Frontiers and Applications of Big Data Science

Program Booklet

cikm2016.org
General Chairs’ Welcome

Warm welcome to the 2016 ACM International Conference on Information and Knowledge Management (CIKM 2016)! CIKM is held annually at locations all over the world. The last two years it has been in Australia and China. In 2016 it returns to the United States in Indianapolis and is held Oct. 24-28.

CIKM 2016 will be the 25th running of the conference, which remains the only conference that inherently focuses on the need for users to have unified systems that access structured and unstructured data. With a new theme on Frontier and Applications of Big Data Science, CIKM 2016 further broadens the scope of topics of CIKM to include all topics in the general area of data science, including both foundational work and applications and encouraging interdisciplinary topics across multiple relevant communities such as information retrieval, databases, data mining, and knowledge management. CIKM has traditionally emphasized on relevance to industry and government applications; CIKM 2016 continues this tradition and further increases this emphasis by including a CIKM Cup Workshop and expanding its Industry Track to include a variety of submissions including long papers, short papers, and software system demonstrations in addition to its invited talks. Furthermore, to facilitate recruiting and encourage interactions between students and industry recruiters, CIKM 2016 introduced a new Career Day event. The conference also has a special session “CIKM at its 25th Anniversary” to celebrate its growth and success. The whole conference program includes eight tutorials, seven workshops, 182 oral presentations of full papers, 116 poster presentations of short papers, 20 system demonstrations, and three keynote talks.

We are honored and grateful to have three outstanding members of National Academy of Engineering as our keynote speakers, Rakesh Agrawal, Andrei Broder, and Susan Dumais.

Organizing a conference such as CIKM requires the dedication and effort of many volunteers. We are grateful to an outstanding group of colleagues for serving on the organization team and for generously contributing their time and energy, including especially our Local Organization Chair and Treasurer, Mohammad Al Hasan; Research Track Program Chairs, Elisa Bertino, Fabio Crestani, Javed Mostafa, Jie Tang, Luo Si, and Xiaofang Zhou; Industry Track Program Chairs, Yi Chang, Yunyao Li, and Parikshit Sondhi; Tutorial Chairs, Hui Fang, Latifur Khan, and Charles Nicholas; Workshop Chairs, Iadh Ounis, Huawei Shen, and Min Song; CIKM Cup Chairs, Pavel Izhutov and Nikita Spirin; Sponsorship Chairs, Sarath Janga (who also Co-Chaired the Career Day with Snehasis), Dou Shen, and Yi Zhang; Publication Chairs, Kavita Ganesan, Chase Geigle, and Xia Ning; Publicity Chairs, David Grossman, Yan Qu, and Hongning Wang; and Event Coordinator, Maureen Bowling who has taken care of most logistics of the conference. The conference would be impossible to organize without their volunteer spirit. Throughout the entire period of conference organization, we have also benefited from advice and mentoring from the CIKM Steering Committee, especially the past chair David Grossman and the current chair Jimi Shanahan, past CIKM Chairs Alistair Moffat, James Bailey, and Qi He, and other conference chairs, James Allan, Ricardo Baeza-Yates, Charles Clarke, Bruce Croft, Hang Li, Mark Sanderson, and Marianne Winslett. We want to thank all of them, especially Jimi and Alistair who not only provided mentoring but also offered detailed help for the conference budget and multiple planning decisions. Finally, we want to thank the two hosting universities, IUPUI and UIUC, ACM sponsors, SIGIR and SIGWEB, special sponsors, Professor Ram Kumar Memorial Foundation, AMiner and Whova, and all our corporate sponsors, especially our Platinum Sponsor Alibaba, Gold Sponsor Baidu, Silver Sponsor Huawei, and Bronze (CIKM Cup) Sponsor DCA for their generous support for CIKM 2016.

We hope that you will enjoy the conference and the opportunity to share and discuss ideas in Big Data Science with other researchers and practitioners around the world.

Snehasis Mukhopadhyay
CIKM’16 General CoChair
Indiana University Purdue
University Indianapolis, USA

ChengXiang Zhai
CIKM’16 General CoChair
University of Illinois at Urbana-Champaign, USA
Research Track Program Chairs’ Welcome

It is our great pleasure to welcome you to the 25th ACM International Conference on Information and Knowledge Management (CIKM 2016), held in Indianapolis, USA from October 24–28, 2016. CIKM is a top-tier ACM conference in the areas of information retrieval, knowledge management and databases. Since 1992, it has successfully brought together leading researchers and developers from the three communities, with the purpose of identifying challenging problems facing the development of advanced knowledge and information systems, and shaping future research directions through the publication of high quality, applied and theoretical research findings. CIKM 2016 continues to create an outstanding technical program including research paper presentations; poster presentations for extended short papers and short papers; and industry track presentations arranged by Industry Track Chairs. In total, we have received 701 full paper (10 pages) submissions and 234 short paper (4 pages) submissions reviewed by a group of 60 senior program committee members and 464 program committee members. The research track accepted 160 full papers with acceptance rate as 23%. 54 full paper submissions were accepted as extended short papers (6 pages) and 55 short submissions were accepted as short papers (4 pages) with acceptance rate as 24%.

We are grateful to everyone who made this technical program possible. First of all, we would like to thank all the contributing authors who submitted their excellent work. The conference would not exist without your contributions. Second, we would in particular like to recognize the immense efforts of the senior program committee members, program committee members, and external reviewers, who selected the technical program with their great expertise. Finally, we hope to thank all our corporate sponsors for their generous support.

We hope you enjoy the program and the opportunity to interact and collaborate with friends and colleagues from around the world.

CIKM 2016 PC Chairs:

Elisa Bertino
Fabio Crestani
Javed Mostafa
Jie Tang
Luo Si
Xiaofang Zhou
Industry Track Program Chairs' Welcome

It is our great pleasure to welcome you to industry track of the 25th ACM International Conference on Information and Knowledge Management (CIKM 2016). Our primary emphasis is on papers that advance the understanding of, and show how to deal with, practical issues related to the deployment of information and knowledge management systems that address the need for integration of structured and unstructured data in real-world applications.

CIKM being a premier conference in information retrieval, data mining and databases, the acceptance rates historically tend to be low. CIKM 2016 Industry Track attracted a significant number of submissions from countries all over the world: 111 long paper (10-page) submissions, 22 accepted (19.8% acceptance rate); 26 short paper (4-page) submissions, 7 accepted (26.9% acceptance rate); and 58 demo paper (4-page) submissions, 20 accepted (34.5% acceptance rate).

Each paper was first independently reviewed by at least three program committee members, and then discussed among them in a discussion moderated by a program chair. Borderline papers were thoroughly reviewed by the program chairs before final decisions were made.

All accepted long papers will be presented at the talk sessions, all short papers at the poster sessions, and all demo papers at the demo sessions. We have invited outstanding keynotes to be given by internationally renowned, distinguished scientists and practitioners: Rong Jin (Alibaba), Marc Nojork (Google), Shivakumar Vaithyanathan (IBM Research), and Haifeng Wang (Baidu).

We are grateful to everyone whose hard work made the industry program possible. Most of all, we would like to thank all authors who submitted their work to this program. We extend special thanks the program committee and external reviewers for their dedication and hard work in reviewing papers and participating in the extensive discussions. Finally, we are grateful for all our corporate sponsors for their generous support.

We hope that you will find this program interesting, and that this industry track will provide you with a valuable opportunity to learn and exchange ideas with other researchers and practitioners from academia, industry and governments around the world.

CIKM 2016 Industry Track Chairs:

Yi Chang (Yahoo Research): yichang@acm.org

Yunyao Li (IBM Research – Almaden): yunyaoli@us.ibm.com

Parikshit Sondhi (WalmartLabs): psondhi@walmartlabs.com
Conference Organization Committee

General Chairs: Snehasis Mukhopadhyay (Indiana University Purdue University Indianapolis, USA)
               ChengXiang Zhai (University of Illinois at Urbana-Champaign, USA)

Local Organization Chair and Treasurer: Mohammad Al Hasan (Indiana University – Purdue University Indianapolis, USA)

Research Track Program Chairs: Elisa Bertino (Purdue University, USA)
                                Fabio Crestani (Università della Svizzera italiana, Switzerland)
                                Javed Mostafa (University of North Carolina, USA)
                                Luo Si (Alibaba Group Inc and Purdue University, USA)
                                Jie Tang (Tsinghua University, China)
                                Xiaofang Zhou (University of Queensland, Australia)

Industry Track Program Chairs: Yi Chang (Yahoo Research, USA)
                               Yunyao Li (IBM Almaden Research Center, USA)
                               Parikshit Sondhi (WalmartLabs, USA)

Proceedings Chair: Kavita Ganesan, (Github Inc., USA)
                   Chase Geigle (University of Illinois at Urbana-Champaign, USA)
                   Xia Ning (Indiana University Purdue University Indianapolis, USA)

Tutorial Chairs: Hui Fang (University of Delaware, USA)
                  Latifur Khan (University of Texas at Dallas, USA)
                  Charles Nicholas (University of Maryland, Baltimore County, USA)

Workshop Chairs: Iadh Ounis (University of Glasgow, UK)
                 Huawei Shen (Chinese Academy of Sciences, China)
                 Min Song (Yonsei University, Korea)
**Publicity Chairs:**  
David Grossman (Georgetown University, USA)  
Yan Qu (eBay Inc, USA)  
Hongning Wang (University of Virginia, USA)

**Sponsorship Chairs:**  
Sarath Janga (Indiana University Purdue University Indianapolis, USA)  
Dou Shen (Baidu, China)  
Yi Zhang (University of California, Santa Cruz, USA)

**CIKM Cup Chairs:**  
Pavel Izhutov (Stanford University, USA)  
Nik Spirin (University of Illinois at Urbana-Champaign, USA)

**CIKM Career Day Chairs:**  
Sarath Janga (Indiana University Purdue University Indianapolis, USA)  
Snehasis Mukhopadhyay (Indiana University Purdue University Indianapolis, USA)

**Event Coordinator:**  
Maureen Bowling (IUPUI Event and Conference Services)

**Steering Committee:**  
James G. Shanahan (University of California Berkeley and Church and Duncan Group Inc), **Chair**  
Ricardo Baeza-Yates, (Universidad de Chile)  
Nick Craswell (Microsoft Bing)  
Marianne Winslett (University of Illinois at Urbana-Champaign)  
Phillip Yu (University of Illinois at Chicago)  
Rakesh Agrawal (Data Insights Laboratories)  
Jimmy Lin (University of Waterloo), **ACM SIGIR Liaison**  
Charles Nicholas (University of Maryland, Baltimore County), **ACM SIGWeb Liaison**
Sponsors & Supporters

Sponsors

Platinum Supporter

Gold Supporter

Silver Supporter

Bronze Supporter

General Supporters

CIKM Cup Supporter

Career Day Supporter

Academic Supporters
Alibaba Group’s mission is to make it easy to do business anywhere. We operate China’s leading online and mobile marketplaces in retail and wholesale trade, as well as cloud computing and other services. We provide the technologies and services to enable consumers, merchants, and other participants to do business in our ecosystem. We aim to build the future infrastructure of commerce. We envision that our customers will meet, work and live at Alibaba, and that we will be a company that lasts at least 102 years. http://alibabagroup.com

Baidu, Inc. is the leading Chinese language Internet search provider. As a technology-based media company, Baidu aims to provide the best and most equitable way for people to find they are looking for. In addition to serving individual Internet search users, Baidu provides an effective platform for businesses to reach potential customers. Baidu’s ADSs trade on the NASDAQ Global Select Market under the symbol “BIDU”. http://baidu.com

The Noah’s Ark Lab is a research lab of Huawei Technologies, located in Hong Kong and Shenzhen. The mission of the lab is to make significant contributions to both the company and society by innovating in data mining, artificial intelligence, and related fields. Founded in 2012, the lab has now grown to be a research organization with many significant achievements. We welcome talented researchers and engineers to join us to realize their dreams. http://www.noahlab.com.hk/

DCA (Data-Centric Alliance) is a Russian company specializing in Big Data and heavily loaded systems. Having one of the largest arrays of anonymous data about internet users, DCA develops and delivers a lineup of software products providing various up-to-date business solutions. http://datacentric.ru/en/

Yahoo Research serves as Yahoo’s incubator for bold new ideas and laboratory for rigorous experimentation. We work very closely with product and engineering partners to power the company’s most critical products with innovative science. The forward-looking insights from our research scientists also help position Yahoo as an industry and research leader. https://research.yahoo.com/

Google Research’s mission is to deliver cutting-edge innovation that improves Google products and enriches the lives of all who use them. We publish innovation through industry standards, and our researchers are often helping to define not just today’s products but also tomorrow’s. For more information, visit http://research.google.com

Professor Ram Kumar Memorial Foundation seeks to accelerate societal advancements through scientific breakthroughs, technological innovations, and the democratization of knowledge acquisition and diffusion. The foundation commemorates the life and achievements of Prof. Ram Kumar. http://ramkumarfoundation.org/
Alibaba Group's mission is to make it easy to do business anywhere.

We operate China’s leading online and mobile marketplaces in retail and wholesale trade, as well as cloud computing and other services. We provide the technologies and services to enable consumers, merchants, and other participants to do business in our ecosystem.

Our vision
We aim to build the future infrastructure of commerce. We envision that our customers will meet, work and live at Alibaba, and that we will be a company that lasts at least 102 years.

Meet @ Alibaba
We enable billions of commercial and social interactions among our users, between consumers and merchants, and among businesses every day.

Work @ Alibaba
We empower our customers with the fundamental infrastructure for commerce and data technology, so that they can build businesses and create value that can be shared among our ecosystem participants.

Live @ Alibaba
We strive to expand our products and services to become central to the everyday lives of our customers.

102 years
For a company that was founded in 1999, lasting at least 102 years means we will have spanned three centuries, an achievement that few companies can claim. Our culture, business models and systems are built to last, so that we can achieve sustainability in the long run.

If you are interested in Alibaba technology or looking for opportunities to cooperate with us, please do not hesitate to reach out to us at technology@service.alibaba.com
The Noah’s Ark Lab is a research lab of Huawei Technologies, located in Hong Kong and Shenzhen. The mission of the lab is to make significant contributions to both the company and society by innovating in data mining, artificial intelligence, and related fields. Founded in 2012, the lab has now grown to be a research organization with many significant achievements. We welcome talented researchers and engineers to join us to realize their dreams.

Data-Centric Alliance provides digital services based on Big Data. We operate massive behavioral datasets, cutting-edge technology and expertise to drive top efficiency for the businesses we work with. Our full-stack programmatic ecosystem Exebid DCA drives digital-marketing for hundreds of clients in Russia and globally. Data-Centric Alliance is the one-stop-shop for Big Data. Our team consists of data scientists, skilled developers and successful managers aiming to transform the understanding, acceptance and real-time application of Big Data into useful innovation and economic growth tools for our clients. We operate massive datasets resulting from tracking anonymous consumer behavior. These datasets are crunched by data management applications and real-time marketing solutions we create for targeted advertising and audience relationship management. DCA has developed a real-time key-value storage that integrates all user data available in the set to create an integrated user profile (non-personal identification entity) with easy and fast access for business integrations of any level. Our exceptional expertise in operating huge-size datasets, combined with non-stop investing in technological excellence lead to full-stack hi-end proprietary platforms, that are ready-to-use or easily customizable.
VENUE INFORMATION

Hyatt Regency
One South Capitol Avenue
Indianapolis, IN 46204, USA

T: +317 632 1234
F: +317 616 6299
E: indri.rfp@hyatt.com
http://indianapolis.hyatt.com

All conference activities are on the second floor of the hotel.

Hyatt Second Floor Map

System Demonstration
Network
Concept
Theory
Tutorials

Workshops & Parallel Sessions

Poster Presentations

Career Day

All Plenary Sessions
(Regency Ballroom = Regency A+B+C+D)
Conference Schedule: Overview

**Oct. 23, 2016, Sunday**
3:00pm-7:00pm: Registration

**Oct. 24, 2016, Monday**
8:00am-6:00pm: Registration
9:00am-5:30pm: Tutorials & Career Day *(lunch provided)*
6:00pm-7:30pm: Welcome Reception

**Oct. 25, 2016, Tuesday**
8:00am-6:00pm: Registration
9:00am-10:10am: Conference Welcome
9:20am-10:30am: Keynote Talk *(Rakesh Agrawal)*
11:00am-5:30pm: Research & Industry Track presentations *(lunch on your own)*
5:30pm-7:00pm: Poster Session 1 & Demo Session 1

**Oct. 26, 2016, Wednesday**
8:00am-6:00pm: Registration
9:00am-10:10am: Keynote Talk *(Susan Dumais)*
10:40am-3:30pm: Research & Industry Track presentations *(lunch on your own)*
4:00pm-5:15pm: CIKM 25th Anniversary Special Session
5:30pm-7:00pm: Poster Session 2 & Demo Session 2
7:00pm-9:00pm: Banquet

**Oct. 27, 2016, Thursday**
8:00am-5:30pm: Registration
9:00am-10:10am: Keynote Talk *(Andrei Broder)*
10:40am-12:10pm: Research & Industry Track presentations
12:40pm-2:00pm: CIKM business meeting *(lunch provided)*
2:30pm-6:00pm: Research & Industry Track presentations

**Oct. 28, 2016, Friday**
8:00am-2:00pm: Registration
9:00am-5:30pm: Parallel Workshops *(lunch provided)*
Keynote, Tuesday, 9:20am-10:30am

Toward Data-Driven Education

Rakesh Agrawal (Data Insights Laboratories)

Room: Regency Ball Room

Abstract: A program of study can be viewed as a knowledge graph consisting of learning units and relationships between them. Such a knowledge graph provides the core data structure for organizing and navigating learning experiences. We address three issues in this talk. First, how can we synthesize the knowledge graph, given a set of concepts to be covered in the study program. Next, how can we use data mining to identify and correct deficiencies in a knowledge graph. Finally, how can we use data mining to form study groups with the goal of maximizing overall learning. We conclude by pointing out some open research problems.

Bio: Rakesh Agrawal is the President and Founder of the Data Insights Laboratories. He is a member of the National Academy of Engineering, a Fellow of ACM, and a Fellow of IEEE. He has been both an IBM Fellow and a Microsoft Fellow. ACM SIGKDD awarded him its inaugural Innovations Award and ACM SIGMOD the Edgar F. Codd Award. He was named to the Scientific American’s First list of top 50 Scientists. Rakesh has been granted 80+ patents and published 200+ papers, including the 1st and 2nd highest cited in databases and data mining. Four of his papers have received “test-of-time” awards. His research formed the nucleus of IBM Intelligent Miner that led the creation of data mining as a new software category. Besides Intelligent Miner, several other commercial products incorporate his work, including IBM DB2 and WebSphere and Microsoft Bing.
Keynote, Wednesday, 9:00am-10:10am

**Personalized Search: Potential and Pitfalls**

Susan Dumais (Microsoft Research)

**Room:** Regency Ball Room

**Abstract:** Traditionally search engines have returned the same results to everyone who asks the same question. However, using a single ranking for everyone in every context at every point in time limits how well a search engine can do in providing relevant information. In this talk I present a framework to quantify the “potential for personalization” which we use to characterize the extent to which different people have different intents for the same query. I describe several examples of how we represent and use different kinds of contextual features to improve search quality for individuals and groups. Finally, I conclude by highlighting important challenges in developing personalized systems at Web scale including privacy, transparency, serendipity, and evaluation.

**Bio:** Susan Dumais a Distinguished Scientist and Deputy Managing Director of the Microsoft Research Lab in Redmond, and an adjunct professor at the University of Washington. Prior to joining Microsoft, she was at Bell Labs where she worked on Latent Semantic Analysis, techniques for combining search and browsing, and organizational impacts of new technology. Her current research focuses on user modeling and personalization, context and search, and temporal dynamics of information. She has worked closely with several Microsoft groups (Bing, Windows Desktop Search, SharePoint, and Office Online Help) on search-related innovations. Susan has published widely in the fields of information science, human-computer interaction and cognitive science, and holds several patents on novel retrieval algorithms and interfaces. She is Past-Chair of ACM’s Special Interest Group in Information Retrieval (SIGIR), and serves on editorial boards, technical program committees, and government panels. She was elected to the CHI Academy in 2005, an ACM Fellow in 2006, received the SIGIR Gerard Salton Award for Lifetime Achievement in 2009, was elected to the National Academy of Engineering (NAE) in 2011, received the ACM Athena Lecturer and Tony Kent Strix Awards in 2014, was elected to the American Academy of Arts and Sciences (AAAS) in 2015, and received the Lifetime Achievement Award from Indiana University Department of Psychological and Brain Science in 2016.
Keynote, Thursday, 9:00am-10:10am

A Personal Perspective and Retrospective on Web Search Technology

Andrei Broder (Google Research)

Room: Regency Ball Room

Abstract: This talk is a review of some Web research and predictions that I co-authored over the last two decades: both what turned out gratifyingly right and what turned out embarrassingly wrong. Topics will include near-duplicates, the Web graph, query intent, inverted indices efficiency, and others. While this seems a completely idiosyncratic collection there are in fact concealed connections that offer good clues to the big question: what will happen next?

Bio: Broder is a Distinguished Scientist at Google where he leads a multidisciplinary research team located across three continents. From 2005 to 2012 he was a Fellow and VP for Computational Advertising at Yahoo. Previous positions include Distinguished Engineer at IBM and VP for Research and Chief Scientist at AltaVista. He was graduated Summa cum Laude from Technion and obtained his M.Sc. and Ph.D. in Computer Science at Stanford under Don Knuth. Broder has authored more than a hundred papers and was awarded fifty US patents. His current research interests are focused on user understanding, computational advertising, context-driven information supply, and randomized algorithms. He is a member of the US National Academy of Engineering and a Fellow of ACM and of IEEE. Other honors include the ACM Paris Kanellakis Theory and Practice Award and a doctorate Honoris Causa from Technion.
Industry Track Keynote, Tuesday, 11:00am

Building Industry-specific Knowledge Bases
Shivakumar Vaithyanathan (IBM Research)

Room: Network

Abstract: Building industry-specific knowledge bases relies heavily on collecting and representing domain knowledge over time. Domain knowledge includes: (1) the logical schema, constraints and domain vocabulary of the application, (2) the models and algorithms to populate instances of that schema, and (3) the data necessary to build and maintain those models and algorithms. In IBM Watson we are using an ontology-driven approach for the creation and consumption of industry-specific knowledge bases. The creation of such knowledge bases involves well known building blocks: natural language processing, entity resolution, data transformation, etc. It is critical that the models and algorithms that implement these building blocks be transparent and optimizable for efficient execution.

In this talk, I will describe the design of domain-specific languages (DSL) with specialized constructs that serve as target languages for learning these models and algorithms, and the generation of training data for scaling up the learning.

Industry Track Keynote, Tuesday, 2:00pm

Duer: Intelligent Personal Assistant
Haifeng Wang (Baidu, Inc.)

Room: Network

Abstract: Intelligent personal assistant is widely recognized as a more natural and efficient way of human-computer interaction, which has attracted extensive interests from both academia and industry. In this talk, I describe Duer, Baidu's intelligent personal assistant. In particular, I would like to focus on the following three features. Firstly, Duer comprehensively understands people's requirements via multiple channels, including not only explicit utterances, but also user models and rich contexts. Duer's user models are learnt from users' interaction history, and the rich contexts consist of temporal and geographical information, as well as the foregoing dialogues. Secondly, Duer meets diverse requirements with a range of instruments, such as chatting, information provision, reminder service, etc. These instruments are implemented based on mining the big data of web pages, applications, and user logs, which are then seamlessly integrated in the dialogue flow. Thirdly, Duer features multi-modal interaction, which allows people to interact with it by means of
texts, speech, and images. We believe the above features will enable Duer to become a better and distinguished intelligent assistant for each of you.

**Industry Track Keynote, Wednesday, 10:40am**

*Using Machine Learning to Improve the Email Experience*

Marc Najork (Google Research)

**Room:** Network

**Abstract:** Email is an essential communication medium for billions of people, with most users relying on web-based email services. Two recent trends are changing the email experience: smartphones have become the primary tool for accessing online services including email, and machine learning has come of age. Smartphones have a number of compelling properties (they are location-aware, usually with us, and allow us to record and share photos and videos), but they also have a few limitations, notably limited screen size and small and tedious virtual keyboards. Over the past few years, Google researchers and engineers have leveraged machine learning to ameliorate these weaknesses, and in the process created novel experiences. In this talk, I will give three examples of machine learning improving the email experience.

The first example describes how we are improving email search. Displaying the most relevant results as the query is being typed is particularly useful on smartphones due to the aforementioned limitations. Combining hand-crafted and machine-learned rankers is powerful, but training learned rankers requires a relevance-labeled training set. User privacy prohibits us from employing raters to produce relevance labels. Instead, we leverage implicit feedback (namely clicks) provided by the users themselves. Using click logs as training data in a learning-to-rank setting is intriguing, since there is a vast and continuous supply of fresh training data. However, the click stream is biased towards queries that receive more clicks -- e.g. queries for which we already return the best result in the top-ranked position. I will summarize our work on neutralizing that bias.

The second example describes how we extract key information from appointment and reservation emails and surface it at the appropriate time as a reminder on the user's smartphone. Our basic approach is to learn the templates that were used to generate these emails, use these templates to extract key information such as places, dates and times, store the extracted records in a personal information store, and surface them at the right time, taking contextual information such as estimated transit time into account.

The third example describes Smart Reply, a system that offers a set of three short responses to those incoming emails for which a short response is appropriate, allowing users to respond quickly with just a few taps, without typing or involving voice-to-text transcription. The basic approach is to learn a model of likely short responses to original emails from the corpus, and then to apply the model whenever a new message arrives. Other considerations include offering a set of responses that are all appropriate and yet
diverse, and triggering only when sufficiently confident that each responses is of high quality and appropriate.

Industry Track Keynote, Thursday, 10:40am

Large-scale Robust Online Matching and Its Application in E-commerce

Rong Jin (Alibaba)

Room: Network

Abstract: This talk will be focused on large-scale matching problem that aims to find the optimal assignment of tasks to different agents under linear constraints. Large-scale matching has found numerous applications in e-commerce. A well-known example is budget aware online advertisement. A common practice in online advertisement is to find, for each opportunity or user, the advertisements that fit best with his/her interests. The main shortcoming with this greedy approach is that it did not take into account the budget limits set by advertisers. Our studies, as well as others, have shown that by carefully taking into budget limits of individual advertisers, we could significantly improve performance of the advertisement system.

Despite of rich literature, two important issues are often overlooked in the previous studies of matching/assignment problem. The first issue arises from the fact that most quantities used by optimization are estimated based on historical data and therefore are likely to be inaccurate and unreliable. The second challenge is how to perform online matching as in many e-commerce problems, tasks are created in an online fashion and algorithm has to make assignment decision immediately when every task emerges. We refer to these two issues as challenges of “robust matching” and “online matching”.

To address the first challenge, I will introduce two different techniques for robust matching. The first approach is based on the theory of robust optimization that takes into account the uncertainties of estimated quantities when performing optimization. The second approach is based on the theory of two-sided matching whose result only depends on the partial preference of estimated quantities. To deal with the challenge of online matching, I will discuss two online optimization techniques, one based on theory of primal-dual online optimization and one based on minimizing dynamic regret under long term constraints. We verify the effectiveness of all these approaches by applying them to real-world projects developed in Alibaba.
# Conference Schedule: Daily Schedule

**Monday Oct. 24, 2016: Tutorials, Career Day, and Welcome Reception**

<table>
<thead>
<tr>
<th>October 24</th>
<th>Regency A&amp;B</th>
<th>Regency C-F</th>
<th>Concept</th>
<th>Network</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-18:00</td>
<td>Registration (Regency Foyer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:30-11:00</td>
<td>Coffee Break (Regency Foyer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:30-14:00</td>
<td>Lunch Provided (Regency Ballroom)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:30-16:00</td>
<td>Coffee Break (Regency Foyer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18:00-19:30</td>
<td>Welcome Reception (Regency Ballroom)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Tutorials, Oct. 24, 2016, 9am-5:30pm**

1. **“Truth discovery for passive and active crowdsourcing”**  
   **Instructors:** Jing Gao (University at Buffalo), Qi Li (University at Buffalo) and Wei Fan (Baidu).  
   **Location:** Concept

2. **“Similarity search on time series data: past, present and future”**  
   **Instructors:** Abdullah Mueen (University of New Mexico).  
   **Location:** Regency A&B

3. **“Big data science in drug discovery and development”**  
   **Instructors:** Ping Zhang (IBM), Xia Ning (IUPUI) and David Wild (Indiana University).  
   **Location:** Theory

4. **“Graph exploration: taking the user into the loop”**  
   **Instructors:** Davide Mottin (Hasso Plattner Institute), Anja Jentzsch (Hasso Plattner Institute) and Emmanuel Muller (Hasso Plattner Institute).  
   **Location:** Network

5. **“Internet of Things (IoT) big data stream mining”**  
   **Instructors:** Gianmarco De Francisci Morales (QCRI), Albert Bifet (Telecom ParisTech), Joao Gama (University of Porto) and Wei Fan (Baidu).  
   **Location:** Concept

6. **“Large scale distributed data science using Apache Spark 2.0”**  
   **Instructors:** James G. Shanahan (Church and Duncan Group and UC Berkeley) and Liang Dai (NativeX)  
   **Location:** Regency A&B

7. **“Data-driven behavioral analytics: observations, representations and models”**  
   **Instructors:** Meng Jiang (UIUC), Peng Cui (Tsinghua University) and Jiawei Han (UIUC).  
   **Location:** Network

8. **“Learning, Prediction and Optimisation in Real-Time Bidding based Display Advertising”**  
   **Instructors:** Weinan Zhang (Shanghai Jiao Tong University) and Jian Xu (TouchPal).  
   **Location:** Theory

**CIKM Career Day, Oct. 24, 2016, 9am-5:30pm**

**Location:** Regency C,D,E,F  
**Chairs:** Sarath Janga, Snehasis Mukhopadhyay

9:00am-9:15am: Opening Remarks (Regency C&D)  
9:15am-10:30am: Company Presentations (Regency C&D)  
**10:30am-11:00am:** Coffee Break (Regency Foyer)  
11:00am-12:30pm: Poster Presentations by Students (Regency E&F)  
**12:30pm-2:00pm:** Lunch provided (Regency Ballroom)  
2:00pm-3:30pm: Flash Interviews (Regency C&D)  
**3:30pm-4:00pm:** Coffee Break (Regency Foyer)  
4:00pm-5:00pm: Panel Discussions by Industry (Regency C&D)  
5:00pm-5:30pm: Non-Traditional Opportunities (Regency E&F)  
**6:00pm-7:30pm:** Conference Welcome Reception (Regency Ballroom)
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-18:00</td>
<td>Registration (Regency Foyer)</td>
<td>Regency A&amp;B</td>
</tr>
<tr>
<td>09:00-09:20</td>
<td>Conference Welcome (Regency Ballroom)</td>
<td>Regency C</td>
</tr>
<tr>
<td>09:20-10:30</td>
<td>Plenary Keynote Talk: Rakesh Agarwal (Regency Ballroom) (Chair: ChengXiang Zhai)</td>
<td>Regency D</td>
</tr>
<tr>
<td>10:30-11:00</td>
<td>Coffee Break (Regency Foyer)</td>
<td>Concept</td>
</tr>
<tr>
<td>11:00-12:30</td>
<td>Session 1c: Document Classification and Labeling (Chair: Yi Fang)</td>
<td>Theory</td>
</tr>
<tr>
<td></td>
<td>Session 1d: Better Queries (Chair: Davood Rafiei)</td>
<td>Network</td>
</tr>
<tr>
<td></td>
<td>Session 1e: Better Search (Chair: Fabio Crestani)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 1b: Deep Learning Applications (Chair: Elisa Bertino)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 1a: Recommendation (Chair: Fazli Can)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 1f: Industry Session I (Chair: Yunyao Li)</td>
<td></td>
</tr>
<tr>
<td>12:30-14:00</td>
<td>Lunch On Your Own</td>
<td></td>
</tr>
<tr>
<td>14:00-15:30</td>
<td>Session 2d: Clustering (Chair: Xia Ning)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 2e: Understanding Text (Chair: Zhiguo Gong)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 2a: Learning to Rank (Chair: Elisa Bertino)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 2c: Wikipedia (Chair: Fazli Can)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 2b: Question Answering (Chair: Jing Gao)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 2f: Industry Session II (Chair: Parikshit Sondhi)</td>
<td></td>
</tr>
<tr>
<td>15:30-16:00</td>
<td>Coffee Break (Regency Foyer)</td>
<td></td>
</tr>
<tr>
<td>16:00-17:30</td>
<td>Session 3a: Graph Analytics (Chair: Mohammad Hasan)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 3b: Event Detection and Analytics (Chair: Fazli Can)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 3e: Social Networks---Links and Trust (Chair: James Caverlee)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 3d: Mobile (Chair: Fabio Crestani)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 3c: Crowdsourcing (Chair: Kyumin Lee)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Session 3f: Industry Session III (Chair: Parikshit Sondhi)</td>
<td></td>
</tr>
<tr>
<td>17:30-19:00</td>
<td>Poster Session I (Regency E&amp;F) + Demo (Starts at 6pm, Network)</td>
<td></td>
</tr>
</tbody>
</table>
Tuesday Morning Sessions, Oct 25, 2016

Session 1a: Recommendation  
Time: 11:00 AM - 12:30 PM  
Location: Theory  
Chair: Fazli Can

Paper 6: Social Recommendation with Strong and Weak Ties., Xin Wang(Zhejiang University, China), Wei Lu(University of British Columbia, Canada), Martin Ester(Simon Fraser University, Canada), Can Wang(Zhejiang University, China), Chun Chen(Zhejiang University, China)

Paper 55: Learning Graph-based POI Embedding for Location-based Recommendation, Min Xie(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, China), Hongzhi Yin(The University of Queensland, School of Information Technology and Electrical Engineering, Australia), Hao Wang(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, China), Fanjiang Xu(State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, China, China), Weitong Chen(The University of Queensland, School of Information Technology and Electrical Engineering, Australia), Sen Wang(School of Information and Communication Technology, Griffith University, Australia)

Paper 281: Improving Personalized Trip Recommendation by Avoiding Crowds, Xiaoting Wang(University of Melbourne, Australia), Christopher Leckie(University of Melbourne, Australia), Jeffrey Chan(RMIT University, Australia), Kwan Hui Lim(University of Melbourne, Australia), Tharshan Vaithianathan(Chisholm Institute, Australia)

Paper 949: Memory-based Recommendations of Entities for Web Search Users, Ignacio Fernández-Tobías(University Autónoma de Madrid, Spain), Roi Blanco(University of A Coruña, Spain)

Session 1b: Deep Learning Application  
Time: 11:00 AM - 12:30 PM  
Location: Concept  
Chair: Elisa Bertino

Paper 278: LICON: A Linear Weighting Scheme for the Contribution of Input Variables in Deep Artificial Neural Networks, Gjergji Kasneci(SCHUFA Holding AG, Germany), Thomas Gottron(SCHUFA Holding AG, Germany)

Paper 417: A Deep Relevance Matching Model for Ad-hoc Retrieval, Jiafeng Guo(CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences, China), Yixing Fan(CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences, China), Qingyao Ai(Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA), W. Bruce Croft(Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA)

Paper 587: A Neural Network Approach to Quote Recommendation in Writings, Jiwei Tan(Peking University, China), Xiaojun Wan(Peking University, China), Jianguo Xiao(Peking University, China)

Paper 754: Retweet Prediction with Attention-based Deep Neural Network, Qi Zhang(Fudan University, China), Yeyun Gong(Fudan University, China), Jindou Wu(Fudan University, China), Haoran Huang(Fudan University, China), Xuanjing Huang(Fudan University, China)

Session 1c: Document Classification and labeling  
Time: 11:00 AM - 12:30 PM  
Location: Regency A&B  
Chair: Yi Yang

Paper 129: Effective Document Labeling with Very Few Seed Words: A Topic Model Approach, Chenliang Li(Wuhan University, China), Jian Xing(Wuhan University, China), Aixin Sun(Nanyang Technological University, Singapore), Zongyang Ma(Nanyang Technological University, Singapore)

Paper 217: Cross-lingual Text Classification via Model Translation with Limited Dictionaries, Ruochen Xu(Carnegie Mellon University, USA), Yiming Yang(Carnegie Mellon University, USA), Hanxiao Liu(Carnegie Mellon University, USA), Andrew Hsi(Carnegie Mellon University, USA)

Paper 308: Semi-supervised Multi-Label Topic Models for Document Classification and Sentence Labeling, Hossein Soleimani(Pennsylvania State University, USA), David Miller(Pennsylvania State University, USA)

Paper 327: Linked Document Embedding for Classification, Suhang Wang(Computer Science & Engineering, USA), Jiliang Tang(Michigan State University, USA), Charu Aggarwal(IBM T.J. Watson, USA), Huan Liu(Arizona State University, USA)
Session 1d: Better Query
Time: 11:00 AM - 12:30 PM  Location: Regency C  Chair: Davood Rafiei

Paper 46: Detecting Promotion Campaigns in Query Auto Completion, Yuli LIU(Tsinghua University, China), Yiqun LIU(Tsinghua University, China), Ke Zhou(Yahoo! Research, England UK), Min Zhang(Tsinghua University, China), Shaoping Ma(Tsinghua University, China), Yue Yin(Samsung R&D Institute China, China), Hengliang LUO(Samsung R&D Institute China, China)

Paper 298: A Unified Index for Spatio-Temporal Keyword Queries, Tuan-Anh Hoang-Vu(New York University, USA), Huy Vo(The City College of New York, USA), Juliana Freire(New York University, USA)

Paper 669: Privacy-Preserving Reachability Query Services for Massive Networks, Jiaxin Jiang(Department of Computer Science Hong Kong Baptist University, Hong Kong), Peipei Yi(Department of Computer Science Hong Kong Baptist University, Hong Kong), Byron Choi(Department of Computer Science Hong Kong Baptist University, Hong Kong), Zhiwei Zhang(Department of Computer Science Hong Kong Baptist University, Hong Kong), Xiaohui Yu(School of Computer Science and Technology Shandong University, China)

Paper 1275: Sequential Query Expansion using Concept Graph, Saeid Balaneshin-kordan(Wayne State University, USA), Alexander Kotov(Wayne State University, USA)

Session 1e: Better Search
Time: 11:00 AM - 12:30 PM  Location: Regency D  Chair: Fabio Crestani

Paper 13: Learning Latent Vector Spaces for Product Search, Christophe Van Gysel(University of Amsterdam, Netherlands), Maarten de Rijke(University of Amsterdam, Netherlands), Evangelos Kanoulas(University of Amsterdam, Netherlands)

Paper 980: Incorporating Clicks, Attention and Satisfaction into a Search Engine Result Page Evaluation Model, Aleksandr Chuklin(Google Research Europe & University of Amsterdam, Switzerland), Maarten de Rijke(University of Amsterdam, Netherlands)

Paper 1102: The Role of Relevance in Sponsored Search, Luca Aiello(Bell Labs, United Kingdom), Ioannis Arapakis(Eurecat, Spain), Ricardo Baeza-Yates(NTENT, USA), Xiao Bai(Yahoo Research, USA), Nicola Barbieri(Tumblr, USA), Amin Mantrach(Yahoo Research, USA), Fabrizio Silvestri(Facebook, United Kingdom)

Paper 63: PowerWalk: Scalable Personalized PageRank via Random Walks with Vertex-Centric Decomposition, Qin Liu(The Chinese University of Hong Kong, China), Zhenguo Li(Huawei Noah's Ark Lab, China), John Lui(The Chinese University of Hong Kong, China), Jiefeng Cheng(Huawei Noah's Ark Lab, China)

Session 1f: Industry Session I
Time: 11:00 AM - 12:30 PM  Location: Network  Chair: Yunyao Li

Keynote: Building Industry-specific Knowledge Bases, Shivakumar Vaithyanathan (IBM Research)

Paper 1476: Reuters Tracer: A Large Scale System of Detecting & Verifying Real-Time News Events from Twitter, Xiaomo Liu(Thomson Reuters, USA), Quanzhi Li(Thomson Reuters, USA), Armineh Nourbakhsh(Thomson Reuters, USA), Rui Fang(Thomson Reuters, USA), Merine Thomas(Thomson Reuters, USA), Kajsa Anderson(Thomson Reuters, USA), Russ Kociuba(Thomson Reuters, USA), Mark Vedder(Thomson Reuters, USA), Steven Pomerville(Thomson Reuters, USA), Ramdev Wudali(Thomson Reuters, USA), Robert Martin(Thomson Reuters, USA), John Duprey(Thomson Reuters, USA), Arun Vachher(Thomson Reuters, USA), William Keenan(Thomson Reuters, USA), Sameena Shah(Thomson Reuters, USA)

Paper 1366: Structural Clustering of Machine-Generated Mail, Noa Avigdor-Elgrabli(Yahoo research, Israel), Mark Cwalinski(Yahoo inc., USA), Dotan Di Castro(Yahoo research, Israel), Ittah Gamzu(Yahoo Research, Israel), Irena Grabovitch-Zuev(Yahoo Research, Israel), Liane Lewin-Eytan(Yahoo Research, Israel), Yoelle Maarek(Yahoo Research, Israel)
Tuesday Afternoon Sessions, Oct 25, 2016

Session 2a: Learning to Rank
Time: 2:00 PM - 3:30 PM  Location: Regency D  Chair: Elisa Bertino

Paper 340: LambdaFM: Learning Optimal Ranking with Factorization Machines Using Lambda Surrogates, Fajie Yuan (University of Glasgow, United Kingdom), Guibing Guo (Northeastern University, China), Joemon Jose (University of Glasgow, United Kingdom), Long Chen (University of Glasgow, United Kingdom), Haitao Yu (yuhaitao@slis.tsukuba.ac.jp, Japan), Weinan Zhang (Shanghai Jiao Tong University, China)

Paper 382: Plackett-Luce Regression Mixture Model for Heterogeneous Rankings, Maksim Tkachenko (Singapore Management University, Singapore), Hady Lauw (Singapore Management University, Singapore)

Paper 826: Compression-Based Selective Sampling for Learning to Rank, Rodrigo Silva (Federal University of Minas Gerais, Brazil), Guilherme Gomes (Federal University of Minas Gerais, Brazil), Mário Alvim (Federal University of Minas Gerais, Brazil), Marcos Gonçalves (Federal University of Minas Gerais, Brazil)

Paper 606: Incorporating Risk-Sensitiveness into Feature Selection for Learning to Rank, Daniel Sousa (Universidade Federal de Minas Gerais, Brazil), Sérgio Canuto (Universidade Federal de Minas Gerais, Brazil), Thierson Rosa (Universidade Federal de Goiás, Brazil), Wellington Martins (Universidade Federal de Goiás, Brazil), Marcos Gonçalves (Universidade Federal de Minas Gerais, Brazil)

Session 2b: Question Answering
Time: 2:00 PM - 3:30 PM  Location: Theory  Chair: Jing Gao

Paper 442: Answering Twitter Questions: a Model for Recommending Answerers through Social Collaboration, Laure Soulier (Sorbonne Univ. UPMC LIP6, France), Lynda Tamine (Université de Toulouse, UPS - IRIT, France), Gia-Hung Nguyen (Université de Toulouse, UPS - IRIT, France)

Paper 476: Learning to Extract Conditional Knowledge for Question Answering using Dialogue, Pengwei Wang (South China University of Technology, China), Lei Ji (Microsoft Research Asia, China), Jun Yan (Microsoft Research Asia, China), Lianwen Jin (South China University of Technology, China), Wei-Ying Ma (Microsoft Research Asia, China)

Paper 904: aNMM: Ranking Short Answer Texts with Attention-Based Neural Matching Model, Liu Yang (University of Massachusetts Amherst, USA), Qingyao Ai (University of Massachusetts Amherst, USA), Jiafeng Guo (Chinese Academy of Sciences, China), W. Bruce Croft (University of Massachusetts Amherst, USA)

Paper 905: Medical Question Answering for Clinical Decision Support, Travis Goodwin (University of Texas at Dallas, USA), Sandra Harabagiu (University of Texas at Dallas, USA)

Session 2c: Wikipedia
Time: 2:00 PM - 3:30 PM  Location: Concept  Chair: Fazil Can

Paper 26: Error Link Detection and Correction in Wikipedia, Chengyu Wang (East China Normal University, China), Rong Zhang (East China Normal University, China), Xiaofeng He (East China Normal University, China), Aoying Zhou (East China Normal University, China)

Paper 162: Using Prerequisites to Extract Concept Maps from Textbooks, Shuting Wang (Pennsylvania State University, USA), Qiuying Ao (University of Massachusetts Amherst, USA), Jiao Guo (Chinese Academy of Sciences, China), W. Bruce Croft (University of Massachusetts Amherst, USA), Kaye Williams (Pennsylvania State University, USA), Jie Li (Pennsylvania State University, USA), Bart Purcell (Pennsylvania State University, USA), Lee Giles (Pennsylvania State University, USA)

Paper 257: Vandalism Detection in Wikidata, Stefan Heindorf (Paderborn University, Germany), Martin Potthast (Bauhaus-Universität Weimar, Germany), Benno Stein (Bauhaus-Universität Weimar, Germany), Gregor Engels (Paderborn University, Germany)

Paper 752: Finding News Citations for Wikipedia, Besnik Fetahu (L3S Research Center, Leibniz University of Hannover, Germany), Katja Markert (Institute of Computational Linguistics, Heidelberg University, Germany), Wolfgang Nejdl (L3S Research Center, Leibniz University of Hannover, Germany), Avishek Anand (L3S Research Center, Leibniz University of Hannover, Germany)
Session 2d: Clustering
Time: 2:00 PM - 3:30 PM  
Location: Regency A&B  
Chair: Xia Ning

Paper 37: SemiNMF-PCA framework for Sparse Data Co-clustering, Kais Allab(University of Paris Descartes, France), Lazhar Labiod(University of Paris Descartes, France), Mohamed Nadif(University of Paris Descartes, France)

Paper 43: Effective and Efficient Spectral Clustering on Text and Link Data, Zhiqiang Xu(Institute for Infocomm Research, Singapore), Yiping Ke(Nanyang Technological University, Singapore)

Paper 275: Robust Spectral Ensemble Clustering, Zhiqiang Tao(Northeastern University, USA), Hongfu Liu(Northeastern University, USA), Sheng Li(Northeastern University, USA), Yun Fu(Northeastern University, USA)

Paper 225: Hybrid Indexing for Versioned Document Search with Cluster-based Retrieval, Xin Jin(University of California, Santa Barbara, USA), Daniel Agun(University of California, Santa Barbara, USA), Tao Yang(University of California, Santa Barbara, USA), Qinghao Wu(University of California, Santa Barbara, USA), Yifan Shen(University of California, Santa Barbara, USA), Susen Zhao(University of California, Santa Barbara, USA)

Session 2e: Understand Text
Time: 2:00 PM - 3:30 PM  
Location: Regency C  
Chair: Zhiguo Gong

Paper 52: Time-aware Multi-Viewpoint Summarization of Multilingual Social Text Streams, Zhaochun Ren(University of Amsterdam, Netherlands), Oana Inel(VU University Amsterdam, Netherlands), Lora Aroyo(VU University Amsterdam, Netherlands), Maarten de Rijke(University of Amsterdam, Netherlands)

Paper 249: Data Summarization with Social Contexts, Hao Zhuang(École polytechnique fédérale de Lausanne (EPFL), Switzerland), Rameez Rahman(École polytechnique fédérale de Lausanne (EPFL), Switzerland), Xia Hu(Texas A&M University, USA), Tian Guo(École polytechnique fédérale de Lausanne (EPFL), Switzerland), Pan Hui(Hong Kong University of Science and Technology, China), Karl Aberer(École polytechnique fédérale de Lausanne (EPFL), Switzerland)

Paper 404: Understanding Sparse Topical Structure of Short Text via Stochastic Variational-Gibbs Inference, Tianyi Lin(The Chinese University of Hong Kong, Hong Kong), Siyuan Zhang(The Chinese University of Hong Kong, Hong Kong), Hong Cheng(The Chinese University of Hong Kong, Hong Kong)

Paper 1190: Annotating Points of Interest with Geo-tagged Tweets, Kaiqi Zhao(Nanyang Technological University, Singapore), Gao Cong(Nanyang Technological University, Singapore), Aixin Sun(Nanyang Technological University, Singapore)

Session 2f: Industry Session II
Time: 2:00 PM - 3:30 PM  
Location: Network  
Chair: Parikshit Sondhi

Keynote: Duer: Intelligent Personal Assistant, Haifeng Wang (Baidu Inc.)

Paper 1392: Measuring Metrics, Pavel Dmitriev(Microsoft Corporation, USA), Xian Wu(Microsoft Corporation, China)

Paper 1318: City-Scale Localization with Telco Big Data, Hong Su(Oracle, USA), Mohamed Zait(Oracle, USA), Vladimir Barrière(Oracle, USA), Joseph Torres(oracle, USA), Andre Menck(oracle, USA)

Session 3a: Graph Analytics
Time: 4:00 PM - 5:30 PM  
Location: Regency A&B  
Chair: Mohammad Hasan

Paper 405: Approximating Graph Pattern Queries Using Views, Jia Li(Beihang University, China), Yang Cao(University of Edinburgh & Beihang University, United Kingdom), Xudong Liu(Beihang University, China)

Paper 419: Group-Aware Weighted Bipartite B-Matching, Cheng Chen(University of Victoria, Canada), Kui Wu(University of Victoria, Canada), Alex Thomo(University of Victoria, Canada)

Paper 970: Growing Graphs from Hyperedge Replacement Graph Grammars, Salvador Aguiñaga(University of Notre Dame, USA), Rodrigo Palacios(California State University at Fresno, USA), David Chiang(University of Notre Dame, USA), Tim Weninger(University of Notre Dame, USA)
Session 3b: Event Detection & Analytics  
**Time:** 4:00 PM - 5:30 PM  
**Location:** Regency C  
**Chair:** Fazil Can

Paper 273: Graph Topic Scan Statistic for Spatial Event Detection, Yu Liu(The University of Hong Kong, Hong Kong), Baojian Zhou(University at Albany-SUNY, USA), Feng Chen(University at Albany-SUNY, USA), David W. Cheung(The University of Macau, Macao), Zhiguo Gong(University of Macau, Macao)

Paper 927: A Multiple Instance Learning Framework for Identifying Key Sentences and Detecting Events, Wei Wang(Virginia Tech, USA), Yue Ning(Virginia Tech, USA), Huzefa Rangwala(George Mason University, USA), Naren Ramakrishnan(Virginia Tech, USA)

Paper 1047: PairFac: Event Analytics through Discriminant Tensor Factorization, Xidao Wen(University of Pittsburgh, USA), Yu-Ru Lin(University of Pittsburgh, USA), Konstantinos Pelechrinis(University of Pittsburgh, USA)

Session 3c: Crowdsourcing  
**Time:** 4:00 PM - 5:30 PM  
**Location:** Theory  
**Chair:** Kyumin Lee

Paper 106: Active Content-Based Crowdsourcing Task Selection, Piyush Bansal(International Institute of Information Technology Hyderabad, India), Carsten Eickhoff(ETH Zurich, Switzerland), Thomas Hofmann(ETH Zurich, Switzerland)

Paper 982: CrowdSelect: Increasing Accuracy of Crowdsourcing Tasks through Behavior Prediction and User Selection, Chenxi Qiu(Pennsylvania State University, USA), Anna Squicciarini(Pennsylvania State University, USA), Barbara Carminati(Universiy of Insubria, Italy), James Caverlee(Texas A&M University, USA), Dev Kharai(Pennsylvania State University, USA)

Paper 985: Attribute-based Crowd Entity Resolution, Asif Khan(Stanford University, USA), Hector Garcia-Molina(Stanford University, USA)

Paper 336: Efficient Processing of Location-Aware Group Preference Queries, Miao Li(Northeastern University, China), Lisi Chen(Hong Kong Baptist University, China), Gao Cong(Nanyang Technological University, Singapore), Yu Gu(Northeastern University, China), Ge Yu(Northeastern University, China)

Session 3d: Mobile  
**Time:** 4:00 PM - 5:30 PM  
**Location:** Concept  
**Chair:** Fabio Crestani

Paper 691: Mining Shopping Patterns for Divergent Urban Regions by Incorporating Mobility Data, Tianran Hu(University of Rochester, USA), Ruihua Song(Microsoft Research, China), Yingzi Wang(University of Science and Technology of China, China), Xing Xie(Microsoft Research, China), Jiebo Luo(University of Rochester, USA)

Paper 1139: Large-Scale Analysis of Viewing Behavior: Towards Measuring Satisfaction with Mobile Proactive Systems, Qi Guo(Google, USA), Yang Song(Google, USA)

Paper 1132: Where Did You Go: Personalized Annotation of Mobility Records, Fei Wu(Penn state, USA), Zhenhui Li(Penn State, USA)

Paper 1234: Understanding Mobile Searcher Attention with Rich Ad Formats, Dmitry Lagun(Google, USA), Donal McMahon(Indeed, USA), Vidhya Navalpakkam(Google, USA)

Session 3e: Social Network – Link & Trust  
**Time:** 4:00 PM - 5:30 PM  
**Location:** Regency D  
**Chair:** James Caverlee

Paper 134: Link Prediction in Heterogeneous Social Networks, Sumit Negi(Indian Institute of Technology, Delhi, India), Santanu Chaudhury(Indian Institute of Technology, Delhi, India)

Paper 705: Who are My Familiar Strangers? Revealing Hidden Friend Relations and Common Interests from Smart Card Data, Fusang Zhang(Institute of Software, Chinese Academy of Sciences; University of Chinese Academy of Sciences, China)
Session 3f: Industry Session III
Time: 4:00 PM - 5:30 PM  
Location: Network  
Chair: Parikshit Sondhi

Paper 1427: "Shall I Be Your Chat Companion" Towards an Online Human-Computer Conversation System, Rui Yan(Peking University, China), Yiping Song(Peking University, China), Xiangyang Zhou(Baidu Inc., China), Hua Wu(Baidu Inc., China)

Paper 1349: To Click or Not To Click: Automatic Selection of Beautiful Thumbnails from Videos, Yale Song(Yahoo Research, USA), Miriam Redi(Yahoo Research, United Kingdom), Jordi Vallmitjana(Yahoo Research, USA), Alejandro Jaimes(Yahoo Research, USA)

Paper 1333: User Response Learning for Directly Optimizing Campaign Performance in Display Advertising, Kan Ren(Shanghai Jiao Tong University, China), Weinan Zhang(Shanghai Jiao Tong University, China), Yong Yu(Shanghai Jiao Tong University, China), Jun Wang(University College London, United Kingdom)

Poster Session I: Short Papers
Time: 5:30 PM-7:00 PM  
Location: Regency E&F

SP1. Skipping Word: A Character-Sequential Representation based Framework for Question Answering. Lingxun Meng (Sogou Inc.); Yan Li, Mengyi Liu (Key Lab of Intelligent Information Processing of CAS, Institute of Computing Technology); Peng Shu (Sogou Inc.)

SP2. Towards Time-Discounted Influence Maximization. Arijit Khan (Nanyang Technological University)

SP3. Quantifying Query Ambiguity with Topic Distributions. Yuki Yano, Yukihiro Tagami, Akira Tajima (Yahoo Japan Corporation)

SP4. ASNets : A Benchmark Dataset of Aligned Social Networks for Cross-Platform User Modeling. Xuezhi Cao, Yong Yu (Shanghai Jiao Tong University)

SP5. Data Locality in Graph Engines: Implications and Preliminary Experimental Results. Yong-Yeon Jo, Jiwon Hong, Myung-Hwan Jang, Jae-Geun Bang, Sang-Wook Kim (Hanyang University)

SP6. Active Zero-Shot Learning. Sihong Xie (Lehigh University); Shaoxiong Wang (Tsinghua University); Philip S. Yu (University of Illinois at Chicago)

SP7. Learning to Account for Good Abandonment in Search Success Metrics. Madian Khabsa, Aidan Crook, Ahmed Hassan Awadallah, Imed Zitouni, Tasos Anastasakos (Microsoft); Kyle Williams (The Pennsylvania State University)


SP9. Incorporate Group Information to Enhance Network Embedding. Jifan Chen, Qi Zhang, Xuanjing Huang (Fudan University)

SP10. Exploiting Cluster-based Meta Paths for Link Prediction in Signed Networks. Jiangfeng Zeng, Ke Zhou, Xiao Ma, Fuhao Zou, Hua Wang (Huazhong University of Science and Technology)

SP11. Predicting Importance of Historical Persons using Wikipedia. Adam Jatowt, Daisuke Kawai, Katsumi Tanaka (Kyoto University)
SP12. Noise-Contrastive Estimation for Answer Selection with Deep Neural Networks. Jinfeng Rao, Hua He (University of Maryland College Park); Jimmy Lin (University of Waterloo)

SP13. Global and Local Influence-based Social Recommendation. Qinze Zhang, Jia Wu (University of Technology Sydney); Hong Yang (MathWorks, Beijing); Weixue Lu (Chinese Academy of Sciences); Guodong Long, Chengqi Zhang (University of Technology, Sydney)

SP14. Tag-Aware Personalized Recommendation Using a Deep-Semantic Similarity Model with Negative Sampling. Zhenhua Xu (University of Oxford); Cheng Chen (Beijing University of Posts and Telecommunications); Thomas Lukasiewicz, Yishu Miao (University of Oxford); Xiangwu Meng (Beijing University of Posts and Telecommunications)

SP15. Personalized Semantic Word Vectors. Javid Ebrahimi, Dejing Dou (University of Oregon)

SP16. Query Expansion Using Word Embeddings. Saar Kuzi, Anna Shtok, Oren Kurland (Technion --- Israel Institute of Technology)

SP17. Efficient Distributed Regular Path Queries on RDF Graphs Using Partial Evaluation. Xin Wang (Tianjin University); Junhu Wang (Griffith University); Xiaowang Zhang (Tianjin University)

SP18. Webpage Depth-level Dwell Time Prediction. Chong Wang (New Jersey Institute of Technology); Achir Kalra (Forbes Inc.); Cristian Borcea, Yi Chen (New Jersey Institute of Technology)

SP19. Collaborative Social Group Influence for Event Recommendation. Li Gao (Chinese Academy of Sciences); Jia Wu (University of Technology Sydney); Zhi Qiao (IBM Research - China); Chuan Zhou (Chinese Academy of Sciences); Hong Yang (MathWorks); Yue Hu (Chinese Academy of Sciences)

SP20. Graph-Based Multi-Modality Learning for Clinical Decision Support. Ziwei Zheng, Xiaojun Wan (Peking University)

SP21. Where are You Tweeting? A Context and User Movement Based Approach. Zhi Liu, Yan Huang (University of North Texas)

SP22. Ensemble Learned Vaccination Uptake Prediction using Web Search Queries. Niels Dalum Hansen (University of Copenhagen / IBM Denmark); Christina Lioma (University of Copenhagen); Kåre Mølbak (Statens Serum Institut)

SP23. Location-aware Friend Recommendation in Event-based Social Networks: A Bayesian Latent Factor Approach. Yao Lu (Chinese Academy of Sciences); Zhi Qiao (IBM Research - China); Chuan Zhou, Yue Hu, Li Guo (Chinese Academy of Sciences)

SP24. Extracting Skill Endorsements from Personal Communication Data. Darshan M. Shankaralingappa (Aalto University / Helsinki Institute for Information Technology); Gianmarco De Francisci Morales (Qatar Computing Research Institute); Aristides Gionis (Aalto University / Helsinki Institute for Information Technology)

SP25. A Self-Organizing Map for Identifying Influential Communities in Speech-based Networks. Sameen Mansha, Faisal Kamiran (Information Technology University of Punjab); Asim Karim (Lahore University of Management Sciences); Aizaz Anwar (Information Technology University of Punjab)

SP26. Crowdsourcing-based Urban Anomaly Prediction System for Smart Cities. Chao Huang, Xian Wu, Dong Wang (University of Notre Dame)

SP27. Near Real-time Geolocation Prediction in Twitter Streams via Matrix Factorization Based Regression. Nghia Duong-Trung, Nicolas Schilling, Lars Schmidt-Thieme (University of Hildesheim)


SP30. Improving Search Results with Prior Similar Queries. Yashar Moshfeghi, Kristiyan Velinov, Peter Triantafillou (University of Glasgow)

SP31. The Solitude of Relevant Documents in the Pool. Aldo Lipani, Mihai Lupu (Vienna University of Technology); Evangelos Kanoulas (University of Amsterdam); Allan Hanbury (Vienna University of Technology)

SP32. Scarcity Feature Topic Mining for Video Recommendation. Wei Lu, Fu-lai Chung (Hong Kong Polytechnic University); Kunfeng Lai (Tencent)

SP33. Learning to Re-Rank Questions in Community Question Answering Using Advanced Features. Giovanni Da San Martino, Alberto Barrón Cedeño, Salvatore Romeo (Hamad bin Khalifa University); Antonio Uva (University of Trento); Alessandro Moschitti (Hamad bin Khalifa University)

SP34. Learning to Rank System Configurations. Romain Deveaud, Josiane Mothe (Université de Toulouse); Jian-Yun Nia (Université de Montréal)

SP35. Adaptive Distributional Extensions to DFR Ranking. Casper Petersen, Jakob Grue Simonsen (University of Copenhagen); Kalervo Järvelin (University of Tampere); Christina Lioma (University of Copenhagen)
SP36. CyberRank- Knowledge Elicitation for Risk Assessment of Database Security. Hagit Grushka - Cohen (Ben-Gurion University of the Negev); Oded Sofer (IBM Security Division); Ofer Biller (IBM Cyber Security Center of Excellence, Beer Sheva); Bracha Shapira, Lior Rokach (Ben-Gurion University of the Negev)

SP37. Online Food Recipe Title Semantics: Combining Nutrient Facts and Topics. Tomasz Kusmierczyk, Kjetil Nørvåg (Norwegian University of Science and Technology)


SP39. Forecasting Seasonal Time Series Using Weighted Gradient RBF Network Based Autoregressive Model. Wenjie Ruan, Quan Z. Sheng (The University of Adelaide); Peipei Xu (UESTC); Nguyen Khoi Tran, Nickolas J.G. Falkner (The University of Adelaide); Xue Li (The University of Queensland); Wei Emma Zhang (The University of Adelaide)

SP40. When Sensor Meets Tensor: Filling Missing Sensor Values Through a Tensor Approach. Wenjie Ruan (The University of Adelaide); Peipei Xu (UESTC); Quan Z. Sheng, Nguyen Khoi Tran, Nickolas J.G. Falkner (The University of Adelaide); Xue Li (The University of Queensland); Wei Emma Zhang (The University of Adelaide)

SP41. PEQ: An Explainable, Specification-based, Aspect-oriented Product Comparator for E-commerce. Abhishek Sikchi, Pawan Goyal (Indian Institute of Technology Kharagpur); Samik Datta (Flipkart)

SP42. Forecasting Geo-sensor Data with Participatory Sensing Based on Dropout Neural Network. Jyun-Yu Jiang (University of California, Los Angeles); Cheng-Tie Li (National Cheng Kung University)

SP43. Iterative Search using Query Aspects. Manmeet Singh, W. Bruce Croft (University of Massachusetts)

SP44. A Preference Approach to Reputation in Sponsored Search. Aritra Ghosh, Dinesh Gaurav, Rahul Agrawal (Microsoft India)

SP45. Clustering Speed in Multi-lane Traffic Networks. Bing Zhang, Goce Trajcevski, Feiyiye Li (Northwestern University)

SP46. Learning to Rank Non-Factoid Answers: Comment Selection in Web Forums. Kateryna Tymoshenko, Daniele Bonadiman (University of Trento); Alessandro Moschitti (Qatar Computing Research Institute, HBKU)

SP47. A Theoretical Framework on the Ideal Number of Classifiers for Online Ensembles in Data Streams. Hamed R. Bonab, Fazli Can (Bilkent University)

SP48. User Modeling on Twitter with WordNet Synsets and DBpedia Concepts for Personalized Recommendations. Guangyuan Piao, John G. Breslin (Insight Centre for Data Analytics, National University of Ireland, Galway)

SP49. Improving Entity Ranking for Keyword Queries. John Foley, Brendan O'Connor, James Allan (University of Massachusetts Amherst)

SP50. The Healing Power of Poison: Helpful Non-relevant Documents in Feedback. Mostafa Dehghani, Samira Abnar, Jaap Kamps (University of Amsterdam)

SP51. Probabilistic Approaches to Controversy Detection. Myungha Jang, John Foley, Shiri Dori-Hacohen, James Allan (University of Massachusetts Amherst / CIIR)

SP52. Evaluating Document Retrieval Methods for Resource Selection in Clustered P2P IR. Rami Suleiman Alkhawaldeh, Joemon M. Jose (University of Glasgow); Deepak P. (Queen's University)

SP53. Detecting and Ranking Conceptual Links between Texts Using a Knowledge Base. Martin Tutek, Goran Glavas, Jan Šnajder (Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia); Nataša Mile-Frayling (School of Computer Science, University of Nottingham); Bojana Dalbelo Basic (Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia)

SP54. DePP: A System for Detecting Pages to Protect in Wikipedia. Kelsey Suyehira, Francesca Spezzano (Boise State University)

SP55. Hashtag Recommendation Based on Topic Enhanced Embedding, Tweet Entity Data and Learning to Rank. Quanzhi Li, Sameena Shah, Armineh Nourbakhsh, Xiaomo Liu, Rui Fang (Thomson Reuters)
Industry Track Demo Papers (all demos will be presented twice)

Session 1: 6:00 PM - 7:00 PM, Tuesday, Oct. 25, 2016

Session 2: 5:30 PM – 7:00 PM, Wednesday, Oct. 26 2016

Location: Network

D1. XKnowSearch! Exploiting Knowledge Bases for Entity-based Cross-lingual Information Retrieval. Lei Zhang, Michael Färber, Achim Rettinger (Karlsruhe Institute of Technology (KIT))


D3. PARC: Privacy-Aware Data Cleaning. Dejun Huang, Dhruv Gairola, Yu Huang, Zheng Zheng, Fei Chiang (McMaster University)


D5. FIN10K: A Web-based Information System for Financial Report Analysis and Visualization. Yu-Wen Liu (National Chengchi University); Liang-Chih Liu (National Chiao Tung University); Chuan-Ju Wang (Academia Sinica); Ming-Feng Tsai (National Chengchi University)


D7. Deola: A System for Linking Author Entities in Web Document with DBLP. Yinan Liu, Wei Shen, Xiaojie Yuan (Nankai University)

D8. ConHub: A Metadata Management System for Docker Containers. Chris Xing Tian (National University of Singapore); Aditya Pan (Amity University); Yong Chiang Tay (National University of Singapore)

D9. BIGtensor: Mining Billion-Scale Tensor Made Easy. Namyoung Park, Byungsoo Jeon, Jungwoo Lee, U Kang (Seoul National University)

D10. eGraphSearch: Effective Keyword Search in Graphs. Mehdi Kargar (University of Windsor); Lukasz Golab (University of Waterloo); Jaroslav Szlichta (University of Ontario Institute of Technology)

D11. EnerQuery: Energy-Aware Query Processing. Amine Roukh (University of Mostaganem); Ladjel Bellatreche (LIAS/ISAE-ENSMA); Carlos Ordonez (University of Houston)

D12. TGraph: A Temporal Graph Data Management System. Haixing Huang, Jinghe Song, Xuelian Lin, Shuai Ma, Jinping Huai (SKLSDE Lab, Beihang University & Beijing Advanced Innovation Center for Big Data and Brain Computing)


D14. Thymeflow, A Personal Knowledge Base with Spatio-temporal Data. David Montoya (Engie Ineo); Thomas Pellissier Tanon (École Normale Supérieure de Lyon); Serge Abiteboul (Institut national de recherche en informatique et en automatique); Fabian M. Suchanek (Télécom ParisTech)

D15. Inferring Traffic Incident Start Time with Loop Sensor Data. Mingxuan Yue, Liyue Fan, Cyrus Shahabi (University of Southern California)

D16. TEAMOPT: Interactive Team Optimization in Big Networks. Liangyue Li, Hanghang Tong (Arizona State University); Nan Cao (NYU Shanghai); Kate Ehrlich (IBM Research); Yu-Ru Lin (University of Pittsburgh); Norbou Buchler (US Army Research Laboratory)

D17. GStreamMiner: A GPU-accelerated Data Stream Mining Framework. Chandima HewaNadungodage, Yuni Xia, John Jaehwan Lee (IUPUI)

D18. QART: A Tool for Quality Assurance in Real-Time in Contact Centers. Ragunathan Mariappan, Balaji Peddamuthu, Preethi R. Raajaratnam, Sandipan Dandapat, Neeta Pande, Shourya Roy (Xerox Research Centre India)

D19. A Fatigue Strength Predictor for Steels Using Ensemble Data Mining. Ankit Agrawal, Alok Choudhary (Northwestern University)
# Wednesday Oct. 26, 2016: Main Conference Day Two

<table>
<thead>
<tr>
<th>Time</th>
<th>Location</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-18:00</td>
<td>Regency Foyer</td>
<td>Registration (Regency Foyer)</td>
</tr>
<tr>
<td>09:00-10:10</td>
<td>Regency Ballroom</td>
<td>Plenary Keynote Talk: Susan Dumais (Chair: ChengXiang Zhai)</td>
</tr>
<tr>
<td>10:10-10:40</td>
<td>Regency Foyer</td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>10:40-12:10</td>
<td>Regency A&amp;B</td>
<td>Session 4a: Information Retrieval I (Chair: Elisa Bertino)</td>
</tr>
<tr>
<td></td>
<td>Regency C</td>
<td>Session 4b: User Behavior and Interfaces (Chair: Hady Wirawan Lauw)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td>Session 4c: Documents (Chair: Davood Rafiei)</td>
</tr>
<tr>
<td></td>
<td>Concept</td>
<td>Session 4e: Truth Discovery (Chair: Xia Ning)</td>
</tr>
<tr>
<td></td>
<td>Theory</td>
<td>Session 4d: Knowledge Mining and Management (Chair: De-Ning Yang)</td>
</tr>
<tr>
<td></td>
<td>Network</td>
<td>Session 4f: Industry Session IV (Chair: Parikshit Sondhi)</td>
</tr>
<tr>
<td>12:10-14:00</td>
<td>Regency Foyer</td>
<td>Lunch (on your own)</td>
</tr>
<tr>
<td>14:00-15:30</td>
<td>Regency C</td>
<td>Session 5a: Sentiment and Opinion Mining (Chair: Yi Fang)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td>Session 5b: Time Series (Chair: Goce Trajcevski)</td>
</tr>
<tr>
<td></td>
<td>Concept</td>
<td>Session 5d: Social Media (Chair: Maarten de Rijke)</td>
</tr>
<tr>
<td></td>
<td>Theory</td>
<td>Session 5c: Learning for Classification and Prediction (Chair: Wei Liu)</td>
</tr>
<tr>
<td></td>
<td>Network</td>
<td>Session 5e: Queries and Search (Chair: Davood Rafiei)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Session 5f: Industry Session V (Chair: Yi Chang)</td>
</tr>
<tr>
<td>15:30-16:00</td>
<td>Regency Foyer</td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>16:00-17:15</td>
<td>Regency Ballroom</td>
<td>CIKM 25th Anniversary Plenary Session (Chair: ChengXiang Zhai)</td>
</tr>
<tr>
<td>17:30-19:00</td>
<td>Regency E&amp;F</td>
<td>Poster Session II (Regency E&amp;F) + Demo (Network)</td>
</tr>
<tr>
<td>19:00-21:00</td>
<td>Regency Ballroom</td>
<td>Banquet Dinner (Regency Ballroom)</td>
</tr>
</tbody>
</table>
Wednesday Morning Sessions, Oct 26, 2016

Session 4a: Information Retrieval I
Time: 10:40 AM – 12:10 PM  Location: Regency A&B  Chair: Elisa Bertino

Paper 157: Query Variations and their Effect on Comparing Information Retrieval Systems, Guido Zuccon(Queensland University of Technology, Australia), Joao Palotti(Vienna University of Technology, Austria), Allan Hanbury(Vienna University of Technology, Austria)

Paper 416: Semantic Matching by Non-Linear Word Transportation for Information Retrieval, Jiafeng Guo(CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences, China), Yixing Fan(CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology, Chinese Academy of Sciences, China), Qingyao Ai(Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA), W. Bruce Croft(Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA)

Paper 989: Generalizing Translation Models in the Probabilistic Relevance Framework, Navid Rekabsaz(Vienna University of Technology, Austria), Mihai Lupu(Vienna University of Technology, Austria), Allan Hanbury(Vienna University of Technology, Austria), Guido Zuccon(Queensland University of Technology, Australia)

Paper 19: Axiomatic Result Re-Ranking, Matthias Hagen(Bauhaus-Universität Weimar, Germany), Michael Völkle(Bauhaus-Universität Weimar, Germany), Steve Göring(Bauhaus-Universität Weimar, Germany), Benno Stein(Bauhaus-Universität Weimar, Germany)

Session 4b: User Behavior and Interface
Time: 10:40 AM – 12:10 PM  Location: Regency C  Chair: Hady Wirawan Lauw

Paper 713: Agents, Simulated Users and Humans, David Maxwell(University of Glasgow, United Kingdom), Leif Azzopardi(University of Strathclyde, United Kingdom)

Paper 910: Inspiration or Preparation? Explaining Creativity in Scientific Enterprise, Xinyang Zhang(Lehigh University, USA), Dashun Wang(The Australian National University, Australia), Paul Thomas(Microsoft, Australia), Ramesh Sankaranarayana(The Australian National University, Australia), Tom Gedeon(The Australian National University, Australia), Hwan-Jin Yoon(The Australian National University, Australia)

Paper 947: Studying the Dark Triad of Personality through Twitter Behavior, Daniel Preotiuc-Pietro(University of Pennsylvania, USA), Jordan Carpenter(University of Pennsylvania, USA), Salvatore Giorgi(University of Pennsylvania, USA), Lyle Ungar(University of Pennsylvania, USA)

Session 4c: Document Analysis
Time: 10:40 AM – 12:10 PM  Location: Regency D  Chair: Davood Rafiel

Paper 177: Document Filtering for Long-tail Entities, Ridho Reinanda(University of Amsterdam, Netherlands), Edgar Meij(Bloomberg L.P., United Kingdom), Maarten de Rijke(University of Amsterdam, Netherlands)

Paper 689: Estimating Time Models for News Article Excerpts, Arunav Mishra(Max Planck Institute for Informatics, Germany), Klaus Berberich(Max Planck Institute for Informatics, Germany)

Paper 711: A Framework for Task-specific Short Document Expansion, Ramakrishna Bairi(Indian Institute of Technology, Bombay, India), Raghavendra Udupa(Microsoft Research India, India), Ganesh Ramakrishnan(Indian Institute of Technology, Bombay, India)

Paper 769: Beyond Clustering: Sub-DAG Discovery for Categorising Documents, Ramakrishna Bairi(Indian Institute of Technology, Bombay, India), Mark Carman(Monash University, Australia), Ganesh Ramakrishnan(Indian Institute of Technology, Bombay, India)

Session 4d: Knowledge Mining & Management
Time: 10:40 AM – 12:10 PM  Location: Theory  Chair: De-Ning Yang

Paper 190: On Transductive Classification in Heterogeneous Information Networks, Xiang Li(The University of Hong Kong, Hong Kong), Ben Kao(The University of Hong Kong, Hong Kong), Yudian Zheng(The University of Hong Kong, Hong Kong), Zhipeng Huang(The University of Hong Kong, Hong Kong)
Paper 638: Efficient Hidden Trajectory Reconstruction from Sparse Data, Ning Yang(Sichuan University, China), Philip Yu(University of Illinois at Chicago, USA)

Paper 176: Quark-X: An Efficient Top-K Processing Framework for RDF Quad Stores, Jyoti Leeka(Institute of Information Technology, Delhi, India), Srikantha Bedathur(IBM-IRL, India), Debajyoti Bera(Institute of Information Technology, Delhi, India), Medha Atre(Institute of Information Technology, Kanpur, India)

Paper 953: Reenactment for Read-Committed Snapshot Isolation, Bahareh Sadat Arab(Illinois Institute of Technology, USA), Dieter Gawlick(Oracle, USA), Vasudha Krishnaswamy(Oracle, USA), Venkatesh Radhakrishnan(LinkedIn, USA), Boris Glavic(Illinois Institute of Technology, USA)

Session 4e: Truth Discovery
Time: 10:40 AM – 12:10 PM  Location: Concept  Chair: Xia Ning

Paper 567: Influence-Aware Truth Discovery, Hengtong Zhang(SUNY Buffalo, USA), Qi Li(SUNY Buffalo, USA), Fenglong Ma(SUNY Buffalo, USA), Houping Xiao(SUNY Buffalo, USA), Yaliang Li(SUNY Buffalo, USA), Jing Gao(SUNY Buffalo, USA), Lu Su(SUNY Buffalo, USA)

Paper 605: Truth Discovery via Exploiting Implications from Multi-Source Data, Xianzhi Wang(University of New South Wales, Australia), Quan Sheng(University of Adelaide, Australia), Lina Yao(University of New South Wales, Australia), Xue Li(University of Queensland, Australia), Xiu Fang(University of Adelaide, Australia), Xiaofei Xu(Harbin Institute of Technology, China), Boualem Benatallah(University of New South Wales, Australia)

Paper 979: FacetGist: Collective Extraction of Document Facets in Large Technical Corpora, Tarique Siddiqui(University of Illinois at Urbana Champaign, USA), Xiang Ren(University of Illinois at Urbana Champaign, USA), Aditya Parameswaran(University of Illinois at Urbana Champaign, USA), Jiawei Han(University of Illinois at Urbana Champaign, USA)

Paper 412: Empowering Truth Discovery with Multi-Truth Prediction, Xianzhi Wang(University of New South Wales, Australia), Quan Sheng(University of Adelaide, Australia), Lina Yao(University of New South Wales, Australia), Xue Li(University of Queensland, Australia), Xiu Fang(University of Adelaide, Australia), Xiaofei Xu(Harbin Institute of Technology, China), Boualem Benatallah(University of New South Wales, Australia)

Session 4f: Industry Session IV
Time: 10:40 AM – 12:10 PM  Location: Network  Chair: Parikshit Sondhi

Keynote: Using Machine Learning to Improve the Email Experience, Marc Najork(Google Inc.)

Paper 1481: Hashtag Recommendation for Enterprise Applications, Dhruv Mahajan(Microsoft Corporation, USA), Vishwajit Kolathur(Microsoft Corporation, India), Chetan Bansal(Microsoft Research, USA), Suresh Parthasarathy(Microsoft Research, India), Sundararajan Sellamanickam(Microsoft Research, India), Sathiya Keerthi(Microsoft Corporation, USA), Johannes Gehrke(Microsoft Corporation, USA)

Paper 1368: Survival Analysis based Framework for Early Prediction of Student Dropouts, Sattar Ameri(Wayne State University, USA), Mahtab J. Fard(Wayne State University, USA), Ratna Chinnam(Wayne State University, USA)

Wednesday Afternoon Sessions, Oct 26, 2016

Session 5a: Sentiment & Opinion Mining
Time: 2:00 PM – 3:30 PM  Location: Regency A&B  Chair: Yi Fang

Paper 179: Generative Feature Language Models for Mining Implicit Features from Customer Reviews, Shubhra Kanti Karmaker Santu(University of Illinois at Urbana-Champaign, USA), Parikshit Sondhi(WalmartLabs, USA), ChengXiang Zhai(University of Illinois at Urbana-Champaign, USA)

Paper 607: Data-Driven Contextual Valence Shifter Quantification for Multi-Theme Sentiment Analysis, Hongkun Yu(University of Illinois at Urbana-Champaign, USA), Jingbo Shang(University of Illinois at Urbana-Champaign, USA), Meichun Hsu(HPE Vertica, USA), Malu Castellanos(HPE Vertica, USA), Jiawei Han(University of Illinois at Urbana-Champaign, USA)
Paper 1202: Sentiment Domain Adaptation with Multi-Level Contextual Sentiment Knowledge, Fangzhao Wu(Tsinghua University, China), Sixing Wu(Tsinghua University, China), Yongfeng Huang(Tsinghua University, China), Songfang Huang(IBM Research - China, China), Yong Qin(IBM Research - China, China)

Paper 1123: Mobile App Retrieval for Social Media Users via Inference of Implicit Intent in Social Media Text, Dae Hoon Park(Yahoo! Inc., USA), Yi Fang(Santa Clara University, USA), Mengwen Liu(Drexel University, USA), ChengXiang Zhai(University of Illinois at Urbana-Champaign, USA)

**Session 5b: Time Series**
**Time:** 2:00 PM – 3:30 PM  
**Location:** Regency C  
**Chair:** Goce Trajcevski

Paper 80: Derivative Delay Embedding: Online Modeling of Streaming Time Series, Zhifei Zhang(The University of Tennessee, USA), Yang Song(The University of Tennessee, USA), Wei Wang(The University of Tennessee, USA), Hairong Qi(The University of Tennessee, USA)

Paper 459: PISA: an Index for Aggregating Big Time Series Data, Xiangdong Huang(Tsinghua University, China), Jianmin Wang(Tsinghua University, China), Raymond Wong(University of New South Wales, Australia), Jinrui Zhang(Tsinghua University, China), Chen Wang(Tsinghua University, China)

Paper 510: Multi-View Time Series Classification: A Discriminative Bilinear Projection Approach, Sheng Li(Northeastern University, USA), Yaliang Li(SUNY Buffalo, USA), Yun Fu(Northeastern University, USA)

Paper 1247: Semi-Supervision Dramatically Improves Time Series Clustering under Dynamic Time Warping, Hoang Anh Dau(University of California, Riverside, USA), Nurjahan Begum(University of California, Riverside, USA), Eamonn Keogh(University of California, Riverside, USA)

**Session 5c: Learning for Classification & Prediction**
**Time:** 2:00 PM – 3:30 PM  
**Location:** Concept  
**Chair:** Wei Liu

Paper 551: Model-Based Oversampling for Imbalanced Sequence Classification, Zhichen Gong(University of Science and Technology of China, China), Huanhuan Chen(University of Science and Technology of China, China)

Paper 495: CRISP: Consensus Regularized Selection based Prediction, Ping Wang(Virginia Tech, USA), Karthik Padthe(Wayne State University, USA), Bhanukiran Vinzamuri(Wayne State University, USA), Chandan Reddy(Virginia Tech, USA)

Paper 1309: Regularizing Structured Classifier with Conditional Probabilistic Constraints for Semi-supervised Learning, Vincent Zheng(Advanced Digital Sciences Center, Singapore), Kevin Chang(University of Illinois at Urbana-Champaign, USA)

Paper 468: Scalability of Continuous Active Learning for Reliable High-Recall Text Classification, Gordon Cormack(University of Waterloo, Canada), Maura Grossman(University of Waterloo, Canada)

**Session 5d: Social Media**
**Time:** 2:00 PM – 3:30 PM  
**Location:** Regency D  
**Chair:** Maarten de Rijke

Paper 279: Towards the Effective Linking of Social Media Contents to Products in E-Commerce Catalogs, Henry Vieira(Federal University of Amazonas & FPF Tech, Brazil), Altigran da Silva(Federal University of Amazonas, Brazil), Pavel Calado(INESC-ID & Universidade de Lisboa, Portugal), Marco Cristo(Federal University of Amazonas, Brazil), Edlone de Moura(Federal University of Amazonas, Brazil)

Paper 682: Tracking Virality and Susceptibility in Social Media, Tuan-Anh Hoang(L3S Research Center, Germany), Ee-Peng Lim(Singapore Management University, Singapore)

Paper 784: Feature Driven and Point Process Approaches for Popularity Prediction, Swapnil Mishra(The Australian National University and Data61, Australia), Marian-Andrei Rizoiu(The Australian National University and Data61, Australia), Lexing Xie(The Australian National University and Data61, Australia)

Paper 360: Adaptive Evolutionary Filtering in Real-Time Twitter Stream, Feifan Fan(Peking University, China), Yansong Feng(Peking University, China), Lili Yao(Peking University, China), Dongyan Zhao(Peking University, China)
Session 5e: Query & Search
Time: 2:00 PM – 3:30 PM  Location: Theory  Chair: Davood Rafieil

Paper 886: Multiple Queries as Bandit Arms, Cheng Li(University of Michigan, USA), Paul Resnick(University of Michigan, USA), Qiaozhu Mei(University of Michigan, USA)
Paper 878: Off the Beaten Path: Let's Replace Term-Based Retrieval with k-NN Search, Leonid Boytsov(Carnegie Mellon University, USA), David Novak(Masaryk University, Czech Rep), Yury Malkov(Institute of Applied Physics RAS, Russian Fed.), Eric Nyberg(Carnegie Mellon University, USA)
Paper 263: Scalability and Total Recall with Fast CoveringLSH, Ninh Pham(IT University of Copenhagen, Denmark), Rasmus Pagh(IT University of Copenhagen, Denmark)
Paper 30: Query-Biased Partitioning for Selective Search, Zhuyun Dai(Carnegie Mellon University, USA), Chenyan Xiong(Carnegie Mellon University, USA), Jamie Callan(Carnegie Mellon University, USA)

Session 5f: Industry Session V
Time: 2:00 PM – 3:30 PM  Location: Network  Chair: Yi Chang

Invited Talk: Neural Network Models for Information Accessing, Zhengdong Lu (Huawei Inc.)
Invited Talk: Building a Medical Diagnosis Chatbot using Artificial Intelligence Technologies, Wei Fan (Baidu Inc.)

Paper 1474: Characterizing Diseases from Unstructured Text: A Vocabulary Driven Word2vec Approach, Saurav Ghosh(Virginia Tech, USA), Prithwish Chakraborty(Virginia Tech, USA), Emily Cohn(Boston Children’s Hospital, USA), John Brownstein(Harvard Medical School, USA), Naren Ramakrishnan(Virginia Tech, USA)
Paper 1464: Network-Efficient Distributed Word2vec Training System for Large Vocabularies, Erik Ordentlich(Yahoo, Inc., USA), Lee Yang(Yahoo, Inc., USA), Andy Feng(Yahoo, Inc., USA), Peter Cnudde(Yahoo, Inc., USA), Mihajlo Grbovic(Airbnb, Inc., USA), Nemanja Djuric(Uber ATC, USA), Vladan Radosavljevic(Uber ATC, USA), Gavin Owens(Deco Software, USA)

CIKM 25th Anniversary Session: CIKM and Big Data

Time: 4:00 PM -- 5:15 PM  Location: Regency Ballroom

This is a special plenary session for celebrating the 25th Anniversary of CIKM with a Panel on CIKM and Big Data and announcement of the Inaugural CIKM Test of Time Award.

Panel on CIKM and Big Data (Time: 4:00 PM – 5:00 PM, Location: Regency Ballroom)

Panelists (current/past CIKM Steering Committee chairs): David Grossman (Georgetown University), Charles Nicholas (University of Maryland, Baltimore County), James G. Shanahan (University of California Berkeley and Church and Duncan Group Inc)

Announcement of Inaugural CIKM Test of Time Award (5:00 PM – 5:15PM, Regency Ballroom)
Poster Session II: Extended Short Papers and Industry Track Short Papers

Time: 5:30 PM – 7:00 PM   Location: Regency E&F

ESP1. An Experimental Comparison of Iterative MapReduce Frameworks. Haejoon Lee, Minseong Kang (KAIST), Sun-Bum Youn (Netmarble Games); Jae-Gil Lee (KAIST); YongChul Kwon (Microsoft Corp.)

ESP2. A Density-Based Approach to the Retrieval of Top-K Spatial Textual Clusters. Dingming Wu (Shenzhen University); Christian S. Jensen (Aalborg University)

ESP3. Top-N Recommendation on Graphs. Zhao Kang, Chong Peng, Ming Yang, Qiang Cheng (Southern Illinois University)

ESP4. KB-Enabled Query Recommendation for Long-Tail Queries. Zhipeng Huang (The University of Hong Kong); Bogdan Cautis (Huawei Noah’s Ark Lab); Reynold Cheng, Yudian Zheng (The University of Hong Kong)

ESP5. RAP: Scalable RPCA for Low-rank Matrix Recovery. Chong Peng, Zhao Kang, Ming Yang, Qiang Cheng (Southern Illinois University Carbondale)

ESP6. Query Answering Efficiency in Expert Networks Under Decentralized Search. Liang Ma, Mudhakar Srivatsa (IBM T. J. Watson Research); Derya Cansever (Army CERDEC); Xifeng Yan (University of California); Sue Kase, Michelle Vanni (Army Research Laboratory)

ESP7. A Study of Real-time Summarization Metrics. Matthew Ekstrand-Abueg (Northeastern University); Richard McCreadie (University of Glasgow); Virgil Pavlu (Northeastern University); Fernando Diaz (Microsoft Research)

ESP8. Framing Mobile Information Needs: An Investigation of Hierarchical Query Sequence Structure. Shuguang Han (University of Pittsburgh); Xing Yi, Zhen Yue, Zhigeng Geng, Alyssa Glass (Yahoo! Research)


ESP10. Combining Powers of Two Predictors in Optimizing Real-Time Bidding Strategy under Constrained Budget. Chi-Chun Lin (National Taiwan University); Kun-Ta Chuang (National Cheng Kung University); Wush Chi-Hsuan Wu, Ming-Syan Chen (National Taiwan University)

ESP11. Attractiveness versus Competition: Towards an Unified Model for User Visitation. Thanh-Nam Doan, Ee-Peng Lim (Singapore Management University)

ESP12. OptMark: A Toolkit for Benchmarking Query Optimizers. Zhan Li, Olga Papaemmanouil, Mitch Cherniack (Brandeis University)

ESP13. Multi-Dueling Bandits and Their Application to Online Ranker Evaluation. Brian Brost, Yevgeny Seldin (University of Copenhagen); Ingemar J. Cox (University College London); Christina Lioma (University of Copenhagen)


ESP15. Credibility Assessment of Textual Claims on the Web. Kashyap Popat, Subhabrata Mukherjee, Jannik Ströglén, Gerhard Weikum (Max Planck Institute for Informatics)

ESP16. Collective Traffic Prediction with Partially Observed Traffic History using Location-Based Social Media. Xinyue Liu, Xiangnan Kong, Yanhua Li (Worcester Polytechnic Institute)

ESP17. Recommendations For Streaming Data. Karthik Subbian (University of Minnesota); Charu Aggarwal (IBM Watson Research Center); Kshitij Hegde (University of Minnesota)

ESP18. PRO: Preference-Aware Recurring Query Optimization. Zhongfang Zhuang (Worcester Polytechnic Institute); Chuan Lei (NEC Labs. America); Elke Rundensteiner, Mohamed Eltabakh (Worcester Polytechnic Institute)

ESP19. Discovering Temporal Purchase Patterns with Different Responses to Promotions. Ling Luo (University of Sydney); Bin Li (CSIRO); Irena Koprinska (University of Sydney); Shlomo Berkovsky, Fang Chen (CSIRO)

ESP20. ZEST: a Hybrid Model on Predicting Passenger Demand for Chauffeured Car Service. Hua Wei, Yuanlong Wang, Tianyu Wo (Beihang University); Yaxiao Liu (Tsinghua University); Jie Xu (Beihang University)


ESP22. Reuse-based Optimization for Pig Latin. Jesús Camacho-Rodríguez (Hortonworks); Dario Colazzo (Université Paris-Dauphine, PSL Research University, CNRS, LAMSADE); Melanie Herschel (IPVS, University of Stuttgart); Ioana Manolescu (INRIA & LIX, Ecole Polytechnique, CNRS); Souhip Roy Chowdhury (Fractal Analytics)
ESP23. Discriminative View Learning for Single View Co-Training. Joseph St.Amand, Jun Huan (University of Kansas)
ESP24. Learning Points and Routes to Recommend Trajectories. Dawei Chen, Cheng Soon Ong, Lexing Xie (The Australian National University, Data61)
ESP25. Towards Representation Independent Similarity Search Over Graph Databases. Yodsawalai Chodpathumwan, Amirhossein Aleyasen (University of Illinois); Arash Termehchy (Oregon State University); Yizhou Sun (Northeastern University)
ESP27. Anomalies in the Peer-review system: A Case Study of the Journal of High Energy Physics. Sandipan Sikdar (Indian Institute of Technology); Matteo Marsili (International Centre for Theoretical Physics); Niloy Ganguly, Animesh Mukherjee (Indian Institute of Technology)
ESP28. Multi-source Hierarchical Prediction Consolidation. Chenwei Zhang (University of Illinois at Chicago); Sihong Xie (Lehigh University); Yaliang Li, Jing Gao (SUNY Buffalo); Wei Fan (Baidu Research Big Data Lab); Philip S. Yu (University of Illinois at Chicago & Tsinghua University)
ESP29. Probabilistic Knowledge Graph Construction: Compositional and Incremental Approaches. Dongwoo Kim (Australian National University); Lexing Xie (Australian National University & Data to Decision CRC); Cheng Soon Ong (Data61 & Australian National University)
ESP30. Explaining Sentiment Spikes in Twitter. Anastasia Giachanou, Ida Mele, Fabio Crestani (Università della Svizzera italiana)
ESP31. Qualitative Cleaning of Uncertain Data. Henning Koehler (Massey University); Sebastian Link (The University of Auckland)
ESP33. OrientStream: A Framework for Dynamic Resource Allocation in Distributed Data Stream Management Systems. Chunkei Wang, Xiaofeng Meng (Renmin University of China); Qi Guo (Chinese Academy of Sciences); Zujian Weng, Chen Yang (Renmin University of China)
ESP34. Tag2Word: Using Tags to Generate Words for Content Based Tag Recommendation. Yong Wu, Yuan Yao, Feng Xu (State Key Laboratory for Novel Software Technology, Nanjing University, China); Hanghang Tong (Arizona State University); Jian Lu (State Key Laboratory for Novel Software Technology, Nanjing University, China)
ESP35. Digging Multilingual Reader Comments via Latent Discussion Topics with Commonality and Specificity. Bei Shi, Wai Lam (The Chinese University of Hong Kong); Lidong Bing (Carnegie Mellon University); Yinqing Xu (The Chinese University of Hong Kong)
ESP36. Digging News Reader Comments via Fine-Grained Associations with Event Facets and News Contents. Bei Shi, Wai Lam (The Chinese University of Hong Kong)
ESP37. Efficient Algorithms for the Two Locus Problem in Genome-Wide Association Study. Sanguthevar Rajasekaran, Subrata Saha (University of Connecticut)
ESP38. FolkTrails: Interpreting Navigation Behavior in a Social Tagging System. Thomas Niebler, Martin Becker, Daniel Zoller (Universität Würzburg); Stephan Doerfel (Universität Kassel); Andreas Hotho (Universität Würzburg)
ESP39. Memory-Optimized Distributed Graph Processing through Novel Compression Techniques. Panagiotis Liakos, Katia Papakonstantinopoulou, Alex Delis (University of Athens)
ESP40. Tracking the Evolution of Congestion in Dynamic Urban Road Networks. Tarique Anwar, Chengfei Liu, Hai L. Vu, Md. Saiful Islam (Swinburne University of Technology)
ESP41. The Rich and the Poor: A Markov Decision Process Approach to Optimizing Taxi Driver Revenue Efficiency. Huiguí Rong (Hunan University); Xun Zhou (University of Iowa); Chang Yang (Hunan University); Zubair Shafigh (University of Iowa); Alex Liu (Michigan State University)
ESP42. Ensemble of Anchor Adapters for Transfer Learning. Fuzhen Zhuang, Ping Luo (Institute of Computing Technology, CAS); Sinno Jialin Pan (Nanyang Technological University); Hui Xiong (MSIS Department, Rutgers University); Qing He (Institute of Computing Technology, CAS)
ESP43. Incremental Mining of High Utility Sequential Patterns in Incremental Databases. Jun-Zhe Wang, Jian-Long Huang (National Chiao Tung University)
ESP44. Understanding Stability of Noisy Networks through Centrality Measures and Local Connectivity. Vladimir Ufimtsev (University of Nebraska Omaha); Soumya Sarkar, Animesh Mukherjee (IIT Kharagpur); Sanjukta Bhowmick (University of Nebraska Omaha)
ESP45. Online Adaptive Topic Focused Tweet Acquisition. Mehdi Sadri, Sharad Mehrotra, Yaming Yu (University of California, Irvine)

ESP46. Optimizing Nugget Annotations with Active Learning. Gaurav Baruah, Haotian Zhang, Rakesh Guttikonda, Jimmy Lin, Mark D. Smucker, Olga Vechtomova (University of Waterloo)

ESP47. Uncovering Fake Likers in Online Social Networks. Prudhvi Ratna Badri Satya, Kyumin Lee (Utah State University); Dongwon Lee (The Pennsylvania State University); Thanh Tran (Utah State University); Jason (Jiasheng) Zhang (The Pennsylvania State University)

ESP48. Where to Place Your Next Restaurant? Optimal Restaurant Placement via Leveraging User-Generated Reviews. Feng Wang, Li Chen (Hong Kong Baptist University); Weike Pan (Shenzhen University)

ESP49. Leveraging the Implicit Structure within Social Media for Emergent Rumor Detection. Justin Sampson, Fred Morstatter, Liang Wu, Huan Liu (Arizona State University)

ESP50. Automatical Storyline Generation with Help from Twitter. Ting Hua, Xuchao Zhang, Wei Wang, Chang-Tien Lu, Naren Ramakrishnan (Virginia Tech)

ESP51. A Comparative Study of Query-biased and Non-redundant Snippets for Structured Search on Mobile Devices. Nikita V. Spirin (University of Illinois at Urbana-Champaign); Alexander S.Kotov (Wayne State University); Karrie G. Karahalios, Vassil Mladenov (University of Illinois at Urbana-Champaign); Pavel A. Izhutov (Stanford University)

ESP52. Content-Agnostic Malware Detection in Heterogeneous Malicious Distribution Graph. Ibrahim Alabdulmohsin (King Abdullah University of Science & Technology); YuFei Han, Yun Shen (Symantec Research Labs); XiangLiang Zhang (King Abdullah University of Science & Technology)

Industry Track Short Papers

ISP1. Improving Advertisement Recommendation by Enriching User Browser Cookie Attributes. Liang Wang, Kuang-chih Lee, Quan Lu (Yahoo! Inc.)

ISP2. Balanced Supervised Non-Negative Matrix Factorization for Childhood Leukaemia Patients. Ali Braytee (University of Technology, Sydney); Daniel R. Catchpoole (The Children’s Hospital at Westmead); Paul J. Kennedy, Wei Liu (University of Technology, Sydney)

ISP3. SoLSCSum: A Linked Sentence-Comment Dataset for Social Context Summarization. Minh-Tien Nguyen, Chien-Xuan Tran, Duc-Vu Tran, Minh-Le Nguyen (Japan Advanced Institute of Science and Technology (JAIST))

ISP4. Distributed Deep Learning for Question Answering. Minwei Feng, Bing Xiang, Bowen Zhou (IBM Watson)

ISP5. Bus Routes Design and Optimization via Taxi Data Analytics. Seong Ping Chuah, Huayu Wu, Yu Lu, Liang Yu (Institute for Infocomm Research); Stephane Bressan (National University of Singapore)

ISP6. Routing an Autonomous Taxi with Reinforcement Learning. Miyoung Han, Pierre Senellart (Télécom ParisTech); Stéphane Bressan (IPAL, NUS); Huayu Wu (I2R, A*STAR)

Industry Track Demo Papers (all demos will be presented twice)

Session 1: 6:00 PM - 7:00 PM, Tuesday, Oct. 25, 2016

Session 2: 5:30 PM – 7:00 PM, Wednesday, Oct. 26 2016

Location: Network
### October 27, 2016: Main Conference Day Three

<table>
<thead>
<tr>
<th>Time</th>
<th>Location</th>
<th>Session Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-17:30</td>
<td>Regency Foyer</td>
<td>Registration (Regency Foyer)</td>
</tr>
<tr>
<td>09:00-10:10</td>
<td>Regency C</td>
<td>Plenary Keynote Talk: Andrei Broder (Regency Ballroom) (Chair: ChengXiang Zhai)</td>
</tr>
<tr>
<td>10:10-10:40</td>
<td>Regency Foyer</td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>10:40-12:10</td>
<td>Regency B</td>
<td><strong>Session 6a:</strong> Learning Algorithms (Chair: Murat Dundar)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 6b:</strong> Databases and Data Processing (Chair: Alfredo Cuzzucrea)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 6d:</strong> Information Retrieval II (Chair: Maarten de Rijke)</td>
</tr>
<tr>
<td></td>
<td>Regency C</td>
<td><strong>Session 6c:</strong> Large Graph Processing (Chair: Mohammad Hasan)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 6e:</strong> Entity Detection and Analysis (Chair: Yi Fang)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 6f:</strong> Industry Session VI (Chair: Yi Chang)</td>
</tr>
<tr>
<td>12:40-14:00</td>
<td>Regency Ballroom</td>
<td>CIKM Business Meeting (Lunch Provided, Regency Ballroom)</td>
</tr>
<tr>
<td>14:30-16:30</td>
<td>Regency A&amp;B</td>
<td><strong>Session 7b:</strong> Query Analytics (Chair: Xia Ning)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 7c:</strong> Information Retrieval III (Chair: Maarten de Rijke)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 7e:</strong> Network Analytics (Chair: Mohammad Hasan)</td>
</tr>
<tr>
<td></td>
<td>Regency C</td>
<td><strong>Session 7a:</strong> Advertising and Ranking (Chair: Yi Fang)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 7d:</strong> Data Mining (Chair: Zhenhui Li)</td>
</tr>
<tr>
<td></td>
<td>Regency Foyer</td>
<td><strong>Session 7f:</strong> Industry Session VII (Chair: Jiliang Tang)</td>
</tr>
<tr>
<td>16:00-16:30</td>
<td>Regency Foyer</td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>16:30-18:00</td>
<td>Regency C</td>
<td><strong>Session 8a:</strong> Learning (Chair: Murat Dundar)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 8b:</strong> Social Networks---Diffusion and Cascades (Chair: Elisa Bertino)</td>
</tr>
<tr>
<td></td>
<td>Regency A&amp;B</td>
<td><strong>Session 8c:</strong> Applications (Chair: Xiao Ning)</td>
</tr>
<tr>
<td></td>
<td>Regency D</td>
<td><strong>Session 8d:</strong> Data Mining Algorithms (Chair: Lei Liu)</td>
</tr>
<tr>
<td></td>
<td>Regency C</td>
<td><strong>Session 8f:</strong> Industry Session VIII (Chair: Parikshit Sondhi)</td>
</tr>
</tbody>
</table>

---
Thursday Morning Sessions, Oct 27, 2016

Session 6a: Learning Algorithms
Time: 10:40 AM – 12:10 PM  Location: Regency A&B  Chair: Murat Dundar

Paper 253: Scalable Spectral k-Support Norm Regularization for Robust Low Rank Subspace Learning, Yiu-ming Cheung(Hong Kong Baptist University, Hong Kong), Jian Lou(Hong Kong Baptist University, Hong Kong)

Paper 570: Online Adaptive Passive-Aggressive Methods for Non-Negative Matrix Factorization and Its Applications, Chenghao Liu(Zhejiang University, China), Steven C.H. Hoi(Singapore Management University, Singapore), Peilin Zhao(Ant Financial, China), Jianling Sun(Zhejiang University, China), Ee-Peng Lim(Singapore Management University, Singapore)

Paper 547: aptMTVL: Nailing Interactions in Multi-Task Multi-View Multi-Label Learning using Adaptive-basis Multilinear Factor Analyzers , Xiaoli Li(The University of Kansas, USA), Jun Huan(The university of Kansas, USA)

Paper 1116: An Adaptive Framework for Multistream Classification, Swarup Chandra(University of Texas at Dallas, USA), Ahsanul Haque(University of Texas at Dallas, USA), Latifur Khan(University of Texas at Dallas, USA), Charu Aggarwal(IBM T. J. Watson Research Center, USA)

Session 6b: Database & Data Processing
Time: 10:40 AM – 12:10 PM  Location: Regency C  Chair: Alfredo Cuzzucrea

Paper 118: Optimizing Update Frequencies for Decaying Information, Simon Razniewski(Free University of Bozen-Bolzano, Italy)

Paper 736: Cutty: Aggregate Sharing for User-Defined Windows, Paris Carbone(KTH Royal Institute of Technology, Sweden), Jonas Traub(technische Universitat Berlin & DFKI, Germany), Asterios Katsifodimos(technische Universitat Berlin & DFKI, Germany), Seif Haridi(KTH Royal Institute of Technology, Sweden), Volker Markl(technische Universitat Berlin & DFKI, Germany)

Paper 685: Relational Database Schema Design for Uncertain Data, Sebastian Link(The University of Auckland, New Zealand), Henri Prade(IRIT, CNRS, and Universite de Toulouse III, France)

Paper 110: BICP: Block-Incremental CP Decomposition with Update Sensitive Refinement, Shengyu Huang(Arizona State University, USA), K. Selçuk Candan(ARIZONA STATE UNIVERSITY, USA), Maria Luisa Sapino(University of Torino, Italy)

Session 6c: Large Graph Processing
Time: 10:40 AM – 12:10 PM  Location: Concept  Chair: Mohammad Hasan

Paper 237: Topological Graph Sketching for Incremental and Scalable Analytics, Bortik Bandyopadhyay(The Ohio State University, USA), David Fuhry(The Ohio State University, USA), Aniket Chakrabarti(The Ohio State University, USA), Srinivasan Parthasarathy(The Ohio State University, USA)

Paper 280: Querying Minimal Steiner Maximum-Connected Subgraphs in Large Graphs, Jiafeng Hu(The University of Hong Kong, Hong Kong), Xiaowei Wu(The University of Hong Kong, Hong Kong), Reynold Cheng(The University of Hong Kong, Hong Kong), Siqiang Luo(The University of Hong Kong, Hong Kong), Yixiang Fang(The University of Hong Kong, Hong Kong)

Paper 1168: Efficient Estimation of Triangles in Very Large Graphs, Roohollah Etemadi(University of Windsor, Canada), Yung H. Tsin(University of Windsor, Canada)

Paper 734: Efficient Batch Processing for Multiple Keyword Queries on Graph Data, Lu Chen(Swinburne University of Technology, Australia), Chengfei Liu(Swinburne University of Technology, Australia), Xiaochun Yang(Northeastern University, China), Bin Wang(Northeastern University, China), Jianxin Li(University of Western Australia, Australia), Rui Zhou(Victoria University, Australia)
Session 6d: Information Retrieval II
Time: 10:40 AM – 12:10 PM    Location: Regency D    Chair: Maarten de Rijke

Paper 272: Supervised Robust Discrete Multimodal Hashing for Cross-Media Retrieval, Ting-Kun Yan(Shandong University, China), Xin-Shun Xu(Shandong University, China), Shanqing Guo(Shandong University, China), Zi Huang(The University of Queensland, Australia), Xiao-Lin Wang(Shandong University, China)

Paper 286: Word Vector Compositionality based Relevance Feedback using Kernel Density Estimation, Dwaipayan Roy(Indian Statistical Institute, India), Debasis Ganguly(Dublin City University, Ireland), Mandar Mitra(Indian Statistical Institute, India), Gareth Jones(Dublin City University, Ireland)

Paper 399: Q+Tree: An Efficient Quad Tree based Data Indexing for Parallelizing Dynamic and Reverse Skylines, Md. Saiful Islam(La Trobe University, Australia), Chengfei Liu(Swinburne University of Technology, Australia), Wenny Rahayu(La Trobe University, Australia), Tarique Anwar(Swinburne University of Technology, Australia)

Paper 830: Luhn Revisited: Significant Words Language Models, Mostafa Dehghani(University of Amsterdam, Netherlands), Hosein Azarbonyad(University of Amsterdam, Netherlands), Jaap Kamps(University of Amsterdam, Netherlands), Maarten Marx(University of Amsterdam, Netherlands)

Session 6e: Entities Detection & Analysis
Time: 10:40 AM – 12:10 PM    Location: Theory    Chair: Yi Fang

Paper 482: ESPRESSO: Explaining Relationships between Entity Sets, Stephan Seufert(Max Planck Institute for Informatics, Germany), Klaus Berberich(Max Planck Institute for Informatics, Germany), Srikanta Bedathur(IBM Research, India), Patrick Ernst(Max Planck Institute for Informatics, Germany), Gerhard Weikum(Max Planck Institute for Informatics, Germany)

Paper 635: Geotagging Named Entities in News and Online Documents, Jiangwei Rafiei(University of Alberta, Canada), Davood Rafiei(University of Alberta, Canada)

Paper 657: Discovering Entities with Just a Little Help from You, Jaspreet Singh(L3S Research Center, Leibniz University Hannover, Germany), Johannes Hoffart(MPII, Germany), Avishek Anand(L3S Research Center, Leibniz University Hannover, Germany)

Paper 78: Bayesian Non-Exhaustive Classification A Case Study: Online Name Disambiguation using Temporal Record Stream, Baichuan Zhang(Indiana University Purdue University Indianapolis, USA), Murat Dundar(Indiana University Purdue University Indianapolis, USA), Mohammad Al Hasan(Indiana University Purdue University Indianapolis, USA)

Session 6f: Industry Session VI
Time: 10:40 AM – 12:10 PM    Location: Network    Chair: Yi Chang

Keynote: Large-scale Robust Online Matching and Its Application in E-commerce, Rong Jin (Alibaba Inc.)

Paper 1387: A Distributed Graph Algorithm for Discovering Unique Behavioral Groups from Large-Scale Telco Data, Qirong Ho(Institute for Infocomm Research, Singapore), Wenchao Lin(Qatar Computing Research Institute, Qatar), Eran Shaham(Institute for Infocomm Research, Singapore), Shonal Krishna(Indian Institute of Technology Bombay, India), Thinh Hoang(Indian Institute of Technology Bombay, India), Debashis Kr. Pramanik(SingTel, Singapore), Isabel Zhongyan(DataSpark, Singapore), Amy She-Nash(Commonwealth Bank of Australia, Australia)

Paper 1396: Urban Traffic Prediction through the Second Use of Inexpensive Big Data from Buildings, Zimu Zheng(The Hong Kong Polytechnic University, China), Dan Wang(The Hong Kong Polytechnic University, China), Jian Pei(Simon Fraser University, Canada), Yi Yuan(Tencent Co. Ltd., China), Cheng Fan(The Hong Kong Polytechnic University, China), Fu Xiao(The Hong Kong Polytechnic University, China)
Thursday Afternoon Sessions, Oct 27, 2016

Session 7a: Advertising & Ranking
Time: 2:30 PM – 4:00 PM  Location: Concept  Chair: Yi Fang

Paper 582: A Probabilistic Multi-Touch Attribution Model for Online Advertising, Wendi Ji(East China Normal University, China), Xiaoling Wang(East China Normal University, China), Dell Zhang(Birkbeck, University of London, United Kingdom)

Paper 1006: Optimizing Ad Allocation in Social Advertising, Shaojie Tang(University of Texas at Dallas, USA), Jing Yuan(University of Texas at Dallas, USA)

Paper 1082: Joint Collaborative Ranking with Social Relationships in Top-N Recommendation, Dimitrios Rafailidis(Aristotle University of Thessaloniki, Greece), Fabio Crestani(University della Svizzera italiana (USI), Switzerland)

Paper 1236: Modeling Customer Engagement from Partial Observations, Jelena Stojanovic(Temple University, USA), Djordje Gligorijevic(Temple University, USA), Zoran Obradovic(Temple University, USA)

Session 7b: Query Analytics
Time: 2:30 PM – 4:00 PM  Location: Regency A&B  Chair: Xia Ning

Paper 1211: On the Effectiveness of Query Weighting for Adapting Rank Learners to New Unlabelled Collections, Pengfei Li(RMIT University, Australia), Mark Sanderson(RMIT University, Australia), Mark Carman(Monash University, Australia), Falk Scholer(RMIT University, Australia)

Paper 1255: One Query, Many Clicks: Analysis of Queries with Multiple Clicks by the Same User, Elad Kravi(Israel Institute of Technology, Israel), Ido Guy(Ben-Gurion University of the Negev, Israel), Avihai Mejer(Yahoo Research, Israel), David Carmel(Yahoo Research, Israel), Yoelle Maarek(Yahoo Research, Israel), Dan Pelleg(Yahoo Research, Israel), Gilad Tsur(Yahoo Research, Israel)

Paper 952: Precision-Oriented Query Facet Extraction, Weize Kong(University of Massachusetts Amherst, USA), James Allan(University of Massachusetts Amherst, USA)

Paper 1027: Learning to Rewrite Queries, Yunlong He(Yahoo! Research, USA), Jiliang Tang(Michigan State University, USA), Hua Ouyang(Yahoo! Research, USA), Changsung Kang(Yahoo! Research, USA), Dawei Yin(Yahoo! Research, USA), Yi Chang(Yahoo! Research, USA)

Session 7c: Information Retrieval III
Time: 2:30 PM – 4:00 PM  Location: Regency C  Chair: Maarten de Rijke

Paper 1282: When is the Time Ripe for Natural Language Processing for Patent Passage Retrieval?, Linda Andersson(Vienna University Of Technology, Austria), Mihai Lupu(Vienna University Of Technology, Austria), Joao Palotti(Vienna University Of Technology, Austria), Allan Hanbury(Vienna University Of Technology, Austria), Andreas Rauber(Vienna University Of Technology, Austria)

Paper 256: A Probabilistic Fusion Framework, Yael Anava(Israel Institute of Technology, Israel), Anna Shtok(Israel Institute of Technology, Israel), Oren Kurland(Israel Institute of Technology, Israel), Ella Rabinovich(IBM Haifa Research Labs, Israel)

Paper 252: Selective Cluster-Based Document Retrieval, Or Levi(Israel Institute of Technology, Israel), Fiana Raiber(Israel Institute of Technology, Israel), Oren Kurland(Israel Institute of Technology, Israel), Ido Guy(Yahoo Research & Ben-Gurion University of the Negev, Israel)

Paper 1130: Pseudo-Relevance Feedback Based on Matrix Factorization, Hamed Zamani(University of Massachusetts Amherst, USA), Javid Dadashkarimi(University of Tehran, Iran), Azadeh Shakery(University of Tehran, Iran), W. Bruce Croft(University of Massachusetts Amherst, USA)

Session 7d: Data Mining
Time: 2:30 PM – 4:00 PM  Location: Theory  Chair: Zhenhui Li

Paper 530: Uncovering the Spatio-Temporal Dynamics of Memes in the Presence of Incomplete Information, Hancheng Ge(Texas A&M University, USA), James Caverlee(Texas A&M University, USA), Nan Zhang(Fudan University, China), Anna Squicciarini(Pennsylvania State University, USA)
Paper 974: From Recommendation to Profile Inference (Rec2PI): A Value-added Service to Wi-Fi Data Mining, Cheng Chen(University of Victoria, Canada), Fang Dong(University of Victoria, Canada), Kui Wu(University of Victoria, Canada), Venkatesh Srinivasan(University of Victoria, Canada), Alex Thomo(University of Victoria, Canada)

Paper 234: On Backup Battery Data in Base Stations of Mobile Networks: Measurement, Analysis, and Optimization, Xiaoyi Fan(Simon Fraser University, Canada), Feng Wang(The University of Mississippi, USA), Jiangchuan Liu(Simon Fraser University, Canada)

Paper 649: Automatic Generation and Validation of Road Maps from GPS Trajectory Data Sets, Hengfeng Li(The University of Melbourne, Australia), Lars Kulik(The University of Melbourne, Australia), Kotagiri Ramamohanarao(The University of Melbourne, Australia)

Session 7e: Network Analysis
Time: 2:30 PM – 4:00 PM     Location: Regency D     Chair: Mohammad Hasan

Paper 202: Fully Dynamic Shortest-Path Distance Query Acceleration on Massive Networks, Takanori Hayashi(The University of Tokyo, Japan), Takuya Akiba(Preferred Networks, Inc, Japan), Ken-ichi Kawarabayashi(National Institute of Informatics, Japan)

Paper 58: Hierarchical and Dynamic k-Path Covers, Takuya Akiba(Preferred Networks, Inc, Japan), Yosuke Yano(Recruit Holdings Co., Ltd., Japan), Naoto Mizuno(The University of Tokyo, Japan)

Paper 1037: Efficient Computation of Importance Based Communities in Web-Scale Networks Using a Single Machine, Shu Chen(University of Victoria, Canada), Ran Wei(University of Victoria, Canada), Diana Popova(University of Victoria, Canada), Alex Thomo(University of Victoria, Canada)

Paper 326: Collective Classification via Discriminative Matrix Factorization on Sparsely Labeled Networks, Daokun Zhang(University of Technology Sydney, Australia), Jie Yin(CSIRO, Australia), Xingquan Zhu(Florida Atlantic University, USA), Chengqi Zhang(University of Technology Sydney, Australia)

Session 7f: Industry Session VII
Time: 2:30 PM – 4:00 PM     Location: Network     Chair: Jiliang Tang

Paper 1406: LogMine: Fast Pattern Recognition for Log Analytics, Hossein Hamooni(University of New Mexico, USA), Biplob Debnath(NEC Laboratories America, USA), Jianwu Xu(NEC Laboratories America, USA), Hui Zhang(NEC Laboratories America, USA), Guofei Jiang(NEC Laboratories America, USA), Abdullah Mueen(University of New Mexico, USA)

Paper 1479: Scaling Up Factorization Machines with Parameter Server, Erheng Zhong(Baidu Research, USA), Yue Shi(Yahoo! Labs, USA), Nathan Liu(Google Inc., USA), Suju Rajan(Criteo Labs, USA)

Paper 1390: DI-DAP: An Efficient Disaster Information Delivery and Analysis Platform in Disaster Management, Tao Li(Florida International University, USA), Wubai Zhou(Florida International University, USA), Chunqiu Zeng(Florida International University, USA), Qing Wang(Florida International University, USA), Qifeng Zhou(Xiamen University, China), Dingding Wang(Florida Atlantic University, USA), Jia Xu(Nanjing University of Posts and Telecommunications, China), Yue Huang(Nanjing University of Posts and Telecommunications, China), Wentao Wang(Florida International University, USA), Minjing Zhang(Florida International University, USA), Steve Lu(Florida International University, USA), Shu-Ching Chen(Florida International University, USA), Naphtali Rishe(Florida International University, USA)

Paper 1381: Approximate Aggregates in Oracle 12C, Hong Su(Oracle, USA), Mohamed Zait(Oracle, USA), Vladimir Barrière(Oracle, USA), Joseph Torres(Oracle, USA), Andre Menck(Oracle, USA)

Session 8a: Learning
Time: 4:30 PM – 6:00 PM     Location: Theory     Chair: Murat Dundar

Paper 378: Supervised Feature Selection by Preserving Class Correlation, Jun Wang(Nankai University, China), Jinmao Wei(Nankai University, China), Zhenglu Yang(Nankai University, China)

Paper 588: CGMOS: Certainty Guided Minority OverSampling, Xi Zhang(Illinois Institute of Technology, USA), Di Ma(Illinois Institute of Technology, USA), Lin Gan(Illinois Institute of Technology, USA), Shanshan Jiang(Illinois Institute of Technology, USA), Gady Agam(Illinois Institute of Technology, USA)
Session 8b: Social Network - Diffusion & Cascade
Time: 4:30 PM – 6:00 PM
Location: Concept
Chair: Elisa Bertino

Paper 112: A Model-Free Approach to Infer the Diffusion Network from Event Cascade, Yu Rong(The Chinese University of Hong Kong, Hong Kong), Qiankun Zhu(The Chinese University of Hong Kong, Hong Kong), Hong Cheng(The Chinese University of Hong Kong, Hong Kong)
Paper 897: Multiple Infection Sources Identification with Provable Guarantees, Hung Nguyen(Virginia Commonwealth University, USA), Preetam Ghosh(Virginia Commonwealth University, USA), Michael Mayo(US Army Engineer RD Center, USA), Thang Dinh(Virginia Commonwealth University, USA)
Paper 1165: Information Diffusion at Workplace, Jiawei Zhang(University of Illinois at Chicago, USA), Philip S. Yu(University of Illinois at Chicago, USA), Yuanhua Lv(Microsoft Research, USA), Qianyi Zhan(Nanjing University, China)
Paper 159: Targeted Influence Maximization in Social Networks, Chonggang Song(National University of Singapore, Singapore), Mong Li Lee(National University of Singapore, Singapore)

Session 8c: Applications
Time: 4:30 PM – 6:00 PM
Location: Regency C
Chair: Xia Ning

Paper 15: Updating an Existing Social Graph Snapshot via a Limited API, Norases Vesdapunt(Stanford University, USA), Hector Garcia-Molina(Stanford University, USA)
Paper 449: Making Sense of Entities and Quantities in Web Tables, Yusra Ibrahim(Max Planck Institute for Informatics, Germany), Mirek Riedewald(Northeastern University, USA), Gerhard Weikum(Max Planck Institute for Informatics, Germany)
Paper 258: Influence Maximization for Complementary Goods: Why Parties Fail to Cooperate? Han-Ching Ou(National Taiwan University, Taiwan R.O.C.), Chung-Kuang Chou(National Taiwan University, Taiwan R.O.C.), Ming-Syan Chen(National Taiwan University, Taiwan R.O.C.)
Paper 1054: Fast Spelling Correction for Eye-based Typing using domain-specific Information about Error Distribution, Raiza Hanada(Institute of Mathematical Sciences and Computing / University of Sao Paulo, Brazil), Maria Pimentel(Institute of Mathematical Sciences and Computing / University of Sao Paulo, Brazil), Marco Cristo(Institute of Computing / Federal University of Amazonas, Brazil), Fernando Lores(Institute of Computing / Federal University of Amazonas, Brazil)

Session 8d: Data Mining Algorithms
Time: 4:30 PM – 6:00 PM
Location: Regency D
Chair: Lei Liu

Paper 350: Computing and Summarizing the Negative Skycube, Nicolas Hanusse(University of Bordeaux - CNRS, France), Patrick Kamnang Wanko(University of Bordeaux - CNRS, France), Sofian Maabout(University of Bordeaux - CNRS, France)
Paper 372: Efficient Orthogonal Non-negative Matrix Factorization over Stiefel Manifold, Wei Emma Zhang(The University of Adelaide, Australia), Mingkui Tan(South China University of Technology, China), Quan Z. Sheng(The University of Adelaide, Australia), Lina Yao(UNSW Australia, Australia), Qingfeng Shi(The University of Adelaide, Australia)
Paper 333: Paired Restricted Boltzmann Machine for Linked Data, Suhang Wang(Arizona State University, USA), Jiliang Tang(Michigan State University, USA), Fred Morstatter(Arizona State University, USA), Huan Liu(Arizona State University, USA)
Paper 616: LDA Revisited: Entropy, Prior and Convergence, Jianwei Zhang(School of Computer Science and Technology, Soochow University, Suzhou 215006, China, China), Jia Zeng(Huawei Noah's Ark Lab, Hong Kong), Mingxuan Yuan(Huawei Noah's Ark Lab, Hong Kong), Weixiong Rao(Tongji University, China), Jianfeng Yan(School of Computer Science and Technology, Soochow University, Suzhou 215006, China, China)
Session 8e: High Performance Towards Big Data
Time: 4:30 PM – 6:00 PM  Location: Regency A&B  Chair: Alfredo Cuzzocrea

Paper 451: Cost-Effective Stream Join Algorithm on Cloud System, Junhua Fang(East China Normal University, China), Rong Zhang(East China Normal University, China), Xiaotong Wang(East China Normal University, China), Tom Z.J. Fu(Advanced Digital Sciences Center, Singapore), Zhenjie Zhang(Advanced Digital Sciences Center, Singapore), Aoying Zhou(East China Normal University, China)
Paper 987: Leveraging Multiple GPUs and CPUs for Graphlet Counting in Large Networks, Ryan Rossi(Palo Alto Research Center, USA), Rong Zhou(Palo Alto Research Center, USA)
Paper 1110: Scalable Local-Recoding Anonymization using Locality Sensitive Hashing for Big Data Privacy Preservation, Xuyun Zhang(The University of Auckland, New Zealand), Christopher Leckie(The University of Melbourne, Australia), Wanchun Dou(Nanjing University, China), Jinjun Chen(University of Technology, Sydney, Australia), Ramamohanarao Kotagiri(The University of Melbourne, Australia), Zoran Salic(The University of Auckland, New Zealand)
Paper 525: Approximate Discovery of Functional Dependencies for Large Datasets, Tobias Bleifuß(Hasso Plattner Institute, Germany), Susanne Bülow(Hasso Plattner Institute, Germany), Johannes Frohnhofen(Hasso Plattner Institute, Germany), Julian Risch(Hasso Plattner Institute, Germany), Georg Wiese(Hasso Plattner Institute, Germany), Sebastian Kruse(Hasso Plattner Institute, Germany), Felix Naumann(Hasso Plattner Institute, Germany)

Session 8f: Industry Session VIII
Time: 4:30 PM – 6:00 PM  Location: Network  Chair: Parikshit Sondhi

Paper 1413: On Structural Health Monitoring Using Tensor Analysis and Support Vector Machine with Artificial Negative Data, Prasad Cheema(CSIRO, Australia), Nguyen Lu Dang Khoa(CSIRO, Australia), Mehrisadat Makki Alamdari(CSIRO, Australia), Wei Liu(University of Technology Sydney, Australia), Yang Wang(CSIRO, Australia), Fang Chen(CSIRO, Australia), Peter Runcie(CSIRO, Australia)
Paper 885: A Self-learning and Online Algorithm for Time Series Anomaly Detection, with Application in CPU Manufacturing, Xing Wang(George Mason University, USA), Jessica Lin(George Mason University, USA), Nitish Patel(Intel Corporation, USA)
Paper 1369: Deep Match between Geology Reports and Well Logs Using Spatial Information, Bin Tong(Central Research Laboratory, Hitachi, Ltd., Japan), Martin Klinkigt(Central Research Laboratory, Hitachi, Ltd., Japan), Makoto Iwayama(Central Research Laboratory, Hitachi, Ltd., Japan), Yoshiyuki Kobayashi(Central Research Laboratory, Hitachi, Ltd., Japan), Anshuman Sahu(Big Data Laboratory, Hitachi America, Ltd., USA), Ravigopal Vennelakanti(Big Data Laboratory, Hitachi America, Ltd., USA)
Paper 1325: MIST: Missing Person Intelligence Synthesis Toolkit, Elham Shaabani(Arizona State University, USA), Hamidreza Alvandari(Arizona State University, USA), Paulo Shakarian(Arizona State University, USA), J.E. Kelly Snyder(Find Me Group, USA)
## Friday Oct. 28, 2016: Workshops

<table>
<thead>
<tr>
<th>Time</th>
<th>Theory</th>
<th>Concept</th>
<th>Regency C&amp;D</th>
<th>Regency E&amp;F</th>
<th>Regency A&amp;B</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00-14:00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Registration (Regency Foyer)</td>
</tr>
<tr>
<td>09:00-10:30</td>
<td>The First International Workshop on Computational Methods for CyberSafety</td>
<td>The Fourth International Workshop on Social Web for Disaster Management</td>
<td>The first Workshop on Big Network Analytics</td>
<td>The Workshop on Data-Driven Talent Acquisition</td>
<td>CIKM Cup Workshop</td>
<td>The Tenth ACM International Workshop on Data and Text Mining in Biomedical Informatics</td>
</tr>
<tr>
<td>10:30-11:00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>11:00-12:30</td>
<td>The First International Workshop on Computational Methods for CyberSafety</td>
<td>The Fourth International Workshop on Social Web for Disaster Management</td>
<td>The first Workshop on Big Network Analytics</td>
<td>The Workshop on Data-Driven Talent Acquisition</td>
<td>CIKM Cup Workshop</td>
<td>The Tenth ACM International Workshop on Data and Text Mining in Biomedical Informatics</td>
</tr>
<tr>
<td>12:30-14:00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lunch Provided (Regency Ballroom)</td>
</tr>
<tr>
<td>14:00-15:30</td>
<td>The First International Workshop on Computational Methods for CyberSafety</td>
<td>The Fourth International Workshop on Social Web for Disaster Management</td>
<td>The first Workshop on Big Network Analytics</td>
<td>The Workshop on Data-Driven Talent Acquisition</td>
<td>The 2nd International Workshop on Data mining meets Visual Analytics at Big Data Era</td>
<td>The Tenth ACM International Workshop on Data and Text Mining in Biomedical Informatics</td>
</tr>
<tr>
<td>15:30-16:00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Coffee Break (Regency Foyer)</td>
</tr>
<tr>
<td>16:00-17:30</td>
<td>The First International Workshop on Computational Methods for CyberSafety</td>
<td>The Fourth International Workshop on Social Web for Disaster Management</td>
<td>The first Workshop on Big Network Analytics</td>
<td>The Workshop on Data-Driven Talent Acquisition</td>
<td>The 2nd International Workshop on Data mining meets Visual Analytics at Big Data Era</td>
<td>The Tenth ACM International Workshop on Data and Text Mining in Biomedical Informatics</td>
</tr>
</tbody>
</table>
Workshops, Oct. 28, 2016, 9am-5:30pm
(Visit each workshop’s website to see its detailed schedule)

   Chairs: Shivakant Mishra (University of Colorado), Qin Lv (University of Colorado), Richard Han (University of Colorado), Jeremy Blackburn (Telefonica Research)
   Location: Theory

   Chairs: Carlos Castillo (EURECAT), Fernando Diaz (Microsoft Research), Yu-Ru Lin (University of Pittsburgh), Jie Yin (CSIRO)
   Location: Concept

   Chairs: Jie Tang (Tsinghua University), Keke Cai (IBM, China Research Lab), Zhong Su (IBM, China Research Lab), Hanghang Tong (Arizona State University), Michalis Vazirgiannis (École Polytechnique), Yang Yang (Zhejiang University)
   Location: Regency C&D

4. DDTA 2016 - The Workshop on Data-Driven Talent Acquisition  https://sites.google.com/site/ddta2016cikm/
   Chairs: Yi Fang (Santa Clara University), Maarten de Rijke (University of Amsterdam), Huangming Xie (LinkedIn)
   Location: Regency E&F

5. ACM DAVA’16: 2nd International Workshop on DAta mining meets Visual Analytics at Big Data Era  http://vis.ios.ac.cn/DAVA16/
   Chairs: Lei Shi (Chinese Academy of Sciences), Hanghang Tong (Arizona State University), Chaoli Wang (University of Notre Dame), Leman Akoglu (Stony Brook University)
   Location: Regency A&B

6. DTMBIO 2016: The Tenth International Workshop on Data and Text Mining in Biomedical Informatics  http://dtmbio.net/dtmbio2016/index.html
   Chairs: Sangwoo Kim (Yonsei University College of Medicine), Jake Chen (University of Alabama at Birmingham), Vincenzo Cutello (University of Catania), Doheon Lee (Korea Advanced Institute of Science and Technology)
   Location: Network

7. CIKM Cup Workshop  http://cikmcup.org/
   Chairs: Nik Spirin (University of Illinois at Urbana-Champaign), Pavel Izhutov, (Stanford University)
   Location: Regency A&B
General Information

**Internet Connection:** Your meeting room High Speed Internet Access is being provided by Presentation Services. To connect, please follow these instructions:

1. Connect to the wireless network called **HYATT-MEETING**.
2. Open your internet browser and you should be transferred to the PSAV login webpage. If you are not transferred, please type google.com into your address bar and click go. This will transfer you to the PSAV login webpage.
3. Please enter the code **TRACK2016** and click on the connect button. Your access code is valid **10-23-16 through 10-28-16**

If you will be connecting to a VPN, please follow the instructions above before connecting to your VPN. If you need assistance, please dial zero on any house phone and ask the operator to radio PSAV.

**Food Ideas:** Downtown Indy's diverse restaurants and bars allow you to choose from nearly 300 establishments including Indy originals and well-known chains. Menus vary from home-grown ingredients infused into healthy entrees to fiery shrimp cocktail to gourmet burgers and huge tenderloins to abundant sushi and everything in between.

**Indianapolis City Market** – Market hours are 7 am - 6 pm Monday through Friday; October 26th is the Original Farmers Market.

**The Rathskeller** – Offers a biergarten for those 21 and older; has traditional German food; live music offered Thursday, October 27th: Green River Ordinance performing from 7 – 11pm; $13 for tickets purchased in advance, $18 tickets day-of show.

**Activities and Attractions in Indianapolis:**

**Indianapolis Motor Speedway** – IMS Museum ticket rates- Adults (16+) $10, Youth (6-15) $5, Children (5 and under) FREE; Museum hours: 9am – 5pm, Monday – Friday

**Indianapolis Museum of Art** – Discounted ticket rates for 15 people or more, self-guided adult group tours: $14 per person (must go through their website [http://imamuseum.org](http://imamuseum.org) and fill out ADULT GROUP TOUR REQUEST FORM); free parking

**Indianapolis Zoo** – Discounted ticket rates to groups of 25 or more; must place ticket order 2 weeks in advance; Dolphin presentations included in admission fee. **Monday – Friday:** Adult (ages 13 – 61) $14; Youth (ages 2 – 12) $10; Senior (age 62+) $13; Parking: $6 per regular vehicle, $12 per large vehicles (including RVs and campers)

**NCAA Hall of Champions** – The NCAA Hall of Champions has two-levels of interactive exhibits to engage visitors and create a true-to-life understanding of what it takes to succeed

* **A Bicentennial Celebration Exhibition:** To help celebrate the Bicentennial of our state, we are highlighting memorable moments in collegiate athletics that have taken place at our member schools in Indiana. It is a collection of photos that help tell the stories throughout the years such as breaking down barriers in diversity, women in sports, celebrated teams and incredible individual achievements.
Discounted rates for groups of 15 or more; $4 for adults, $2 for seniors (65+), $3 for youth (ages 6-17), free for ages 5 and under; parking located in White River State Park Underground Garage.

**Easley Winery** – **DAILY WINE TASTINGS:** $5.00, Includes Sampling of 7 Wines; For groups larger than 8 people, please call ahead at (317) 636-4516; free parking available in surrounding parking lots and street; Open Mon-Sat 11am – 6pm

**FREE EVENTS**

**Butler University:**

**Oct 25 – Duckwall Artist Series: Larry Shapiro, Violin 7:30pm showtime** at Eidson-Duckwall Recital Hall

“The Duckwall Artist Series presents intimate recitals featuring Butler University School of Music faculty and guest artists. This event is free and open to the public.”

**Oct 25 – Visiting Writers Series: Robert Wrigley 7:30pm showtime** at Atherton Union, Reilly Room

“His poems are concerned with rural Western landscapes and humankind’s place within the natural world, and he aims to “tell all the truth, but make it sing.” He has been awarded fellowships from the National Endowment for the Arts, the Rockefeller Foundation, the Guggenheim Foundation, and the Idaho Commission on the Arts. His poems have been widely anthologized, twice included in Best American Poetry, and featured on NPR’s The Writer’s Almanac. Free and open to the public.”

**The Indianapolis Public Library – The Indiana History Club: How We All Save History**

This special exhibit in the Nina Mason Pulliam Indianapolis Special Collections Room highlights how Indiana's history has been recorded and disseminated through various types of media and the creators behind them. The exhibit will include examples of personal items, such as photographs, letters and media and how those have been preserved. It also highlights the role of such institutions as the Indiana State Library, Indiana Historical Society and Indiana State Archives.

**Hours:**

- **Monday – Wednesday:** 10am – 8pm
- **Thursday:** 10am – 6pm
- **Friday:** 10am – 5pm

**Soldiers' and Sailors Monument** – The State Soldiers and Sailors Monument is in the center of downtown on Monument Circle. The monument's original purpose was to honor Hoosiers who were veterans of the American Civil War; however, it is also a tribute to Indiana's soldiers who served during the American Revolutionary War, the War of 1812, the Mexican–American War, and the Spanish–American War. The monument is the first in the United States to be dedicated to the common soldier. The Colonel Eli Lilly Civil War Museum on the interior of the monument is open for free to the public 10:30 AM - 5:30 PM, Wednesday - Sunday (between May and October).

**Gallery 924 at the Arts Council** – Susan Tennant Solo Exhibition

**Hours:** Monday – Friday, 9am – 5pm

**Free admission**

Check out the CIKM’16 website for more information about many interesting activities you can do and attractions you can visit in Indianapolis: http://www.cikm2016.org/
Hyatt Second Floor Map

System Demonstrations

Network

Concept

Workshops & Parallel Sessions

Poster Presentations

Career Day

Tutorials

Theory

Regency Ballroom

All Plenary Sessions
(Regency Ballroom = Regency A+B+C+D)