Fourth ACM International Conference on Web Search and Data Mining

wsdm

2011 Hong Kong

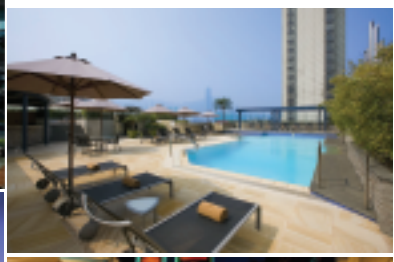


Table of Contents

Welcome Message

On behalf of the Organizing Committee, I would like to extend a warm welcome to the Fourth ACM International Conference on Web Search and Data Mining (WSDM2011), Hong Kong.

Following the great successes of WSDM2008 in Stanford University, USA, WSDM2009 in Barcelona, Spain, and WSDM2010 in New York, USA, WSDM2011, Hong Kong has been chosen as the premier annual international ACM conference for presenting the very best research innovations and other significant work covering areas in search and data mining on the Web and the Social Web, with an emphasis on practical but principled novel models of search, retrieval and data mining, algorithm design and analysis, economics implications, and in-depth experimental analysis of accuracy and performance. The purpose and goal of WSDM2011 is to provide a forum for researchers, practitioners, and professionals alike to identify challenging issues in the areas, to share the latest ideas and results, to shape research directions, and to inspire future breakthroughs.

As in the past WSDM conferences, WSDM2011 features top-quality and excellent technical program from a record number of submissions. In particular, WSDM2011's emphasis this year is on Social Computing and related topics. Moreover, here are some of the highlights of the program we have planned at WSDM2011:

- Focused single-track session
- All accepted papers will be full-length with poster-presentations to increase interaction
- A small set of selected oral presentations in concentrated areas
- Two excellent keynote speakers
- Outstanding pre-conference workshops and tutorials

To put together a conference is not an easy task. I would like to thank all the members of the Conference Committee for their dedication and hard work in ensuring an excellent technical program with social activities. In particular, I would like to thank our Program Co-Chairs, Wolfgang Nejdl and Hang Li, the 19 Senior PC members, the 134 PC members, the Best Paper Award Committee, and other reviewers for their great efforts in providing WSDM2011 with an excellent technical program. Last, but not least, I thank all the authors of submitted papers for their contributions and all the delegates for their participation.

My most sincere thanks go to the following Organizing Committee members for their dedication and support in making WSDM2011 a great success: Workshop Co-Chairs, Evgeniy Gabrilovich and Wai Lam; Tutorial Co-Chairs, Marc Najork and Qing Li; Finance Chair, Michael R. Lyu; Publicity Chair, Qiang Yang; Registration Chair, Michael Chau; Publication Chair, Hong Cheng; Local Arrangement Co-Chairs, Dan Hong and Tak-Lam Wong; Poster Session Chair, Jie Tang; Event Coordinator, Samuel Tam; Secretariat, Baichuan Li and Yi Zhu. Without their selfless support, this conference could not have been a success!

I am deeply grateful to our two distinguished speakers Christos Falousos and Harry Shum for their insightful keynote talks that highlight the current research work in web search and mining.

I would also like to thank the sponsors of WSDM2011: ACM, SIGIR, SIGKDD, SIGMOD, SIGWEB, the K.C. Wong Education Foundation, Microsoft, Google, Yahoo, and Yandex. Moreover, we are also grateful to Arnetminer and VeriGuide for their technical assistance. Our special thanks go to Microsoft and Google for sponsoring the Best Paper Award and the Best Student Paper Award, respectively. In addition, Google once again sponsored the Conference Travel Award for Women this year. Their generous sponsorship and support contributed immensely to the success of the conference.

Finally, I thank the WSDM Steering Committee for their vote of confidence to hold WSDM2011 in Hong Kong, and for their guidance throughout the planning and organizing of the conference.

Hong Kong is a world-class metropolis full of dynamism and character. Aside from the technical and social program at the conference, I invite you to take some time to explore the Pearl of Orient and its nearby regions. With the celebration of the Chinese Lunar New Year just before the conference, there are a number of festive events specially prepared to show off the ancient tradition in this modern city. I hope you will have a chance to enjoy it as much as I do.

Once again, I welcome you to WSDM2011, and wish you a very stimulating, productive, and fruitful experience in Hong Kong!

Irwin King
General Conference Chair, WSDM2011
The Chinese University of Hong Kong
Hong Kong

Foreword from the Program Co-Chairs

Welcome to the Fourth ACM International Conference on Web Search and Data Mining (WSDM 2011) held on February 9–12, 2011, in Hong Kong. As the premier ACM conference in the field, WSDM 2011 offers a highly competitive forum for reporting the latest developments in web search, social search and data mining. We are pleased to present the proceedings of the conference as its published record.

Although it is only in its fourth year, WSDM has already witnessed significant growth. We received a record 372 submissions, representing a 22% increase compared to WSDM 2010. 19 Senior PC members and 134 PC members conducted reviews to the submissions. The conference accepted 83 papers (22.3% acceptance rate). Among these, 32 papers were selected for oral and poster presentations and 51 papers were selected for poster only presentations. The authors of submitted papers were from 35 countries and regions, authors of accepted papers are from 13 countries and regions. The quality of accepted papers is very high, making WSDM a first tier conference in computer science.

The conference program represents the great efforts of many people. We want to express our genuine gratitude to all the members of the program committee, and the external reviewers for their hard work in reviewing the submissions. Many thanks also go to the general chair of the conference, Irwin King, who provided with us support in many ways. We are also grateful to the two invited speakers, Christos Faloutsos and Harry Shum, for sharing their thoughts and insights. Finally, we thank all the authors and the participants! The conference would not be possible without the contributions from them.

We hope that this program will further stimulate research in the field, and provide better technologies to industry and practitioners. We are honored and privileged to serve the field of WSDM through this exciting program.

Please enjoy the conference!

Wolfgang Nejdl
L3S and University of Hannover
Germany

Hang Li
Microsoft Research Asia
China

Wolfgang Nejdl
L3S and
University of Hannover
Germany

Hang Li
Microsoft Research Asia
China

Benjamin W. Wah

Provost and
Wei Lun Professor of
Computer Science and
Engineering,
The Chinese University
of Hong Kong

Professor Benjamin Wah was the Franklin W. Woeltge Endowed Professor of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, and is a prominent computer scientist, with expertise in non-linear programming, multimedia signal processing and artificial intelligence. He is a fellow of the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery (ACM), and the American Association for the Advancement of Science (AAAS) and has served as the President of IEEE Computer Society. Professor Wah has received numerous international honours and awards for his distinguished academic and professional achievements. Among these are the W. Wallace McDowell Award, the Tsutomu Kanai Award and the Richard E. Merwin Distinguished Service Award of the IEEE Computer Society.

In 1998-99, Professor Wah was Professor of Computer Science and Engineering at CUHK, and in that year received an Exemplary Teaching Award. His bonds with the University continued afterwards as he served in the capacity of Adjunct Professor in the Department from 1999 to 2003.

Professor Wah has also long been committed to enhancing the development of higher education and research in Hong Kong. He was a member of the Research Grants Council of the University Grants Committee in Hong Kong between 2005 and 2009, and served as the Chairman of its Engineering Panel between 2006 and 2009.

Born and brought up in Hong Kong, Professor Wah graduated from Queen Elizabeth School and pursued further studies in the US. He received his BS and MS in Electrical Engineering and Computer Science from Columbia University, and his MS in Computer Science and PhD in Engineering from the University of California, Berkeley. He began teaching in Purdue University in 1979, and later joined the University of Illinois at Urbana-Champaign in 1985. He also served as Director of the Advanced Digital Sciences Centre established by the University of Illinois in Singapore in 2009, with funding from the Singapore government's Agency for Science, Technology and Research.

General Chair:	Irwin King (*CUHK, Hong Kong*)
Program Co-Chairs:	Wolfgang Nejdl (*L3S and University of Hannover, Germany*) Hang Li (*Microsoft Research Asia, China*)
Workshop Co-Chairs:	Evgeniy Gabrilovich (*Yahoo! Research, USA*) Wai Lam (*CUHK, Hong Kong*)
Tutorial Co-Chairs:	Marc Najork (*Microsoft Research, USA*) Qing Li (*City University of HK, Hong Kong*)
Local Arrangement Co-Chairs:	Dan Hong (*HKUST, Hong Kong*) Tak-Lam Wong (*HKIEd, Hong Kong*)
Finance Chair:	Michael R. Lyu (*CUHK, Hong Kong*)
Publicity Chair:	Qiang Yang (*HKUST, Hong Kong*)
Registration Chair:	Michael Chau (*HKU, Hong Kong*)
Publication Chair:	Hong Cheng (*CUHK, Hong Kong*)
Poster Session Chair:	Jie Tang (*Tsinghua University, China*)
Best Paper Award Committee:	Eytan Adar (*University of Michigan, USA*) James Allan (*University of Massachusetts Amherst, USA*) Brian D. Davison (*Lehigh University, USA*) Mounia Lalmas (*University of Glasgow, UK*) Hang Li (*Microsoft Research Asia, China*) Edleno Silva de Moura (*Universidade Federal do Amazonas, Brazil*) Wolfgang Nejdl (*L3S and University of Hannover, Germany*) Qiang Yang (*HKUST, Hong Kong*)
Steering Committee:	Rakesh Agrawal (*Microsoft, USA*) Ricardo Baeza-Yates (*Yahoo! Research, Spain*) Ziv Bar-Yossef (*Technion, Google, Israel*) Soumen Chakrabarti (*IIT Bombay, India*) Monika Rauch Henzinger (*Google, USA*) Jon Kleinberg (*Cornell University, USA*) Marc Najork (*Microsoft Research, USA*)

Sponsors

SIGIR Special Interest Group on Information Retrieval

ACM SIGMOD

acm KDD.org ACM Special Interest Group on Knowledge Discovery and Data Mining

sig web

Supporters

Platinum Supporter

Microsoft Research

Gold Supporters

Google

YAHOO! LABS

Silver Supporter

Yandex

Special Supporter

王寬誠教育基金會 K. C. WONG EDUCATION FOUNDATION

Technical Supporters

Arnetminer

Wednesday, 9 February 2011

Time	Description
8:30 - 10:00	**WS1** [Tang Room I] 3/F; **WS2** [Ming Room II] 4/F; **T-AM1** [Tang Room II] 3/F; **T-AM2** [Ming Room I] 4/F
10:00 - 10:30	**Coffee Break**
10:30 - 12:00	**WS1** [Tang Room I]; **WS2** [Ming Room II]; **T-AM1** [Tang Room II]; **T-AM2** [Ming Room I]
12:00 - 13:30	**Lunch** [Oyster Bar & Sky Lounge, 18/F]
13:30 - 15:00	**WS1** [Tang Room I] 3/F; **WS2** [Ming Room II] 4/F; **T-PM1** [Tang Room II] 3/F; **T-PM2** [Ming Room I] 4/F
15:00 - 15:30	**Coffee Break**
15:30 - 17:00	**WS1** [Tang Room I]; **WS2** [Ming Room II]; **T-PM1** [Tang Room II]; **T-PM2** [Ming Room I]
18:00 - 20:00	**Welcome Reception** [Ching Room & Ching Terrace, 4/F]

Workshops Information

- WS1: Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011)
- WS2: Workshop on User Modelling for Web Applications

Tutorials Information

- T-AM1: Crowdsourcing 101: Putting the "Wisdom of the Crowd" to Work for You
- T-AM2: Introduction to Display Advertising
- T-PM1: Exploiting Statistical and Relational Information on the Web and in Social Media
- T-PM2: Web Retrieval: The Role of Users

Thursday, 10 February 2011 Venue: Grand Ballroom, 3/F

Time	Description
8:30 - 8:40	**Opening Ceremony** Chair: Irwin King
8:40 - 9:40	**Keynote I: Mining Billion-node Graphs: Patterns, Generators and Tools** Chair: Hang Li
9:40 - 10:30	**Oral Session 1: Web Search I** Chair: Yoelle Maarek
10:30 - 10:50	**Coffee Break**
10:50 - 12:30	**Oral Session 2: Social Search and Mining I** Chair: Eytan Adar
12:30 - 14:00	**Lunch** [Oyster Bar & Sky Lounge, 18/F]
14:00 - 15:40	**Oral Session 3: Web Search II** Chair: David Carmel
15:40 - 16:00	**Coffee Break**
15:40 - 17:15	**Poster Session I** Chair: Jie Tang

Friday, 11 February 2011 Venue: Grand Ballroom, 3/F

Time	Description
8:30 - 9:30	**Keynote II: Bing Dialog Model: Intent, Knowledge and User Interaction** Chair: Wolfgang Nejdl
9:30 - 10:20	**Oral Session 4: Social Search and Mining II** Chair: Ravi Kumar
10:20 - 10:50	**Coffee Break**
10:50 - 12:30	**Oral Session 5: Web Search III** Chair: Nick Craswell
12:30 - 14:00	**Lunch** [Oyster Bar & Sky Lounge, 18/F]
14:00 - 15:40	**Oral Session 6: Social Search and Mining III** Chair: Kevin Chang
15:40 - 16:00	**Coffee Break**
15:40 - 17:15	**Poster Session II** Chair: Jie Tang
17:15 - 19:15	**Excursion** The Peak Tram, The Peak and Sky Terrace
19:15 - 20:00	**Transit to Banquet Venue**
20:00	**WSDM2011 Awards Banquet** [Jumbo Floating Restaurant, Aberdeen]

Saturday, February 12, 2011 Venue: Grand Ballroom, 3/F

Time	Description
9:00 - 10:40	**Oral Session 7: Web Mining** Chair: Qiang Yang
10:40 - 11:00	**Coffee Break**
11:00 - 12:40	**Oral Session 8: Web Search IV** Chair: Ricardo Baeza-Yates
12:40 - 14:00	**Lunch** [Oyster Bar & Sky Lounge, 18/F]
14:00 - 15:00	**WSDM General Meeting** Chair: Ricardo Baeza-Yates
15:00 - 15:20	**Coffee Break**
15:20 - 17:00	**Oral Session 9: Social Search and Mining IV** Chair: Jaime Teevan
17:00	**Closing Ceremony** Chair: Irwin King

Christos Faloutsos

Professor, Department of Computer Science, Carnegie Mellon University

Mining Billion-node Graphs: Patterns, Generators and Tools

Date: 10 February 2011
Time: 8:40 - 9:40
Venue: Grand Ballroom

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 (like, e.g., "eigenSpokes"). For generators, we describe some recent ones, which naturally match all of the known properties of real graphs. Finally, for tools, we present "oddBall" for discovering anomalies and patterns, as well as an overview of the PEGASUS system which is designed for handling 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, the SIGKDD Innovations Award (2010) sixteen "best paper" awards, and four teaching awards. He has served as a member of the executive committee of SIGKDD; he has published over 200 refereed articles, 11 book chapters and one monograph. He holds five patents and he has given over 30 tutorials and over 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.

Bing Dialog Model: Intent, Knowledge and User Interaction

Date: 11 February 2011
Time: 8:30 - 9:30
Venue: Grand Ballroom

Abstract:

The decade-old Internet search outcomes, manifested in the form of "ten blue links," are no longer sufficient for Internet users. Many studies have shown that when users are ushered off the conventional search result pages through blue links, their needs are often partially met at best in a "hit-or-miss" fashion. To tackle this challenge, we have designed Bing, Microsoft's decision engine, to not just navigate users to a landing page through a blue link but to continue engaging with users to clarify intent and facilitate task completion. Underlying this new paradigm is the Bing Dialog Model that consists of three building blocks: an indexing system that comprehensively collects information from the web and systematically harvests knowledge, an intent model that statistically infers user intent and predicts next action, and an interaction model that elicits user intent through mathematically optimized presentations of web information and domain knowledge that matches user needs. In this talk, I'll describe Bing Dialog Model in details and demonstrate it in action through some innovative features since the launch of www.Bing.com

Biography:

Harry Shum is the corporate vice president responsible for search product development at Microsoft Corp. Previously he oversaw the research activities at Microsoft Research Asia and the lab's collaborations with universities in the Asia Pacific region, and was responsible for the Internet Services Research Center, an applied research organization dedicated to long-term and short-term technology investments in search and advertising at Microsoft.

Shum joined Microsoft Research in 1996, as a researcher based in Redmond, Wash. He moved to Beijing as one of the founding members of Microsoft Research China (later renamed Microsoft Research Asia). There he began a nine-year tenure as a research manager, subsequently moving on to become assistant managing director, managing director of Microsoft Research Asia, Distinguished Engineer and corporate vice president. Shum is an Institute of Electrical and Electronics Engineers Fellow and an Association for Computing Machinery Fellow for his contributions on computer vision and computer graphics. He has published more than 100 papers about computer vision, computer graphics, pattern recognition, statistical learning and robotics. He holds more than 50 U.S. patents.

Shum received a doctorate in robotics from the School of Computer Science at Carnegie Mellon University in Pittsburgh. In his spare time he enjoys playing basketball, rooting for the Pittsburgh Steelers and spending time with his family.

Harry Shum

VP of Search Product Development, Microsoft

WS1: Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011)

Wednesday, 9 February 2011, 8:30 - 17:00 [Venue: Tang Room I, 3/F]

Workshop Organizers:
Vitor R. Carvalho, Microsoft Bing, USA
Matthew Lease, School of Information and Department of Computer Science, University of Texas at Austin, USA
Emine Yilmaz, Microsoft Research, UK

Web site:
http://ir.ischool.utexas.edu/csdm2011/

Abstract
The advent of crowdsourcing is revolutionizing data annotation, evaluation, and other traditionally manual-labor intensive processes by dramatically reducing the time, cost, and effort involved. This in turn is driving a disruptive shift in search and data mining methodology in areas such as evaluation, supervised learing, and applications.

Evaluation:
the Cranfield paradigm for evaluating search engines requires manually assessing document relevance to search queries. Recent work on stochastic evaluation has reduced but not removed this dependence on manual assessment.

Supervised Learning:
while traditional costs associated with data annotation have driven recent machine learning work (e.g. Learning to Rank) toward greater use of unsupervised and semi-supervised methods, the emergence of crowdsourcing has made labeled data far easier to acquire, thereby driving a potential resurgence in fully-supervised methods.

Applications:
Crowdsourcing has introduced exciting new opportunities to integrate human labor into automated systems: handling difficult cases where automation fails, exploiting the breadth of backgrounds, geographic dispersion, and real-time crowd response, etc.

Program

Invited Talks

The Smarter Crowd: Active Learning, Knowledge Corroboration, and Collective IQs
Thore Graepel

Crowdsourcing using Mechanical Turk: Quality Management and Scalability
Panagiotis G. Ipeirotis

Individual vs. Group Success in Social Networks
Winter Mason

Accepted Papers

Perspectives on Infrastructure for Crowdsourcing
Omar Alonso

How Crowdsourcable is Your Task?
Carsten Eickhoff and Arjen de Vries

You're Hired! An Examination of Crowdsourcing Incentive Models in Human Resource Tasks
Christopher Harris

Modeling Annotator Accuracies for Supervised Learning
Abhimanu Kumar and Matthew Lease

Crowdsourcing Blog Track Top News Judgments at TREC
Richard McCreadie, Craig Macdonald and Iadh Ounis

Investigating Factors Influencing Crowdsourcing Tasks with High Imaginative Load
Raynor Vliegendhart, Martha Larson, Christoph Kofler, Carsten Eickhoff and Johan Pouwelse

Estimating Completion Time for Crowdsourced Tasks Using Survival Analysis Models
Jing Wang, Siamak Faridani and Panagiotis Ipeirotis

Crowdsourcing Interactions - A proposal for capturing user interactions through crowdsourcing
Guido Zuccon, Teerapong Leelanupab, Stewart Whiting, Joemon Jose and Leif Azzopardi

WS2: Workshop on User Modelling for Web Applications

Wednesday, 9 February 2011, 8:30 - 17:00 [Venue: Ming Room II, 4/F]

Workshop Organizers:
David Carmel, IBM Haifa Research Lab, Israel
Vanja Josifovski, Yahoo! Research Silicon Valley, USA
Yoelle S. Maarek, Yahoo! Research, Israel

Web site:
http://research.yahoo.com/workshops/umwa2011/

Users have taken a more and more central role in the Web. Their role is both explicit, as they become more savvy, they have more expectations, and new interactive features keep appearing, and implicit, as their actions are monitored at various levels of granularity for various needs from live traffic evaluation for usage data mining to improve ranking, spelling etc. In a few years, most Web applications will have the ability to successfully adapt to both the explicit and implicit needs and tastes of their users. Such adaptation requires the ability to model the user's personal goals, interests, preferences and knowledge, and to apply this model while users interact with various applications. While adaptive applications that are based on user modeling have attracted the attention of multiple communities, from AI to UI, there is no forum that specifically focuses on user modeling and adaptive applications in the Web domain.

This workshop will focus on user modeling and the usage of these models in Web applications. The emphasis of the workshop will be on modeling techniques that scale for the Web. User modeling might be based on explicit and implicit user feedback gathered from variety of sources such as sign-on information, clickthrough data, user previous queries, social network, purchases, and real-world activity. Adaptive Web based applications include search personalization, advertisement targeting, recommendation systems, social networks, on-line shopping, etc.

Program

UMWA'2011
David Carmel, Vanja Josifovski and Yoelle Maarek

Inferring Clickthrough Rates on Ads from Click Behavior on Search Results
Sreenivas Gollapudi, Rina Panigrahy and Moises Goldszmidt

Social Media Analytics: Tracking the Flow of Information in Networks
Jure Leskovec

Learning to Predict Web Collaborations
Lilyana Mihalkova, Walaa Eldin, Moustafa and Lise Getoor

Targeting for Computational Market Research
Frank Smadja

Invited talk: Leveraging Temporal Variability of Users and Content
Eytan Adar

Position paper: Towards a Science of User Engagement
Mounia Lalmas, Simon Attfield, Gabriella Kazai and Benjamin Piwowarski

Different Users and Intents: An Eye-tracking Analysis of Web Search
Cristina Gonzalez-Caro and Mari-Carmen Marcos

Web Queries: the Tip of the Iceberg of the User's Intent
Cristina Gonzalez-Caro, Liliana Caldero-Benavides and Ricardo Baeza-Yates

T-AM1: Crowdsourcing 101: Putting the "Wisdom of the Crowd" to Work for You

Wednesday, 9 February 2011, 8:30 - 12:00 [Venue: Tang Room II, 3/F]

Presenters:
Omar Alonso, Bing team at Microsoft, USA
Matthew Lease, School of Information and Department of Computer Science, University of Texas at Austin, USA

Crowdsourcing has emerged in recent years as an exciting new avenue for leveraging the tremendous potential and resources of today's digitally-connected, diverse, distributed workforce. Generally speaking, crowdsourcing describes outsourcing of tasks to a large group of people instead of assigning such tasks to an in-house employee or contractor. Crowdsourcing platforms such as Amazon Mechanical Turk and CrowdFlower have gained particular attention as active online market places for reaching and tapping into this glut of a still largely under-utilized workforce. Crowdsourcing offers intriguing new opportunities for accomplishing different kinds of tasks or achieving broader participation than previously possible, as well as completing standard tasks more accurately in less time and at lower cost. Unlocking the potential of crowdsourcing in practice, however, requires a tri-partite understanding of principles, platforms, and best practices. This tutorial will introduce the opportunities and challenges of crowdsourcing while discussing the three issues above. This will provide attendees with a basic foundation to begin applying crowdsourcing in the context of their own particular tasks.

T-AM2: Introduction to Display Advertising

Wednesday, 9 February 2011, 8:30 - 12:00 [Venue: Ming Room I, 4/F]

Presenters:
Andrei Broder, Yahoo! Research Silicon Valley, USA
Vanja Josifovski, Yahoo! Research Silicon Valley, USA
Jayavel Shanmugasundaram, Yahoo! Research Silicon Valley, USA

Web advertising supports a large swath of the Internet ecosystem. It brings revenue to countless publishers that rent space on their pages for advertising: from small mom-and-pop shops to major search engines. It also provides valuable traffic to numerous commercial Web sites and has fueled the development of Web search engines. Today, Web advertising is increasingly impacting the world outside the Internet by shaping the attitudes of numerous users. Computational advertising is a new scientific discipline that aims to formalize the problem of finding the best ad for a given user in a given context. In traditional advertising, the number of venues is small, the cost per venue is higher, and little or no personalization is possible (as for example in print magazines). In contrast, in online advertising there are billion of opportunities (page views), hundreds of millions of ads and it is possible to provide personalization with quantifiable results. This brings the advertising into the realm of the other "computational" sciences. An overview of the current state of computational advertising can be found in http://msande239.stanford.edu/lectures/lecture-01.pdf. Display advertising is one of the two major advertising channels on the web (in addition to search advertising). Display advertising on the Web is usually done by graphical ads placed on the publishers' Web pages. There is noexplicit user query, and the ad selection is performed based on the page where the ad is placed (contextual targeting) or user's past activities (behavioral targeting). In both cases, sophisticated text analysis and learning algorithms are needed to provide relevant ads to the user.

Display advertising includes both a brand awareness component, where the aim of the advertiser is to promote awareness of a brand or a product, as well as a direct response component, where the aim of the advertiser is a click or conversion that leads to a visit to the advertiser's Web site or other downstream economic activity. In addition, advertisers can also choose one of several payment types: CPM (Cost Per Mille — or 1000 — impressions/user visits), CPC (Cost Per Click), or CPA (Cost Per Action/Conversion, which may involve, for instance, filling out a form or an actual purchase). Dealing with multiple objectives and payment types again requires sophisticated learning algorithms to enable conversion and comparison between the payment types.

Finally, in display advertising, advertisers can choose to buy ads on a guaranteed basis many months in advance (these are typically CPM buys). For instance, an advertiser can request 100 million impressions during Superbowl 2011, and the publisher guarantees these visits ahead of time (even though the users have not actually shown up!). In essence, purchasing on a guaranteed basis is like purchasing goods on a futures market. Advertiser can also choose to buy on a non-guaranteed basis (these can be CPM, CPC or CPA buys), and in this case, they only pay for each impression, click or conversion. Many of the mechanisms required to support these forms of buying, such as traffic forecasting, ad selection, and pricing are just starting to attract the attention of the research community, and there is ample opportunity for impactful research in this area.

T-PM1: Exploiting Statistical and Relational Information on the Web and in Social Media

Wednesday, 9 February 2011, 13:30 - 17:00 [Venue: Tang Room II, 3/F]

Presenters:
Lise Getoor, Computer Science Department, University of Maryland Institute for Advanced Computer Studies, University of Maryland, College Park, USA
Lilyana Mihalkova, LINQS group at the UMD College Park Computer Science Department, University of Maryland College Park, USA

The popularity of Web 2.0, characterized by a proliferation of social media sites, and Web 3.0, with more richly semantically annotated objects and relationships, brings to light a variety of important prediction, ranking, and extraction tasks. The input to these tasks is often best seen as a (noisy) multi-relational graph, such as the graph of the Web itself; the click graph, defined by user interactions with Web sites; and the social graph, defined by friendships and affiliations on social media sites.

This tutorial will provide an overview of statistical relational learning and inference techniques, motivating and illustrating them using web and social media applications. We will start by briefly surveying some of the sources of statistical and relational information on the web and in social media and will then dedicate most of the tutorial time to an introduction to representations and techniques for learning and reasoning with multi-relational information, viewing them through the lens of web and social media domains. We will end with a discussion of current trends and related fields, such as privacy in social networks and probabilistic databases.

T-PM2: Web Retrieval: The Role of Users

Wednesday, 9 February 2011, 13:30 - 17:00 [Venue: Ming Room I, 4/F]

Presenters:
Ricardo A. Baeza-Yates, Yahoo! Research labs at Barcelona, Spain and Santiago, Chile, Spain
Yoelle S. Maarek, Yahoo! Research, Israel

Web retrieval methods have evolved through three major steps in the last decade or so. They started from standard document-centric IR in the early days of the Web, then made a major step forward by leveraging the structure of the Web, using link analysis techniques in both crawling and ranking challenges. A more recent, no less important but maybe more discrete step forward, has been to enter the user in this equation in two ways:

Implicitly, through the analysis of usage data captured by query logs, and session and click information in general; the goal here being to improve ranking as well as to measure user's happiness and engagement.

Explicitly, by offering novel interactive features; the goal here being to better answer users' needs. This half day tutorial will cover the user-related challenges associated with the implicit and explicit role of users in Web retrieval. More specifically, we will review and discuss challenges associated with two types of activities, namely:

Usage data analysis and metrics - It is critical to monitor how users take advantage and interact with Web retrieval systems, as this implicit relevant feedback aggregated at a large scale, can approximate quite accurately the level of success of a given feature. Here we have to consider not only clicks statistics but also the time spent in a page, the number of actions per session, etc.

User interaction - Given the intrinsic problems posed by the Web, the key challenge for the user is to conceive a good query to be submitted to the search system, one that leads to a manageable and relevant answer. The retrieval system must complete search requests fast and give back relevant results, even for poorly formulated queries, as is the common case in the Web. Web retrieval engines thus interact with the user at two key stages:

Expressing a query: Human beings have needs or tasks to accomplish, which are frequently not easy to express as "queries". Queries, even when expressed in a more natural manner, are just a reflection of human needs and are thus, by definition, imperfect. This phenomenon could be compared to Plato's cave metaphor, where shadows are mistaken for reality.

Interpreting results: Even if the user is able to perfectly express a query, the answer might be split over thousands or millions of Web pages or not exist at all. In this context, numerous questions need to be addressed. Examples include: How do we handle a large answer? How do we rank results? How do we select the documents that really are of interest to the user? Even in the case of a single document candidate, the document itself could be large. How do we browse such documents efficiently?

Thursday, 10 February, 2011

08:30 – 08:40 **Opening Ceremony**
Chair: Irwin King

08:40 – 09:40 **Keynote I: Mining Billion-node Graphs: Patterns, Generators and Tools**
Chair: Hang Li

09:40 – 10:30 **Oral Session 1: Web Search I**
Chair: Yoelle Maarek

Who Uses Web Search for What? And How?
Ingmar Weber and Alejandro Jaimes
Keywords: query logs, demographics, session analysis, topic classification.

Personalizing Web Search using Long Term Browsing History
Nicolaas Matthijs and Filip Radlinski
Keywords: Personalized Web Search, Browsing History, Large-scale online evaluation, AlterEgo, Interleaving, Ranking, User Profile.

10:30 – 10:50 **Coffee Break**

10:50 – 12:30 **Oral Session 2: Social Search and Mining I**
Chair: Eytan Adar

TwitterSearch: A Comparison of Microblog Search and Web Search
Jaime Teevan, Daniel Ramage and Meredith Ringel Morris
Keywords: social search, query log analysis, microblog search.

Identifying Topical Authorities in Microblogs
Aditya Pal and Scott Counts
Keywords: Online Social Media, Authority Identification, Recommendation systems, Microblogging and Twitter, Clustering and Ranking.

Correcting for Missing Data in Information Cascades
Eldar Sadikov, Montserrat Medina, Jure Leskovec and Hector Garcia-Molina
Keywords: missing data, information cascades, sampling, social networks, Twitter.

Identifying 'Influencers' on Twitter
Eytan Bakshy, Jake M. Hofman, Winter Mason and Duncan J. Watts
Keywords: Communication networks, Twitter, influence, diffusion, word-of-mouth.

12:30 – 14:00 **Lunch**

14:00 – 15:40 **Oral Session 3: Web Search II**
Chair: David Carmel

A Comparative Analysis of Cascade Measures for Novelty and Diversity
Charles Clarke, Nick Craswell, Ian Soboroff and Azin Ashkan
Keywords: search effectiveness measures, novelty, diversity, ERR, nDCG, editorial assessment.

Understanding and Predicting Personal Navigation
Jaime Teevan, Daniel J. Liebling and Gayathri Ravichandran Geetha
Keywords: navigation, query intent, personalization, query log analysis.

Quality-Biased Ranking of Web Documents
Michael Bendersky, W. Bruce Croft and Yanlei Diao
Keywords: quality-biased ranking, document quality, web search.

Ranking From Pairs and Triplets: Information Quality, Evaluation Methods and Query Complexity
Kira Radinsky and Nir Ailon
Keywords: Search Quality, Search Evaluation, Ranking Evaluations, Preference Analysis.

15:40 – 16:00 **Coffee Break**

15:40 – 17:15 **Poster Session I**
Chair: Jie Tang

Friday, 11 February 2011

08:30 – 09:30 **Keynote II: Bing Dialog Model: Intent, Knowledge and User Interaction**
Chair: Wolfgang Nejdl

09:30 – 10:20 **Oral Session 4: Social Search and Mining II**
Chair: Ravi Kumar

We Feel Fine and Searching the Emotional Web
Sepandar Kamvar and Jonathan Harris
Keywords: Social Media Mining, Sentiment Analysis, Exploratory Sentiment Visualization.

On the Selection of Tags for Tag Clouds
Petros Venetis, Georgia Koutrika and Hector Garcia-Molina
Keywords: search, tag, tag cloud.

10:20 – 10:50 **Coffee Break**

10:50 – 12:30 **Oral Session 5: Web Search III**
Chair: Nick Craswell

Efficient Indexing of Repeated n-Grams
Samuel Huston, Alistair Moffat and W. Bruce Croft
Keywords: Repeated phrase, n-gram, hash filter, text reuse, scalable algorithm, distributed algorithm.

Batch Query Processing for Web Search Engines
Shuai Ding, Josh Attenberg, Ricardo Baeza-Yates and Torsten Suel
Keywords: Web Search Performance, Query Processing, Caching, Algorithms, Measurement, Experimentation.

Detecting Duplicate Web Documents using Clickthrough Data
Filip Radlinski, Paul N. Bennett and Emine Yilmaz
Keywords: Duplication, Redundancy, Novelty, Utility, Web Search.

KMV-Peer: A Robust and Adaptive Peer-Selection Algorithm
Yosi Mass, Yehoshua Sagiv and Michal Shmueli-Scheuer
Keywords: P2P Search, Performance, Adaptive algorithm

12:30 – 14:00 **Lunch**

14:00 – 15:40 **Oral Session 6: Social Search and Mining III**
Chair: Kevin Chang

Understanding Temporal Query Dynamics
Anagha Kulkarni, Jaime Teevan, Krysta Svore and Susan Dumais
Keywords: temporal dynamics, query intent, query popularity, results content.

Patterns of Temporal Variation in Online Media
Jaewon Yang and Jure Leskovec
Keywords: information cascade, time-series clustering, online media.

Using Graded-Relevance Metrics for Evaluating Community QA Answer Selection
Tetsuya Sakai, Daisuke Ishikawa, Noriko Kando, Yohei Seki, Kazuko Kuriyama and Chin-Yew Lin
Keywords: best answers, community question answering, evaluation, graded relevance, NTCIR, test collections.

Mining Social Images with Distance Metric Learning for Automated Image Tagging
Pengcheng Wu, Steven C. H. Hoi, Peilin Zhao and Ying He
Keywords: social images, distance metric learning, inductive learning, transductive learning, automated image tagging.

15:40 – 16:00 **Coffee Break**

15:40 – 17:15 **Poster Session II**
Chair: Jie Tang

17:15 – 19:15 **Excursion: The Peak Tram, The Peak and Sky Terrace**
19:15 – 20:00 **Transit to Banquet Venue**
20:00 **WSDM2011 Awards Banquet** [Venue: Jumbo Floating Restaurant, Aberdeen]

Saturday, 12 February 2011

09:00 – 10:40 **Oral Session 7: Web Mining**
Chair: Qiang Yang

Dynamic Relationship and Event Discovery
Anish Das Sarma, Alpa Jain and Cong Yu
Keywords: entity relationship discovery, event discovery, algorithm.

Joint Training for Open-domain Extraction on the Web: Exploiting Overlap when Supervision is Limited
Rahul Gupta and Sunita Sarawagi
Keywords: Conditional Random Fields, Collective Training, Graphical Models, Open-domain Ad-hoc Extraction, Learning-based Record Extractors.

Scalable Knowledge Harvesting with High Precision and High Recall
Ndapandula Nakashole, Martin Theobald and Gerhard Weikum
Keywords: Knowledge Harvesting, Information Extraction, Scalability.

Mining Named Entities with Temporally Correlated Bursts from Multilingual Web News Streams
Alexander Kotov, ChengXiang Zhai and Richard Sporat
Keywords: text streams, natural language processing, transliteration, probabilistic modeling, dynamic programming.

10:40 – 11:00 **Coffee Break**

11:00 – 12:40 **Oral Session 8: Web Search IV**
Chair: Ricardo Baeza-Yates

Dynamic Ranked Retrieval
Christina Brandt, Thorsten Joachims, Yisong Yue and Jacob Bank
Keywords: Ranking and Machine Learning for Ranking, Diversified Retrieval, Relevance Feedback, Decision Theory, Personalized Search and Ranking.

Optimizing Two-Dimensional Search Results Presentation
Flavio Chierichetti, Ravi Kumar and Prabhakar Raghavan
Keywords: Two-dimensional search results, Markov chains, Placement problem.

Result Enrichment in Commerce Search using Browse Trails
Sreenivas Gollapudi and Debmalya Panigrahi
Keywords: commerce search, browse trails, result enrichment, web domain annotation.

Identifying Task-based Sessions in Search Engine Query Logs
Claudio Lucchese, Salvatore Orlando, Raffaele Perego, Fabrizio Silvestri and Gabriele Tolomei
Keywords: Query Log Mining, Query Log Session Boundary Detection, Task-based Sessions, Query Clustering.

12:40 – 14:00	**Lunch**
14:00 – 15:00	**WSDM General Meeting** Chair: Ricardo Beaza-Yates
15:00 – 15:20	**Coffee Break**
15:20 – 17:00	**Oral Session 9: Social Search and Mining IV** Chair: Jaime Teevan **Recommender Systems with Social Regularization** *Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu and Irwin King* *Keywords: Recommender Systems, Collaborative Filtering, Social Network, Matrix Factorization, Social Regularization.* **Unbiased Offline Evaluation of Contextual-bandit-based News Article Recommendation Algorithms** *Lihong Li, Wei Chu, John Langford and Xuanhui Wang* *Keywords: recommendation, offline evaluation, benchmark dataset, contextual multi-armed bandit.* **Efficient Online Ad Serving in a Display Advertising Meta-Exchange** *LKevin Lang, Joaquin Delgado, Swaroop Jagadish, Amita Gajewar, Dongming Jiang, Michael Bindeberger-Ortega, Bhaskar Ghosh, Shirshanka Das, Arathi Seshan, Chavdar Botev, Sunil Nagaraj and Raymie Stata* *Keywords: Display Advertising, Online Ad Serving, Graph Algorithm, Constrained Path Optimization, Ad Exchange.* **Trend Analysis Model: Trend Consists of Temporal Words, Topics, and Timestamps** *Noriaki Kawamae* *Keywords: Topic Modeling, Latent Variable Modeling, Graphical Models, Trend Analysis, Timestamped data.*
17:00	**Closing Ceremony** Chair: Irwin King

Thursday, 10 February 2011 (15:40 – 17:15)

Topical Semantics of Twitter Links
Michael Welch, Uri Schonfeld, Dan He and Junghoo Cho

Evaluating the Visual Quality of Web Pages Using a Computational Aesthetic Approach
Ou Wu

Clustering Product Features for Opinion Mining
Zhongwu Zhai, Bing Liu, Hua Xu and Peifa Jia

Materializing Multi-Relational Databases from the Web using Taxonomic Queries
Matthew Michelson, Sofus Macskassy, Steven Minton and Lise Getoor

What Blogs Tell Us about Websites: A Demographics Study
Matt Michelson and Sofus Macskassy

Cross Llingual Text Classification by Mining Multilingual Topics from Wikipedia
Xiaochuan Ni and Jian-Tao Sun

Multidimensional Mining of Large-Scale Search Logs: A Topic-Concept Cube Approach
Dongyeop Kang, Daxin Jiang, Jian Pei, Zhen Liao, Xiaohui Sun and Ho-Jin Choi

Optimizing Merchant Revenue with Rebates
Rakesh Agrawal, Samuel Ieong and Raja Velu

Searchable Web Sites Recommendation
Nam Nguyen and Yang Song

Inferring Search Behaviors Using Partially Observable Markov Model with Duration (POMD)
Yin He and Kuansan Wang

Improving Relevance using Web Hierarchies for Contextual Advertising
Pavan Kumar GM, Krishna P. Leela, Mehul Parsana, Sachin Garg and Ashwin Tengli

Action Prediction and Identification From Mining Temporal User Behavior
Gang Wang and Weizhu Chen

Collective Bayesian Extraction from Heterogeneous Web Lists
Ashwin Machanavajjhala, Arun Iyer, Philip Bohannon and Srujana Merugu

CMAP: Effective Fusion of Quality and Relevance for Multi-criteria Recommendation
Xin Xin, Michael R. Lyu and Irwin King

CoBayes: Bayesian Knowledge Corroboration with Assessors of Unknown Areas of Expertise
Gjergji Kasneci, Jurgen Van Gael, David Stern and Thore Graepel

Multi-dimensional Search Result Diversification
Zhicheng Dou, Sha Hu, Kun Chen, Ruihua Song and Ji-Rong Wen

Improving Social Bookmark Search Using Personalised Latent Variable Language Models
Morgan Harvey, Mark Carman and Ian Ruthven

Strength of Social Influence in Trust Networks in Product Review Sites
Ching Man Au Yeung and Tomoharu Iwata

A Combined Topical/Non-topical Approach to Identifying Web Sites for Children
Carsten Eickhoff, Pavel Serdyukov and Arjen de Vries

Bid Generation for Advanced Match in Sponsored Search
Michael Welch, Uri Schonfeld, Dan He and Junghoo Cho

Let Web Spammers Expose Themselves
Zhicong Cheng, Bin Gao, Congkai Sun, Yanbing Jiang and Tie-Yan Liu

Efficient Entity Resolution for Large Heterogeneous Information Spaces
George Papadakis, Ekaterini Ioannou, Claudia Niederée and Peter Fankhauser

Web-Scale Table Census and Classification
Eric Crestan and Patrick Pantel

A Probabilistic Approach for Learning Folksonomies from Structured Data
Anon Plangprasopchok, Kristina Lerman and Lise Getoor

Linking Online News and Social Media
Manos Tsagkias, Maarten de Rijke and Wouter Weerkamp

Document Assignment in Multi-site Search Engines
Ulf Brefeld, B. Barla Cambazoglu and Flavio P. Junqueira

Who Uses Web Search for What? And How?
Ingmar Weber and Alejandro Jaimes

Personalizing Web Search using Long Term Browsing History
Nicolaas Matthijs and Filip Radlinskio

#TwitterSearch: A Comparison of Microblog Search and Web Search
Jaime Teevan, Daniel Ramage and Meredith Ringel Morris

Identifying Topical Authorities in Microblogs
Aditya Pal and Scott Counts

Correcting for Missing Data in Information Cascades
Eldar Sadikov, Montserrat Medina, Jure Leskovec and Hector Garcia-Molina

Identifying 'Influencers' on Twitter
Eytan Bakshy, Jake M. Hofman, Winter Mason and Duncan J. Watts

A Comparative Analysis of Cascade Measures for Novelty and Diversity
Charles Clarke, Nick Craswell, Ian Soboroff and Azin Ashkan

Dynamic Ranked Retrieval
Christina Brandt, Thorsten Joachims, Yisong Yue and Jacob Bank

Quality-Biased Ranking of Web Documents
Michael Bendersky, W. Bruce Croft and Yanlei Diao

Ranking From Pairs and Triplets: Information Quality, Evaluation Methods and Query Complexity
Kira Radinsky and Nir Ailon

Dynamic Relationship and Event Discovery
Anish Das Sarma, Alpa Jain and Cong Yu

Joint Training for Open-domain Extraction on the Web: Exploiting Overlap when Supervision is Limited
Rahul Gupta and Sunita Sarawagi

Scalable Knowledge Harvesting with High Precision and High Recall
Ndapandula Nakashole, Martin Theobald and Gerhard Weikum

Mining Named Entities with Temporally Correlated Bursts from Multilingual Web News Streams
Alexander Kotov, ChengXiang Zhai and Richard Sporat

Understanding and Predicting Personal Navigation
Jaime Teevan, Daniel J. Liebling and Gayathri Ravichandran Geetha

Optimizing Two-Dimensional Search Results Presentation
Flavio Chierichetti, Ravi Kumar and Prabhakar Raghavan

Friday, 11 February 2011 (15:40 – 17:15)

Transient Crowd Discovery on the Real-Time Social Web
Krishna Kamath and James Caverlee

Adaptive Bootstrapping of Recommender Systems Using Decision Trees
Nadav Golbandi, Yehuda Koren and Ronny Lempel

Predicting Future Reviews: Sentiment Analysis Models for Collaborative Filtering
Noriaki KAWAMAE

Learning Similarity Function for Rare Queries
Jingfang Xu and Gu Xu

A Two-View Learning Approach for Image Tag Ranking
Jinfeng Zhuang and Steven Hoi

Supervise Random Walks: Predicting Links in Social Networks
Lars Backstrom and Jure Leskovec

OOLAM: An Opinion Oriented Link Analysis Model for Influence Persona Discovery
Keke Cai, Shenghua Bao, Zi Yang, Jie Tang, Rui Ma, Li Zhang and Zhong Su

eBay: An E-Commerce Marketplace as a Complex Network
Zeqian Shen and Neel Sundaresan

A Framework for Quantitative Analysis of Cascades on Networks
Rumi Ghosh and Kristina Lerman

Scalable Clustering of News Search Results
Srinivas Vadrevu, Choon Hui Teo, Suju Rajan, Kunal Punera, Byron Dom, Alexander Smola, Yi Chang and Zhaohui Zheng

Large-Scale Hierarchical Text Classification without Labelled Data
Viet Ha-Thuc and Jean-Michel Renders

Low-order Tensor Decomposition for Social Tagging Recommendation
Yuanzhe Cai, Miao Zhang, Chris Ding and Sharma Chakravarthy

Shopping for Products You Don't Know You Need
Srikanth Jagabathula, Nina Mishra and Sreenivas Gollapudi

On Composition of a Federated Web Search Result Page: Using Online Users to Provide Pairwise Preference Heterogeneous Verticals
Ashok Ponnuswami, Kumaresh Pattabiraman, Qiang Wu, Ran Gilad-Bachrach and Tapas Kanungo

Online Stochastic Query Covering
Aris Anagnostopoulos, Luca Becchetti, Stefano Leonardi, Ida Mele and Piotr Sankowski

Learning to Re-rank Web Search Results with Multiple Pairwise Features
Changsung Kang, Xuanhui Wang, Belle Tseng, Yi Chang and Zhaohui Zheng

The Tube over Time: Characterizing Popularity Growth of YouTube Videos
Flavio Figueiredo, Fabricio Benevenuto and Jussara Almeida

Citation Recommendation without Author Supervision
Qi He, Daniel Kifer, Jian Pei, Prasenjit Mitra and Lee Giles

Query Suggestion for E-Commerce Sites
Mohammad Hasan, Nish Parikh, Gyanit Singh and Neel Sundaresan

An Algorithmic Treatment of Strong Queries
Ravi Kumar, Silvio Lattanzi and Prabhakar Raghavan

Enhanced Email Spam Filtering through Combining Similarity Graphs
Anirban Dasgupta, Maxim Gurevich and Kunal Punera

Merging the Results of Query Reformulations
Daniel Sheldon, Milad Shokouhi, Martin Szummer and Nick Craswell

Normalizing Web Product Attributes and Discovering Domain Ontology with Minimal Effort
Tak-lam Wong, Wai Lam and Lidong Bing

Aspect and Sentiment Unification Model for Online Review Analysis
Yohan Jo and Alice Oh

Rediscovering the Pattern-Relation Duality: Searching Patterns for Relation Extraction over the Web
Yuan Fang and Kevin Chen-Chuan Chang

We Feel Fine and Searching the Emotional Web
Sepandar Kamvar and Jonathan Harris

On the Selection of Tags for Tag Clouds
Petros Venetis, Georgia Koutrika and Hector Garcia-Molina

Efficient Indexing of Repeated n-Grams
Samuel Huston, Alistair Moffat and W. Bruce Croft

Batch Query Processing for Web Search Engines
Shuai Ding, Josh Attenberg, Ricardo Baeza-Yates and Torsten Suel

Detecting Duplicate Web Documents using Clickthrough Data
Filip Radlinski, Paul N. Bennett and Emine Yilmaz

KMV-Peer: A Robust and Adaptive Peer-Selection Algorithm
Yosi Mass, Yehoshua Sagiv and Michal Shmueli-Scheuer

Understanding Temporal Query Dynamics
Anagha Kulkarni, Jaime Teevan, Krysta Svore and Susan Dumais

Patterns of Temporal Variation in Online Media
Jaewon Yang and Jure Leskovec

Using Graded-Relevance Metrics for Evaluating Community QA Answer Selection
Tetsuya Sakai, Daisuke Ishikawa, Noriko Kando, Yohei Seki, Kazuko Kuriyama and Chin-Yew Lin

Mining Social Images with Distance Metric Learning for Automated Image Tagging
Pengcheng Wu, Steven C.H. Hoi, Peilin Zhao and Ying He

Result Enrichment in Commerce Search using Browse Trails
Sreenivas Gollapudi and Debmalya Panigrahi

Identifying Task-based Sessions in Search Engine Query Logs
Claudio Lucchese, Salvatore Orlando, Raffaele Perego, Fabrizio Silvestri and Gabriele Tolomei

Recommender Systems with Social Regularization
Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu and Irwin King

Unbiased Offline Evaluation of Contextual-bandit-based News Article Recommendation Algorithms
Lihong Li, Wei Chu, John Langford and Xuanhui Wang

Efficient Online Ad Serving in a Display Advertising Meta-Exchange
Lang, Joaquin Delgado, Swaroop Jagadish, Amita Gajewar, Dongming Jiang, Michael Bindeberger-Ortega, Bhaskar Ghosh, Shirshanka Das, Arathi Seshan, Chavdar Botev, Sunil Nagaraj and Raymie Stata

Trend Analysis Model: Trend consists of temporal words, topics, and timestamps
Noriaki Kawamae

Banquet and Excursion

On 11 February, we have arranged an excursion for conference participants to take the Peak Tram, the world's oldest and most famous funicular railways to The Peak to enjoy a 360-degree panoramic view across Hong Kong.

After the excursion, banquet is arranged at Jumbo Floating Restaurant where is ornamented in the style of a gorgeous and exquisite ancient Chinese imperial palace and serving fresh seafood in Cantonese style.

You are advised to come as early as possible in order not to miss any excitement in the excursion and have sufficient time for sightseeing and to be timely as delays will affect the subsequent itinerary.

Itinerary

4:45 PM – 5:15 PM
Gather at Group Lounge, Ground Floor of Sheraton Hotel

5:10 PM, 5:15 PM & 5:30 PM
Coaches depart SHARPLY from Sheraton Hong Kong Hotel & Towers to Peak Tram Station to the Peak Tram Terminus
Take the Peak Tram to The Peak
Sightseeing and shopping at The Peak:
- Sky Terrace - The highest 360-degree viewing Terrace (Admission Ticket included)
- Souvenirs and Gifts shops - At the Peak Tower and the Peak Galleria
- Madame Tussauds Hong Kong - features local & international celebrity wax figures

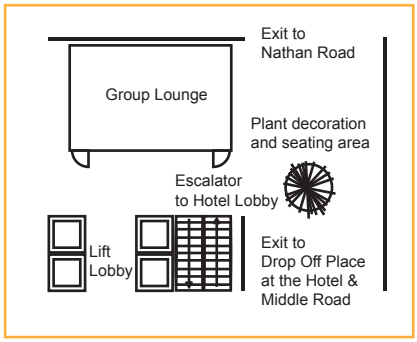
6:50 PM
Gather at the open area outside The Peak Tower and the tour guides will take you to the coach.

7:00 PM and 7:15 PM
Depart SHARPLY from The Peak to Jumbo Floating Restaurant for the Banquet

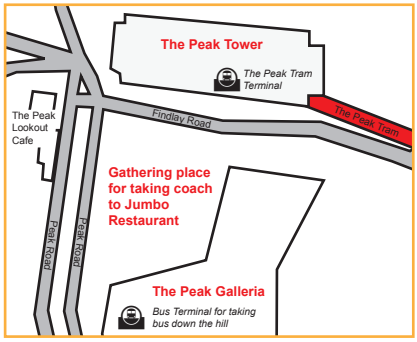
8:00 PM
Banquet at Jumbo Floating Restaurant

10:00 PM
Banquet ends and participants will take the coaches back to Sheraton Hong Kong Hotel & Towers

About The Peak: http://www.thepeak.com.hk/
About Jumbo Floating Restaurant: http://www.jumbo.com.hk/



The Group Lounge at the Ground Floor of Sheraton Hong Kong Hotel & Towers



The Peak

Enquiries:

Our staff will stay at the ***Pacific Coffee at The Peak Tower*** (Shop G08-10, Level G, next to the Peak Tram Terminal).

If you have any query, please go to find them there or call them at: ***(852) 9761 4450***.

Public Transportation to / from the Peak:

Means of Transportation	*Route No.*	*Path*	*Service Time*	*Frequency*	*Fare*
Bus (First Bus)	15	Central Pier 5 to The Peak	10:15 am to 12:15 am (From Central) 9:10 am to 1:00 am (from The Peak)	7 to 15 minutes	HK$ 9.80
Mini Bus	1	Central (Two IFC) to The Peak	6:30am - 12:00am (From Central) 7:10am - 12:25am (From The Peak)	10-12 minutes	HK$ 8.00

Public Transportation to / from Jumbo Floating Restaurant:

Means of Transportation	*Route No.*	*Path*	*Service Time*	*Frequency*	*Fare*
Bus (City Bus)	70	Central (Exchange Square) to Aberdeen *Take off at Aberdeen terminus. Then cross highway via subway (pedestrian tunnel), turn right, walk to dock for free boat ride to floating restaurant*	5:15am to 12:10am (From Central) 5:50am to 11:50 pm (From Aberdeen)	8 to 15 minutes	HK$ 4.70
Bus (City Bus)	75	Central (Exchange Square) to Sham Wan *Take off at Shum Wan Shuttle Ferry Pier for free boat ride to floating restaurant*	5:45am to 12:00am (From Central) 5:10am to 11:30 pm (From Sham Wan)	10 to 15 minutes	HK$ 4.70

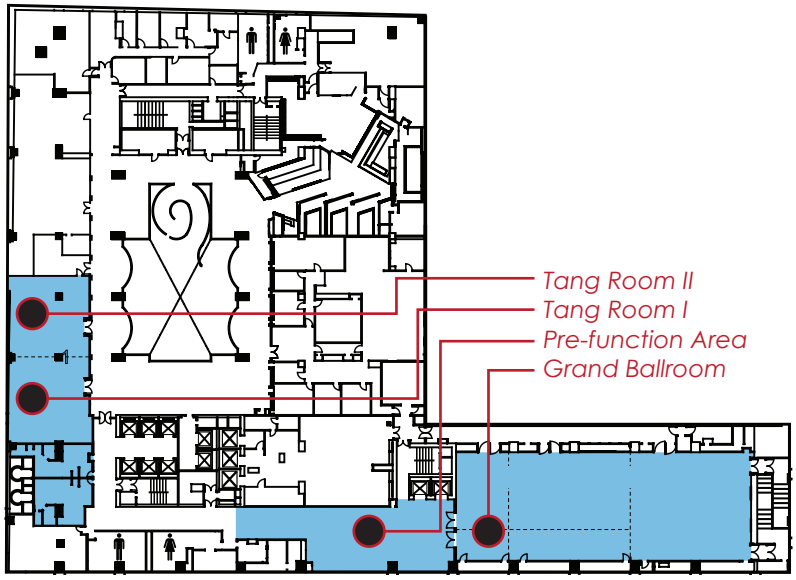
Information about the buses: www.nwfb.com.hk

Tang Room II
- *Tutorial AM1: Crowdsourcing 101: Putting the "Wisdom of the Crowd" to Work for You. (Morning Session, 9 February 2011)*
- *Tutorial PM1: Exploiting Statistical and Relational Information on the Web and in Social Media. (Afternoon Session, 9 February 2011)*

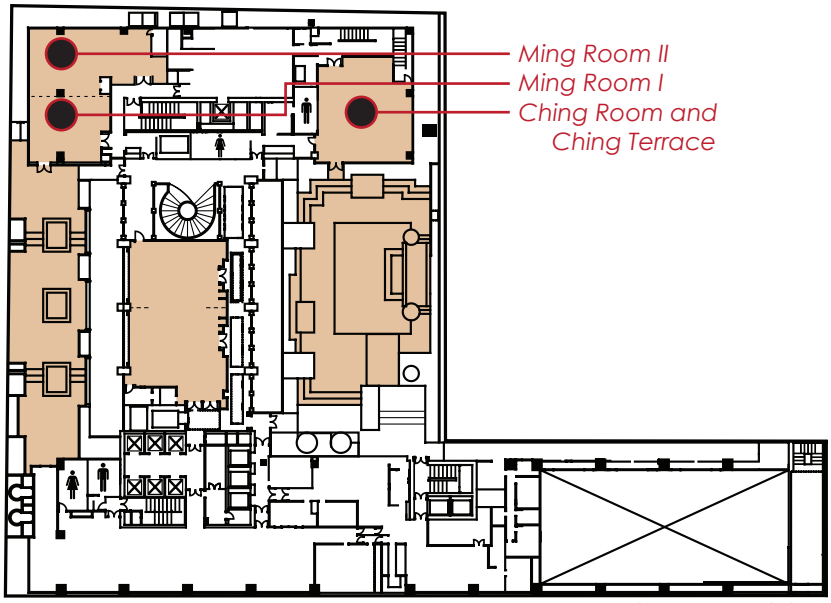
Tang Room I
- *Workshop 1: Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011) (9 February 2011)*

Pre-function Area
- *Registration (10-12 February 2011)*
- *Poster Collection (10-11 February 2011)*
- *Exhibition (10-12 February 2011)*

Grand Ballroom
- *Main Conference (10-12 February 2011)*
- *Poster Session (10-11 February 2011)*



THIRD FLOOR



FOURTH FLOOR

Ming Room II
- *Workshop 2: Workshop on User Modelling for Web Applications (9 February 2011)*

Ming Room I
- *Tutorial AM2: Introduction to Display Advertising. (Morning Session, 9 February 2011)*
- *Tutorial PM2: Web Retrieval: The Role of Users. (Afternoon Session, 9 February 2011)*

Ching Room and Ching Terrace
- *Welcome Reception (9 February 2011)*

wsdm 2011 Hong Kong

http://www.wsdm2011.org/

Registration:

Workshops and Tutorials	9 February 2011 8:00 - 17:00 Outside Tang Room, 3/F
Welcome Reception	9 February 2011 17:30 - 19:00 Outside Ching Room & Ching Terrace, 4/F
Main Conference	10-12 February 2011 8:00 - 16:00 Pre-function Area, 3/F
Posters Collection	10-11 February 2011 10:30 - 16:00 Pre-function Area, 3/F

WIFI Connection:

Please search and enter our network named: WSDM2011.
The network will be available during conference hours in the conference area.
No password is required.

Contact Information:

If there is any problem, please contact us at:

Email: wsdm2011@gmail.com

Phone no.: (852) 9761 4450
(operated during 7-12 February, 2011, 8:00 - 23:00)