect to the groups under consideration, a guide to the degree of separation among the groups can be given in terms of the estimated error rate of a classifier formed to allocate a new observation to one of the groups. Even in this case with labelled training data, care has to be taken with the estimation of the error rate at least for high-dimensional data to avoid an overly optimistic assessment due to selection biases. In the case of unlabelled data, the problem of assessing whether groups identified from some data mining or cluster analytic procedure are genuine can be quite challenging, in particular for a large number of variables. We shall focus on the use of a resampling approach to this problem applied in conjunction with factor analytic models for the generation of the bootstrap samples under the null hypothesis for the number of groups. The proposed methods are to be demonstrated in their application to some high-dimensional data sets from the bioinformatics literature.

Biography

Geoff McLachlan is Professor of Statistics in the Department of Mathematics and a Professorial Research Fellow in the Institute for Molecular Bioscience at the University of Queensland. He is also a chief investigator in the ARC Centre of Excellence in Bioinformatics. He currently holds an ARC Professorial Fellowship and is currently a member of the ARC College of Experts. He has written numerous research articles and six monographs, the last five in the Wiley series in Probability and Statistics. His current research interests are focussed in the fields of machine learning and bioinformatics.

ICDM'10 Social Program

Conference Reception

December 15, 6:30pm – 8:30pm, Lobby, Block D, Building 5, Haymarket Campus, UTS

Excursion

December 16, 1:40pm-6:30pm

Tour	Pick up time	Pick-up point	Drop-off point	Overview
City tour	Thursday 16 December 2010 1:40pm	Ground Floor Foyer, UTS Bldg 5	Campbell Cove (close to Circular Quay) for the Showboat Two dinner cruise (departing @ 7:15pm).	Half day city tour to The Rocks, Harbour Bridge, Mrs Macquarie's Chair, Darling Harbour, Hyde Park, Australian Museum, Bondi Beach.
Opera house tour	Thursday December 16, 2010 1:40pm All of four buses will depart at 1:40pm	Ground Floor Foyer, UTS Bldg 5	Opera House at Circular Quay. Walk to Campbell Cove for the Showboat Two dinner cruise (departing @ 7:15pm).	There will be four Opera House tours: Group A: 2.15pm(40pax) Group B: 3.45pm(40pax) Group C: 2.45pm(40pax) Group D: 4.15pm(40pax) Each group will have one hour in the Opera House. You may choose to visit other scenic or and/or historical places in the area before or after your 1-hour tour.*

Conference Banquet/Award Ceremony

December 16, 6:30pm – 10:00pm, Sydney Harbour Dinner Cruise and Live Show.

Please meet at 6.30pm at Campbell Cove Wharf, Sydney Harbour foreshore. Campbell Cove is opposite the Opera House – walk past Circular Quay and follow the foreshore (waterfront) around past the historic Rocks to Campbell Cove (see map, page 56). Your Showboat Two* banquet dinner (look for the red and white paddle-wheeler) will board at 7.00pm for a 7.15pm departure (please bring your banquet ticket).

* **Note**:   'Showboat Two' is the official name of the company who runs the boat and show. There is only one boat – a red and white paddle-wheeler.

ICDM'10 Organising Committee

Honorary Chair
Ramamohanarao Kotagiri (Melbourne University, Australia)
Conference Co-Chairs:
Chengqi Zhang (University of Technology, Sydney, Australia)
Dimitrios Gunopulos (University of Athens, Greece)
Program Co-chairs:
Geoff Webb (Monash University, Australia)
Bing Liu (University of Illinois at Chicago, USA)
Steering Committee:
Xindong Wu (Chair) (University of Vermont, USA)
David J. Hand (Imperial College, London, UK)
Ramamohanarao Kotagiri (University of Melbourne, Australia)
Vipin Kumar (University of Minnesota, USA)
Heikki Mannila (University of Helsinki, Finland)
Gregory Piatetsky-Shapiro (KDnuggets, USA)
Shusaku Tsumoto (Shimane University, Japan)
Benjamin W. Wah (University of Illinois, Urbana-Champaign, USA)
Philip S. Yu (University of Illinois at Chicago, USA)
Osmar R. Zaiane (University of Alberta, Canada)
Local Arrangements Chair:
Longbing Cao (University of Technology, Sydney, Australia)
Finance Chair:
Xingquan Zhu (University of Technology, Sydney, Australia)
Sponsorship Co-Chairs:
Paul Beinat (NeuronWorks International Pty Ltd, Australia)
Gabor Melli (PredictionWorks)
Workshop Co-Chairs:
Wynne Hsu (National University of Singapore, Singapore)
Wei Fan (IBM T. J. Watson Research Centre, USA)
Tutorial Co-Chairs:
Sanjay Chawla (University of Sydney, Australia)
Myra Spiliopoulou (University of Magdeburg, Germany)
Exhibits and Demos Co-Chairs:
Bart Goethals (University of Antwerp, Belgium)
Yucel Saygin (Sabanci University, Turkey)
Contest Chair:
Demetris Zeinalipour (University of Cyprus, Cyprus)
Publicity Co-Chairs:
Jian Yang (International WIC Institute/BJUT, China)
Taghi M. Khoshgoftaar (Florida Atlantic University, USA)
Wenjie Zhang (University of New South Wales, Australia)
Awards Committee:
Wei Wang (Chair) (University of North Carolina at Chapel Hill, USA)
Panels Chair:
Hillol Kargupta (University of Maryland Baltimore County, USA)

Student Travel Awards Committee Chair:
Wei Ding (University of Massachusetts Boston, USA)

Vice-Chairs:
Deepak Agarwal, Yahoo! Research, USA
Charu Aggarwal, IBM T. J. Watson Research Center, USA
Eugene Agichtein, Emory University, USA
Longbing cao, University of Technology, Sydney, Australia
Alok Choudhary, Northwestern University, USA
Diane Cook, Washington State University, USA
Ian Davidson, University of California, Davis, USA
Tina Eliassi-Rad, Lawrence Livermore National Laboratory, USA
Johannes Fürnkranz, TU Darmstadt, Knowledge Engineering Group, Germany
Bart Goethals, University of Antwerp, Belgium
Robert Grossman, University of Illinois at Chicago, USA
George Karypis, University of Minnesota, Twin Cities, USA
Eamonn Keogh, University of California - Riverside, USA
Masaru Kitsuregawa, the University of Tokyo, Japan
Yehuda Koren, Yahoo! Research, Israel
Ee-Peng Lim, Singapore Management University, Singapore
Olfa Nasraoui, University of Louisville, USA
Zaiqing Nie, Microsoft Research Asia, China
Naren Ramakrishnan, Virginia Tech, USA
Shashi Shekhar, University of Minnesota, USA
Kyuseok Shim, Seoul National University, Korea
buj Singh, University of California, Santa Barbara, USA
Ashok Srivastava, NASA, USA
Hannu Toivonen, University of Helsinki, Finland
Jianyong Wang, Department of Computer Science and Technology, Tsinghua University, Beijing, China
Ke Wang, Simon Fraser University, Canada
Wei Wang, University of North Carolina at Chapel Hill, USA
Stefan Wrobel, Fraunhofer IAIS and Univ. of Bonn, Germany
Hui Xiong, Rutgers University, USA
Qiang Yang, Hong Kong University of Science and Technology, Hong Kong, China
Jeffrey Xu Yu, the Chinese University of Hong Kong, Hong Kong
Osmar Zaiane, University of Alberta, Canada
Yanchun Zhang, Victoria University, Australia
Web Master:
Xinhua Zhu (University of Technology, Sydney, Australia)
Secretary:
Li Liu (University of Technology, Sydney, Australia)

ICDM'10 Program Committee Members

Aijun An, York University, Canada

Aris Anagnostopoulos, Sapienza University of Rome, Italy

Peter Andreae, Victoria University of Wellington, New Zealand

Periklis Andritsos, Thoora.com, Canada

Ioannis Androulakis, Rutgers University, USA

Luiza Antonie, University of Guelph, Canada

Annalisa Appice, University of Bari, Italy

Ira Assent, Aalborg University, Denmark

Vassilis Athitsos, University of Texas at Arlington, USA

Ricardo Baeza-Yates, Yahoo! Research, Spain

Tony Bagnall, University of East Anglia, UK

Jing Bai, Yahoo, USA

James Bailey, The University of Melbourne, Australia

Chris Bailey-Kellogg, Dartmouth College, USA

Arindam Banerjee, University of Minnesota, Twin Cities, USA

Elena Baralis, Politecnico di Torino, Italy

Rohan Baxter, ATO, Australia

Roberto Bayardo, Google, USA

Bettina Berendt, K.U. Leuven, Belgium

Tanya Berger-Wolf, University of Illinois at Chicago, USA

Michael R. Berthold, University of Konstanz, Germany

Elisa Bertino, Purdue University, USA

Kanishka Bhaduri, NASA Ames Research Center & MCT, USA

Indrajit Bhattacharya, IBM Research, Delhi, India

Chiranjib Bhattacharyya, Indian Institute of Science, India

Francesco Bonchi, Yahoo! Research, Spain

Jean-Francois Boulicaut, INSA-Lyon, France

Paul Bradley, Apollo Data Technologies, USA

Michael Burl, JPL, USA

Toon Calders, Eindhoven University of Technology, The Netherlands

Mary Elaine Califf, Illinois State University, USA

Juan P. Valente, Univ. Politecnica de Madrid, Spain

James Caverlee, Texas A&M University, USA

Nick Cercone, York University, Canada

Keith Chan, The Hong Kong Polytechnic University, China

Philip Chan, Florida Institute of Technology, USA

Kevin C. Chang, University of Illinois at Urbana Champaign, USA

Wanpracha Chaovalitwongse, Rutgers University, USA

Sanjay Chawla, The University of Sydney, Australia

Arbee L.P. Chen, National Chengchi University, Taiwan,

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

Ming-Syan Chen, National Taiwan University, Taiwan

Xue-Wen Chen, Univ of Kansas, USA

Yixin Chen, Washington University in St. Louis, USA

Zheng Chen, Microsoft Research Asia, China

Zhiyuan Chen, University of Maryland, Baltimore County, USA

Hong Cheng, The Chinese University of Hong Kong, Hong Kong

David Cheung, The University of HongKong, Hong Kong

Yun Chi, NEC Laboratories America, Inc., USA

Silvia Chiusano, Politecnico di Torino, Italy

Amanda Clare, University of Wales, Aberystwyth, UK

Chris Clifton, Purdue University, USA

frans coenen, University of Liverpool, UK

Gao Cong, Aalborg University, Denmark

Tom Croonenborghs, KHKempen University College, Belgium

Alfredo Cuzzocrea, ICAR Institute and University of Calabria, Italy

Nilesh Dalvi, Yahoo! Research, San Francisco, USA

Gautam Das, University of Texas at Arlington, USA

Sanmay Das, Rensselaer Polytechnic Institute, USA

Anirban Dasgupta, Yahoo! Research, USA

Tamraparni Dasu, AT&T Labs - Research, USA

Jesse Davis, University of Washington, USA

Anne Denton, North Dakota State University, U.S.A.

Alin Dobra, University of Florida, USA

Carlotta Domeniconi, George Mason University, USA

Josep Domingo-Ferrer, Univ. Rovira i Virgili, Catalonia

Guozhu Dong, Wright State University, USA

Dejing Dou, University of Oregon, USA

Margaret Dunham, Southern Methodist University, USA

Haimonti Dutta, Columbia University, New York, USA

Kemal Efe, Louisiana University, USA

Christoph F. Eick, University of Houston, USA

Mohammad El-Hajj, University of Alberta, Canada

Charles Elkan, University of California, San Diego, USA

Tapio Elomaa, Tampere University of Technology, Finland

Floriana Esposito, University of Bari, Italy

Vladimir Estivill-Castro, Griffith University, Australia

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

ICDM'10 Program Committee Members (cont.)

Ad Feelders, Universiteit Utrecht, The Netherlands

Ling Feng, Tsinghua University, China

Hakan Ferhatosmanoglu, The Ohio State University, USA

Andrew Foss, University of Alberta, Canada

Eibe Frank, University of Waikato, New Zealand

Takeshi Fukuda, IBM Yamato Software Laboratory, Japan

Benjamin C. M. Fung, Concordia University, Canada

Efstratios Gallopoulos, University of Patras, Greece

Matjaz Gams, Jozef Stefan Institute, Slovenia

Aryya Gangopadhyay, UMBC, USA

Gemma Garriga, University of Paris VI, France

Johannes Gehrke, Cornell University, USA

Amol Ghoting, IBM TJ Watson, USA

Chris Giannella, The MITRE Corporation, USA

Fosca Giannotti, ISTI-CNR, Italy

Aristides Gionis, Yahoo! Research, Spain

Attilio Giordana, Universita' del Piemonte Orientale, Italy

Shantanu Godbole, IBM Research, India

Ananth Grama, Purdue University, USA

Dennis Groth, Indiana University, USA

Fabrice Guillet, LINA, CNRS 6241, Polytech'Nantes, Nantes university, France

Diansheng Guo, University of South Carolina, USA

Yike Guo, Imperial College, UK

Marios Hadjieleftheriou, AT&T Labs - Research, USA

Maria Halkidi, Univ. of Pireus, Greece

Larry Hall, Univ. of South Florida, USA

Jiawei Han, University of Illinois at Urbana Champaign, USA

Robert Hilderman, University of Regina, Canada

Alexander Hinneburg, Martin-Luther-University Halle-Wittenberg, Germany

Shen-Shyang Ho, California Institute of Technology/Jet Propulsion Laboratory, USA

Tu-Bao Ho, Japan Advanced Institute of Science and Technology, Japan

Larry Holder, Washington State University, USA

Jaakko Hollmén, Aalto University School of Science and Technology, Finland

Se June Hong, IBM Research, emeritus, USA

Andreas Hotho, University of Kassel, Germany

Vagelis Hristidis, Florida International University, USA

Minqing Hu, MySpace, USA

Xiaohua Hu, Drexel University, USA

Ming Hua, Facebook Inc., USA

Luke Huan, University of Kansas, USA

Eyke Huellermeier, Department of Mathematics and Computer Science, Marburg University, Germany

Szymon Jaroszewicz, National Institute of Telecommunications, Poland

David Jensen, University of Massachusetts Amherst, USA

Christopher Jermaine, Rice University, USA

Yan Jia, National University of Defense Technology, China

Ruoming Jin, Kent State University, USA

Vanja Josifovski, Yahoo! Research, USA

Konstantinos Kalpakis, UMBC, USA

Toshihiro Kamishima, National Institute of Advanced Industrial Science and Technology (AIST), Japan

Murat Kantarcioglu, The University of Texas at Dallas, USA

Hillol Kargupta, UMBC, USA

Kamalakar Karlapalem, IIIT, Hyderabad, India

Hiroyuki Kawano, Nanzan University, Japan

John Keane, University of Manchester, United Kingdom

Latifur Khan, The University of Texas at Dallas, USA

Walter Kosters, LIACS, Leiden University, The Netherlands

Petra Kralj Novak, Jozef Stefan Institute, Slovenia

Stefan Kramer, Technische Universitaet Muenchen, Germany

Raghu Krishnapuram, IBM India Research Lab, India

Deept Kumar, Virginia Tech, USA

Ravi Kumar, Yahoo! Research, USA

Vipin Kumar, University of Minnesota, USA

Lotfi Lakhal, Aix-Marseille Université, France

Wai Lam, The Chinese University of Hong Kong, Hong Kong

Longin Jan Latecki, Temple University, USA

Anne Laurent, Polytech'Montpellier - Univ. Montpellier 2, France

Hady Lauw, Institute for Infocomm Research, Singapore

Srivatsan Laxman, Microsoft Research, India

Aleksandar Lazarevic, United Technology Research Center, USA

Chengkai Li, University of Texas at Arlington, USA

Feifei Li, Florida State University, USA

Hang Li, Microsoft Research Asia, China

Ming Li, University of Waterloo, Canada

Ping Li, Cornell University, USA

Qi Li, Western Kentucky University, USA

Tao Li, Florida International University, USA

Xiaoli Li, Institute for Infocomm Research, Singapore

Xue Li, The University of Queensland, Australia

Jessica Lin, George Mason University, U.S.

Song Lin, Google, USA

Charles Ling, the University of Western Ontario, Canada

Huan Liu, Arizona State University, USA

Jinze Liu, University of Kentucky, USA

ICDM'10 Program Committee Members (cont.)

Kun Liu, Yahoo! Labs, USA

Ling Liu, Georgia Institute of Technology, USA

Tie-Yan Liu, Microsoft Research Asia, China

Xiaohui Liu, Brunel University, UK

Yan Liu, IBM Research, USA

Chang-Tien Lu, Virginia Tech, USA

Claudio Lucchese, H.P.C. Lab. I.S.T.I.-C.N.R., Italy

Ashwin Machanavajjhala, Cornell University, USA

Richard Maclin, University of Minnesota-Duluth, USA

Sofus Macskassy, Fetch Technologies, USA

Masoud Makrehchi, Thomson Reuters, USA

Bradley Malin, Vanderbilt University, USA

Mark Maloof, Georgetown University, USA

Nikos Mamoulis, Hong Kong University, Hong Kong, China

Giuseppe Manco, ICAR-CNR, Italy

Murali Mani, WPI, USA

Pedro Jose Marron, University of Duisburg-Essen, Germany

Keith Marsolo, CCHMC, USA

Florent Masseglia, INRIA, France

Francesco Masulli, University of Genova, Italy

Stan Matwin, University of Ottawa, Canada

Sameep Mehta, IBM India Research Lab, India

Prem Melville, IBM Research, USA

Ernestina Menasalvas, Facultad de Informatica. Universidad Politecnica de Madrid, Spain

Rosa Meo, University of Torino, Italy

Taneli Mielikainen, Nokia, USA

Silvia Miksch, Vienna University of Technology, Austria

Prasenjit Mitra, The Pennsylvania State University, USA

Fabian Moerchen, Siemens Corporate Research, USA

Shinichi Morishita, University of Tokyo, Japan

Tsuyoshi Murata, Tokyo Institute of Technology, Japan

Mirco Nanni, ISTI-CNR - Pisa, Italy

Svetlozar Nestorov, University of Chicago, USA

Jennifer Neville, Purdue University, USA

See-Kiong Ng, Institute for Infocomm Research, Singapore

Siegfried Nijssen, KU Leuven, Belgium

Tim Oates, University of Maryland Baltimore County, USA

Salvatore Orlando, Ca' Foscari University of Venice, Italy

Gerhard Paass, Fraunhofer IAIS, Germany

Balaji Padmanabhan, The Wharton School, University of Pennsylvania, USA

Rina Panigrahy, Microsoft Research, USA

Spiros Papadimitriou, IBM Research, USA

Rajesh Parekh, Yahoo, USA

Douglass S. Parker, University of California, Los Angeles, USA

Srinivasan Parthasarathy, Ohio State University, USA

Jian Pei, Simon Fraser University, Canada

Chang-shing Perng, IBM T.J. Watson Research Center, USA

Pascal Poncelet, LIRMM lab, France

William Pottenger, Rutgers, State University of New Jersey, USA

Kunal Punera, Yahoo! Labs, USA

Kai Puolamaki, Aalto University School of Science and Technology, Finland

Davood Rafiei, University of Alberta, Canada

Jan Ramon, Katholieke Universiteit Leuven, Belgium

Zbigniew Ras, University of North Carolina at Charlotte, USA

Rajeev Rastogi, Yahoo! Labs Bangalore, India

Chotirat "Ann" Ratanamahatana, Chulalongkorn University, Thailand

Jan Rauch, University of Economics, Prague, Czech Republic

S.S. Ravi, University at Albany - State University of New York, USA

Chandan Reddy, Wayne State University, USA

Christophe Rigotti, INSA Lyon / LIRIS, France

Celine Rouveirol, LIPN, France

Stefan Rueping, Fraunhofer IAIS, Germany

Sourav S Bhowmick, Nanyang Technological University, Singapore

Lorenza Saitta, Università del Piemonte Orientale, Italy

Joerg Sander, University of Alberta, Canada

Tamas Sarlos, Yahoo! Research, USA

Yucel Saygin, Sabanci University, Italy

Martin Scholz, HP Labs, USA

Rudy Setiono, National University of Singapore, Singapore

Dou Shen, Microsoft, China

Jialie Shen, Singapore Management University, Singapore

Qian Shen, Aberystwyth University, UK

Olivia Sheng, The University of Utah, USA

Shirish Shevade, Indian Institute of Science, INDIA

Luo Si, Purdue University, US

Arno Siebes, Utrecht University, Netherlands

Fabrizio Silvestri, ISTI-CNR, Italy

Dan Simovici, University of Massachusetts at Boston, USA

Lisa Singh, Georgetown University, Washington DC, USA

Mike Sips, Stanford University, USA

Krishnamoorthy Sivakumar, Washington State University, USA

ICDM'10 Program Committee Members (cont.)

Andrzej Skowron, University of Warsaw, Poland

Kate Smith-Miles, Monash University, Australia

Min Song, New Jersey Institute of Technology, USA

Nandit Soparkar, Ubiquiti Inc., USA

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

Michael Steinbach, University of Minnesota, USA

W. Nick Street, University of Iowa, USA

Jimeng Sun, IBM, USA

Einoshin Suzuki, Kyushu University, Japan

Wojciech Szpankowski, Purdue University, USA

Atsuhiro Takasu, National Institute of Informatics, Japan

Ah-Hwee Tan, Nanyang Technological University, Singapore

Pang-Ning Tan, Michigan State University, USA

Jie Tang, Tsinghua University, China

David Taniar, Monash University, Australia

Dacheng Tao, Nanyang Technological University, Singapore

Maguelonne Teisseire, TETIS - Cemagref, France

Masahiro Terabe, Mitsubishi Research Institute, Inc., Japan

Takao Terano, Tokyo Institute of Technology, Japan

Alexandre Termier, LIG, University of Grenoble, France

Evimaria Terzi, Boston University, USA

Yannis Theodoridis, University of Piraeus, Greece

Kai Ming Ting, Monash University, Australia

Masashi Toyoda, University of Tokyo, Japan

Panayiotis Tsaparas, Microsoft Research, USA

Vincent S. Tseng, National Cheng Kung University, Taiwan

Shusaku Tsumoto, Shimane University, Japan

Alexander Tuzhilin, New York University, USA

Jeffrey D. Ullman, Gradiance Corp., USA

Lyle Ungar, University of Pennsylvania, USA

K.P. Unnikrishnan, General Motors Research, USA

Jaideep Vaidya, Rutgers University, USA

Koen Vanhoof, Hasselt University, Belgium

Michalis Vazirgiannis, Athens University of Economics and Business, Greece

Vassilios Verykios, University of Thessaly, Greece

Ricardo Vilalta, University of Houston, USA

Michail Vlachos, IBM Research,

Kiri Wagstaff, Jet Propulsion Laboratory, USA

Haixun Wang, Microsoft, China

Jason T. L. Wang, New Jersey Institute of Technology, USA

Lipo Wang, Nanyang Technological University, Singapore

Wei Wang, Fudan University, China

Takashi Washio, Osaka University, Japan

Li Wei, Google Inc., USA

Gary Weiss, Fordham University, USA

Graham Williams, Australian Taxation Office, Australia

Ran Wolff, University of Haifa, Israel

Wensheng Wu, UNC Charlotte, USA

Xindong Wu, University of Vermont, USA

Xintao Wu, University of North Carolina at Charlotte, USA

Dong Xin, Microsoft, USA

Zhengzheng Xing, Simon Fraser University, Canada

Xiaowei Xu, University of Arkansas at Little Rock, USA

Katsutoshi Yada, Kansai University, Japan

Xifeng Yan, University of California, Santa Barbara, USA

Hui Yang, The Ohio State University, USA

Jiong Yang, CWRU, USA

Ying Yang, Australian Taxation Office, Australia

Jieping Ye, Arizona State University, USA

Xiaoxin Yin, Microsoft, USA

Hwanjo Yu, Pohang University of Science and Technology, Korea

Lei Yu, Binghamton University, USA

Philip Yu, University of Illinois at Chicago, USA

Bianca Zadrozny, Universidade Federal Fluminense, Brazil

Mohammed Zaki, Rensselaer Polytechnic Institute, USA

Carlo Zaniolo, UCLA, USA

Gerson Zaverucha, COPPE/PESC-Federal University of Rio de Janeiro(UFRJ), Brazil

Demetris Zeinalipour, University of Cyprus, Cyprus

Aidong Zhang, State University of New York at Buffalo, USA

Donghui Zhang, Northeastern University, USA

Harry Zhang, University of New Brunswick, Canada

Kun Zhang, Xavier University of Louisiana, USA

Shichao Zhang, University of Technology, Sydney, Australia

Zhongfei Zhang, State University of New York at Binghamton, USA

Kaidi Zhao, Amazon.com Inc, USA

Ying Zhao, Tsinghua University, China

Zijian Zheng, Microsoft, USA

Ning Zhong, Maebashi Institute of Technology, Japan

Zhi-Hua Zhou, Nanjing University, China

Shenghuo Zhu, NEC Laboratories America, Inc., USA

Xingquan Zhu, University of Technology, Sydney, Australia

ICDM'10 Non-Program Committee Reviewers

Amrudin Agovic	Christian Desrosiers	Chris Kauffman
Mohammad Salim Ahmed	Sandipan Dey	Jaya Kawale
Tahseen Al-khateeb	Nicola Di Mauro	Arijit Khan
Samir Al-stouhi	Wei Ding	Mohammad Khoshneshin
Noor Alaydie	Zhongli Ding	Jinha Kim
Bilal Alsallakh	Jun Du	Sungchul Kim
Xiangdong An	Lian Duan	Taehoon Kim
Rajul Anand	Ana Luisa Duboc	Taehyong Kim
Sitaram Asur	Timo Duchrow	Youngdae Kim
Gowtham Atluri	James Faghmous	Suzan Koknar-Tezel
Gowtham Atluri	Gang Fang	Kashyap Kolipaka
Xiang Bai	Yi Fang	Xiangnan Kong
Marenglen Biba	Nicola Fanizzi	John Korecki
Albert Bifet	Fabio Fassetti	Yehuda Koren
Julien Blanchard	Qinyuan Feng	Da Kuang
Arnold Boedihardjo	Stefano Ferilli	Onur Kucuktunc
Ben Bornstein	Natalja Friesen	Avinava Kumar Dubey
Marco Botta	Lawrence Fu	Gautam Kunapuli
Karl Branting	Qiang Fu	Mayank Lahiri
Ivan Brugere	Zhouyu Fu	Shibamouli Lahiri
Fabian Buchwald	Pui Cheong Fung	Tim Lammarsch
Luca Cagliero	Liang Gan	Nikolay Laptev
Toon Calders	Bin Gao	Juan Alfonso Lara Torralbo
Juana Canul-Reich	Paolo Garza	Ngoc Tu Le
Cornelia Caragea	Liang Ge	Wangchao Le
Michelangelo Ceci	Cibin George	Haden Hooyeon Lee
Tania Cerquitelli	Nikos Giatrakos	Jaeyong Lee
Kaushik Chakrabarti	Aris Gkoulalas-Divanis	Rémi Lehn
Mithun Chakraborty	Sujatha Gollapalli	Hui Li
Chun-Sheng Chen	Zhenmei Gu	Jing Li
Feng Chen	Ziyu Guan	Jun Li
Jianhui Chen	Naiyang Guang	Ming Li
Rui Chen	Tias Guns	Nan Li
You Chen	Rohit Gupta	Runheng Li
Tao Cheng	Sara Hajian	Shenzhi Li
Weiwei Cheng	Yi Han	Tao Li
Meenal Chhabra	Andreas Hapfelmeier	Yu-Feng Li
Bin-Hui Chou	Qi He	Dan Liu
Kit Chui Chun	zhenying He	Fei Tony Liu
Gianni Costa	Teemu Heinimäki	Haishan Liu
Amanda Crick	Thanh Lam Hoang	Ji Liu
Claudia d'Amato	Dino Ienco	Jun Liu
Kamalika Das	Jeffrey Jestes	Wei Liu
Kaustav Das	Rachsuda Jiamthapthaksin	Xin Liu
Puja Das	Xing Jiang	Rohit Lotlikar
Kurt De Grave	Li Jiye	
Krzysztof Dembczynski	Nobuhiro Kaji	
Engin Demir	Peter Karsmakers	

ICDM'10 Non-Program Committee Reviewers (cont.)

Grigorios Loukides
Zhenyu Lu
Kathy Macropol
Arun Maiya
Rajarshi Mallik
Mark Manasse
Claudia Marinica
Gerasimos Marketos
Ana Martinez
Elio Masciari
Elio Masciari
Mohammad Masud
Andreas Maunz
Kyle McQuisten
Glauber Menezes
Shicong Meng
Aditya Menon
Srujana Merugu
Gengxin Miao
Matthew Michelson
Ruth Miller
Noman Mohammed
Hamid Mousavi
Barzan Mozafari
Mummoorthy Murugesan
Ramanathan Narayanan
Sumit Negi
Ercan Nergiz
Wee Siong Ng
Xuan Bach Ngo
Khoa Nguyen
Luong Dong Nguyen
Eileen A. Ni
Alex Nikolov
Benjamin Négrevergne
Omar Odibat
Jinoh Oh
Riccardo Ortale
Michelle Osmond
Aline Paes
Balaji Palanisamy
Sun Park
Jonathon Parker
Ruggero G. Pensa
Anja Pilz
Fabio Pinelli
Cristiano Pitangui
Luigi Pontieri
Miao Qiao

Bo Qin
Han Qin
Alessandra Raffaetà
Shyamsundar Rajaram
Ganesh Ramesh
Ganesh Ramesh
Mike Rechenthin
Frank Reichartz
Lothar Richter
Alexander Rind
Mathieu Roche
Sudip Roy
Eduardo Ruiz
Reza Sadoddin
Kameshwaran Sampath
Boon-Siew Seah
Madeleine Seeland
Hanhuai Shan
Hao Shao
Madhu Shashanka
bin shen
Lei Shi
Larry Shoemaker
Manu Shukla
Claudio Silvestri
Nickolas Skapura
Manas Somaiya
Dongjin Song
Damon Sotoudeh
Michal Sramka
Keith Sullivan
Liang Sun
Hana Tai
Frank Takes
James Tan
Mingwang Tang
Yuzhe Tang
Chayant Tantipathananandh
S. Tatikonda
Bin Tong
Kong-Wah WAN
Hao Wang
Huahua Wang
Lei Wang
Pu Wang
Steven Wang
Tao Wang
Ting Wang
Xiaowei Wang
Xiaoyue Wang

Xufei Wang
Zhuang Wang
Zhuang Wang
Dennis Wegener
Ben Wellner
Tim Weninger
Jörg Wicker
Daya Wimalasuriya
Tianyi Wu
Wei Wu
Yanghua Xiao
Bo Xie
Hua Xu
Jingfang Xu
Jun Xu
Lisa Jean Yan
Chen Yang
Jianmei Yang
Xingwei Yang
Zhenglu Yang
Bin Yao
Jieping Ye
Jun Yin
Xiaowei Ying
Naoki Yoshinaga
Xiao Yu
Yang Yu
Reza Zafarani
Dimitris Zeimpekis
Kai Zeng
De-Chuan Zhan
Chao Zhang
Chi Zhang
Chongsheng Zhang
Dan Zhang
Gong Zhang
Nan Zhang
Shiming Zhang
Wei Zhang
Zhang Zhang
Peixiang Zhao
Yuchen Zhao
Jiayu Zhou
Tianyi Zhou
haofeng zhou
Huisheng Zhu
Lei Zhu
Xianshu Zhu

ICDM'10 Volunteers

Xinhua Zhu

Can Wang

Chayapol (Paul) Moemeng

Chao Luo

Lin Zhu

Xuhui Fan

Lei Zheng

Yiling Zeng

Yin Song

Yong Yang

Xia Cui

Bo Liu

Yanshan Xiao

Guohua Liang

Barbara Munday

Ling Chen

Mehrnaz Natoori

Yifan Fu

Wei Wang

Tara Babaie

Georgina Wilcox

Khoa Nguyen

Diego Silva

Tianyi Zhou

Ava Bargi

Junfu Yin

Conference Room Layout

CITY MARKETS
HAYMARKET, BUILDING 5
LEVEL 01

University of Technology, Sydney

Facilities Management Unit

AS BUILT   JUNE 1999

CM0501AP-04584_00

CITY MARKETS
HAYMARKET, BUILDING 5
LEVEL 02

University of Technology, Sydney
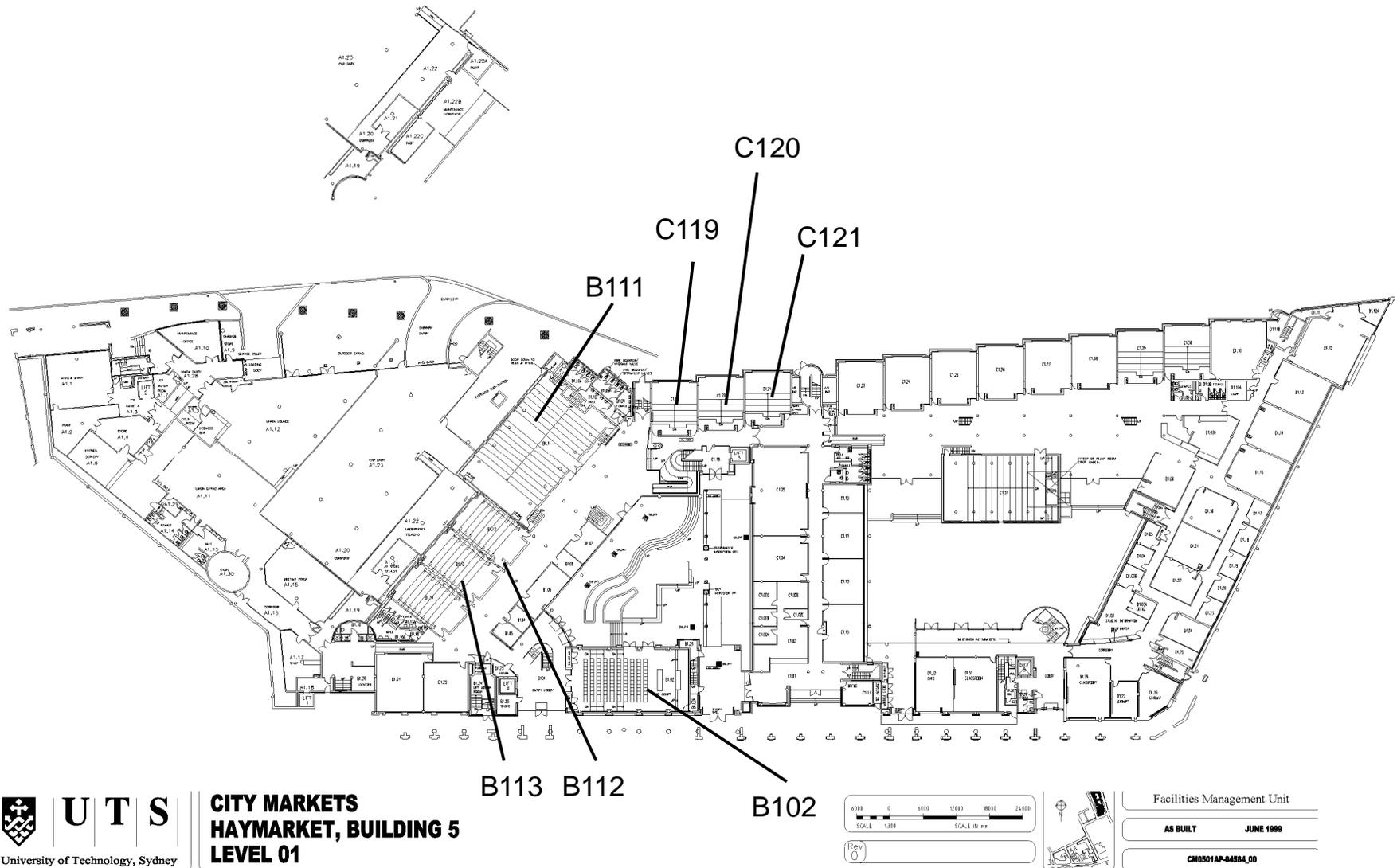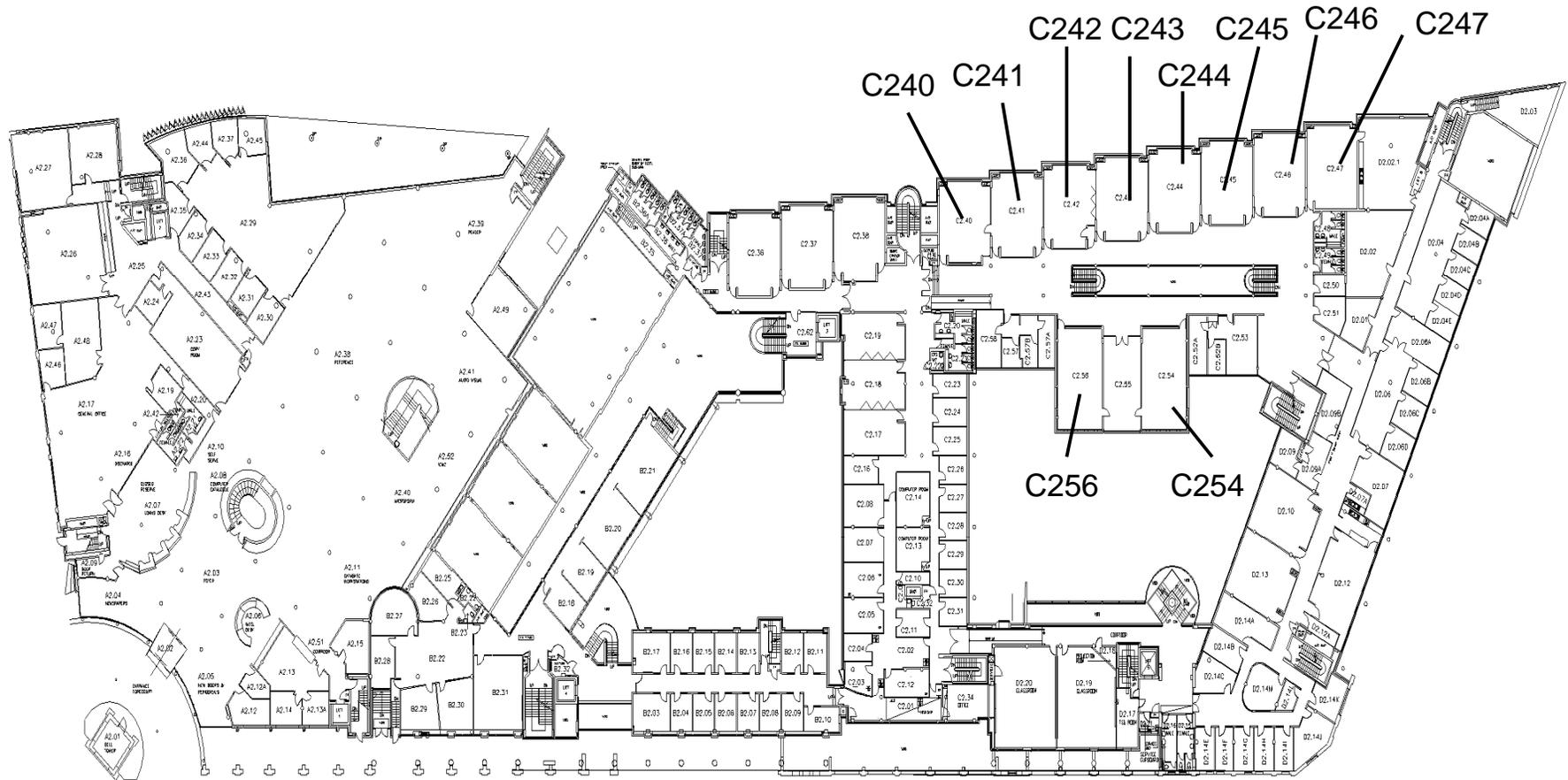
Facilities Management Unit

AS BUILT   JUNE 1999

CM0502AP-04585_00

Useful Links

About Sydney:

- City of Sydney: http://www.cityofsydney.nsw.gov.au/
- About Sydney: http://www.sydney.com.au/info.htm
- NSW government: http://www.nsw.gov.au/
- Current events: http://www.nsw.gov.au/events.asp
- Sydney CityRail: http://www.cityrail.com.au/
- Sydney map: http://citymap.cityofsydney.nsw.gov.au/Website/basemap/default.htm
- Visitor guide and information:
  http://www.cityofsydney.nsw.gov.au/AboutSydney/VisitorGuidesInformation/Default.asp
- Public transport:
  http://www.cityofsydney.nsw.gov.au/AboutSydney/ParkingAndTransport/PublicTransport.asp
- Sydney bus: http://www.sydneybuses.info/
- Tansport information 131500: http://www.131500.info/realtime/default.asp
- Tourist services: http://www.sydneybuses.info/tourist-services.htm
- Sydney ferries: http://www.sydneyferries.info/
- T-Way system: http://www.t-way.nsw.gov.au/

About the University of Technology, Sydney:

(The following units are the host of this conference)

- Data Sciences & Knowledge Discovery Lab, QCIS: http://datamining.it.uts.edu.au/
- Centre for Quantum Computation & Intelligent Systems: www.qcis.uts.edu.au
- Faculty of Engineering and Information Technology: http://www.feit.uts.edu.au/
- Advanced Analytics Institute: http://www.analytics.uts.edu.au
- University of Technology Sydney: http://www.uts.edu.au

About the ICDM'10:

- ICDM'10 Website: http://datamining.it.uts.edu.au/icdm10/

Sydney: Points of Interest
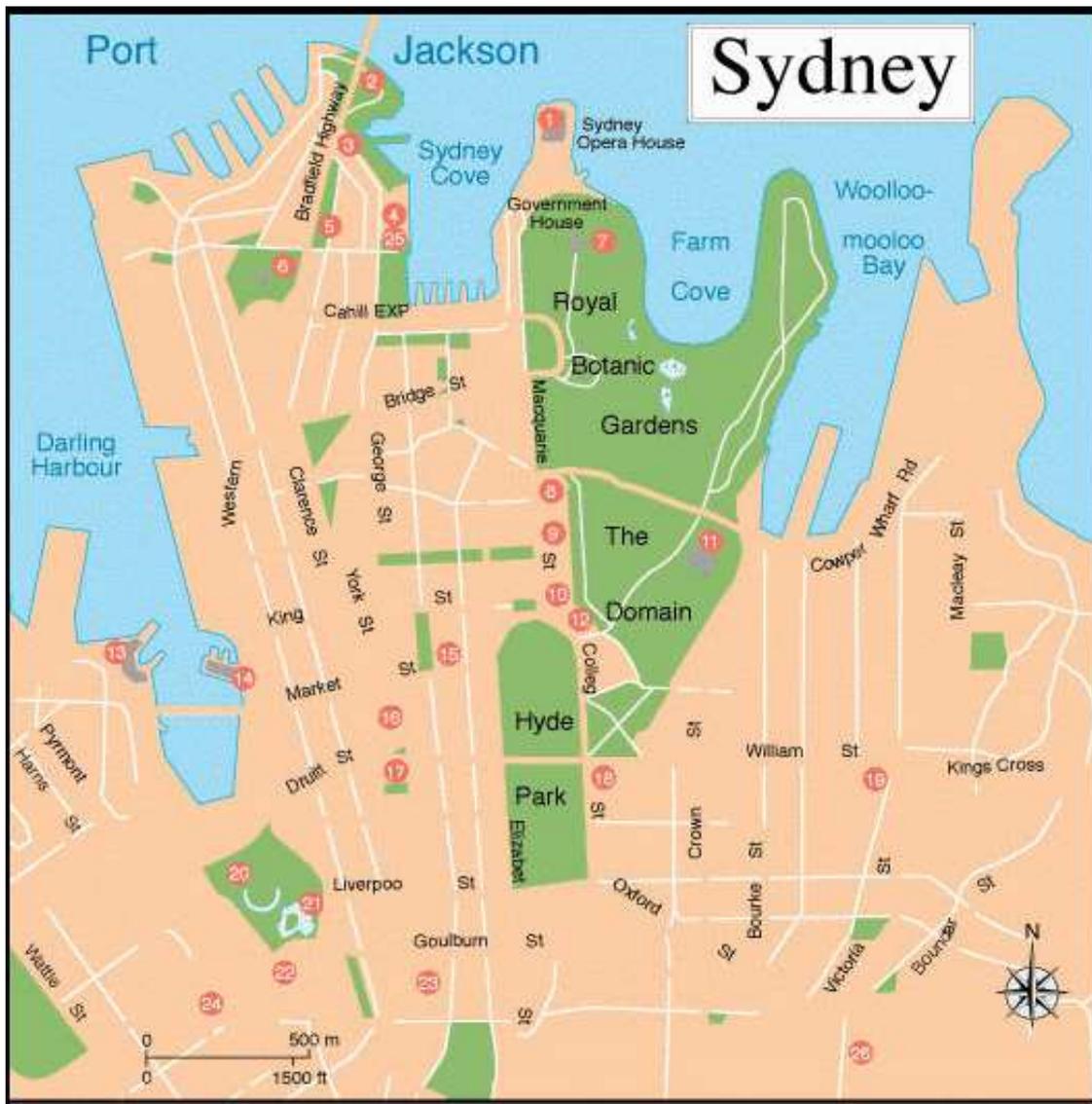
Points of Interest:

1. Sydney Opera House	2. Sydney Harbour Bridge	3. Earth Exchange	4. The Story of Sydney
5. The Rocks	6. Sydney Observatory	7. Government House	8. NSW State Library
9. NSW Parlient House	10. Mint Museum	11. NSW Art Gallery	12. Hyde Park Barracks
13. National Maritime Museum	14. Sydney Aquarium	15. Sydney Tower	16.Queen Victoria Building
17. Sydney Town Hall	18. Australian Museum	19. Kings Cross	20. Darling Harbour
21. Chinese Garden	22. Pumphouse Brewery Tavern	23. Chinatown	24 Powerhouse Museum
25. Cadman's Cottage	26. Victoria Barracks		

Map of Sydney Harbour Waterfront

1   Historical Sydney – The Rocks
2   Sydney Harbour Bridge
3   **Campbells Cove** – meet here to board **Showboard Two** for your **banquet dinner**, show and cruise.
4   Circular Quay
5   Sydney Opera House

The Rocks   Sydney Harbour Bridge

Showboat Banquet Cruise   Circular Quay   Sydney Opera House
Board at Campbell's Cove

Showboat Banquet and Harbour Cruise

Please meet at 6.30pm at Campbell Cove Wharf, Sydney Harbour foreshore. Campbell Cove is opposite the Opera House – walk past Circular Quay and follow the foreshore (waterfront) around past the historic Rocks to Campbell Cove, as per above map. Your Showboat Two banquet dinner (look for the red and white paddle-wheeler) will board at 7.00pm for a 7.15pm departure (please bring your banquet ticket).

* **Note**:   'Showboat Two' is the official name of the company who runs the boat and show. There is only one boat – a red and white paddle-wheeler.

Conference Venue, Lunch Venue and UTS Campuses
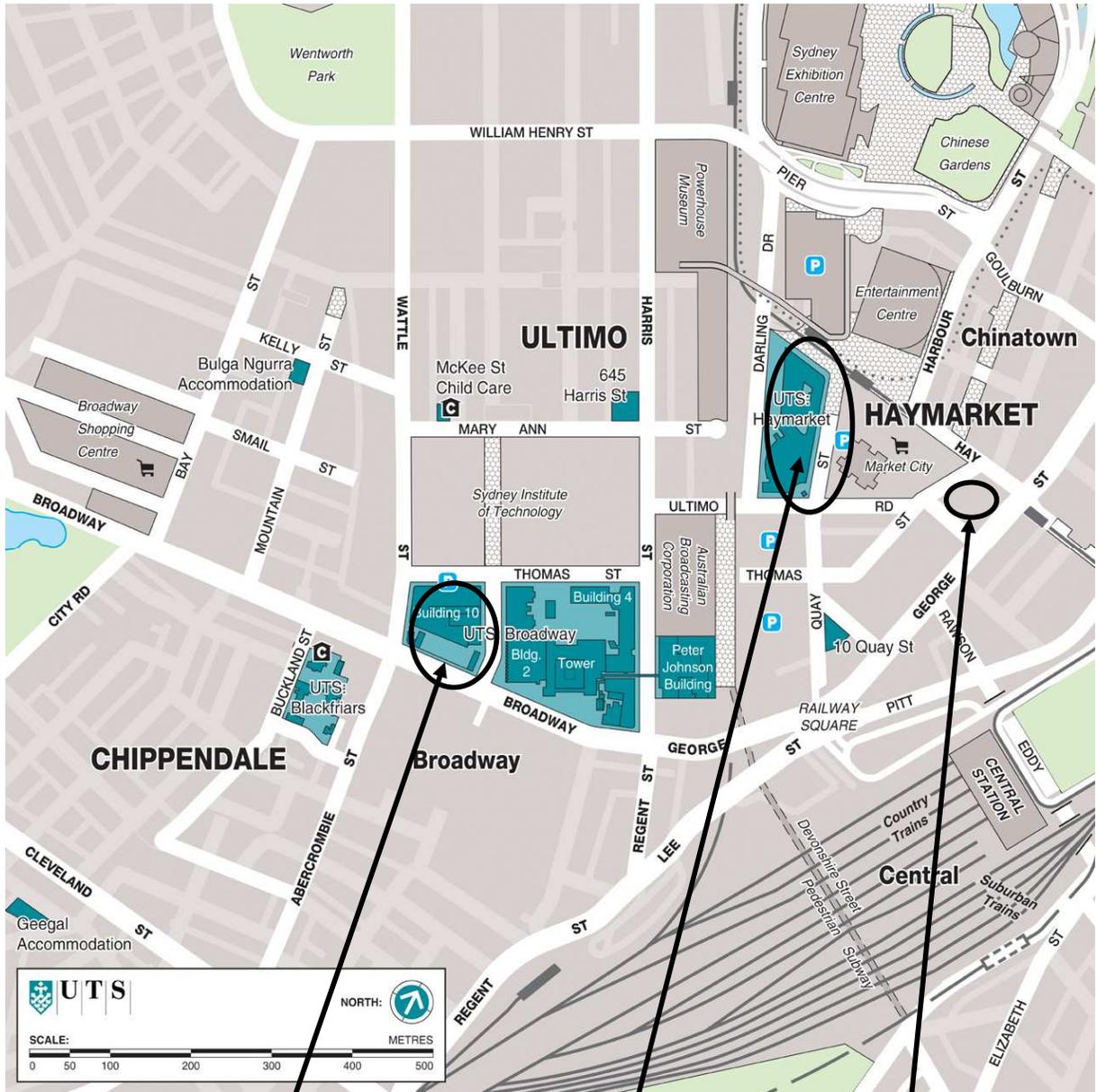
UTS Faculty of Engineering & IT
QCIS

Conference Venue
UTS B5

Lunch Venue
Marigold Restaurant