



42nd International Conference on  
**Very Large Data Bases**

New Delhi, India • September 5-9, 2016



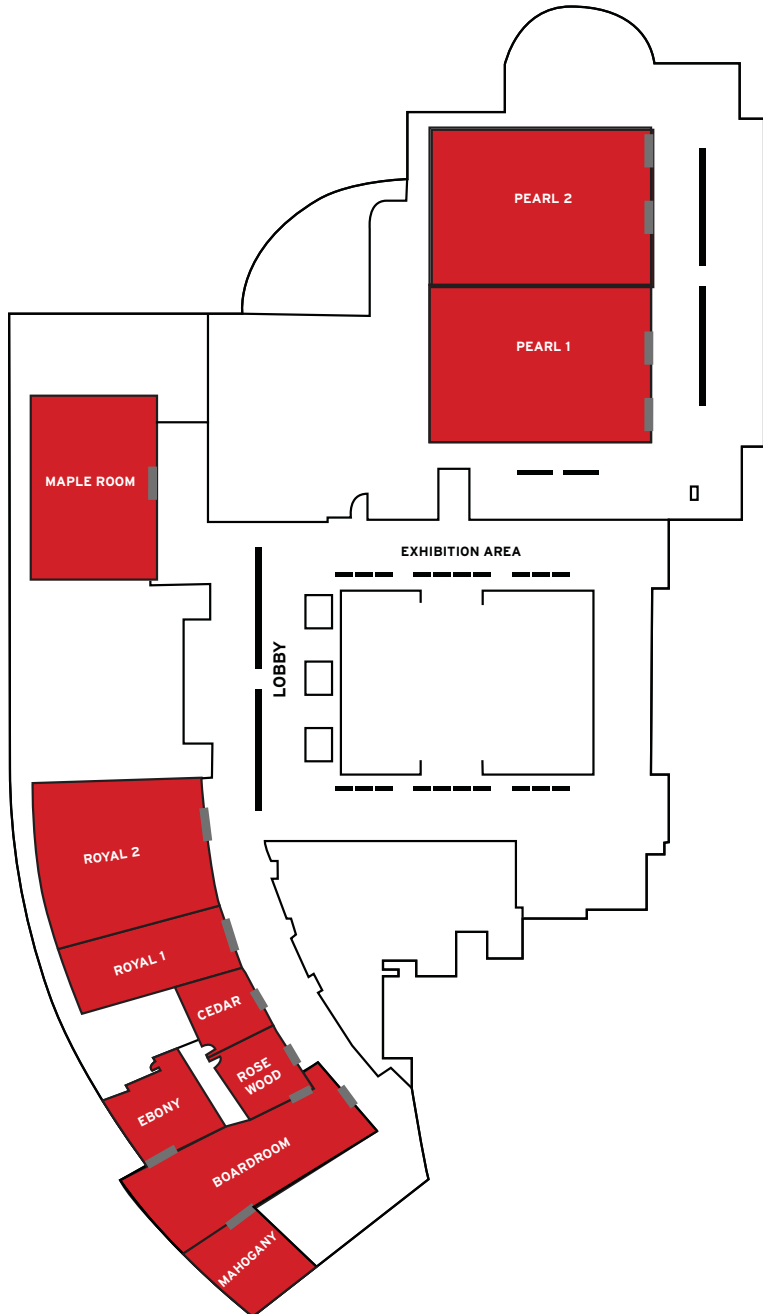
# Table of Contents

---

Conference Hotel Floor Plan .....	3
Welcome Message from the Conference General Chairs .....	4
Message from the Technical Program Chairs .....	5
Conference Organization .....	7
PVLDB Review Board .....	9
Industrial and Demo PC Members .....	11
Awards .....	12
Keynotes .....	15
Program at a Glance .....	16
DAY 1 - Monday, September 5th .....	20
DAY 2 - Tuesday, September 6th .....	24
Day 3 - Wednesday, September 7th .....	32
Day 4 - Thursday, September 8th .....	40
Day 5 - Friday, September 9th.....	48
Tutorial Abstracts .....	53
Research Paper Abstracts .....	57
Industrial Track Paper Abstracts.....	87
Demonstration Track Paper Abstracts.....	93
Sponsors .....	102
Notes.....	123

# Conference Hotel Floor Plan

---



# Welcome Message from the Conference General Chairs

---



As the General Chairs of VLDB 2016, we welcome you to the 42nd International Conference on Very Large Databases (VLDB) and to the capital city of India, New Delhi. The VLDB conferences started in 1975. It took a little over 20 years for a VLDB to be held in India in 1996 in Bombay (now Mumbai) and it had a great impact on the database community in India. And, we have now a VLDB back in India exactly after 20 years! We are all excited to have this in our capital city.

The technical program is the core of such a conference and we would like to thank the Program Committee chairs and their team for putting together an exciting technical program and shouldering the bulk of the work that went into organizing the VLDB2016 conference.

To enhance local participation and to create excitement about data-oriented research we are organizing a one-day Digital India event on September 5th. As a part of this event, we are running a data-centric hackathon where more than 7500 programmers have registered.

Organizing an event of this nature requires effort from a large number of volunteers. Thanks are due to all the conference officers and other volunteers who contributed in organizing the event. We are also grateful to their respective organizations for allowing them to spend their time on the conference activities and in a few cases even the generous use of their resources. Last but not the least, we owe a great thank you to the sponsors for their generous support.

We hope you find the technical program inspiring. We also hope you find the stay in New Delhi enjoyable. New Delhi happens to be a cultural capital too and it has many cultural and tourist offerings. During your stay, we hope you will have a chance to experience them.



**Anand Deshpande and T.M. Vijayaraman**

General Conference Chairs, VLDB 2016

# Message from the Technical Program Chairs

---



On behalf of the technical program committee, we extend you a warm welcome to VLDB 2016 in New Delhi, India. Similar to past VLDB conferences, we received a large number of submissions to the Research, Industrial and Demonstration tracks, clearly showing that the state of our research community continues to be strong. The submissions covered a wide spectrum of topics that included not only traditional concerns of database systems, but also emerging areas such as graph data management, social networks and communities, and hardware-software co-design.

The research track (published in Vol 9 of PVLDB) received a record number of 719 submissions during April 2015 to March 2016, with 104 papers accepted for presentation at this year's conference. In addition, twelve research papers from Vol 8 of PVLDB, whose processing could not be completed in time for the VLDB 2015 conference, will also be presented here. Moreover, apart from the standard stage presentation, all research papers will also be separately displayed in poster sessions, along with recent papers from the VLDB Journal.

VLDB conferences have always included a strong industrial track, testifying to the importance of researchers exchanging ideas with practitioners - we are proud to showcase an excellent industrial program this time too, thanks to our Industrial Chairs and their committee. This track received 53 submissions, of which 17 were accepted for presentation. We also feature three invited talks from leading industrial speakers.

The demonstration track provides a ringside opportunity for attendees to witness novel software prototypes of database technology at different stages of development. Our Demo Chairs and their committee selected 37 attractive demos from the 103 submissions received by this track.

Our conference also hosts an excellent set of tutorials, hand-picked by our Tutorial Chairs, consisting of six lectures covering important emerging areas, and helping researchers keep pace with our fast changing world of database technology. We also incorporate an exciting panel, organized by our Panel Chairs, debating the influence of Artificial Intelligence on contemporary data management issues.

We are delighted to have two excellent keynote speakers spearheading the conference. Professor Ion Stoica from University of California, Berkeley will be sharing his insights on the Big Data revolution, while Dr. Anand Rajaraman, well-known as a successful serial entrepreneur from Silicon Valley, will be speaking on data-driven disruption. We are confident that you will find the talks thought-provoking and educative.

Another highlight of the conference is the awards session on Thursday. In addition to the Best Paper and Best Demonstration awards for the conference, this session will include presentation of three VLDB Endowment awards: Early Career Researcher, Women in Database Research, and 10-year Best Paper. The achievements of these award winners are highlighted in this booklet, and they will be presenting short talks during the session. We warmly congratulate all of them on the impressive technical contributions that have led to these recognitions.

# Message from the Technical Program Chairs

---

In addition to the core conference, our technical program features seven thematic workshops chosen by our Workshop Chairs, covering a range of modern topics. There is also a PhD workshop to provide feedback and guidance to select young researchers on their ongoing work.

With the richness of the technical program, comes the challenge of organizing the sessions in a coherent and manageable framework. In this regard, VLDB 2016 has made some notable changes. For instance, all three-hour tutorials have been moved to the day before the conference. We have leveraged this shift to reduce the number of parallel tracks to five during the core conference days, thus helping the attendees engage with an appreciably larger part of the conference content. In addition, similar to demonstrations, the research poster presentations are also spread across multiple sessions and days, providing ample opportunity to interact with presenters. Finally, all presentations are of at least twenty minutes' duration, ensuring adequate time for the speakers to convey their creative ideas to the audience.

As you can imagine, putting together the technical program for VLDB 2016 required dedicated work by an army of volunteers. We are extremely grateful to our track chairs, as well as the associate editors of PVLDB Vol 9. They bore the burden of decision making in a timely manner, and to reach their decisions, they relied on program committee members and external reviewers, whose collective and painstaking efforts have resulted in the exciting program that you will enjoy over the course of the conference. We are also grateful to the VLDB Endowment Awards Committee and the VLDB 2016 Best Paper Award Committee for their generous and timely support.

As in past years, VLDB 2016 relied on Microsoft's CMT Software Service for handling submissions and the reviewing process. We thank the staff of the CMT for their prompt help in customizing the software to meet our requirements.



**Surajit Chaudhuri and Jayant Haritsa**

VLDB 2016 Program Chairs

# Conference Organization

---

## General Conference Chairs

Anand Deshpande (Persistent Systems, India)

T M Vijayaraman (Persistent Systems, India)

## Organization Committee Chair

Mukesh Mohania (IBM Research, India and Australia)

## Program Committee Chairs

Surajit Chaudhuri (Microsoft Research, USA)

Jayant Haritsa (IISc Bangalore, India)

## Research Track Associate Editors

Daniel Abadi (Yale University, USA)

Peter Boncz (CWI Amsterdam, Netherlands)

Yanlei Diao (University of Massachusetts - Amherst, USA)

Anhai Doan (University of Wisconsin - Madison, USA)

Amr El Abbadi (University of California - Santa Barbara, USA)

Alan Fekete (University of Sydney, Australia)

Minos Garofalakis (TU Crete, Greece)

Alon Halevy (Recruit Institute of Technology, USA)

Ihab Ilyas (University of Waterloo, Canada)

Donald Kossman (Microsoft Research, USA)

Sharad Mehrotra (University of California - Irvine, USA)

Sunita Sarawagi (IIT Bombay, India)

Kyuseok Shim (Seoul National University, Korea)

Kian-Lee Tan (NUS, Singapore)

## Treasury Chair

Rohit Kamat (Persistent Systems, India)

## VLDB Endowment Representative

S Sudarshan (IIT Bombay, India)

## Sponsorship Committee Chairs

Michael Carey (University of California - Irvine, USA)

Anand Deshpande (Persistent Systems, India)

## Publicity Committee Chairs

Rajkumar Sen (MemSQL, USA)

Manish Bhide (IBM Software Lab, India)

## Tutorial Chairs

Chris Re (Stanford University, USA)

S Sudarshan (IIT Bombay, India)

Gerhard Weikum (Max-Planck Institute for Informatics, Germany)

## Industrial Chairs

Gustavo Alonso (ETH Zurich, Switzerland)

Anant Jhingran (Apigee, USA)

S Seshadri (Zettata, India)

## Demo Chairs

Mike Cafarella (University of Michigan - Ann Arbor, USA)

Stratos Idreos (Harvard University, USA)

Anthony Tung (NUS, Singapore)

## Panel Chairs

Juliana Freire (New York University, USA)

Sam Madden (MIT, USA)

## Workshop Chairs

Sihem Amer-Yahia (CNRS, France)

Joe Hellerstein (University of California - Berkeley, USA)

## PhD Workshop Chairs

Torsten Grust (University of Tübingen, Germany)

Kamal Karlapalem (IIIT-Hyderabad/IIT-Gandhinagar, India)

Andy Pavlo (CMU, USA)

## Proceedings Chairs

Srikanta Bedathur (IBM Research, India)

Katja Hose (Aalborg University, Denmark)

## VLDB Best Paper Awards Committee

Jignesh Patel (University of Wisconsin, Madison, USA)

Jeffrey D. Ullman (Stanford University, USA)

Gerhard Weikum (Max-Planck Institute for Informatics, Germany)

## VLDB Endowment Awards Committee

Sunita Sarawagi (IIT Bombay, India)

Surajit Chaudhuri (Microsoft Research, USA)

Alon Halevy (Recruit Institute of Technology, USA)

Christian Jensen (Aarhus University, Denmark)

Gerhard Weikum (Max-Planck Institute for Informatics, Germany)

## Organization Committee

Rashmi Tambe (Persistent Systems, India)

Pratyush Panangipalli (Persistent Systems, India)

# Conference Organization

---

## **Student Grant Coordinators**

Kamal Karlapalem (IIIT-Hyderabad/IIIT-Gandhinagar, India)

Rekha Singhal (TCS, India)

Srinath Srinivasa (IIIT-Bangalore)

## **Volunteers Coordinator**

Shyam Gupta (IIT Delhi, India)

## **Registration Coordinators**

Rajeev Gupta (IQLECT, India)

Rema Ananthanarayanan (IBM Research - India)

## **Travel and Transportation Coordinator**

Ramesh Agrawal (JNU, India)

Subhash Bhalla (University of Aizu - Japan)

## **Sessions Coordinators**

Vasudha Bhatnagar (Delhi University, India)

Danish Contractor (IBM Research - India)

## **Cultural Program Coordinator**

Nitendra Rajput (IBM Research - India)

## **Food Coordinator**

Naveen Gupta (Delhi University, India)



# PVLDB Review Board

---

Ashraf Aboulmaga (Qatar Computing Research Institute, Qatar)

Marcos K. Aguilera (VMware Research, USA)

Yael Amsterdamer (Tel-aviv University, Israel)

Walid Aref (Purdue University, USA)

Marcelo Arenas (Pontifical Catholic University, Chile)

Shivnath Babu (Duke University, USA)

Magdalena Balazinska (University of Washington, USA)

Zohra Bellahsene (University of Montpellier, France)

Arnab Bhattacharya (IIT Kanpur, India)

Francesco Bonchi (Yahoo Labs Barcelona, Spain)

Nico Bruno (Google, USA)

Michael Cafarella (University of Michigan, USA)

Andrea Cali (Birkbeck University of London, UK)

Reynold C. K. Cheng (University of Hong Kong, Hong Kong)

Venkatesan Chakaravarthy (IBM Research, India)

Kaushik Chakrabarti (Microsoft Research, USA)

Arbee L P Chen (National Chengchi University, Taiwan)

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

Shimin Chen (Inst.of Computing Technology, China)

Rada Chirkova (North Carolina State University, USA)

Tyson Condie (University of California, Los Angeles, USA)

Isabel Cruz (University of Illinois at Chicago, USA)

Bin Cui (Peking University, China)

Gautam Das (University of Texas, Arlington, USA)

Mahashweta Das (HP Labs, USA)

Sudipto Das (Microsoft Research, USA)

Antonios Deligiannakis (Technical University of Crete, Greece)

Prasad Deshpande (IBM Research, India)

Daniel Deutch (Tel Aviv University, Israel)

Jens Dittrich (Saarland University, Germany)

Gillian Dobbie (University of Auckland, NewZealand)

Xin Luna Dong (Google, USA)

Aaron Elmore (University of Chicago, USA)

Wenfei Fan (University of Edinburgh, UK)

Hakan Ferhatosmanoglu (Bilkent University, Turkey)

Florin Geerts (University of Antwerp, USA)

Johannes Gehrke (Microsoft, USA)

Aristides Gionis (Aalto University, Finland)

Sergio Greco (University of Calabria, Italy)

Haryadi Gunawi (University of Chicago, USA)

Wook-Shin Han (Pohang University of Science and Tech., Korea)

Stavros Harizopoulos (Amazon, USA)

Herodotos Herodotou (Cyprus University of Technology, Cyprus)

Melanie Herschel (University of Stuttgart, Germany)

Zachary Ives (University of Pennsylvania, USA)

Christopher Jermaine (Rice University, USA)

Ryan Johnson (LogicBlox, USA)

Theodore Johnson (AT&T Labs, USA)

Murat Kantarcioglu (University of Texas, Dallas, USA)

Raghav Kaushik (Microsoft Research, USA)

Ken-ichi Kawarabayashi (National Inst.of Informatics, Japan)

Bettina Kemme (McGill University, Canada)

Eamonn Keogh (University of California, Riverside, USA)

Florian Kerschbaum (SAP, Germany)

Dan Kifer (Pennsylvania State University, USA)

Masaru Kitsuregawa (Tokyo University, Japan)

Nick Koudas (University of Toronto, Canada)

Tim Kraska (Brown University, USA)

Ravi Kumar (Google, USA)

Laks V.S. Lakshmanan (University of British Columbia, Canada)

Justin Levandoski (Microsoft Research, USA)

Feifei Li (University of Utah, USA)

Guoliang Li (Tsinghua University, China)

Jimmy Lin (University of Maryland, USA)

Qiong Luo (Hong Kong University of Science and Technology, Hong Kong)

## PVLDB Review Board

---

Ashwin Machanavajjhala (Duke University, USA)  
Nikos Mamoulis (University of Hong Kong, Hong Kong)

Stefan Manegold (Centrum Wiskunde and Informatica, Netherlands)

Ioana Manolescu (Inria Saclay, France)

Norman May (SAP AG Walldorf, Germany)

Sergey Melnik (Google, USA)

Gerome Miklau (University of Massachusetts, Amherst, USA)

Mohamed Mokbel (University of Minnesota, USA)

Bongki Moon (Seoul National University, Korea)

Barzan Mozafari (University of Michigan, USA)

Arnab Nandi (The Ohio State University, USA)

Vivek Narasayya (Microsoft Research, USA)

Mario A. Nascimento (University of Alberta, Canada)

Rimma Nehme (Microsoft Gray Systems Lab, USA)

Thomas Neumann (Technical University Munich, Germany)

Fatma Ozcan (IBM Research, Almaden, USA)

Themis Palpanas (Paris Descartes University, France)

Hweehwa Pang (Singapore Management University, Singapore)

Olga Papaemmanouil (Brandeis University, USA)

Aditya Parameswaran (University of Illinois, Urbana-Champaign, USA)

Srinivasan Parthasarathy (The Ohio State University, USA)

Jignesh Patel (University of Wisconsin, Madison, USA)

Andy Pavlo (Carnegie Mellon University, USA)

Jian Pei (Simon Fraser University, Canada)

Neoklis Polyzotis (University of California, Santa Cruz, USA)

Dan Ports (University of Washington, USA)

Rachel Pottinger (University of British Columbia, Canada)

Ganesh Ramakrishnan (IIT Bombay, India)

Vijayshankar Raman (IBM Research, USA)

Rajeev Rastogi (Amazon, India)

Matthias Renz (Ludwig Maximilians University, Germany)

Kenneth Ross (Columbia University, USA)

Pierangela Samarati (University of Milan, Italy)

Atish Das Sarma (eBay Research, USA)

Bernhard Seeger (University of Marburg, Germany)

Sudipta Sengupta (Microsoft Research, USA)

Cyrus Shahabi (University of Southern California, USA)

Mohamed Soliman (Pivotal, USA)

Torsten Suel (New York University, USA)

Letizia Tanca (Politecnico di Milano, Italy)

Sandeep Tata (Google, USA)

Nesime Tatbul (Intel Labs and MIT, USA)

Doug Terry (Samsung Research, USA)

Evimaria Terzi (Boston University, USA)

Jens Teubner (TU Dortmund University, Germany)

Martin Theobald (Ulm University, Germany)

Srikanta Tirthapura (Iowa State University, USA)

Kristin Tufte (Portland State University, USA)

Jeffrey Ullman (Stanford University, USA)

Jaideep Vaidya (Rutgers University, USA)

Vasilis Vassalos (Athens University of Economics and Business, Greece)

Haixun Wang (Google, USA)

Wei Wang (University of New South Wales, Australia)

Steven Whang (Google, USA)

Sai Wu (Zhejiang University, China)

Chuan Xiao (Nagoya University, Japan)

Jeffrey Xu Yu (Chinese University of Hong Kong, Hong Kong)

Cong Yu (Google, USA)

Nan Zhang (George Washington University, USA)

Xiaofang Zhou (University of Queensland, Australia)

Lei Zou (Peking University, China)

# Industrial and Demo PC Members

---

## Industrial Track PC Members

Cagri Balkesen (Oracle Labs, Switzerland)  
Sang K. Cha (Seoul National University, South Korea)  
Shimin Chen (ICT, China)  
Wei Fan (Huawei Labs, Hong Kong)  
Ioana Giurgiu (IBM Research, Switzerland)  
Gernot Heiser (NICTA, Australia)  
Wolfgang Lehner (TU Dresden, Germany)  
Ippokratis Pandis (Amazon, USA)  
Shankar Pasupathy (NetApp, USA)  
Marta Patiño (UPM, Spain)  
Vishy Poosala (Facebook, USA)  
Jun Rao (Confluent, USA)  
Prasan Roy (Sclera, India)  
Tudor-Ioan Salomie (Google, Switzerland)  
Sunil Sayyaparaju (Aerospike, India)  
Liuba Shrira (Brandeis University, USA)  
Swami Sivasubramanian (Amazon, USA)  
Ruihua Song (Microsoft Research, China)  
Jun Tatemura (NEC Labs, USA)  
Yuanyuan Tian (IBM Research, USA)  
Philipp Unterbrunner (Snowflake Computing, USA)  
Lan Jun Wang (IBM Research, China)

## Demo Track PC Members

Ioannis Alagiannis (Microsoft, USA)  
Manos Athanassoulis (Harvard University, USA)  
Zhifeng Bao (RMIT University, Australia)  
Carsten Binnig (Brown University, USA)  
Vanessa Braganholo (Universidade Federal de Minas Gerais, Brazil)  
Gao Cong (Nanyang Technological University, Singapore)  
Jing Gao (SUNY Buffalo, USA)  
Todd Green (LogicBlox, USA)  
Jan Hidders (University of Delft, Netherlands)  
Alekh Jindal (Microsoft, USA)  
Panos Kalnis (King Abdullah University of Science and Technology, Saudi Arabia)  
Verena Kantere (University of Geneva, Switzerland)

Zoi Kaoudi (Qatar Computing Research Institute, Qatar)  
Panagiotis Karras (Skoltech University, Russia)  
Oliver Kennedy (SUNY Buffalo, USA)  
Min-Soo Kim (DGIST, South Korea)  
Arnd Christian König (Microsoft Research, USA)  
Georgia Koutrika (HP Labs, USA)  
Harumi Kuno (HP Labs, USA)  
Jae-Gil Lee (KAIST, South Korea)  
Erietta Liarou (Harvard University, USA)  
Eric Lo (Polytechnic University of Hong Kong, China)  
Guy Lohman (IBM Research, USA)  
Mirella Moro (Universidade Federal de Minas Gerais, Brazil)  
Wilfred Ng (Hong Kong University of Science and Technology, China)  
Kjetil Nørvåg (Norwegian University of Science and Technology, Norway)  
Tamer Ozsu (University of Waterloo, Canada)  
Michalis Petropoulos (Amazon, USA)  
Mohammad Sadoghi (IBM Research, USA)  
Yasushi Sakurai (Kumamoto University, Japan)  
Mohamed Sharaf (University of Queensland, Australia)  
Lidan Shou (Zhejiang University, China)  
Manish Singh (IIT Hyderabad, India)  
Kostas Stefanidis (ICS-FORTH, Greece)  
Julia Stoyanovich (Drexel University, USA)  
Danai Symeonidou (INRIA, France)  
Ernest Teniente (Universidade Politècnica de Catalunya, Spain)  
Pinar Tözün (IBM Research, USA)  
Xiaoli Wang (Xiamen University, China)  
Raymond Wong (Hong Kong University of Science and Technology, China)  
Demetris Zeinalipour (University of Cyprus, Cyprus)  
Ce Zhang (University of Wisconsin, USA)  
Feida Zhu (Singapore Management University, Singapore)

# Awards

Session Chair: Sunita Sarawagi (IIT Bombay)

## VLDB Early Career Research Contribution Award

**Xin Luna Dong (Amazon) for "Advancing the state of the art of knowledge fusion"**



**Xin Luna Dong** is a Principal Scientist at Amazon, leading the efforts of constructing Amazon Product Graph. She was one of the major contributors to the Knowledge Vault project, and has led the Knowledge-based Trust project, which is called the "Google Truth Machine" by Washington's Post. She has co-authored book "Big Data Integration", published 65+ papers in top conferences and journals, given 20+ keynotes/invited-talks/tutorials, and got the Best Demo award in Sigmod 2005. She is the PC co-chair for Sigmod 2018 and WAIM 2015, and serves as an area chair for Sigmod 2017, Sigmod 2015, ICDE 2013, and CIKM 2011

## VLDB Women in Database Research Award

**Magdalena Balazinska (University of Washington, Seattle) for "Inspirational research on scalable distributed data systems"**



**Magdalena Balazinska** is the Associate Jean Loup Baer Professor of Computer Science and Engineering at the University of Washington. She's the director of the UW IGERT PhD Program in Big Data and Data Science and the associated PhD Option in Advanced Data Science. She's also a Senior Data Science Fellow of the UW eScience Institute. Magdalena holds a Ph.D. from the Massachusetts Institute of Technology (2006). She is a Microsoft Research New Faculty Fellow (2007), received an NSF CAREER Award (2009), a 10-year most influential paper award (2010), a Google Research Award (2011), an HP Labs Research Innovation Award (2009 and 2010), a Rogel Faculty Support Award (2006), a Microsoft Research Graduate Fellowship (2003-2005), and multiple best-paper awards.

## VLDB 10-year Best Paper Award

**The New Casper: Query Processing for Location Services without Compromising Privacy by Mohamed F. Mokbel, Chi-Yin Chow, Walid G. Aref (VLDB 2006) for "Inspiring research to support privacy in spatial query processing"**



**Mohamed F. Mokbel** (Ph.D., Purdue University, MS, B.Sc., Alexandria University) is an Associate Professor in the Department of Computer Science and Engineering, University of Minnesota. His research interests include the interaction of GIS and location-based services with database systems and cloud computing. His research work has been recognized by five Best Paper Awards and by the NSF CAREER award. Mohamed is/was the program co-chair of SIGMOD 2018, ACM SIGSPATIAL from 2008 to 2010, and IEEE MDM 2011 and 2014. He is an Associate Editor for ACM TODS, ACM TSAS, VLDB journal, and GeoInformatica. Mohamed is currently serving as the elected Chair of ACM SIGSPATIAL 2014-2017. For more information, please visit: [www.cs.umn.edu/~mokbel](http://www.cs.umn.edu/~mokbel)



**Chi-Yin Chow** received the B.A. and M.Phil. degrees from The Hong Kong Polytechnic University, Hong Kong in 2002 and 2005, respectively. He also received the M.S. and Ph.D. degrees from the University of Minnesota-Twin Cities, USA in 2008 and 2010, respectively. He is currently an assistant professor in Department of Computer Science, City University of Hong Kong. His research interests include big data analytics, data management, GIS, mobile computing, location-based services, and data privacy. He is the co-founder and co-chair of the ACM SIGSPATIAL MobiGIS 2012 to 2016, and the editor of the ACM SIGSPATIAL Newsletter. Dr. Chow received the best paper awards in ICA3PP 2015 and IEEE MDM 2009.

# Awards



**Walid G. Aref** is a professor of computer science at Purdue. His research interests are in extending the functionality of database systems in support of emerging applications, e.g., spatial, spatio-temporal, multimedia, biological, and sensor databases. He is also interested in query processing, indexing, data mining, and geographic information systems (GIS). Professor Aref's research has been supported by the National Science Foundation, the National Institute of Health, Purdue Research Foundation, Qatar National Research Foundation, CERIAs, Panasonic, and Microsoft Corp. In 2001, he received the CAREER Award from the National Science Foundation and in 2004, he received a Purdue University Faculty Scholar award. Professor Aref is a member of Purdue's CERIAs. He is an associate editor of the ACM Transactions of Database Systems (ACM TODS) and the ACM Transactions of Spatial Algorithms and Systems (TSAS), and has been an editor of the VLDB Journal. He is a senior member of the IEEE, and a member of the ACM. Professor Aref is an executive committee member, co-founder, and the past chair of the ACM Special Interest Group on Spatial Information (SIGSPATIAL).

## VLDB Best Paper Award

**Compressed Linear Algebra for Large-Scale Machine Learning by Ahmed Elgohary (University of Maryland), Matthias Boehm (IBM Research - Almaden), Peter Haas (IBM Research), Fred Reiss (IBM Research - Almaden), and Berthold Reinwald (IBM Research, Almaden) for "Proposing a comprehensive solution to the challenge of compressing matrices for linear algebra operations in the context of machine learning tasks".**



**Ahmed Elgohary** is a second year PhD student at the computer science department of the University of Maryland, College Park. He got his master's degree at the University of Waterloo, Canada and bachelor's degree at Alexandria University, Egypt both in Computer Engineering. Ahmed is interested in designing algorithms for natural language processing and building scalable infrastructure for machine learning workloads.



**Matthias Boehm** is a Research Staff Member at IBM Research - Almaden, where he is working since 2012 on optimization and runtime techniques for declarative, large-scale machine learning in SystemML. He received his Ph.D. from Technische Universitaet Dresden in 2011 with a dissertation on cost-based optimization of integration flows under the supervision of Prof. Wolfgang Lehner. His previous research also includes systems support for time series forecasting as well as in-memory indexing and query processing.



**Peter J. Haas** has been a Research Staff Member at the IBM Almaden Research Center since 1987 where he has pursued research at the interface of information management, applied probability, statistics, and computer simulation. He is also a Consulting Professor in the Department of Management Science and Engineering at Stanford University, teaching and conducting research in stochastic modeling and simulation. He was designated an IBM Master Inventor in 2012, and his ideas have been incorporated into products including IBM's DB2 database system. He is an ACM Fellow, and has received a number of awards from IBM and both the Simulation and Computer Science communities, including an ACM SIGMOD 10-year Best Paper award for his work on sampling-based exploration of massive datasets. Other work has included the Splash platform for collaborative modeling and simulation, techniques for massive-scale matrix completion, Monte Carlo methods for scalable querying and machine learning over massive uncertain data, automated relationship discovery in databases, query optimization methods, and autonomic computing. He serves on the editorial boards of ACM TODS, Operations Research and ACM Transactions on Modeling and Computer Simulation, and was an Associate Editor for the VLDB Journal from 2007 to 2013. He is the author of over 100 conference publications, journal articles, and books.

# Awards

---



**Fred Reiss** is the Chief Architect at the IBM Spark Technology Center in San Francisco and is one of the founding employees of the Center. Fred received his Ph.D. from UC Berkeley in 2006, then worked for IBM Research -- Almaden for the next nine years. At Almaden, Fred worked on the SystemML and SystemT projects, as well as on the research prototype of DB2 with BLU Acceleration.



**Berthold Reinwald** is a Principal RSM at the IBM Almaden Research Center. His research areas include scalable analytics platforms and database technology. He is the technical lead for SystemML. He holds a Ph.D. in Computer Science from the University of Erlangen-Nuernberg, Germany.

# Keynotes

## Keynote 1 : Ion Stoica

### Trends and Challenges in Big Data Processing

**Session Chair: Surajit Chaudhuri (Microsoft Research)**

Almost six years ago we started the Spark project at UC Berkeley. Spark is a cluster computing engine that is optimized for in-memory processing, and unifies support for a variety of workloads, including batch, interactive querying, streaming, and iterative computations. Spark is now the most active big data project in the open source community, and is already being used by over one thousand organizations.

One of the reasons behind Spark's success has been our early bet on the continuous increase in the memory capacity and the feasibility to fit many realistic workloads in the aggregate memory of typical production clusters. Today, we are witnessing new trends, such as Moore's law slowing down, and the emergence of a variety of computation and storage technologies, such as GPUs, FPGAs, and 3D Xpoint. In this talk, I'll discuss some of the lessons we learned in developing Spark as a unified computation platform, and the implications of today's hardware and software trends on the development of the next generation of big data processing systems.



**Bio: Ion Stoica** is a Professor in the EECS Department at University of California at Berkeley. He does research on cloud computing and networked computer systems. Past work includes the Dynamic Packet State (DPS), Chord DHT, Internet Indirection Infrastructure (i3), declarative networks, replay-debugging, and multi-layer tracing in distributed systems. He is an ACM Fellow and has received numerous awards, including the SIGCOMM Test of Time Award (2011), and the ACM doctoral dissertation award (2001). In 2006, he co-founded Conviva, a startup to commercialize technologies for large scale video distribution, and in 2013, he co-founded Databricks a startup to commercialize Apache Spark.

## Keynote 2 : Anand Rajaraman

### Data-Driven Disruption: The View from Silicon Valley

**Session Chair: Jayant Haritsa (IISc, Bangalore)**

We live in an era where software is transforming industries, the sciences, and society as a whole. This exciting phenomenon has been described by the phrase "software is eating the world." It is becoming increasingly apparent that data is the fuel powering software's conquests. Data is the new disruptor.

It's hard to believe that the first decade of the Big Data era is already behind us. Silicon Valley has been at the forefront of developing and applying data-driven approaches to create disruption at many levels: infrastructure (e.g., Hadoop and Spark), capabilities (e.g., image recognition and machine translation), and killer apps (e.g., self-driving cars and messaging bots). In this talk, we first look back on the past decade and share learnings from the frontlines of data-driven disruption. Looking ahead, we then describe challenges and opportunities for the next decade. Since this has also been a personal journey, we will use examples drawn from personal experience to illustrate each point.



**Bio: Anand Rajaraman** is a Founding Partner of two Silicon Valley venture capital funds focused on early-stage technology companies: Milliways Ventures and RocketshipVC. He was the co-founder of two successful startups: Junglee (acquired by Amazon.com) and Kosmix (acquired by Walmart). At Walmart, he created and led WalMartLabs (as its Senior Vice President). As an academic, Anand's research has focused at the intersection of database systems, the World-Wide Web, and social media. His research publications have won several awards at prestigious academic conferences, including three retrospective 10-year Best Paper awards at ACM SIGMOD and VLDB, and ICDT. His textbook "Mining of Massive Datasets", co-authored with Jeff Ullman and Jure Leskovec, has been published.



# Program



# Program at a Glance

DAY 1: MONDAY [Sep 5]					
Time	Track 1 (Pearl 1)	Track 2 (Maple)	Track 3 (Royal 1)	Track 4 (Royal 2)	Track 5 (Boardroom)
9:00-10:30	Tutorial 1: Differential Privacy in the Wild	Tutorial 2: Operational Analytics Data Management Systems	Workshop 1: TPCTC	Workshop 2: BIRTE	Workshop 3: QDB
10:30-11:00	Tea Break (30 mts)				
11:00-12:30	Tutorial 1 (contd): Differential Privacy in the Wild	Tutorial 2 (contd): Operational Analytics Data Management Systems	Workshop 1 (contd): TPCTC	Workshop 2 (contd): BIRTE	Workshop 3 (contd): QDB
12:30-2:00	Lunch Break (90 mts)				
2:00-3:30	Tutorial 3: Modern Main-memory DBMS	Tutorial 4: Machine Learning in the Real World	Workshop 1 (contd): TPCTC	Workshop 2 (contd): BIRTE	Workshop 3 (contd): QDB
3:30-4:00	Tea Break (30 mts)				
4:00-5:30	Tutorial 3 (contd): Modern Main-memory DBMS	Tutorial 4 (contd): Machine Learning in the Real World	Workshop 1 (contd): TPCTC	Workshop 2 (contd): BIRTE	Workshop 3 (contd): QDB
5:30 -6:00	High Tea Break (30 mts)				
6:30-8:30	DIGITAL INDIA RECEPTION (Pearl)				
8:30 onwards	Mini Dinner				

DAY 2: TUESDAY [Sep 6]						
Time	Track 1 (Pearl 1)	Track 2 (Pearl 2)	Track 3 (Royal 1)	Track 4 (Royal 2)	Track 5 (Maple)	
8:30-9:15	Conference Inauguration					
9:15-10:30	Invited Keynote 1 - Ion Stoica: Trends and Challenges in Big Data Processing					
10:30-11:15	Tea Break (45 mts)					
11:15-12:45	Research 1 : Transaction Processing	Research 2: RDF Data Systems	Research 3: Distributed and Cloud Systems - 1	Industrial 1: Distributed Data Analytics Platforms - 1	Demo 1A: Data Engines and Analytics	Research Posters - 1
12:45-2:00	Lunch Break (75 mts)					
2:00-3:30	Research 4: Memory Management	Research 5: Data Cleaning - 1	Research 6: Graph Processing - 1	Industrial 2: Distributed Data Analytics Platforms - 2	Demo 2A: Interactive and Exploratory Systems	Research Posters - 2
3:30-4:00	Tea Break (30 mts)					
4:00-5:30	Research 7: Query Execution - 1	Research 8: Data Security and Privacy	Research 9: Ranking Queries	Industrial 3: Data Engine Architectures - 1	Demo 3A: Graph and Semistructured Data	Research Posters - 3
5:30-6:30	High Tea Break (60 mts)					
6:30-8:30	VLDB CONFERENCE RECEPTION and QUIZ PROGRAM (Pearl)					

# Program at a Glance

DAY 3: WEDNESDAY [Sep 7]						
Time	Track 1 (Pearl 1)	Track 2 (Pearl 2)	Track 3 (Royal 1)	Track 4 (Royal 2)	Track 5 (Maple)	
9:00-9:15	Conference Announcements					
9:15-10:30	Invited Keynote 2 - Anand Rajaraman: Data-Driven Disruption: The View from Silicon Valley					
10:30-11:15	Tea Break (45 mts)					
11:15-12:45	Research 10 : Query Optimization - 1	Research 11: Spatial Data and Queries - 1	Research 12: Distributed and Cloud Systems - 2	Industrial 4: Data Engine Architectures - 2	Demo 3B: Graph and Semistructured Data	Research Posters - 4
12:45-2:00	Lunch Break (75 mts)					
2:00-3:30	Research 13: Query Execution - 2	Panel Will AI Eat Us All?	Research 14: Graph Processing - 2	Research 15: Data Cleaning - 2	Demo 1B: Data Engines and Analytics	Research Posters - 5
3:30-4:00	Tea Break (30 mts)					
4:00-5:30	Research 16: Data and Query Models - 1	Research 17: Scalable Analytics - 1	Research 18: Database Hardware- Software Codesign	Industrial 5: Graph Systems and Analytics	Demo 2B: Interactive and Exploratory Systems	Research Posters - 6
5:45-6:00	Depart for Banquet Venue					
7:30-10:00	VLDB CONFERENCE BANQUET					

DAY 4: THURSDAY [Sep 8]					
Time	Track 1 (Pearl 1)	Track 2 (Pearl 2)	Track 3 (Royal 1)	Track 4 (Royal 2)	Track 5 (Maple)
9:00-10:30	<b>VLDB 2016 Conference Awards</b> <b>VLDB Endowment Awards</b> <b>VLDB 10-year Best Paper Award Talk</b> <b>VLDB Early Career Research Contribution Award Talk</b>				
10:30-11:15	Tea Break (45 mts)				
11:15-12:45	Research 19: Query Optimization - 2	Research 20: Spatial Data and Queries - 2	Research 21: Distributed and Cloud Systems - 3	Tutorial 5: Human Factors in Crowdsourcing	Research 22: Entity Matching - 1
12:45-2:00	Lunch Break (75 mts)				
2:00-3:30	Research 23: Query Execution-3	Research 24: Social Networks and Crowdsourcing	Research 25: Graph Processing - 3	Tutorial 6: Qualitative Data Cleaning	Research 26: Data and Query Models - 2
3:30-4:00	Tea Break (30 mts)				
4:00-5:30	Research 27: Data and Query Models -3	Research 28: Entity Matching - 2	Research 29: Community Search and Mining		Research 30: Scalable Analytics - 2
6:30-9:30	ENTERTAINMENT PROGRAM AND GALA DINNER (PEARL)				

## Program at a Glance

DAY 5: FRIDAY [Sep 9]					
Time	Track 1 (Pearl 1)	Track 2 (Pearl 2)	Track 3 (Royal 1)	Track 4 (Royal 2)	Track 5 (Maple)
9:00-10:30	Workshop 4 : PhD Workshop	Workshop 5: ADMS/IMDM	Workshop 6: DMAH	Workshop 7: SoDAM	Workshop 8: BOSS/Polyglot
10:30-11:00	Tea Break (30 mts)				
11:00-12:30	Workshop 4 (contd): PhD Workshop	Workshop 5 (contd): ADMS/IMDM	Workshop 6 (contd): DMAH	Workshop 7 (contd): SoDAM	Workshop 8 (contd): BOSS/Polyglot
12:30-2:00	Lunch Break (90 mts)				
2:00-3:30	Workshop 4 (contd): PhD Workshop	Workshop 5 (contd): ADMS/IMDM	Workshop 6 (contd): DMAH		Workshop 8 (contd): BOSS/Polyglot
3:30-4:00	Tea Break (30 mts)				
4:00-5:30	Workshop 4 (contd): PhD Workshop	Workshop 5 (contd): ADMS/IMDM	Workshop 6 (contd): DMAH		Workshop 8 (contd): BOSS/Polyglot

## Digital India Event, 5th Sept Schedule

Time	Venue	Activity
9:00 am - 10:30 am	Pearl 2	Session 1: India Stack
10:30 am - 11:00 am		Tea Break
11:00 am - 12:30 pm	Pearl 2	Session 2: Data and India Scale Challenges
12:30 pm - 1:30 pm		Lunch
1:30 pm - 4:00 pm	Pearl 2	Digital India Hackathon Finale
4:00 pm - 4:30 pm		Break for Judges' decision
4:30 pm - 5:30 pm	Pearl 2	Hackathon Winner announcement and prize distribution
5:30 pm - 6:30 pm		High Tea Break
6:30 pm - 8:30 pm	Pearl	Digital India - Keynote and Panel Discussion
8:30 pm onwards		Mini Dinner



**DAY 1 - Monday, September 5th**

## Monday Sept 5th 9:00 - 10:30

### Tutorial 1: Differential Privacy in the Wild - Part 1

#### Location: Pearl 1

Ashwin Machanavajjhala (Duke University), Xi He (Duke University), Michael Hay (Colgate University)

### Tutorial 2: Operational Analytics Data Management Systems - Part 1

#### Location: Maple

Alexander Boehm (SAP), Jens Dittrich (University of Saarland), Niloy Mukherjee (LinkedIn), Ippokratis Pandis (Amazon), Rajkumar Sen (MemSQL)

### Workshop 1: TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC)

#### Location: Royal 1

Raghunath Nambiar (Cisco, USA), Meikel Poess (Oracle, USA)

### Workshop 2: Business Intelligence for the Real Time Enterprise (BIRTE)

#### Location: Royal 2

Meichun Hsu (Hewlett-Packard), Malu Castellanos (Hewlett-Packard), Panos K Chrysanthis (University of Pittsburgh)

### Workshop 3: Quality in Databases

#### Location: Boardroom

Christoph Quix (Fraunhofer FIT, Germany), Rihan Hai (RWTH Aachen University, Germany), Hongzhi Wang (Harbin Inst. of Tech., China), Venkat N. Gudivada (East Carolina University, USA), Laure Berti (QCRI, Qatar)

## Monday Sept 5th 10:30 - 11: 00

### Tea Break

## Monday Sept 5th 11:00 - 12:30

### Tutorial 1: Differential Privacy in the Wild - Part 2

#### Location: Pearl 1

Ashwin Machanavajjhala (Duke University), Xi He (Duke University), Michael Hay (Colgate University)

### Tutorial 2: Operational Analytics Data Management Systems - Part 2

#### Location: Maple

Alexander Boehm (SAP), Jens Dittrich (University of Saarland), Niloy Mukherjee (LinkedIn), Ippokratis Pandis (Amazon), Rajkumar Sen (MemSQL)

### Workshop 1: TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC) - contd.

#### Location: Royal 1

Raghunath Nambiar (Cisco, USA), Meikel Poess (Oracle, USA)

### Workshop 2: Business Intelligence for the Real Time Enterprise (BIRTE) - contd.

#### Location: Royal 2

Meichun Hsu (Hewlett-Packard), Malu Castellanos (Hewlett-Packard), Panos K Chrysanthis (University of Pittsburgh, USA)

### Workshop 3: Quality in Databases - contd.

#### Location: Boardroom

Christoph Quix (Fraunhofer FIT, Germany), Rihan Hai (RWTH Aachen University, Germany), Hongzhi Wang (Harbin Inst. of Tech., China), Venkat N. Gudivada (East Carolina University, USA), Laure Berti (QCRI, Qatar)

## Monday Sept 5th 12:30 - 14:00

### Lunch Break

## Monday Sept 5th 14:00 - 15:30

### Tutorial 3: Modern Main-memory DBMS - Part 1

#### Location: Pearl 1

Paul Larson (Microsoft Research), Justin Levandoski (Microsoft Research)

### Tutorial 4: Machine Learning in the Real-world - Part 1

#### Location: Maple

Vineet Chaoji (Amazon, India), Rajeev Rastogi (Amazon, India), Gourav Roy (Amazon, India)

### Workshop 1: TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC) - contd.

#### Location: Royal 1

Raghunath Nambiar (Cisco, USA), Meikel Poess (Oracle, USA)

### Workshop 2: Business Intelligence for the Real Time Enterprise (BIRTE) - contd.

#### Location: Royal 2

Meichun Hsu (Hewlett-Packard, USA), Malu Castellanos (Hewlett-Packard, USA), Panos K Chrysanthis (University of Pittsburgh, USA)

### Workshop 3: Quality in Databases - contd.

#### Location: Boardroom

Christoph Quix (Fraunhofer FIT, Germany), Rihan Hai (RWTH Aachen University, Germany), Hongzhi Wang (Harbin Inst. of Tech., China), Venkat N. Gudivada (East Carolina University, USA), Laure Berti (QCRI, Qatar)

## Monday Sept 5th 15:30 - 16:00

### Tea Break

## Monday Sept 5th 16:00 - 17:30

### Tutorial 3: Modern Main-memory DBMS - Part 2

#### Location: Pearl 1

Paul Larson (Microsoft Research), Justin Levandoski (Microsoft Research)

### Tutorial 4: Machine Learning in the Real-world - Part 2

#### Location: Maple

Vineet Chaoji (Amazon, India), Rajeev Rastogi (Amazon, India), Gourav Roy (Amazon, India)

### Workshop 1: TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC) - contd.

#### Location: Royal 1

Raghunath Nambiar (Cisco, USA), Meikel Poess (Oracle, USA)

### Workshop 2: Business Intelligence for the Real Time Enterprise (BIRTE) - contd.

#### Location: Royal 2

Meichun Hsu (Hewlett-Packard, USA), Malu Castellanos (Hewlett-Packard, USA), Panos K Chrysanthis (University of Pittsburgh, USA)

### **Workshop 3: Quality in Databases - contd.**

#### **Location: Boardroom**

Christoph Quix (Fraunhofer FIT, Germany), Riham Hai (RWTH Aachen University, Germany), Hongzhi Wang (Harbin Inst. of Tech., China), Venkat N. Gudivada (East Carolina University, USA), Laure Berti (QCRI, Qatar)

### **Monday Sept 5th 17:30 - 18:00**

High Tea Break

### **Monday Sept 5th 18:30 - 20:30**

Digital India Reception

Location: Pearl

### **Monday Sept 5th 20:30 onwards**

Mini Dinner



**DAY 2 - Tuesday, September 6th**



## Tuesday Sept 6th 8:30 - 9:15

### Conference Inauguration

Anand Deshpande (Persistent Systems); T. M. Vijayaraman (Persistent Systems);  
Surajit Chaudhuri (Microsoft Research); Jayant Haritsa (Indian Institute of Science);  
Mukesh Mohania (IBM Research)

## Tuesday Sept 6th 9:15 - 10:30

### Keynote 1 by Ion Stoica

**Location:** Pearl

**Session Chair:** Surajit Chaudhuri (Microsoft Research)

#### Trends and Challenges in Big Data Processing

Almost six years ago we started the Spark project at UC Berkeley. Spark is a cluster computing engine that is optimized for in-memory processing, and unifies support for a variety of workloads, including batch, interactive querying, streaming, and iterative computations. Spark is now the most active big data project in the open source community, and is already being used by over one thousand organizations.

One of the reasons behind Spark's success has been our early bet on the continuous increase in the memory capacity and the feasibility to fit many realistic workloads in the aggregate memory of typical production clusters. Today, we are witnessing new trends, such as Moore's law slowing down, and the emergence of a variety of computation and storage technologies, such as GPUs, FPGAs, and 3D Xpoint. In this talk, I'll discuss some of the lessons we learned in developing Spark as a unified computation platform, and the implications of today's hardware and software trends on the development of the next generation of big data processing systems.



**Bio:** Ion Stoica is a Professor in the EECS Department at University of California at Berkeley. He received his PhD from Carnegie Mellon University in 2000. He does research on cloud computing and networked computer systems. Past work includes the Dynamic Packet State (DPS), Chord DHT, Internet Indirection Infrastructure (i3), declarative networks, replay-debugging, and multi-layer tracing in distributed systems. His current research focuses on resource management and scheduling for data centers, cluster computing frameworks, and network architectures. He is an ACM Fellow and has received numerous awards, including the SIGCOMM Test of Time Award (2011), and the ACM doctoral dissertation award (2001). In 2006, he co-founded Conviva, a startup to commercialize technologies for large scale video distribution, and in 2013, he co-founded Databricks a startup to commercialize technologies for Big Data processing.

## Tuesday Sept 6th 10:30 - 11:15

### Tea Break

## Tuesday Sept 6th 11:15 - 12:45

### Research 1: Transaction Processing

**Location:** Pearl 1

**Session Chair:** Daniel Abadi (Yale University)

#### Leveraging Lock Contention to Improve OLTP Application Performance

Cong Yan (University of Washington); Alvin Cheung (University of Washington)

#### BCC: Reducing False Aborts in Optimistic Concurrency Control with Low Cost for In-Memory Databases

Yuan Yuan (The Ohio State University); Kaibo Wang (The Ohio State University); Rubao Lee (The Ohio State University); Xiaoning Ding (New Jersey Institute of Technology); Jing Xing (Institute of Computing Technology, Chinese Academy of Sciences); Spyros Blanas (The Ohio State University); Xiaodong Zhang (The Ohio State University)

#### Multi-Version Range Concurrency Control in Deuteronomy

Ryan Stutsman (Microsoft Research); Justin Levandoski (Microsoft Research); David Lomet (Microsoft Research); Sudipta Sengupta (Microsoft Research); Rui Wang (Microsoft Research)

#### S-Store: Streaming Meets Transaction Processing

John Meehan (Brown University); Nesime Tatbul (Intel Labs and MIT); Stan Zdonik (Brown University); Cansu Aslantas (Brown University); Ugur Cetintemel (Brown University); Jiang Du (University of Toronto); Tim Kraska (Brown University); Samuel Madden (MIT); David Maier (Portland State University); Andrew Pavlo (CMU); Michael Stonebraker (MIT); Kristin Tufte (Portland State University); Hao Wang (MIT)

## Research 2: RDF Data Systems

**Location:** Pearl 2

**Session Chair:** Atsuyuki Morishima (University of Tsukuba)

### **S2RDF: RDF Querying with SPARQL on Spark**

Alexander Schätzle (University of Freiburg); Martin Przyjaciel-Zablocki (University of Freiburg); Simon Skilevic (University of Freiburg); Georg Lausen (University of Freiburg)

### **Inferray: fast in-memory RDF inference**

Julien Subercaze (Laboratoire Hubert Curien); Christophe Gravie (Laboratoire Hubert Curien); Jules Chevalier (Laboratoire Hubert Curien); Frederique Laforest (Laboratoire Hubert Curien)

### **RDF Graph Alignment with Bisimulation**

Peter Buneman (University of Edinburgh); Sławek Staworko (University of Edinburgh, University of Lille, and LINKS, INRIA Nord-Europe)

### **Semantic SPARQL Similarity Search Over RDF Knowledge Graphs**

Weiguo Zheng (Peking University); Lei Zou (Peking University); Wei Peng (Peking University); Xifeng Yan (University of California at Santa Barbara); Shaoxu Song (Tsinghua University); Dongyan Zhao (Peking University)

## Research 3: Distributed and Cloud Systems - 1

**Location:** Royal 1

**Session Chair:** Ashraf Aboulmaga (Qatar Computing Research Institute)

### **Clash of the Titans: MapReduce vs. Spark for Large Scale Data Analytics**

Juweji Shi (Renmin University of China); Yunjie Qiu (IBM Research - China); Umar Farooq Minhas (IBM Research - Almaden); Limei Jiao (IBM Research - China); Chen Wang (Tsinghua University); Berthold Reinwald (IBM Research - Almaden); Fatma Ozcan (IBM Research - Almaden)

### **Compressed Linear Algebra for Large-Scale Machine Learning [Best Paper Award Winner]**

Matthias Boehm (IBM Research - Almaden); Ahmed Elgohary (University of Maryland); Peter Haas (IBM Research - Almaden); Fred Reiss (IBM Research - Almaden); Berthold Reinwald (IBM Research - Almaden)

### **PIXIDA: Optimizing Data Parallel Jobs in Wide-Area Data Analytics**

Konstantinos Kloudas (INESC-ID); Rodrigo Rodrigues (INESC-ID/IST at University of Lisbon); Nuno Preguica (NOVA LINCS / NOVA at University of Lisbon); Margarida Mamede (NOVA LINCS / NOVA at University of Lisbon)

### **Cümülön: Matrix-Based Data Analytics in the Cloud with Spot Instances**

Botong Huang (Duke University); Nicholas Jarrett (Duke University); Shivnath Babu (Duke University); Sayan Mukherjee (Duke University); Jun Yang (Duke University)

## Industrial 1: Distributed Data Analytics Platforms - 1

**Location:** Royal 2

**Session Chair:** Gustavo Alonso (ETH Zurich)

### **Data Infrastructure at Flipkart (Invited talk)**

Sharad Agarwal (Flipkart)

### **Database System Support of Simulation Data**

Hermano Lustosa (LNCC); Fabio Porto (LNCC); Pablo Blanco (LNCC); Patrick Valduriez (INRIA)

### **Hybrid Row-Column Partitioning in Teradata**

Mohammed Al-Kateb (Teradata Labs); Paul Sinclair (Teradata Labs); Grace Au (Teradata Labs); Carrie Ballinger (Teradata Labs)

### **Consistent Regions: Guaranteed Tuple Processing in IBM Streams**

Gabriela Jacques da Silva (IBM Research - T.J. Watson); Fang Zheng (IBM Research - T.J. Watson); Daniel Debrunner (IBM); Kun-Lung Wu (IBM Research - T.J. Watson); Victor Dogaru (IBM); Eric Johnson (IBM); Michael Spicer (IBM); Ahmet Erdem Sariyuce (Sandia National Labs)

## Demo 1A: Data Engines and Analytics

### **Research Posters 1 (Research 10, Research 11, Research 12, Research 13 and Research 14)**

**Location:** Maple

### **Vita: A Versatile Toolkit for Generating Indoor Mobility Data for Real-World Buildings**

Huan Li (Zhejiang University); Hua Lu (Aalborg University); Xin Chen (Zhejiang University); Gang Chen (Zhejiang University); Ke Chen (Zhejiang University); Lidan Shou (Zhejiang University)

### **JexLog: A Sonar for the Abyss**

Tobias Scheuer (SAP SE); Norman May (SAP SE); Alexander Böhm (SAP SE); Daniel Scheibli (SAP SE)

### **S12P: A Restaurant Recommendation System Using Preference Queries over Incomplete Information**

Xiaoye Miao (Zhejiang University); Yunjun Gao (Zhejiang University); Gang Chen (Zhejiang University); Huiyong Cui (Zhejiang University); Chong Guo (Zhejiang University); Weida Pan (Zhejiang University)

### **Mixed-instance querying: a lightweight integration architecture for data journalism**

Raphaël Bonaque (INRIA); Tien Cao (INRIA); Bogdan Cautis (Université Paris-Sud); François Goasdoué (Université de Rennes 1); Javier Letellier (INRIA); Ioana Manolescu (INRIA); Oscar Mendoza (INRIA); Swen Ribeiro (Université Paris-Sud); Xavier Tannier (Université Paris-Sud); Michaël Thomazo (INRIA)

### **Precision Performance Surgery for PostgreSQL – LLVM-based Expression Compilation, Just in Time**

Dennis Butterstein (University of Tuebingen); Torsten Grust (Universität Tübingen)

### **Partial Marking for Automated Grading of SQL Queries**

Bikash Chandra (IIT Bombay); Mathew Joseph (IIT Bombay); Bharath Radhakrishnan (IIT Bombay); Shreevidhya Acharya (IIT Bombay); S. Sudarshan (IIT Bombay)

### **Squall: Scalable Real-time Analytics**

Aleksandar Vitorovic (EPFL); Mohammed Elseidy (EPFL); Khayyam Guliyev (EPFL); Khue Vu Minh (EPFL); Daniel Espino Timón (EPFL); Mohammad Dashti (EPFL); Yannis Klonatos (EPFL); Christoph Koch (EPFL)

### **LocationSpark: A Distributed In-Memory Data Management System for Big Spatial Data**

Mingjie Tang (Purdue University); Yongyang Yu (Purdue University); Qutaibah Malluhi (Qatar University); Mourad Ouzzani (Qatar Computing Research Institute); Walid Aref (Purdue University)

### **Amoeba: A Shape changing Storage System for Big Data**

Anil Shanbhag (MIT); Samuel Madden (MIT); Alekh Jindal (Microsoft); Yi Lu (MIT)

### **F: Regression Models over Factorized Views**

Dan Olteanu (University of Oxford); Maximilian Schleich (University of Oxford)

### **Magellan: Toward Building Entity Matching Management Systems over Data Science Stacks**

Pradap Konda (University of Wisconsin Madison); Sanjib Das (University of Wisconsin Madison); Paul Suganthan G. C. (University of Wisconsin Madison); Anhui Doan (University of Wisconsin Madison); Adel Ardalan (University of Wisconsin Madison); Jeffrey Ballard (University of Wisconsin Madison); Han Li (University of Wisconsin Madison); Fatemah Panahi (University of Wisconsin Madison); Haojun Zhang (University of Wisconsin Madison); Jeffrey Naughton (University of Wisconsin Madison); Shishir Prasad (University of Wisconsin Madison); Ganesh Krishnan (WalmartLabs); Rohit Deep (WalmartLabs); Vijay Raghavendra (WalmartLabs)

### **Schema Independent and Scalable Relational Learning By Castor**

Jose Picado (Oregon State University); Parisa Ataei (Oregon State University); Arash Termehchy (Oregon State University); Alan Fern (Oregon State University)

### **AD-WIRE: Add-on for Web Item Reviewing System**

Rajeshkumar Kannapalli (University of Texas at Arlington); Azade Nazi (University of Texas at Arlington); Mahashweta Das (Hewlett Packard Labs); Gautam Das (University of Texas at Arlington)

**Research Posters 1:** Papers from Research 10, Research 11, Research 12, Research 13 and Research 14

## **Tuesday Sept 6th 12:45 - 14:00**

Lunch Break

## **Tuesday Sept 6th 14:00 - 15:30**

### **Research 4: Memory Management**

**Location:** Pearl 1

**Session Chair:** Alfons Kemper (Technische Universität München)

#### **RUMA has it: Rewired User-space Memory Access is Possible!**

Felix Martin Schuhknecht (Saarland University); Jens Dittrich (Saarland University); Ankur Sharma (Saarland University)

#### **Asynchronous Memory Access Chaining**

Onur Kocberber (EPFL); Babak Falsafi (EPFL); Boris Grot (University of Edinburgh)

#### **WarpLDA: a Cache Efficient O(1) Algorithm for Latent Dirichlet Allocation**

Jianfei Chen (Tsinghua University); Kaiwei Li (Tsinghua University); Jun Zhu (Tsinghua University); Wenguang Chen (Tsinghua University)

## Lifetime-Based Memory Management for Distributed Data Processing Systems

Lu Lu (Huazhong University of Science and Technology); Xuanhua Shi (Huazhong University of Science and Technology); Yongluan Zhou (University of Southern Denmark); Xiong Zhang (Huazhong University of Science and Technology); Hai Jin (Huazhong University of Science and Technology); Cheng Pei (Huazhong University of Science and Technology); Ligang He (University of Warwick); Yuanzhen Geng (Huazhong University of Science and Technology)

## Research 5: Data Cleaning - 1

**Location:** Pearl 2

**Session Chair:** Ihab Ilyas (University of Waterloo)

### Cleaning Timestamps with Temporal Constraints

Shaoxu Song (Tsinghua University); Yue Cao (Tsinghua University); Jianmin Wang (Tsinghua University)

### Repairing Data through Regular Expressions

Zeyu Li (Harbin Institute of Technology); Hongzhi Wang (Harbin Institute of Technology); Wei Shao (Harbin Institute of Technology); Jianzhong Li (Harbin Institute of Technology); Hong Gao (Harbin Institute of Technology)

### Combining Quantitative and Logical Data Cleaning

Nataliya Prokoshyna (University of Toronto); Jaroslaw Szlichta (University of Ontario Institute of Technology); Fei Chiang (McMaster University); Renée Miller (University of Toronto); Divesh Srivastava (AT&T Labs Research)

### Temporal Rules Discovery for Web Data Cleaning

Ziawasch Abedjan (MIT); Cuneyt Akcora (Qatar Computing Research Institute); Mourad Ouzzani (Qatar Computing Research Institute); Paolo Papotti (Qatar Computing Research Institute); Michael Stonebraker (MIT)

## Research 6: Graph Processing - 1

**Location:** Royal 1

**Session Chair:** Bin Cui (Peking University)

### I/O Efficient ECC Graph Decomposition via Graph Reduction

Long Yuan (The University of New South Wales); Lu Qin (University of Technology, Sydney); Xuemin Lin (The University of New South Wales); Lijun Chang (The University of New South Wales); Wenjie Zhang (The University of New South Wales)

### LEOPARD: Lightweight Edge-Oriented Partitioning and Replication for Dynamic Graphs

Jiewen Huang (Yale University); Daniel Abadi (Yale University)

### Towards Maximum Independent Sets on Massive Graphs

Yu Liu (Renmin University of China); Jiaheng Lu (University of Helsinki); Hua Yang (Renmin University of China); Xiaokui Xiao (Nanyang Technological University); Zhewei Wei (Renmin University of China)

### Weaver: A High-Performance, Transactional Graph Database Based on Refinable Timestamps

Ayush Dubey (Cornell University); Greg Hill (Stanford University); Robert Escriva (Cornell University); Emin Sirer (Cornell University)

## Industrial 2: Distributed Data Analytics Platforms - 2

**Location:** Royal 2

**Session Chair:** Justin Levandoski (Microsoft Research)

### Big Data Analysis as a Service on Public Clouds (Invited talk)

Rajat Venkatesh (Qubole)

### dmapply: A functional primitive to express distributed machine learning algorithms in R

Edward Ma (Hewlett-Packard Enterprise Vertica); Vishrut Gupta (Hewlett-Packard Enterprise Vertica); Meichun Hsu (Hewlett-Packard Enterprise Vertica); Indrajit Roy (Hewlett Packard Labs)

### SystemML: Declarative Machine Learning on Spark

Matthias Boehm (IBM Research - Almaden); Michael Dusenberry (IBM Spark Technology Center); Deron Eriksson (IBM Spark Technology Center); Alexandre Evfimievski (IBM Research - Almaden); Faraz Makari Manshadi (IBM Research - Almaden); Niketan Pansare (IBM Research - Almaden); Berthold Reinwald (IBM Research - Almaden); Frederick Reiss (IBM Research - Almaden); Prithviraj Sen (IBM Research - Almaden); Arvind Surve (IBM Spark Technology Center); Shirish Tatikonda (IBM Research - Almaden)

### Not for the Timid: On the Impact of Aggressive Over-booking in the Cloud

Willis Lang (Microsoft); Karthik Ramachandra (Microsoft); David DeWitt (Microsoft); Shize Xu (Microsoft); Qun Guo (Microsoft); Ajay Kalhan (Microsoft); Peter Carlin (Microsoft)

**Research Posters 2 (Research 15, Research 16, Research 17, Research 18, Research 19 and Research 20)**

**Location:** Maple

**ArchimedesOne: Query Processing over Probabilistic Knowledge Bases**

Xiaofeng Zhou (University of Florida); Yang Chen (University of Florida); Daisy Zhe Wang (University of Florida)

**Rudolf: Interactive Rule Refinement System for Fraud Detection**

Tova Milo (Tel-Aviv University); Slava Novgorodov (Tel-Aviv University); Wang-Chiew Tan (University of California at Santa Cruz)

**Ziggy: Characterizing Query Results for Data Explorers**

Thibault Sellam (CWI); Martin Kersten (CWI)

**Blaeu: Mapping and Navigating Large Tables with Cluster Analysis**

Thibault Sellam (CWI); Robin Cijvat (MonetDB Solutions); Richard Koopmanschap (MonetDB Solutions); Martin Kersten (CWI)

**A Demonstration of VisDPT: Visual Exploration of Differentially Private Trajectories**

Xi He (Duke University); Nisarg Raval (Duke University); Ashwin Machanavajjhala (Duke University)

**YASK: A Why-Not Question Answering Engine for Spatial Keyword Query Services**

Lei Chen (Hong Kong Baptist University); Jianliang Xu (Hong Kong Baptist University); Christian Jensen (Aalborg University); Yafei Li (Hong Kong Baptist University)

**Exploring Databases via Reverse Engineering Ranking Queries with PALEO**

Kiril Panev (TU Kaiserslautern); Sebastian Michel (TU Kaiserslautern); Evica Milchevski (TU Kaiserslautern); Koninika Pal (TU Kaiserslautern)

**ExRank: An Exploratory Ranking Interface**

Ramon Bespinyowong (National University of Singapore); Wei Chen (Zhejiang University); H. V. Jagadish (University of Michigan); Yuxin Ma (Zhejiang University)

**NLProv: Natural Language Provenance**

Daniel Deutch (Tel Aviv University); Nave Frost (Tel Aviv University); Amir Gilad (Tel Aviv University)

**Towards Personalized Maps: Mining User Preferences from Geo-textual Data**

Kaiqi Zhao (Nanyang Technological University); Yiding Liu (Nanyang Technological University); Quan Yuan (Nanyang Technological University); Lisi Chen (Hong Kong Baptist University); Zhida Chen (Nanyang Technological University); Gao Cong (Nanyang Technological University)

**A System for Region Search and Exploration**

Kaiyu Feng (Nanyang Technological University); Kaiqi Zhao (Nanyang Technological University); Yiding Liu (Nanyang Technological University); Gao Cong (Nanyang Technological University)

**SigmaKB: Multiple Probabilistic Knowledge Base Fusion**

Miguel Rodriguez (University of Florida); Sean Goldberg (University of Florida); Daisy Zhe Wang (University of Florida)

**Collaborative Crowdsourcing with Crowd4U**

Kosetsu Ikeda (University of Tsukuba); Atsuyuki Morishima (University of Tsukuba); Habibur Rahman (UT Arlington); Senjuti Basu Roy (NJIT); Saravanan Thirumuruganathan (OCRI, HBKU); Sihem Amer-Yahia (CNRS LIG); Gautam Das (UT Arlington)

**Research Posters 2:** Papers from Research 15, Research 16, Research 17, Research 18, Research 19 and Research 20

**Tuesday Sept 6th 15:30 - 16:00**

Tea Break

**Tuesday Sept 6th 16:00 - 17:30**

**Research 7: Query Execution - 1**

**Location:** Pearl 1

**Session Chair:** Anastasia Ailamaki (EPFL)

**Parallel Evaluation of Multi-Semi-Joins**

Jonny Daenen (Hasselt University); Frank Neven (Hasselt University); Tony Tan (National Taiwan University); Stijn Vansummeren (Université Libre de Bruxelles)

**Lightning Fast and Space Efficient Inequality Joins**

Zuhair Khayyat (King Abdullah University of Science and Technology); William Lucia (Qatar Computing Research Institute); Meghna

Singh (Qatar Computing Research Institute), Mourad Quzzani (Qatar Computing Research Institute); Paolo Papotti (Qatar Computing Research Institute); Jorge-Arnulfo Quijano-Ruiz (Qatar Computing Research Institute); Nan Tang (Qatar Computing Research Institute); Panos Kalnis (King Abdullah University of Science and Technology)

#### **MQJoin: Efficient Shared Execution of Main-Memory Joins**

Darko Makreshanski (ETH Zurich); Georgios Giannikis (Oracle Labs); Gustavo Alonso (ETH Zurich); Donald Kossman (Microsoft Research)

#### **Incremental Computation of Common Windowed Holistic Aggregates**

Richard Wesley (Tableau Software); Fei Xu (Tableau Software)

### **Research 8: Data Security and Privacy**

**Location: Pearl 2**

**Session Chair: Lei Chen (Hong Kong University of Science and Technology)**

#### **Oblivious RAM: A Dissection and Experimental Evaluation**

Zhao Chang (University of Utah); Dong Xie (University of Utah); Feifei Li (University of Utah)

#### **Design of Policy-Aware Differentially Private Algorithms**

Samuel Haney (Duke University); Ashwin Machanavajjhala (Duke University); Bolin Ding (Microsoft Research)

#### **Behavior Query Discovery in System-Generated Temporal Graphs**

Bo Zong (University of California at Santa Barbara); Xusheng Xiao (NEC Labs America, Inc.); Zhichun Li (NEC Labs America, Inc.); Zhenyu Wu (NEC Labs America, Inc.); Zhiyun Qian (University of California at Riverside); Xifeng Yan (University of California at Santa Barbara); Ambuj Singh (University of California at Santa Barbara); Guofei Jiang (NEC Labs America, Inc.)

### **Research 9: Ranking Queries**

**Location: Royal 1**

**Session Chair: Renée Miller (University of Toronto)**

#### **Query Reranking As A Service**

Abolfazl Asudeh (UT Arlington); Nan Zhang (George Washington University); Gautam Das (UT Arlington)

#### **Finding Pareto Optimal Groups: Group-based Skyline**

Jinfei Liu (Emory University); Li Xiong (Emory University); Jian Pei (Simon Fraser University); Jun Luo (Lenovo); Haoyu Zhang (Emory University)

#### **k-Regret Queries with Nonlinear Utilities**

Taylor Kessler Faulkner (Denison University); Will Brackenbury (Denison University); Ashwin Lall (Denison University)

#### **Discovering the Skyline of Web Databases**

Abolfazl Asudeh (UT Arlington); Saravanan Thirumuruganathan (UT Arlington); Nan Zhang (George Washington University); Gautam Das (UT Arlington)

### **Industrial 3: Data Engine Architectures - 1**

**Location: Royal 2**

**Session Chair: Wolfgang Lehner (Technische Universität Dresden)**

#### **Hardware-Software Co-design for Data Management (Invited talk)**

Weiwei Gong (Oracle); Shasank Chavan (Oracle)

#### **Nitro: A Fast, Scalable In-Memory Storage Engine for NoSQL Global Secondary Index**

Sarath Lakshman (Couchbase Inc.); Sriram Melkote (Couchbase Inc.); John Liang (Couchbase Inc.); Ravi Mayuram (Couchbase Inc.)

#### **Aerospike: Architecture of a Real-Time Operational DBMS**

V. Srinivasan (Aerospike Inc.); Brian Bulkowski (Aerospike Inc.); Wei-Ling Chu (Aerospike Inc.); Sunil Sayyaparaju (Aerospike Inc.); Andrew Gooding (Aerospike Inc.); Rajkumar Iyer (Aerospike Inc.); Ashish Shinde (Aerospike Inc.); Thomas Lopatic (Aerospike Inc.)

#### **Comdb2 - Bloomberg's Highly Available Relational Database System**

Alex Scotti (Bloomberg LP); Mark Hannum (Bloomberg LP); Michael Ponomarenko (Bloomberg LP); Dorin Hoguea (Bloomberg LP); Akshat Sikarwar (Bloomberg LP); Mohit Khullar (Bloomberg LP); Adi Zaimi (Bloomberg LP); James Leddy (Bloomberg LP); Fabio Angius (Bloomberg LP); Rivers Zhang (Bloomberg LP); Lingzhi Deng (Bloomberg LP)

**Location: Maple**

**GARUDA: A System for Large-Scale Mining of Statistically Significant Connected Subgraphs**

Satyajit Bhadange (IIT Kanpur); Akhil Arora (Xerox Research Centre India); Arnab Bhattacharya (IIT Kanpur)

**Generating Flexible Workloads for Graph Databases**

Guillaume Bagan (CNRS LIRIS); Angela Bonifati (University of Lyon 1 & CNRS LIRIS); Radu Ciucanu (University of Oxford); George Fletcher (TU Eindhoven); Aurélien Lemay (University of Lille 3 & INRIA); Nicky Advokaat (TU Eindhoven)

**Graph databases in the browser: using LevelGraph to explore New Delhi**

Antonio Maccioni (Roma Tre University); Matteo Collina (NearForm)

**Sapphire: Querying RDF Data Made Simple**

Ahmed El-Roby (University of Waterloo); Khaled Ammar (University of Waterloo); Ashraf Aboulnaga (Qatar Computing Research Institute); Jimmy Lin (University of Waterloo)

**December: A Declarative Tool for Crowd Member Selection**

Yael Amsterdamer (Bar Ilan University); Tova Milo (Tel-Aviv University); Amit Somech (Tel-Aviv University); Brit Youngmann (Tel-Aviv University)

**AutoG: A Visual Query Autocompletion Framework for Graph Databases**

Peipei Yi (Hong Kong Baptist University); Byron Choi (Hong Kong Baptist University); Sourav Bhowmick (Nanyang Technological University); Jianliang Xu (Hong Kong Baptist University)

**Exploratory Querying of Extended Knowledge Graphs**

Mohamed Yahya (Max-Planck Institute for Informatics, Germany); Klaus Berberich (Max-Planck Institute for Informatics, Germany); Maya Ramanath (IIT Delhi); Gerhard Weikum (Max-Planck Institute for Informatics, Germany)

**SPARQLByE: Querying RDF data by example**

Gonzalo Diaz (University of Oxford); Marcelo Arenas (PUC Chile); Michael Benedikt (University of Oxford)

**Graph-based Exploration of Non-graph Datasets**

Udayan Khurana (IBM Research - T.J. Watson); Srinivasan Parthasarathy (IBM Research - T.J. Watson); Deepak Turaga (IBM Research - T.J. Watson)

**Rogas: A Declarative Framework for Network Analytics**

Minjian Liu (The Australian National University); Qing Wang (The Australian National University)

**Large-scale Complex Analytics on Semi-structured Datasets using AsterixDB and Spark**

Wail Alkawaileet (CCEE at KACST and MIT); Sattam Alsubaiee (CCEE at KACST and MIT); Michael Carey (University of California at Irvine); Till Westmann (Couchbase Inc.); Yingyi Bu (Couchbase Inc.)

**Research Posters 3:** Papers from Research 21, Research 22, Research 23, Research 24 and Research 25

**Tuesday Sept 6th 17:30 - 18:30**

High Tea Break

**Tuesday Sept 6th 18:30 - 20:30**

VLDB Conference Reception and Quiz Program

Location: Pearl



**Day 3 - Wednesday, September 7th**



## Wednesday Sept 7th 9:00 - 9:15

### Conference Announcements

## Wednesday Sept 7th 9:15 - 10:30

### Keynote 2 by Anand Rajaraman

**Location:** Pearl

**Session Chair:** Jayant Haritsa (IISc, Bangalore)

#### Trends and Challenges in Big Data Processing

We live in an era where software is transforming industries, the sciences, and society as a whole. This exciting phenomenon has been described by the phrase "software is eating the world." It is becoming increasingly apparent that data is the fuel powering software's conquests. Data is the new disruptor.

It's hard to believe that the first decade of the Big Data era is already behind us. Silicon Valley has been at the forefront of developing and applying data-driven approaches to create disruption at many levels: infrastructure (e.g., Hadoop and Spark), capabilities (e.g., image recognition and machine translation), and killer apps (e.g., self-driving cars and messaging bots).

In this talk, we first look back on the past decade and share learnings from the frontlines of data-driven disruption. Looking ahead, we then describe challenges and opportunities for the next decade. Since this has also been a personal journey, we will use examples drawn from personal experience to illustrate each point.



**Bio:** Anand is a Founding Partner of two Silicon Valley venture capital funds focused on early-stage technology companies: Milliways Ventures and RocketshipVC. He was the co-founder of two successful startups: Jungle (acquired by Amazon.com) and Kosmix (acquired by Walmart). At Walmart, he created and led WalMartLabs (as its Senior Vice President). As an academic, Anand's research has focused at the intersection of database systems, the World-Wide Web, and social media. His research publications have won several awards at prestigious academic conferences, including three retrospective 10-year Best Paper awards at ACM SIGMOD and VLDB, and ICDT. His textbook "Mining of Massive Datasets", co-authored with Jeff Ullman and Jure Leskovec, has been published.

## Wednesday Sept 7th 10:30 - 11:15

### Tea Break

## Wednesday Sept 7th 11:15 - 12:45

### Research 10: Query Optimization - 1

**Location:** Pearl 1

**Session Chair:** S Sudarshan (IIT Bombay)

#### Faster Plan Generation through Consideration of Functional Dependencies and Keys

Marius Eich (University of Mannheim); Pit Fender (Oracle Labs); Guido Moerkotte (University of Mannheim)

#### Exploiting Soft and Hard Correlations in Big Data Query Optimization

Hai Liu (Worcester Polytechnic Institute); Dongqing Xiao (Worcester Polytechnic Institute); Pankaj Didwania (Worcester Polytechnic Institute); Mohamed Eltabakh (Worcester Polytechnic Institute)

#### How Good Are Query Optimizers, Really?

Viktor Leis (Technische Universität München); Andrey Gubichev (Technische Universität München); Atanas Mirchev (Technische Universität München); Peter Boncz (CWI); Alfons Kemper (Technische Universität München); Thomas Neumann (Technische Universität München)

#### Teaching an RDBMS about ontological constraints

Damian Bursztyjn (INRIA & LIX); François Goasdoué (University of Rennes 1 and INRIA); Ioana Manolescu (INRIA & LIX)

### Research 11: Spatial Data and Queries - 1

**Location:** Pearl 2

**Session Chair:** Amr El Abbadi (University of California at Santa Barbara)

#### Processing and Optimizing Main Memory Spatial-Keyword Queries

Taesung Lee (Yonsei University); Jin-Woo Park (POSTECH); Sanghoon Lee (POSTECH); Seung-won Hwang (Yonsei University); Sameh Elnikety (Microsoft Research); Yuxiong He (Microsoft Research)

### **Spatial Online Sampling and Aggregation**

Lu Wang (Hong Kong University of Science and Technology); Robert Christensen (University of Utah) ; Feifei Li (University of Utah); Ke Yi (Hong Kong University of Science and Technology)

### **Online Minimum Matching in Real-Time Spatial Data: Experiments and Analysis**

Yongxin Tong (Beihang University); Jieying She (Hong Kong University of Science and Technology); Bolin Din (Microsoft Research); Lei Chen (Hong Kong University of Science and Technology); Tianyu Wo (Beihang University); Ke Xu (Beihang University)

### **AQWA: Adaptive Query-Workload-Aware Partitioning of Big Spatial Data**

Ahmed M. Aly (Purdue University); Ahmed R. Mahmood (Purdue University); Mohamed S. Hassan (Purdue University); Walid G. Aref (Purdue University); Mourad Ouzzani (Qatar Computing Research Institute); Hazem Elmeleegy (Turn Inc.); Thamir Qadah (Purdue University)

## **Research 12: Distributed and Cloud Systems - 2**

**Location: Royal 1**

**Session Chair: Krithi Ramamritham (IIT Bombay)**

### **High-Speed Query Processing over High-Speed Networks**

Wolf Roediger (Technische Universität München); Tobias Muehlbauer (Technische Universität München); Alfons Kemper (Technische Universität München); Thomas Neumann (Technische Universität München)

### **The End of Slow Networks: It's Time for a Redesign**

Carsten Binnig (Brown University); Andrew Crotty (Brown University); Alex Galakatos (Brown University); Tim Kraska (Brown University); Erfan Zamanian (Brown University)

### **Tempo: Robust and Self-Tuning Resource Management in Multi-tenant Parallel Databases**

Zilong Tan (Duke University); Shivnath Babu (Duke University)

### **Measuring and Optimizing Distributed Array Programs**

Mingxing Zhang (Tsinghua University); Yongwei Wu (Tsinghua University); Kang Chen (Tsinghua University); Teng Ma (Tsinghua University); Weimin Zheng (Tsinghua University)

## **Industrial 4: Data Engine Architectures - 2**

**Location: Royal 2**

**Session Chair: Spyros Blanas (The Ohio State University)**

### **Kodiak: Leveraging Materialized Views For Very Low-Latency Analytics Over High-Dimensional Web-Scale Data**

Shaosu Liu (Turn Inc.); Bin Song (Turn Inc.); Sriharsha Gangam (Turn Inc.); Lawrence Lo (Turn Inc.); Khaled Elmeleegy (Turn Inc.)

### **Accelerating Analytics with Dynamic In-Memory Expressions**

Auroshish Mishra (Oracle America); Shasank Chavan (Oracle America); Allison Holloway (Oracle America); Tirthankar Lahiri (Oracle America); Zhen Hua Liu (Oracle America); Sunil Chakkappen (Oracle America); Dennis Lui (Oracle America); Vinita Subramanian (Oracle America); Ramesh Kumar (Oracle America); Maria Colgan (Oracle America); Jesse Kamp (Oracle America); Niloy Mukherjee (LinkedIn America); Vineet Marwah (Oracle America)

### **TrafficDB: HERE's High Performance Shared-Memory Data Store**

Ricardo Fernandes (HERE Global B.V.); Piotr Zaczekowski (HERE Global B.V.); Bernd Göttler (HERE Global B.V.); Conor Ettinoffe (HERE Global B.V.); Anis Moussa (HERE Global B.V.)

### **The MemSQL Query Optimizer: A modern optimizer for real-time analytics in a distributed database**

Jack Chen (MemSQL Inc.); Samir Jindel (MemSQL Inc.); Robert Walzer (MemSQL Inc.); Rajkumar Sen (MemSQL Inc.); Nika Jimshelishvili (MemSQL Inc.); Michael Andrews (MemSQL Inc.)

## **Demo 3B: Graph and Semistructured Data**

### **Research Posters 4 (Research 1, Research 2, Research 3, Research 4 and Research 5)**

**Location: Maple**

**(Same as in Demo 3A)**

**Research Posters 4:** Papers from Research 1, Research 2, Research 3, Research 4 and Research 5

**Wednesday Sept 7th 12:45 - 14:00**

**Lunch Break**

## Wednesday Sept 7th 14:00 - 15:30

### Research 13: Query Execution - 2

Location: Pearl 1

Session Chair: Wolfgang Lehner (Technische Universität Dresden)

#### A Seven-Dimensional Analysis of Hashing Methods and its Implications on Query Processing

Stefan Richter (Saarland University); Victor Alvarez (TU Braunschweig); Jens Dittrich (Saarland University)

#### Query-Aware Locality-Sensitive Hashing for Approximate Nearest Neighbor Search

Qiang Huang (Sun Yat-sen University); Jianlin Feng (Sun Yat-sen University); Yikai Zhang (Sun Yat-sen University); Qiong Fang (South China University of Technology); Wilfred Ng (Hong Kong University of Science and Technology)

#### Neighbor-Sensitive Hashing

Yongjoo Park (University of Michigan); Michael Cafarella (University of Michigan); Barzan Mozafari (University of Michigan)

#### LSH Ensemble: Internet Scale Domain Search

Erkang Zhu (University of Toronto); Fatemeh Nargesian (University of Toronto); Ken Pu (UOIT); Renée Miller (University of Toronto)

### Panel: Will AI Eat Us All?

Location: Pearl 2

#### Abstract

Big data, data science, AI and deep learning are the latest craze in industry and academia, with huge numbers of dollars being devoted to related applications, including face recognition, machine translation, digital assistants, self-driving cars, ad-serving, chat-bots, and more. Although these emerging applications rely on massive amounts of evolving data that needs to be cleaned, integrated, and analyzed for modeling purposes, data management issues are not usually perceived as central to these applications. In this panel, we will explore the key data management challenges and opportunities for data management in this new world, and discuss how the database community can best contribute to these exciting domains.



#### Ihab F. Ilyas, University of Waterloo

Ihab Ilyas is a professor in the Cheriton School of Computer Science at the University of Waterloo. He received his PhD in computer science from Purdue University, West Lafayette. His main research is in the area of database systems, with special interest in data quality, data integration, managing uncertain data, rank-aware query processing, and information extraction. Ihab is a recipient of the Ontario Early Researcher Award (2009), a Cheriton Faculty Fellowship (2013), an NSERC Discovery Accelerator Award (2014), and a Google Faculty Award (2014). He is an ACM Distinguished Scientist, an elected member of the VLDB Endowment Board of Trustees, and an associate editor of the ACM Transactions of Database Systems (TODS). Ihab is a co-founder of Tamr, a startup focusing on large-scale data integration and cleaning.

#### Panelists



#### H. V. Jagadish, University of Michigan

H. V. Jagadish is Bernard A. Galler Collegiate Professor of Electrical Engineering and Computer Science, and Distinguished Scientist at the Institute for Data Science, at the University of Michigan in Ann Arbor. Prior to 1999, he was Head of the Database Research Department at AT&T Labs, Florham Park, NJ.

Professor Jagadish is well known for his broad-ranging research on information management, and has approximately 200 major papers and 37 patents. He is a fellow of the ACM, "The First Society in Computing," (since 2003) and serves on the board of the Computing Research Association (since 2009).

He has been an Associate Editor for the ACM Transactions on Database Systems (1992-1995), Program Chair of the ACM SIGMOD annual conference (1996), Program Chair of the ISMB conference (2005), a trustee of the VLDB (Very Large DataBase) foundation (2004-2009), Founding Editor-in-Chief of the Proceedings of the VLDB Endowment (2008-2014), and Program Chair of the VLDB Conference (2014). Since 2016, he is Editor of the Morgan & Claypool "Synthesis" Lecture Series on Data Management. Among his many awards, he won the ACM SIGMOD Contributions Award in 2013 and the David E Liddle Research Excellence Award (at the University of Michigan) in 2008.



### **Aditya Parameswaran, University of Illinois at Urbana-Champaign**

Aditya Parameswaran is an Assistant Professor in Computer Science at the University of Illinois. He spent a year as a PostDoc at MIT in 2013-14 following his PhD at Stanford University. He develops systems and algorithms for “human-in-the-loop” data analytics, synthesizing techniques from data mining, database systems, and human computation. He has received multiple dissertation awards (from SIGMOD, SIGKDD, and Stanford), an “Excellent” Lecturer award from Illinois, a Google Faculty award, the Key Scientific Challenges award from Yahoo!, four best-of-conference citations, and a Gold Medal from IIT Bombay. His research group is supported with funding from the Siebel Energy Institute, the NIH, the NSF, and Google.



### **Sihem Amer-Yahia, CNRS**

Sihem Amer-Yahia is DRI CNRS at LIG in Grenoble where she leads the SLIDE team. Her research interests are in social data management and crowdsourcing. She previously held research positions at QCRI, Yahoo! Research and at&t Labs. Sihem served on the ACM SIGMOD Executive Board and the VLDB Endowment. She is Editor-in-Chief of the VLDB Journal and is on the editorial boards of TODS and Information Systems. She is currently chairing the VLDB Workshops 2016 and will chair VLDB 2018. Sihem completed her Ph.D. in Computer Science at INRIA in 1999, and her Diplôme d’Ingénieur at INI, Algeria.



### **Sunita Sarawagi, IIT Bombay**

Sunita Sarawagi researches in the fields of databases and machinelearning. She is professor at IIT Bombay. She got her PhD in databases from the University of California at Berkeley and a bachelor’s degree from IIT Kharagpur. She has also worked at Google Research (2014-2016), CMU (2004), and IBM Almaden Research Center (1996-1999). She has several publications including best paper awards at ACM SIGMOD, VLDB, ICDM, NIPS, and ICML conferences. She has served on the board of directors of the ACM SIGKDD and VLDB foundation. She was program chair for the ACM SIGKDD 2008 conference, research track co-chair for the VLDB 2011 conference and has served as program committee member for SIGMOD, VLDB, SIGKDD, ICDE, and ICML conferences, and on the editorial boards of the ACM TODS and ACM TKDD journals.

## **Research 14: Graph Processing - 2**

**Location: Royal 1**

**Session Chair: Kyuseok Shim (Seoul National University)**

### **G-SQL: Fast Query Processing via Graph Exploration**

Hongbin Ma (eBay); Bin Shao (Microsoft Research Asia); Yanghua Xiao (Fudan University); Liang Jeff Chen (Microsoft Research Asia); Haixun Wang (Facebook)

### **Fully Dynamic Betweenness Centrality Maintenance on Massive Networks**

Takanori Hayashi (The University of Tokyo); Takuya Akiba (National Institute of Informatics); Yuichi Yoshida (National Institute of Informatics)

### **Parallel Local Graph Clustering**

Julian Shun (University of California at Berkeley); Farbod Roosta-Khorasani (University of California at Berkeley); Kimon Fountoulakis (University of California at Berkeley); Michael Mahoney (University of California at Berkeley)

### **KCore Decomposition for Large Networks on a Single PC**

Wissam Khaoûd (University of Victoria); Marina Barsky (University of Victoria); Venkatesh Srinivasan (University of Victoria); Alex Thomo (University of Victoria)

## **Research 15: Data Cleaning - 2**

**Location: Royal 2**

**Session Chair: Panos Chrysanthis (University of Pittsburgh)**

### **Messing Up with BART: Error Generation for Evaluating Data-Cleaning Algorithms**

Patricia Arocena (University of Toronto); Boris Glavic (Illinois Institute of Technology); Giansalvatore Mecca (Università di Basilicata); Renée Miller (University of Toronto); Paolo Papotti (Qatar Computing Research Institute); Donatello Santoro (Università di Basilicata)

### **Detecting Data Errors: Where are we and what needs to be done?**

Ziawasch Abedjan (MIT); Xu Chu (University of Waterloo); Dong Deng (Tsinghua University); Raul Castro Fernandez (MIT); Ihab Ilyas (University of Waterloo); Mourad Ouzzani (Qatar Computing Research Institute); Paolo Papotti (Arizona State University); Michael Stonebraker (MIT); Nan Tang (Qatar Computing Research Institute)

### **ActiveClean: Interactive Data Cleaning for Statistical Modeling**

Sanjay Krishnan (University of California at Berkeley); Jiannan Wang (Simon Fraser University); Eugene Wu (Columbia University); Michael Franklin (University of California at Berkeley); Ken Goldberg (University of California at Berkeley)

## **META: An Efficient Matching-Based Method for Error-Tolerant Autocompletion**

Dong Deng (Tsinghua University); Guoliang Li (Tsinghua University); He Wen (Tsinghua University); H. V. Jagadish (University of Michigan); Feng Jianhua (Tsinghua University)

### **Demo 1B: Data Engines and Analytics**

#### **Research Posters 5 (Research 6, Research 7, Research 8, Research 9 and VLDB Journal)**

**Location: Maple**

**Same as Demo 1A**

**Research Posters 5: Papers from Research 6, Research 7, Research 8, Research 9 and Posters for the following VLDB Journal papers**

#### **Decorating the Cloud: Enabling Annotation Management in MapReduce Infrastructure**

Yue Lu (Worcester Polytechnic Institute); Yuguang Li (Worcester Polytechnic Institute); Mohamed Eltabakh (Worcester Polytechnic Institute)

#### **S<sup>3</sup>-TM: scalable streaming short text matching**

Fuat Basik (Bilkent University); Bugra Gedik (Bilkent University); Hakan Ferhatosmanoglu (Bilkent University); Mert Emin Kalender (Bilkent University)

#### **Elite: an Elastic Infrastructure for Big Spatiotemporal Trajectories**

Xike Xie (Aalborg University); Benjin Mei (Renmin University of China); Jinchuan Chen (Renmin University of China); Xiaoyong Du (Renmin University of China); Christian Jensen (Aalborg University)

#### **Accelerating SPARQL Queries by Exploiting Hash-based Locality and Adaptive Partitioning**

Razen Harbi (King Abdullah University of Science and Technology); Ibrahim Abdelaziz (King Abdullah University of Science and Technology); Panos Kalnis (King Abdullah University of Science and Technology); Nikos Mamoulis (University of Ioannina); Yasser Ebrahim (King Abdullah University of Science and Technology); Majed Sahli (King Abdullah University of Science and Technology);

#### **Know Your Customer: Computing k-Most Promising Products for Targeted Marketing**

Md. Saiful Islam (Swinburne University of Technology); Chengfei Liu (Swinburne University of Technology)

#### **Bitwise Dimensional Co-Clustering for Analytical Workloads**

Stephan Baumann (Technische Universität Ilmenau); Peter Boncz (CWI); Kai-Uwe Sattler (Technische Universität Ilmenau)

## **Wednesday Sept 7th 15:30 - 16:00**

**Tea Break**

## **Wednesday Sept 7th 16:00 - 17:30**

### **Research 16: Data and Query Models - 1**

**Location: Pearl 1**

**Session Chair: Yanlei Diao (University of Massachusetts Amherst)**

#### **SlimShot: In-Database Probabilistic Inference for Knowledge Bases**

Eric Gribkoff (University of Washington); Dan Suciu (University of Washington)

#### **Scalable Package Queries in Relational Database Systems**

Matteo Brucato (University of Massachusetts Amherst); Juan Beltran (New York University Abu Dhabi); Azza Abouzied (New York University Abu Dhabi); Alexandra Meliou (University of Massachusetts Amherst)

#### **Explaining Query Answers with Explanation-Ready Databases**

Sudeepa Roy (Duke University); Laurel Orr (University of Washington); Dan Suciu (University of Washington)

#### **A framework for annotating CSV-like data**

Marcelo Arenas (PUC Chile); Francisco Maturana (PUC Chile); Cristian Riveros (PUC Chile); Domagoj Vrgoc (PUC Chile)

### **Research 17: Scalable Analytics - 1**

**Location: Pearl 2**

**Session Chair: Matthias Boehm (IBM Research - Almaden)**

#### **A General-Purpose Query-Centric Framework for Querying Big Graphs**

Da Yan (The Chinese University of Hong Kong); James Cheng (The Chinese University of Hong Kong); Tamer Özsu (University of Waterloo); Fan Yang (The Chinese University of Hong Kong); Yi Lu (The Chinese University of Hong Kong); John C.S. Lui (The Chinese University of Hong Kong); Qizhen Zhang (The Chinese University of Hong Kong); Wilfred Ng (The Hong Kong University of Science and Technology)

### Distance-based Outlier Detection in Data Streams

Luan Tran (University of Southern California); Lijue Fan (University of Southern California); Cyrus Shahabi (University of Southern California)

### SEEDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics

Manasi Vartak (MIT); Sajjadur Rahman (University of Illinois); Samuel Madden (MIT); Aditya Parameswaran (University of Illinois at Urbana-Champaign); Neoklis Polyzotis (Google)

### Explaining Outputs in Modern Data Analytics

Zaheer Chothia (ETH Zurich); John Liagouris (ETH Zurich); Frank McSherry (ETH Zurich); Timothy Roscoe (ETH Zurich)

## Research 18: Database Hardware-Software Codesign

**Location: Royal 1**

**Session Chair: Ryan Stutsman (University of Utah)**

### YourSQL: A High-Performance Database System Leveraging In-Storage Computing

Insoon Jo (Samsung Electronics Co.); Duck-Ho Bae (Samsung Electronics Co.); Andre Yoon (Samsung Electronics Co.); Jeong-Uk Kang (Samsung Electronics Co.); Sangyeun Cho (Samsung Electronics Co.); Daniel Lee (Samsung Electronics Co.); Jaeheon Jeong (Samsung Electronics Co.)

### Cheap Data Analytics using Cold Storage Devices

Renata Borovica-Gajic (EPFL); Raja Appuswamy (EPFL); Anastasia Ailamaki (EPFL)

### An Experimental Evaluation of Datacenter Workloads On Low-Power Embedded Micro Servers

Yiran Zhao (University of Illinois at Urbana-Champaign); Shen Li (University of Illinois at Urbana-Champaign); Shaohan Hu (University of Illinois at Urbana-Champaign); Hongwei Wang (University of Illinois at Urbana-Champaign); Shuochao Yao (University of Illinois at Urbana-Champaign); Huajie Shao (University of Illinois at Urbana-Champaign); Tarek Abdelzaher (University of Illinois at Urbana-Champaign)

## Industrial 5: Graph Systems and Analytics

**Location: Royal 2**

**Session Chair: Prasad Deshpande (Consultant)**

### LDDB Graphalytics: A Benchmark for Large-Scale Graph Analysis on Parallel and Distributed Platforms

Alexandru Iosup (Delft University of Technology); Tim Hegeman (Delft University of Technology); Wing Lung Ngai (Delft University of Technology); Stijn Heldens (Delft University of Technology); Arnau Prat-Pérez (Universitat Politècnica de Catalunya); Thomas Manhardt (Oracle Labs); Hassan Chafi (Oracle Labs); Mihai Capotă (Intel Labs); Narayanan Sundaram (Intel Labs); Michael Anderson (Intel Labs); Ilie Gabriel Tanase (IBM Research); Yinglong Xia (Huawei Research America); Lifeng Nai (Georgia Tech); Peter Boncz (CWI)

### GraphJet: Real-Time Content Recommendations at Twitter

Aneesh Sharma (Twitter Inc.); Jerry Jiang (Twitter Inc.); Praveen Bommannavar (Twitter Inc.); Brian Larson (Twitter Inc.); Jimmy Lin (University of Waterloo)

### Using Domain-Specific Languages for Analytic Graph Databases

Martin Sevenich (Oracle Labs); Sungpack Hong (Oracle Labs); Oskar Van Rest (Oracle Labs); Zhe Wu (Oracle); Jayanta Banerjee (Oracle); Hassan Chafi (Oracle Labs)

### Cubrick: Indexing Millions of Records per Second for Interactive Analytics

Pedro Pedreira (Facebook Inc); Chris Crosswhite (Facebook Inc.); Luis Bona (Federal University of Parana)

## Demo 2B: Interactive and Exploratory Systems

### Research Posters 6 (Research 26, Research 27, Research 28, Research 29 and Research 30)

**Location: Maple**

**Same as Demo 2A**

**Research Posters 6:** Papers from Research 26, Research 27, Research 28, Research 29 and Research 30

## Wednesday Sept 7th 17:45 - 18:00

Depart for Banquet Venue

## Wednesday Sept 7th 19:30 - 22:00

VLDB Conference Banquet





**Day 4 - Thursday, September 8th**



### VLDB Awards Session

**Location: Pearl**

#### VLDB 2016 Awards: Best Paper and Best Demonstration awards

##### Best Paper Award -

Compressed Linear Algebra for Large-Scale Machine Learning

Ahmed Elgohary (University of Maryland), Matthias Boehm (IBM Research - Almaden), Peter Haas (IBM Research), Fred Reiss (IBM Research - Almaden), and Berthold Reinwald (IBM Research, Almaden) for "Proposing a comprehensive solution to the challenge of compressing matrices for linear algebra operations in the context of machine learning tasks".

### VLDB Endowment Awards

#### VLDB 10-year Best Paper Award -

The New Casper: Query Processing for Location Services without Compromising Privacy

Mohamed F. Mokbel, Chi-Yin Chow, Walid G. Aref VLDB 2006 for inspiring research to support privacy in spatial query processing

#### VLDB Early Career Research Contribution Award -

Xin Luna Dong for advancing the state of the art of knowledge fusion

#### VLDB Women in Database Research Award -

Magdalena Balazinska for her inspirational research record on scalable distributed data systems

### VLDB 10-year Best Paper Award Talk: Mohamed Mokbel, Chi-Yin Chow, and Walid Aref

#### Session Chair: Sunita Sarawagi (IIT Bombay)

#### Location Data Management: A Tale of Two Systems and the "Next Destination"!

In early 2000, we had the vision of ubiquitous location services, where each object is aware of its location, and continuously sends its location to a designated database server. This flood of location data opened the door for a myriad of location-based services that were considered visionary at that time, yet today they are a reality and have become ubiquitous. To realize our early vision, we identified two main challenges that needed to be addressed, namely, scalability and privacy. We have addressed these challenges through two main systems, PLACE and Casper. PLACE, developed at Purdue University from 2000 to 2005, set up the environment for built-in database support of scalable and continuous location-based services. The Casper system, developed at University of Minnesota from 2005 to 2010, was built inside the PLACE server allowing it to provide its high quality scalable service, while maintaining the privacy of its users' locations. This talk will take you through a time journey of location services from 2000 until today, and beyond, highlighting the development efforts of the PLACE and Casper systems, along with their impact on current and future research initiatives in both academia and industry.

#### Bios:



**Mohamed F. Mokbel** (Ph.D., Purdue University, MS, B.Sc., Alexandria University) is an Associate Professor in the Department of Computer Science and Engineering, University of Minnesota. His research interests include the interaction of GIS and location-based services with database systems and cloud computing. His research work has been recognized by five Best Paper Awards and by the NSF CAREER award. Mohamed is/was the program co-chair of SIGMOD 2018, ACM SIGSPATIAL from 2008 to 2010, and IEEE MDM 2011 and 2014. He is an Associate Editor for ACM TODS, ACM TSAS, VLDB journal, and GeoInformatica. Mohamed is currently serving as the elected Chair of ACM SIGSPATIAL 2014-2017. For more information, please visit: [www.cs.umn.edu/~mokbel](http://www.cs.umn.edu/~mokbel).



**Chi-Yin Chow** received the B.A. and M.Phil. degrees from The Hong Kong Polytechnic University, Hong Kong in 2002 and 2005, respectively. He also received the M.S. and Ph.D. degrees from the University of Minnesota-Twin Cities, USA in 2008 and 2010, respectively. He is currently an assistant professor in Department of Computer Science, City University of Hong Kong. His research interests include big data analytics, data management, GIS, mobile computing, location-based services, and data privacy. He is the co-founder and co-chair of the ACM SIGSPATIAL MobiGIS 2012 to 2016, and the editor of the ACM SIGSPATIAL Newsletter. Dr. Chow received the best paper awards in ICA3PP 2015 and IEEE MDM 2009.



**Walid G. Aref** is a professor of computer science at Purdue. His research interests are in extending the functionality of database systems in support of emerging applications, e.g., spatial, spatio-temporal, multimedia, biological, and sensor databases. He is also interested in query processing, indexing, data mining, and geographic information systems (GIS). Professor Aref's research has been supported by the National Science Foundation, the National Institute of Health, Purdue Research Foundation, Qatar National Research Foundation, CERIAS, Panasonic, and Microsoft Corp. In 2001, he received the CAREER Award from the National Science Foundation and in 2004, he received a Purdue University Faculty Scholar award. Professor Aref is a member of Purdue's CERIAS. He is an associate editor of the ACM Transactions of Database Systems (ACM TODS) and the ACM Transactions of Spatial

Algorithms and Systems (TSAS), and has been an editor of the VLDB Journal. He is a senior member of the IEEE, and a member of the ACM. Professor Aref is an executive committee member, co-founder, and the past chair of the ACM Special Interest Group on Spatial Information (SIGSPATIAL).

## VLDB Early Career Research Contribution Award Talk: Xin Luna Dong

**Session Chair: Sunita Sarawaqi (IIT Bombay)**

### Leave No Valuable Data Behind: The Crazy Ideas and the Business

With the mission "leave no valuable data behind", we developed techniques for knowledge fusion to guarantee the correctness of the knowledge. This talk starts with describing a few crazy ideas we have tested. The first, known as "Knowledge Vault", used 15 extractors to automatically extract knowledge from 1B+ Webpages, obtaining 3B+ distinct (subject, predicate, object) knowledge triples and predicting well-calibrated probabilities for extracted triples. The second, known as "Knowledge-Based Trust", estimated the trustworthiness of 119M webpages and 5.6M websites based on the correctness of their factual information. We then present how we bring the ideas to business in filling the gap between the knowledge at Google Knowledge Graph and the knowledge in the world.

**Bio:**



**Xin Luna Dong** is a Principal Scientist at Amazon, leading the efforts of constructing Amazon Product Graph. She was one of the major contributors to the Knowledge Vault project, and has led the Knowledge-based Trust project, which is called the "Google Truth Machine" by Washington's Post. She has co-authored book "Big Data Integration", published 65+ papers in top conferences and journals, given 20+ keynotes/invited-talks/tutorials, and got the Best Demo award in Sigmod 2005. She is the PC co-chair for Sigmod 2018 and WAIM 2015, and serves as an area chair for Sigmod 2017, Sigmod 2015, ICDE 2013, and CIKM 2011.

## Thursday Sept 8th 10:30 - 11:15

Tea Break

## Thursday Sept 8th 11:15 - 12:45

### Research 19: Query Optimization - 2

**Location: Pearl 1**

**Session Chair: Johann Christoph-Freytag (Humboldt University)**

#### Optimization of Conjunctive Predicates for Main Memory Column Stores

Fisnik Kastrati (University of Mannheim); Guido Moerkotte (University of Mannheim)

#### Parallelizing Query Optimization on Shared-Nothing Architectures

Immanuel Trummer (EPFL); Christoph Koch (EPFL)

#### Multiple Query Optimization on the D-Wave 2X Adiabatic Quantum Computer

Immanuel Trummer (EPFL); Christoph Koch (EPFL)

### Research 20: Spatial Data and Queries - 2

**Location: Pearl 2**

**Session Chair: Walid G. Aref (Purdue University)**

#### k-Nearest Neighbors on Road Networks: A Journey in Experimentation and In-Memory Implementation

Tenindra Abeywickrama (Monash University); Muhammad Cheema (Monash University); David Tanian (Monash University)

#### SKYPE: Top-k Spatial-keyword Publish/Subscribe Over Sliding Window

Xiang Wang (University of New South Wales); Ying Zhang (University of Technology Sydney); Wenjie Zhang (University of New South Wales); Xuemin Lin (University of New South Wales); Zengfeng Huang (University of New South Wales)

#### Maximizing Bichromatic Reverse Spatial and Textual k Nearest Neighbor Queries

Farhana Choudhury (RMIT University); Shane Culpepper (RMIT University); Timos Sellis (RMIT University); Xin Cao (Queen's University, Belfast)

#### Cache locality is not enough: High-Performance Nearest Neighbor Search with Product Quantization Fast Scan

Fabien André (Technicolor); Anne-Marie Kermarrec (INRIA); Nicolas Le Scouarnec (Technicolor)

## Research 21: Distributed and Cloud Systems - 3

**Location:** Royal 1

**Session Chair:** Mike Carey (University of California at Irvine)

**WiSeDB: A Learning-based Workload Management Advisor for Cloud Databases**

Ryan Marcus (Brandeis University); Olga Papaemmanouil (Brandeis University)

**Husky: Towards a More Efficient and Expressive Distributed Computing Framework**

Fan Yang (The Chinese University of Hong Kong); Jinfeng Li (The Chinese University of Hong Kong); James Cheng (The Chinese University of Hong Kong)

**Titian: Data Provenance Support in Spark**

Matteo Interlandi (University of California at Los Angeles); Kshitij Shah (University of California at Los Angeles); Sai Tetali (University of California at Los Angeles); Muhammad Gulzar (University of California at Los Angeles); Seunghyun Yoo (University of California at Los Angeles); Miryung Kim (University of California at Los Angeles); Todd Millstein (University of California at Los Angeles); Tyson Condie (University of California at Los Angeles)

## Tutorial 5: Human Factors in Crowdsourcing

**Location:** Royal 2

**Siheem Amer-Yahia (CNRS); Senjuti Basu Roy (New Jersey Institute of Technology)**

## Research 22: Entity Matching - 1

**Location:** Maple

**Session Chair:** Ioana Manolescu (INRIA)

**Online Entity Resolution Using an Oracle**

Donatella Firmani (University of Rome "Tor Vergata"); Barna Saha (University of Massachusetts Amherst); Divesh Srivastava (AT&T Labs Research)

**BLAST: a Loosely Schema-aware Meta-blocking Approach for Entity Resolution**

Giovanni Simonini (Università di Modena e Reggio Emilia); Sonia Bergamaschi (Università di Modena e Reggio Emilia); H. V. Jagadish (University of Michigan)

**QuERy: A Framework for Integrating Entity Resolution with Query Processing**

Hotham Altawjry (University of California at Irvine); Sharad Mehrotra (University of California at Irvine); Dmitri Kalashnikov (AT&T Labs Research)

**Comparative Analysis of Approximate Blocking Techniques for Entity Resolution**

George Papadakis (University of Athens); Jonathan Svirsky (Technion); Avigdor Gal (Technion); Themis Palpanas (Paris Descartes University)

## Thursday Sept 8th 12:45 - 14:00

Lunch Break

## Thursday Sept 8th 14:00 - 15:30

### Research 23: Query Execution - 3

**Location:** Pearl 1

**Session Chair:** Peter Boncz (CWI)

**An Empirical Evaluation of Set Similarity Join Techniques**

Willi Mann (University of Salzburg); Nikolaus Augsten (University of Salzburg); Panagiotis Bouras (Aarhus University)

**Streaming Similarity Self-Join**

Gianmarco De Francisci Morales (Qatar Computing Research Institute); Aristides Gionis (Aalto University)

**An Efficient Partition Based Method for Exact Set Similarity Joins**

Dong Deng (Tsinghua University); Guoliang Li (Tsinghua University); He Wen (Tsinghua University); Jianhua Feng (Tsinghua University)

**Fast Queries Over Heterogeneous Data Through Engine Customization**

Manos Karpapathiotakis (EPFL); Ioannis Alagiannis (EPFL); Anastasia Ailamaki (EPFL)

## Research 24: Social Networks and Crowdsourcing

**Location:** Pearl 2

**Session Chair:** Sihem Amer-Yahia (CNRS)

### **Crowdsourced Top-k Algorithms: An Experimental Evaluation**

Xiaohang Zhang (Tsinghua University); Guoliang Li (Tsinghua University); Jianhua Feng (Tsinghua University)

### **CLAMShell: Speeding up Crowds for Low-latency Data Labeling**

Daniel Haas (University of California at Berkeley); Jiannan Wang (Simon Fraser University); Eugene Wu (Columbia University); Michael Franklin (University of California at Berkeley)

### **Dynamic Influence Analysis in Evolving Networks**

Naoto Ohsaka (The University of Tokyo); Takuya Akiba (National Institute of Informatics); Yuichi Yoshida (National Institute of Informatics); Ken-ichi Kawarabayashi (National Institute of Informatics)

### **From Competition to Complementarity: Comparative Influence Diffusion and Maximization**

Wei Lu (University of British Columbia); Wei Chen (Microsoft Research); Laks Lakshmanan (University of British Columbia)

## Research 25: Graph Processing - 3

**Location:** Royal 1

**Session Chair:** Mohamed Mokbel (University of Minnesota)

### **Data-driven Visual Graph Query Interface Construction and Maintenance: Challenges and Opportunities**

Sourav Bhowmick (Nanyang Technological University); Byron Choi (Hong Kong Baptist University); Curtis Dyreson (Utah State University)

### **The shortest path is not always a straight line**

Vasiliki Kalavri (KTH Royal Institute of Technology); Tiago Simas (Telefonica Research); Dionysios Logothetis (Facebook)

### **New Lower and Upper Bounds for Shortest Distance Queries on Terrains**

Manohar Kaul (Technische Universität Berlin); Raymond Chi-Wing Wong (The Hong Kong University of Science and Technology); Christian Jensen (Aalborg University)

### **On Measuring the Lattice of Commonalities Among Several Linked Datasets**

Michalis Mountantonakis (FORTH); Yannis Tzitzikas (FORTH)

## Tutorial 6: Qualitative Data Cleaning

**Location:** Royal 2

**Xu Chu (University of Waterloo); Ihab Ilyas (University of Waterloo)**

## Research 26: Data and Query Models - 2

**Location:** Maple

**Session Chair:** Tim Kraska (Brown University)

### **Interactive Browsing and Navigation in Relational Databases**

Minsuk Kahng (Georgia Institute of Technology); Shangkant Navathe (Georgia Institute of Technology); John Stasko (Georgia Institute of Technology); Duen Horng Chau (Georgia Institute of Technology)

### **ATHENA: An Ontology-Driven System for Natural Language Querying over Relational Data Stores**

Diptikalyan Saha (IBM Research - India); Avriilia Floratou (IBM Research - Almaden); Karthik Sankaranarayanan (IBM Research - India); Umar Farooq Minhas (IBM Research - Almaden); Ashish Mittal (IBM Research - India); Fatma Ozcan (IBM Research - Almaden)

### **BlinkFill: Semi-supervised Programming by Example for Syntactic String Transformations**

Rishabh Singh (Microsoft Research)

### **Query from Examples: An Iterative, Data-Driven Approach to Query Construction**

Hao Li (National University of Singapore); Chee-Yong Chan (National University of Singapore); David Maier (Portland State University)

**Thursday Sept 8th 15:30 - 16:00**

**Tea Break**

## Thursday Sept 8th 16:00 - 17:30

### Research 27: Data and Query Models - 3

**Location:** Pearl 1

**Session Chair:** Sudeepa Roy (Duke University)

#### Exploiting Equality Generating Dependencies in Checking Chase Termination

Marco Calautti (University of Calabria); Sergio Greco (University of Calabria); Cristian Molinaro (University of Calabria); Irina Trubitsyna (University of Calabria)

#### The Complexity of Resilience and Responsibility for Self-Join-Free Conjunctive Queries

Cibele Freire (University of Massachusetts Amherst); Wolfgang Gatterbauer (Carnegie Mellon University); Neil Immerman (University of Massachusetts Amherst); Alexandra Meliou (University of Massachusetts Amherst)

#### A Shifting Bloom Filter Framework for Set Queries

Tong Yang (Peking University); Alex X. Liu (Michigan State University); Muhammad Shahzad (North Carolina State University); Yuankun Zhong (Nanjing University); Qiaobin Fu (Boston University); Zi Li (Nanjing University); Gaogang Xie (ICT, CAS); Xiaoming Li (Peking University)

#### Decibel: The Relational Dataset Branching System

Michael Maddox (MIT); David Goehring (MIT); Aaron Elmore (University of Chicago); Samuel Madden (MIT); Aditya Parameswaran (University of Illinois at Urbana-Champaign); Amol Deshpande (University of Maryland)

### Research 28: Entity Matching - 2

**Location:** Pearl 2

**Session Chair:** Themis Palpanas (Paris Descartes University)

#### Magellan: Toward Building Entity Matching Management Systems

Pradap Konda (University of Wisconsin Madison); Sanjib Das (University of Wisconsin Madison); Paul Suganthan G. C. (University of Wisconsin Madison); Anhui Doan (University of Wisconsin Madison); Adel Ardalan (University of Wisconsin Madison); Jeff Ballard (University of Wisconsin Madison); Han Li (University of Wisconsin Madison); Fatemah Panahi (University of Wisconsin Madison); Haojun Zhang (University of Wisconsin Madison); Jeffrey Naughton (University of Wisconsin Madison); Shishir Prasad (Instacart); Ganesh Krishnan (WalmartLabs); Rohit Deep (WalmartLabs); Vijay Raghavendra (WalmartLabs)

#### Distributed Data Deduplication

Xu Chu (University of Waterloo); Ihab Ilyas (University of Waterloo); Paraschos Koutris (University of Wisconsin Madison)

#### The iBench Integration Metadata Generator

Patricia Arocena (University of Toronto); Boris Glavic (Illinois Institute of Technology); Radu Ciucanu (University of Oxford); Renée Miller (University of Toronto)

#### Schema-agnostic vs Schema-based Configurations for Blocking Methods on Homogeneous Data

George Papadakis (University of Athens); George Alexiou (IMIS, "Athena" Research Center); George Papastefanatos (IMIS, "Athena" Research Center); Georgia Koutrika (Hewlett-Packard Labs)

### Research 29: Community Search and Mining

**Location:** Royal 1

**Session Chair:** Tamer Öszu (University of Waterloo)

#### Effective Community Search for Large Attributed Graphs

Yixiang Fang (The University of Hong Kong); Reynold C.K. Cheng (The University of Hong Kong); Siqiang Luo (The University of Hong Kong); Jiafeng Hu (The University of Hong Kong)

#### Approximate Closest Community Search in Networks

Xin Huang (University of British Columbia); Laks Lakshmanan (University of British Columbia); Jeffrey Xu Yu (The Chinese University of Hong Kong); Hong Cheng (The Chinese University of Hong Kong)

#### Ego-net Community Mining Applied to Friend Suggestion

Alessandro Epasto (Brown University); Silvio Lattanzi (Google); Vahab Mirrokni (Google); Ismail Sebe (Google); Ahmed Taei (Google); Sunita Verma (Google)

#### Tracking the Conductance of Rapidly Evolving Topic-Subgraphs

Sainyam Galhotra (Xerox Research Centre India); Amitabha Bagchi (IIT Delhi); Srikanta Bedathur (IBM Research - India); Maya Ramanath (IIT Delhi); Vidit Jain (American Express Big Data Labs)

## Research 30: Scalable Analytics - 2

**Location:** Maple

**Session Chair:** Jonathan Goldstein (Microsoft Research)

### **Streaming Anomaly Detection Using Randomized Matrix Sketching**

Hao Huang (General Electric Global Research); Shiva Kasiviswanathan (Samsung Research America)

### **Index-Assisted Hierarchical Computations in Main-Memory RDBMS**

Robert Brunel (Technische Universität München); Norman May (SAP SE); Alfons Kemper (Technische Universität München)

### **DEXTER: Large-Scale Discovery and Extraction of Product Specifications on the Web**

Disheng Qiu (Università degli Studi Roma Tre); Luciano Barbosa (IBM Research - Brazil); Xin Luna Dong (Google); Yanyan Shen (National University of Singapore); Divesh Srivastava (AT&T Labs - Research)

### **Walking in the Cloud: Parallel SimRank at Scale**

Zhenguo Li (Huawei); Yixiang Fang (The University of Hong Kong); Qin Liu (The Chinese University of Hong Kong); Jiefeng Cheng (Huawei); Reynold Cheng (University of Hong Kong); John Lui (The Chinese University of Hong Kong)

## Thursday Sept 8th 18:30 - 21:30

**Entertainment Program and Gala Dinner**

**Location:** Pearl





**Day 5 - Friday, September 9th**



## Friday Sept 9th 9:00 - 10:30

### Workshop 4: Ph.D. Workshop

**Location: Pearl 1**

### Joint Workshop (5): Accelerating Data Management Systems (ADMS) and In-memory Data Management & Analytics (IMDM)

#### **Location: Pearl 2**

ADMS - Rajesh Bordawekar (IBM Research - T.J. Watson); Tirthankar Lahiri (Oracle) IMDM - Spyros Blanas (The Ohio State University); Justin Levandoski (Microsoft Research); Andy Pavlo (Carnegie Mellon University)

### Workshop 6: Data Management and Analytics for Medicine and Healthcare (DMAH)

#### **Location: Royal 1**

Fusheng Wang (Stony Brook University); Lixia Yao (University of North Carolina at Charlotte); Gang Luo (University of Utah)

### Workshop 7: Social Data Analytics and Management (SoDAM)

#### **Location: Royal 2**

Maheshweta Das (Hewlett Packard Labs); Gautam Das (University of Texas at Arlington)

### Workshop 8: Big Data Open-Source Systems (BOSS) / Polyglot

#### **Location: Maple**

Tilman Rabl (TU Berlin); Sebastian Schelter (Amazon)

## Friday Sept 9th 10:30 - 11:00

Tea Break

## Friday Sept 9th 11:00 - 12:30

### Workshop 4: Ph.D. Workshop (contd..)

**Location: Pearl 1**

### Joint Workshop (5): Accelerating Data Management Systems (ADMS) and In-memory Data Management & Analytics (IMDM) (contd..)

#### **Location: Pearl 2**

ADMS - Rajesh Bordawekar (IBM Research - T.J. Watson); Tirthankar Lahiri (Oracle) IMDM - Spyros Blanas (The Ohio State University); Justin Levandoski (Microsoft Research); Andy Pavlo (Carnegie Mellon University)

### Workshop 6: Data Management and Analytics for Medicine and Healthcare (DMAH) (contd..)

#### **Location: Royal 1**

Fusheng Wang (Stony Brook University); Lixia Yao (University of North Carolina at Charlotte); Gang Luo (University of Utah)

### Workshop 7: Social Data Analytics and Management (SoDAM) (contd..)

#### **Location: Royal 2**

Maheshweta Das (Hewlett Packard Labs); Gautam Das (University of Texas at Arlington)

### Workshop 8: Big Data Open-Source Systems (BOSS) / Polyglot (contd..)

#### **Location: Maple**

Tilman Rabl (TU Berlin); Sebastian Schelter (Amazon)

## Friday Sept 9th 12:30 - 14:00

Lunch Break

## Friday Sept 9th 14:00 - 15:30

Workshop 4: Ph.D. Workshop (contd..)

Location: Pearl 1

Joint Workshop (5): Accelerating Data Management Systems (ADMS) and In-memory Data Management & Analytics (IMDM) (contd..)

Location: Pearl 2

ADMS - Rajesh Bordawekar (IBM Research - T.J. Watson); Tirthankar Lahiri (Oracle) IMDM - Spyros Blanas (The Ohio State University); Justin Levandoski (Microsoft Research); Andy Pavlo (Carnegie Mellon University)

Workshop 6: Data Management and Analytics for Medicine and Healthcare (DMAH) (contd..)

Location: Royal 1

Fusheng Wang (Stony Brook University); Lixia Yao (University of North Carolina at Charlotte); Gang Luo (University of Utah)

Workshop 8: Big Data Open-Source Systems (BOSS) / Polyglot (contd..)

Location: Maple

Tilmann Rabl (TU Berlin); Sebastian Schelter (Amazon)

## Friday Sept 9th 15:30 - 16: 00

Tea Break

## Friday Sept 9th 16:00 - 17:30

Workshop 4: Ph.D. Workshop (contd..)

Location: Pearl 1

Joint Workshop (5): Accelerating Data Management Systems (ADMS) and In-memory Data Management & Analytics (IMDM) (contd..)

Location: Pearl 2

ADMS - Rajesh Bordawekar (IBM Research - T.J. Watson); Tirthankar Lahiri (Oracle) IMDM - Spyros Blanas (The Ohio State University); Justin Levandoski (Microsoft Research); Andy Pavlo (Carnegie Mellon University)

Workshop 6: Data Management and Analytics for Medicine and Healthcare (DMAH) (contd..)

Location: Royal 1

Fusheng Wang (Stony Brook University); Lixia Yao (University of North Carolina at Charlotte); Gang Luo (University of Utah)

Workshop 8: Big Data Open-Source Systems (BOSS) / Polyglot (contd..)

Location: Maple

Tilmann Rabl (TU Berlin); Sebastian Schelter (Amazon)





# Abstracts

# Tutorial Abstracts

---

## Tutorial 1: Differential Privacy in the Wild: A tutorial on current practices & open challenges

**Ashwin Machanavajhala (Duke University), Xi He (Duke University) and Michael Hay (Colgate University)**

Differential privacy has emerged as an important standard for privacy preserving computation over databases containing sensitive information about individuals. Research on differential privacy spanning a number of research areas, including theory, security, database, networks, machine learning, and statistics, over the last decade has resulted in a variety of privacy preserving algorithms for a number of analysis tasks. Despite maturing research efforts, the adoption of differential privacy by practitioners in industry, academia, or government agencies has so far been rare. Hence, in this tutorial, we will first describe the foundations of differentially private algorithm design that cover the state of the art in private computation on tabular data. In the second half of the tutorial we will highlight real world applications on complex data types, and identify research challenges in applying differential privacy to real world applications.

### Bios:

**Ashwin Machanavajhala** is an Assistant Professor in the Department of Computer Science, Duke University and an Associate Director at the Information Initiative@Duke (iID). Previously, he was a Senior Research Scientist in the Knowledge Management group at Yahoo! Research. His primary research interests lie in algorithms for ensuring privacy in statistical databases and augmented reality applications. He is a recipient of the National Science Foundation Faculty Early CAREER award in 2013, and the 2008 ACM SIGMOD Jim Gray Dissertation Award Honorable Mention. Ashwin graduated with a Ph.D. from the Department of Computer Science, Cornell University and a B.Tech in Computer Science and Engineering from the Indian Institute of Technology, Madras. His early work on I-diversity has been very influential in the field of data privacy and has been cited over 2500 times (according to Google Scholar). He also helped design one of the first real data publication powered by formal privacy guarantees in collaboration with the US Census Bureau in 2008. He has published in PODS, SIGMOD, VLDB, ICDE, WWW and WSDM, and has given tutorials on privacy at IEEE SSP 2009, ICDE 2010, and on entity resolution at AAAI 2012, VLDB 2012 and KDD 2013.

**Xi He** is a PhD student at Computer Science Department, Duke University. Her research interests lie in privacy-preserving data analysis and security. She has also received an M.S from Duke University and double degree in Applied Mathematics and Computer Science from University of Singapore. Xi has been working with Prof. Machanavajhala on privacy since 2012, and has published in SIGMOD and VLDB.

**Michael Hay** is an Assistant Professor in the Department of Computer Science, at Colgate University. Before that he was a Computing Innovation Fellow at Cornell University and completed his PhD at UMass Amherst in 2010. His research interests include privacy-preserving data analysis, data management, data mining, social networks, and privacy. His PhD thesis titled "Enabling Accurate Analysis of Private Network Data" is the recipient of the 2011 ACM SIGKDD Dissertation Award. His ICDM 2009 paper titled "Accurate estimation of the degree distribution of private networks" received the Best Student Paper award. He has given a tutorial on privacy and graphs at SIGMOD 2011.

## Tutorial 2: Operational Analytics Data Management Systems

**Alexander Boehm (SAP), Jens Dittrich (University of Saarland), Niloy Mukherjee (LinkedIn), Ippokratis Pandis (Amazon), Rajkumar Sen (Oracle)**

As enterprise businesses become more agile and responsive to trends, sentiments and surges, business analytics applications can no longer rely on the classic data warehousing model of attempting to derive insights from data at rest. The data for analytics is as current as of the last ETL (Extract, Transform and Load) job that moved data from the operational system to the data warehouse, and business can no longer afford missing on real-time insights on the data that is fresh in the operational system. For example, a retail store could run analytics on transactional data to track sales and use the information to offer discounts. There is an increasing demand for database management systems to be able to perform real-time analytics on data that gets ingested and modified in live mainstream operational databases. As a response, many commercial vendors as well as academia have attempted to solve the problem by combining transactional and analytical processing capabilities in the same database system; these systems will be referred to as operational analytics systems.

In this tutorial, we shall present an in-depth overview of operational analytical systems. We shall start with a discussion on the various aspects associated with the design of such a system; ranging from data storage, indexing to query optimization and processing. We shall then present a set of representative systems in detail, highlight their individual architecture and design characteristics, and discuss several key research problems they address. This tutorial is intended for both researchers and practitioners in the industry.

## Bios:

**Alexander Böhm** is a database architect working on SAP's HANA in-memory database management system. His focus is on performance optimization and holistic improvements of enterprise architectures, in particular application server/DBMS co-design. Prior to joining SAP, he received his PhD from the University of Mannheim, Germany, where he worked on the development of efficient and scalable applications using declarative message processing.

**Jens Dittrich** is a full professor of Computer Science in the area of Databases, Data Management, and Big Data at Saarland University, Germany. Previous affiliations include U Marburg, SAP AG, and ETH Zurich. He received an Outrageous Ideas and Vision Paper Award at CIDR 2011, a BMBF VIP Grant in 2011, a best paper award at VLDB 2014, two CS teaching awards in 2011 and 2013, as well as several presentation awards including a qualification for the interdisciplinary German science slam finals in 2012 as well as three presentation awards at CIDR (2011, 2013, and 2015). He likes producing educational database videos (<http://youtube.com/jensdit>) and flipped textbooks (<http://amzn.to/1Ts3rwx>).

**Niloy Mukherjee** is a staff engineer at Distributed Data Systems, LinkedIn Corporation, working on Espresso, LinkedIn's distributed, fault-tolerant, online data store service, which currently powers major LinkedIn applications including member profiles, InMail, member engagements, desktop and mobile applications, etc. Prior to that, he has been a consulting member of technical staff at Oracle RDBMS Data and In-memory Technologies where he was one of the primary architects of Oracle Database In-memory Option, a fully consistent dual-format distributed in-memory RDBMS aimed to provide real-time analytics at scale on traditional OLAP as well as on mixed OLTP workloads. He has published and presented his work at several VLDB, SIGMOD, and ICDE conferences, and has been awarded 20+ granted and pending patents. He has been mentioned in Business Insider as one of '26 Oracle rock star engineers changing the company'. He received his Master's degree from the Media Laboratory, Massachusetts Institute of Technology, and holds a Bachelor degree from the department of Computer Science, Indian Institute of Technology, Kharagpur.

**Ippokratis Pandis** is a principal engineer at Amazon Web Services working on AWS Redshift. AWS Redshift is Amazon's fully managed, petabyte scale data warehouse service. Previously, Ippokratis has held positions as software engineer at Cloudera where he worked on the Impala SQL-on-Hadoop query engine and as member of the research staff at IBM Almaden Research Center. At IBM, he was member of the team that designed and implemented the BLU column-store engine, which currently ships as part of IBM's DB2 LUW v10.5 with BLU Acceleration. Ippokratis received his PhD from Carnegie Mellon University. He is the recipient of best demonstration and paper awards at ICDE 2006, SIGMOD 2011 and CIDR 2013. He has also served as PC chair of DaMoN 2014, DaMoN 2015 and CloudDM 2016.

**Rajkumar Sen** is a Software Development Director at Oracle Inc. responsible for architecting Oracle's Business Intelligence Analytics for the Cloud. Prior to that, he was a Director, Engineering at MemSQL Inc. where he architected the query optimizer and the distributed query processing engine, Principal Engineer at Oracle Inc. where he developed features for the Oracle database query optimizer, and Senior Staff Engineer at Sybase Inc. where he architected the Distributed Object Lock Manager and Metadata Manager for Sybase ASE Cluster Edition. He received his Masters Degree in Computer Science with specialization in Databases from Indian Institute of Technology, Bombay in 2004 and his current research interests are in the areas of query optimization and distributed query processing.

## Tutorial 3: Modern Main-Memory Database Systems

### Per Ake Larson (Microsoft Research), Justin Levandoski (Microsoft Research)

This tutorial provides an overview of recent developments in main-memory database systems. With growing memory sizes and memory prices dropping by a factor of 10 every 5 years, data having a "primary home" in memory is now a reality. Main-memory databases eschew many of the traditional architectural tenets of relational database systems that optimized for disk-resident data. Innovative approaches to fundamental issues such as concurrency control and query processing are required to unleash the full performance potential of main-memory databases. The tutorial is focused around design issues and architectural choices that must be made when building a high performance database system optimized for main-memory: data storage and indexing, concurrency control, durability and recovery techniques, query processing and compilation, support for high availability, and ability to support hybrid transactional and analytics workloads. This will be illustrated by example solutions drawn from four state-of-the-art systems: HStore/VoltdB, Hekaton, HyPeR, and SAP HANA. The tutorial will also cover current and future research trends.

## Bios:

**Per-Ake (Paul) Larson:** Paul has conducted research in the database field for well over 30 years. He served as a Professor in the Department of Computer Science at the University of Waterloo for 15 years and as a Principal Researcher at Microsoft Research for close to 20 years. Paul has worked in a variety of areas: file structures, materialized views, query processing, query optimization, column stores, and main-memory databases. He collaborated closely with the SQL Server development team to drastically improve system performance by adding column store indexes and a new main-memory database engine to SQL Server. Paul is a Fellow of the ACM and Vice-President of the Executive Committee of the VLDB Endowment.

**Justin Levandoski** is a researcher in the database group at Microsoft Research. He is interested in a broad range of topics dealing with large-scale data management systems. His current interests include main-memory databases, database support for new hardware platforms, document-oriented databases, transaction processing, and cloud computing. His research has been commercialized in a number of Microsoft products, including the SQL Server Hekaton main-memory database engine, Azure DocumentDB, and Bing.

#### Tutorial 4: Machine Learning in the Real World

##### **Vineet Chaoji (Amazon India), Rajeev Rastogi (Amazon India), Gourav Roy (Amazon India)**

Machine Learning (ML) has become a mature technology that is being applied to a wide range of business problems such as web search, online advertising, product recommendations, object recognition, and so on. As a result, it has become imperative for researchers and practitioners to have a fundamental understanding of ML concepts and practical knowledge of end-to-end modeling. This tutorial takes a hands-on approach to introducing the audience to machine learning. The first part of the tutorial gives a broad overview and discusses some of the key concepts within machine learning. The second part of the tutorial takes the audience through the end-to-end modeling pipeline for a real-world income prediction problem.

The tutorial includes some hands-on exercises. If you want to follow along, you will need a laptop with at least 2 GB of RAM and Firefox/Google Chrome browser installed. Note that your laptop must be capable of connecting to internet via Wifi or your mobile data connection. We will be using docker containers, so specific software does not need to be installed on laptops.

##### **Bios:**

**Vineet Chaoji** is an Applied Science Manager within the Machine Learning team at Amazon where he leads projects related to econometric models of customer behavior, customer targeting and malware detection. Prior to joining Amazon, he was a Scientist at Yahoo! Labs in Bangalore where his research focused on online advertising and social networks. Vineet obtained a PhD in Computer Science from Rensselaer Polytechnic Institute. He has published at top-tier data mining and database conferences and journals. Vineet has also served on the program committees of leading data and web mining conferences.

**Rajeev Rastogi** is the Director of Machine Learning at Amazon where he directs the development of machine learning platforms and applications such as product classification, product recommendations, customer targeting, and deals ranking. Previously, he was the Vice President of Yahoo! Labs in Bangalore where he was responsible for research programs impacting Yahoo's web search and online advertising products. He was named a Bell Labs Fellow in 2003 for his contributions to Lucent's networking products while he was at Bell Labs Research in Murray Hill, New Jersey. Rajeev was named an ACM Fellow in 2012 for his contributions to large-scale data analysis and management. He has published over 100 papers in top-tier international conferences and 33 papers in international journals. Rajeev has also been a prolific inventor with 57 issued US Patents. He is currently a member of the News editorial board of the CACM, and was previously an Associate editor for TKDE. He has served on over 50 program committees of the leading database and data mining conferences, and was a Program Co-chair for the Applied Data Science track of the KDD conference in 2016, the CIKM conference in 2013 and the ICDM conference in 2005.

**Gourav Roy** is a Software Engineer in the Machine Learning team at Amazon where he builds scalable machine learning platforms and applications. He is interested in streaming approximate algorithms and distributed systems. His work on streaming anomaly detection recently got accepted at the International Conference on Machine Learning. Prior to joining Amazon, he got a bachelors degree in Computer Science at BIT Mesra.

#### Tutorial 5: Human Factors in Crowdsourcing

##### **Sihem Amer-Yahia (CNRS), Senjuti Basu Roy (New Jersey Institute of Technology)**

Today, crowdsourcing is used to "taskify" any job ranging from simple receipt transcription to collaborative editing, fan-subbing, citizen science, and citizen journalism. The crowd is typically volatile, its arrival and departure asynchronous, and its levels of attention and accuracy diverse. Tasks vary in complexity and may necessitate the participation of workers with varying degrees of expertise. Sometimes, workers need to collaborate explicitly and build on each other's contributions to complete a single task. For example, in disaster reporting, CrowdMap allows geographically closed people with diverse and complementary skills, to work together to report details about the course of a typhoon or the aftermath of an earthquake.

This uber-ization of human labor requires the understanding of workers' motivation in completing a task, their ability to work together in collaborative tasks, as well as, helping workers find relevant tasks. For over 40 years, organization studies have thoroughly examined human factors that affect workers in physical workplaces. More recently, computer scientists have developed algorithms that verify and leverage those findings in a virtual marketplace, in this case, a crowdsourcing platform.

The goal of this tutorial is to review those two areas and discuss how their combination may improve workers' experience, task throughput and outcome quality for both microtasks and collaborative tasks. We will start with a coverage of motivation theory, team formation, and learning worker profiles. We will then address open research questions that result from this review.

**Bios:**

**Sihem Amer-Yahia** is DR1 CNRS at LIG in Grenoble where she leads the SLIDE team. Before joining CNRS, she was Principal Scientist at the Qatar Computing Research Institute, Senior Scientist at Yahoo! Research and AT&T Labs. Sihem has served on the SIGMOD Executive Board and on the PVLDB and the EDBT Endowments. She is the Editor-in-Chief of the VLDB Journal for Europe and Africa and is on the editorial boards of TODS and the Information System Journal. She is currently serving as area chair for SIGMOD 2016 and is co-chairing the PVLDB 2016 Workshops. Sihem received her Ph.D. in Computer Science from Paris-Osary and INRIA in 1999.

**Senjuti Basu Roy** is an Assistant Professor at the Institute of Technology at the University of Washington Tacoma. Prior to joining UW in January 2012, she was a postdoctoral fellow at DIMACS at Rutgers University. Senjuti received her Ph.D. in Computer Science from UT Arlington in May 2011. Her research interests lie in the data and content management of web and structured data with a focus on exploration, analytics, and algorithms. Her research has been published in premier database and IR conferences and journals. Her past industrial experience includes working at Microsoft Research and IBM Research. She has co-chaired KDD Cup 2013, ExploreDB 2016, and serves as a guest editor of JMLR special issue on KDD Cup 2013.

**Tutorial 6: Qualitative Data Cleaning****Xu Chu (University of Waterloo), Ihab F. Ilyas (University of Waterloo)**

Data quality is one of the most important problems in data management, since dirty data often leads to inaccurate data analytics results and wrong business decisions. Data cleaning exercise often consist of two phases: error detection and error repairing. Error detection techniques can either be quantitative or qualitative; and error repairing is performed by applying data transformation scripts or by involving human experts, and sometimes both. In this tutorial, we discuss the main facets and directions in designing qualitative data cleaning techniques. We present a taxonomy of current qualitative error detection techniques, as well as a taxonomy of current data repairing techniques. We will also discuss proposals for tackling the challenges for cleaning “big data” in terms of scale and distribution.

**Bios:**

**Xu Chu** is a PhD student in the Cheriton School of Computer Science at University of Waterloo. His main research interests are data quality and data cleaning. He won the prestigious Microsoft Research PhD fellowship in 2015. Xu has also received Cheriton Fellowship from the University of Waterloo 2013-2015.

**Ihab Ilyas** is a professor in the Cheriton School of Computer Science at the University of Waterloo. He received his PhD in computer science from Purdue University, West Lafayette. His main research is in the area of database systems, with special interest in data quality, managing uncertain data, rank-aware query processing, and information extraction. Ihab is a recipient of the Ontario Early Researcher Award (2009), a Cheriton Faculty Fellowship (2013), an NSERC Discovery Accelerator Award (2014), and a Google Faculty Award (2014), and he is an ACM Distinguished Scientist. Ihab is a co-founder of Tamr, a startup focusing on large-scale data integration and cleaning. He serves on the VLDB Board of Trustees, and he is an associate editor of the ACM Transactions of Database Systems (TODS).



# Research Paper Abstracts

---

## Research 1: Transaction Processing

**Session Chair: Daniel Abadi (Yale University)**

### **Leveraging Lock Contention to Improve OLTP Application Performance**

Cong Yan (University of Washington); Alvin Cheung (University of Washington)

Locking is one of the predominant costs in transaction processing. While much work has focused on designing efficient concurrency control mechanisms, not much has been done on understanding how transaction applications issue queries and leveraging application semantics to improve application performance. This paper presents QURO, a query-aware compiler that automatically reorders queries in transaction code to improve performance. Observing that certain queries within a transaction are more contentious than others as they require locking the same tuples as other concurrently executing transactions, QURO automatically changes the application such that contentious queries are issued as late as possible. We have evaluated QURO on various transaction benchmarks, and our results show that QURO-generated implementations can increase transaction throughput by up to 6.53x, while reduce transaction latency by up to 85%.

### **BCC: Reducing False Aborts in Optimistic Concurrency Control with Low Cost for In-Memory Databases**

Yuan Yuan (The Ohio State University); Kaibo Wang (The Ohio State University); Rubao Lee (The Ohio State University); Xiaoning Ding (New Jersey Institute of Technology); Jing Xing (Institute of Computing Technology, Chinese Academy of Sciences); Spyros Blanas (The Ohio State University); Xiaodong Zhang (The Ohio State University)

The Optimistic Concurrency Control (OCC) method has been commonly used for in-memory databases to ensure transaction serializability – a transaction will be aborted if its read set has been changed during execution. This simple criterion to abort transactions causes a large proportion of false positives, leading to excessive transaction aborts. Transactions aborted false-positively (i.e. false aborts) waste system resources and can significantly degrade system throughput (as much as 3.68x based on our experiments) when data contention is intensive. Modern in-memory databases run on systems with increasingly parallel hardware and handle workloads with growing concurrency. They must efficiently deal with data contention in the presence of greater concurrency by minimizing false aborts. This paper presents a new concurrency control method named Balanced Concurrency Control (BCC) which aborts transactions more carefully than OCC does. BCC detects data dependency patterns which can more reliably indicate unserializable transactions than the criterion used in OCC. The paper studies the design options and implementation techniques that can effectively detect data contention by identifying dependency patterns with low overhead. To test the performance of BCC, we have implemented it in Silo and compared its performance against that of the vanilla Silo system with OCC and two-phase locking (2PL). Our extensive experiments with TPC-W-like, TPC-C-like and YCSB workloads demonstrate that when data contention is intensive, BCC can increase transaction throughput by more than 3x versus OCC and more than 2x versus 2PL; meanwhile, BCC has comparable performance with OCC for workloads with low data contention.

### **Multi-Version Range Concurrency Control in Deuteronomy**

Ryan Stutsman (Microsoft Research); Justin Levandoski (Microsoft Research); David Lomet (Microsoft Research); Sudipta Sengupta (Microsoft Research); Rui Wang (Microsoft Research)

The Deuteronomy transactional key value store executes millions of serializable transactions/second by exploiting multi-version timestamp order concurrency control. However, it has not supported range operations, only individual record operations (e.g., create, read, update, delete). In this paper, we enhance our multi-version timestamp order technique to handle range concurrency and prevent phantoms. Importantly, we maintain high performance while respecting the clean separation of duties required by Deuteronomy, where a transaction component performs purely logical concurrency control (including range support), while a data component performs data storage and management duties. Like the rest of the Deuteronomy stack, our range technique manages concurrency information in a latch-free manner. With our range enhancement, Deuteronomy can reach scan speeds of nearly 250 million records/s (more than 27 GB/s) on modern hardware, while providing serializable isolation complete with phantom prevention.

### **S-Store: Streaming Meets Transaction Processing**

John Meehan (Brown University); Nesime Tatbul (Intel Labs and MIT); Stan Zdonik (Brown University); Cansu Aslantas (Brown University); Ugur Cetintemel (Brown University); Jiang Du (University of Toronto); Tim Kraska (Brown University); Samuel Madden (MIT); David Maier (Portland State University); Andrew Pavlo (CMU); Michael Stonebraker (MIT); Kristin Tufte (Portland State University); Hao Wang (MIT)

Stream processing addresses the needs of real-time applications. Transaction processing addresses the coordination and safety of short atomic computations. Heretofore, these two modes of operation existed in separate, stove-piped systems. In this work, we attempt to fuse the two computational paradigms in a single system called S-Store. In this way, S-Store can simultaneously accommodate OLTP and streaming applications. We present a simple transaction model for streams that integrates seamlessly with

a traditional OLTP system, and provides both ACID and stream-oriented guarantees. We chose to build S-Store as an extension of H-Store - an open-source, in-memory, distributed OLTP database system. By implementing S-Store in this way, we can make use of the transaction processing facilities that H-Store already provides, and we can concentrate on the additional features that are needed to support streaming. Similar implementations could be done using other main-memory OLTP platforms. We show that we can actually achieve higher throughput for streaming workloads in S-Store than an equivalent deployment in H-Store alone. We also show how this can be achieved within H-Store with the addition of a modest amount of new functionality. Furthermore, we compare S-Store to two state-of-the-art streaming systems, Esper and Apache Storm, and show how S-Store can sometimes exceed their performance while at the same time providing stronger correctness guarantee.

## Research 2: RDF Data Systems

**Session Chair: Atsuyuki Morishima (University of Tsukuba)**

### **S2RDF: RDF Querying with SPARQL on Spark**

Alexander Schätzle (University of Freiburg); Martin Przyjaciel-Zablocki (University of Freiburg); Simon Skilevic (University of Freiburg); Georg Lausen (University of Freiburg)

RDF has become very popular for semantic data publishing due to its flexible and universal graph-like data model. Thus, the ever-increasing size of RDF data collections raises the need for scalable distributed approaches. We endorse the usage of existing infrastructures for Big Data processing like Hadoop for this purpose. Yet, SPARQL query performance is a major challenge as Hadoop is not intentionally designed for RDF processing. Existing approaches often favor certain query pattern shapes while performance drops significantly for other shapes. In this paper, we introduce a novel relational partitioning schema for RDF data called ExtVP that uses a semi-join based preprocessing, akin to the concept of Join Indices in relational databases, to efficiently minimize query input size regardless of its pattern shape and diameter. Our prototype system S2RDF is built on top of Spark and uses SQL to execute SPARQL queries over ExtVP. We demonstrate its superior performance in comparison to state-of-the-art SPARQL-on-Hadoop approaches.

### **Inferray: fast in-memory RDF inference**

Julien Subercaze (Laboratoire Hubert Curien); Christophe Gravier (Laboratoire Hubert Curien); Jules Chevalier (Laboratoire Hubert Curien); Frederique Laforest (Laboratoire Hubert Curien)

The advent of semantic data on the Web requires efficient reasoning systems to infer RDF and OWL data. The linked nature and the huge volume of data entail efficiency and scalability challenges when designing productive inference systems. This paper presents Inferray, an implementation of RDFS, rdf, and RDFS-Plus inference with improved performance over existing solutions. The main features of Inferray are 1) a storage layout based on vertical partitioning that guarantees sequential access and efficient sort-merge join inference; 2) efficient sorting of pairs of 64-bit integers using ad-hoc optimizations on MSD radix and a custom counting sort; 3) a dedicated temporary storage to perform efficient graph closure computation. Our measurements on synthetic and real-world datasets show improvements over competitors on RDFS-Plus, and up to several orders of magnitude for transitivity closure.

### **RDF Graph Alignment with Bisimulation**

Peter Buneman (University of Edinburgh); Slawek Staworko (University of Edinburgh, University of Lille, and LINKS, INRIA Nord-Europe)

We investigate approaches to the alignment of two RDF triple stores inspired by the classical notion of graph bisimulation and relate this to metrics that describe the accuracy of the alignment. Alignment of RDF stores is essential in the understanding of evolving, curated ontologies. One needs to align nodes in successive versions that are given “blank” names; moreover, a new version may switch terminology from one external ontology to another. This means that some further alignment of the identifiers of the two external ontologies is useful. We first describe a form of bisimulation based on node colorings, in which the colors define equivalence classes, we then go on to describe an estimate of the accuracy of a coloring based on (a) the similarity of literal nodes and (b) on the “structural” similarity of internal nodes (URIs and blank nodes). The effectiveness of these methods is tested on two evolving data sets: an ontology described in OWL and triple store derived from an evolving relational database. Both of these are curated resources for biologists.

### **Semantic SPARQL Similarity Search Over RDF Knowledge Graphs**

Weiguo Zheng (Peking University); Lei Zou (Peking University); Xifeng Yan (University of California at Santa Barbara); Shaoux Song (Tsinghua University); Dongyan Zhao (Peking University)

RDF knowledge graphs have attracted increasing attentions these years. However, due to the schema-free nature of RDF data, it is very difficult for users to have full knowledge of the underlying schema. Furthermore, the same kind of information can be represented in diverse graph fragments. Hence, it is a huge challenge to formulate complex SPARQL expressions by taking the union of all the possible structures. In this paper, we propose an effective framework to access the RDF repository even if users have no full knowledge of the underlying schema. Specifically, given a SPARQL query, the system could return as more answers that match the query based on the semantic similarity as possible. Interestingly, we propose a systematic method to mine diverse semantically equivalent structure patterns. More importantly, incorporating both structural and semantic similarities we are the first to propose a novel similarity

measure, semantic graph edit distance. In order to improve the efficiency performance, we utilize the semantic summary graph to summarize the knowledge graph, which supports both high-level pruning and drill-down pruning. We also devise an effective lower bound based on the TA-style access to each of the candidate sets. Extensive experiments over real datasets confirm the effectiveness and efficiency of our approach.

## Research 3: Distributed and Cloud Systems - 1

### Session Chair: Ashraf Aboulnaga (Qatar Computing Research Institute)

#### Clash of the Titans: MapReduce vs. Spark for Large Scale Data Analytics

Juwei Shi (Renmin University of China); Yunjie Qiu (IBM Research - China); Umar Farooq Minhas (IBM Research - Almaden); Limei Jiao (IBM Research - China); Chen Wang (Tsinghua University); Berthold Reinwald (IBM Research - Almaden); Fatma Ozcan (IBM Research - Almaden)

MapReduce and Spark are two very popular open source clustercomputing frameworks for large scale data analytics. These frameworks hide the complexity of task parallelism and fault-tolerance, by exposing a simple programming API to users. In this paper, we evaluate the major architectural components in MapReduce and Spark frameworks including: shuffle, execution model, and caching, by using a set of important analytic workloads. To conduct a detailed analysis, we developed two profiling tools: (1) We correlate the task execution plan with the resource utilization for both MapReduce and Spark, and visually present this correlation; (2) We provide a breakdown of the task execution time for in-depth analysis. Through detailed experiments, we quantify the performance differences between MapReduce and Spark. Furthermore, we attribute these performance differences to different components which are architected differently in the two frameworks. We further expose the source of these performance differences by using a set of micro-benchmark experiments. Overall, our experiments show that Spark is about 2.5x, 5x, and 5x faster than MapReduce, for Word Count, k-means, and PageRank, respectively. The main causes of these speedups are the efficiency of the hash-based aggregation component for combine, as well as reduced CPU and disk overheads due to RDD caching in Spark. An exception to this is the Sort workload, for which MapReduce is 2x faster than Spark. We show that MapReduce's execution model is more efficient for shuffling data than Spark, thus making Sort run faster on MapReduce.

#### Compressed Linear Algebra for Large-Scale Machine Learning

Matthias Boehm (IBM Research - Almaden); Ahmed Elgohary (University of Maryland); Peter Haas (IBM Research - Almaden); Fred Reiss (IBM Research - Almaden); Berthold Reinwald (IBM Research - Almaden)

Large-scale machine learning (ML) algorithms are often iterative with repeated read-only data access and I/O-bound matrix-vector multiplications to converge to an optimal model. It is crucial for performance to fit the data into single-node or distributed main memory. General-purpose, heavy- and lightweight compression techniques struggle to achieve both good compression ratios and fast decompression speed to enable blockwise uncompressed operations. Hence, we initiate work on compressed linear algebra (CLA), in which lightweight database compression techniques are applied to matrices and then linear algebra operations such as matrix-vector multiplication are executed directly on the compressed representations. We contribute effective column compression schemes, cache-conscious operations, and an efficient sampling-based compression algorithm. Our experiments show that CLA achieves in-memory operation performance close to the uncompressed case and good compression ratios that allow us to fit larger datasets into available memory. We thereby obtain significant end-to-end performance improvements up to 26x or reduced memory requirements.

#### PIXIDA: Optimizing Data Parallel Jobs in Wide-Area Data Analytics

Konstantinos Kloudas (INESC-ID); Rodrigo Rodrigues (INESC-ID/IST at University of Lisbon); Nuno Preguica (NOVA LINCS / NOVA at University of Lisbon); Margarida Mamede (NOVA LINC / NOVA at University of Lisbon)

In the era of global-scale services, big data analytical queries are often required to process datasets that span multiple data centers (DCs). In this setting, cross-DC bandwidth is often the scarcest, most volatile, and/or most expensive resource. However, current widely deployed big data analytics frameworks make no attempt to minimize the traffic traversing these links. In this paper, we present PIXIDA, a scheduler that aims to minimize data movement across resource constrained links. To achieve this, we introduce a new abstraction called SILO, which is key to modeling PIXIDA's scheduling goals as a graph partitioning problem. Furthermore, we show that existing graph partitioning problem formulations do not map to how big data jobs work, causing their solutions to miss opportunities for avoiding data movement. To address this, we formulate a new graph partitioning problem and propose a novel algorithm to solve it. We integrated PIXIDA in Spark and our experiments show that, when compared to existing schedulers, PIXIDA achieves a significant traffic reduction of up to ~4x on the aforementioned links.

#### Cümülon: Matrix-Based Data Analytics in the Cloud with Spot Instances

Botong Huang (Duke University); Nicholas Jarrett (Duke University); Shivnath Babu (Duke University); Sayan Mukherjee (Duke University); Jun Yang (Duke University)

We describe Cumulon, a system aimed at helping users develop and deploy matrix-based data analysis programs in a public cloud. A key feature of Cumulon is its end-to-end support for the so-called spot instances—machines whose market price fluctuates over time but is usually much lower than the regular fixed price. A user sets a bid price when acquiring spot instances, and loses them as soon

as the market price exceeds the bid price. While spot instances can potentially save cost, they are difficult to use effectively, and run the risk of not finishing work while costing more. Cumulon provides a highly elastic computation and storage engine on top of spot instances, and offers automatic cost-based optimization of execution, deployment, and bidding strategies. Cumulon further quantifies how the uncertainty in the market price translates into the cost uncertainty of its recommendations, and allows users to specify their risk tolerance as an optimization constraint.

## Research 4: Memory Management

### Session Chair: Alfons Kemper (Technische Universität München)

#### RUMA has it: Rewired User-space Memory Access is Possible!

Felix Martin Schuhknecht (Saarland University); Jens Dittrich (Saarland University); Ankur Sharma (Saarland University)

Memory management is one of the most boring topics in database research. It plays a minor role in tasks like free-space management or efficient space usage. Here and there we also realize its impact on database performance when worrying about NUMA-aware memory allocation, data compacting, snapshotting, and defragmentation. But, overall, let's face it: the entire topic sounds as exciting as 'garbage collection' or 'debugging a program for memory leaks'. What if there were a technique that would promote memory management from a third class helper thingie to a first class citizen in algorithm and systems design? What if that technique turned the role of memory management in a database system (and any other data processing system) upside-down? What if that technique could be identified as a key for re-designing various core algorithms with the effect of outperforming existing state-of-the-art methods considerably? Then we would write this paper. We introduce RUMA: Rewired User-space Memory Access. It allows for physiological data management, i.e. we allow developers to freely rewire the mappings from virtual to physical memory (in user space) while at the same time exploiting the virtual memory support offered by hardware and operating system. We show that fundamental database building blocks such as array operations, partitioning, sorting, and snapshotting benefit strongly from RUMA.

#### Asynchronous Memory Access Chaining

Oris Kocberber (EPFL); Babak Falsafi (EPFL); Boris Grot (University of Edinburgh)

In-memory databases rely on pointer-intensive data structures to quickly locate data in memory. A single lookup operation in such data structures often exhibits long-latency memory stalls due to dependent pointer dereferences. Hiding the memory latency by launching additional memory accesses for other lookups is an effective way of improving performance of pointer-chasing codes (e.g., hash table probes, tree traversals). The ability to exploit such inter-lookup parallelism is beyond the reach of modern out-of-order cores due to the limited size of their instruction window. Instead, recent work has proposed software prefetching techniques that exploit inter-lookup parallelism by arranging a set of independent lookups into a group or a pipeline, and navigate their respective pointer chains in a synchronized fashion. While these techniques work well for highly regular access patterns, they break down in the face of irregularity across lookups. Such irregularity includes variable-length pointer chains, early exit, and read/write dependencies. This work introduces Asynchronous Memory Access Chaining (AMAC), a new approach for exploiting inter-lookup parallelism to hide the memory access latency. AMAC achieves high dynamism in dealing with irregularity across lookups by maintaining the state of each lookup separately from that of other lookups. This feature enables AMAC to initiate a new lookup as soon as any of the in-flight lookups complete. In contrast, the static arrangement of lookups into a group or pipeline in existing techniques precludes such adaptivity. Our results show that AMAC matches or outperforms state-of-the-art prefetching techniques on regular access patterns, while delivering up to 2.3x higher performance under irregular data structure lookups. AMAC fully utilizes the available micro-architectural resources, generating the maximum number of memory accesses allowed by hardware in both single- and multi-threaded execution modes.

#### WarpLDA: a Cache Efficient O(1) Algorithm for Latent Dirichlet Allocation

Jianfei Chen (Tsinghua University); Kaiwei Li (Tsinghua University); Jun Zhu (Tsinghua University); Wenguang Chen (Tsinghua University)

Developing efficient and scalable algorithms for Latent Dirichlet Allocation (LDA) is of wide interest for many applications. Previous work has developed an O(1) Metropolis-Hastings (MH) sampling method for each token. However, its performance is far from being optimal due to frequent cache misses caused by random accesses to the parameter matrices. In this paper, we first carefully analyze the memory access behavior of existing algorithms for LDA by cache locality at document level. We then develop WarpLDA, which achieves O(1) time complexity per-token and fits the randomly accessed memory per-document in the L3 cache. Our empirical results in a wide range of testing conditions demonstrate that WarpLDA is consistently 5-15x faster than the state-of-the-art MH-based LightLDA, and is faster than the state-of-the-art sparsity aware F+LDA in most settings. Our WarpLDA learns a million topics from 639 millions of documents in only five hours at an unprecedented throughput of 11 billion tokens per second.

#### Lifetime-Based Memory Management for Distributed Data Processing Systems

Lu Lu (Huazhong University of Science and Technology); Xuanhua Shi (Huazhong University of Science and Technology); Yongluan Zhou (University of Southern Denmark); Xiong Zhang (Huazhong University of Science and Technology); Hai Jin (Huazhong University of Science and Technology); Cheng Pei (Huazhong University of Science and Technology); Ligang He (University of Warwick); Yuanzhen Geng (Huazhong University of Science and Technology)

In-memory caching of intermediate data and eager combining of data in shuffle buffers have been shown to be very effective in minimizing the re-computation and I/O cost in distributed data processing systems like Spark and Flink. However, it has also been widely reported that these techniques would create a large amount of long-living data objects in the heap, which may quickly saturate the garbage collector, especially when handling a large dataset, and hence would limit the scalability of the system. To eliminate this problem, we propose a lifetime-based memory management framework, which, by automatically analyzing the user-defined functions and data types, obtains the expected lifetime of the data objects, and then allocates and releases memory space accordingly to minimize the garbage collection overhead. In particular, we present Deca, a concrete implementation of our proposal on top of Spark, which transparently decomposes and groups objects with similar lifetimes into byte arrays and releases their space altogether when their lifetimes come to an end. An extensive experimental study using both synthetic and real datasets shows that, in comparing to Spark, Deca is able to 1) reduce the garbage collection time by up to 99.9%, 2) to achieve up to 22.7x speed up in terms of execution time in cases without data spilling and 41.6x speedup in cases with data spilling, and 3) to consume up to 46.6% less memory.

## Research 5: Data Cleaning - 1

### Session Chair: Ihab Ilyas (University of Waterloo)

#### Cleaning Timestamps with Temporal Constraints

Shaouxu Song (Tsinghua University); Yue Cao (Tsinghua University); Jianmin Wang (Tsinghua University)

Timestamps are often found to be dirty in various scenarios, e.g., in distributed systems with clock synchronization problems or unreliable RFID readers. Without cleaning the imprecise timestamps, temporal-related applications such as provenance analysis or pattern queries are not reliable. To evaluate the correctness of timestamps, temporal constraints could be employed, which declare the distance restrictions between timestamps. Guided by such constraints on timestamps, in this paper, we study a novel problem of repairing inconsistent timestamps that do not conform to the required temporal constraints. Following the same line of data repairing, the timestamp repairing problem is to minimally modify the timestamps towards satisfaction of temporal constraints. This problem is practically challenging, given the huge space of possible timestamps. We tackle the problem by identifying a concise set of promising candidates, where an optimal repair solution can always be found. Repair algorithms with efficient pruning are then devised over the identified candidates. Experiments on real datasets demonstrate the superiority of our proposal compared to the state-of-the-art approaches.

#### Repairing Data through Regular Expressions

Zeyu Li (Harbin Institute of Technology); Hongzhi Wang (Harbin Institute of Technology); Wei Shao (Harbin Institute of Technology); Jianzhong Li (Harbin Institute of Technology); Hong Gao (Harbin Institute of Technology)

Since regular expressions are often used to detect errors in sequences such as strings or date, it is natural to use it for data repair. Motivated by this, we propose a data repair method based on regular expression to make the input sequence data obey the given regular expression with minimal revisions. The proposed method contains two steps, sequence repair and token value repair. For sequence repair, we propose the Regular-expression-based Structural Repair algorithm (RSR in short). RSR algorithm is a dynamic programming algorithm that utilizes Nondeterministic Finite Automata (NFA) to calculate the edit distance between the prefix of the input string and partial pattern regular expression with time complexity  $O(mn^2)$  and space complexity  $O(mn)$  where  $m$  is the number of NFA's edge and  $n$  is the input string length. We also give an optimization to achieve high performance. For token value repair, we combine the edit-distance-based method and associate rules by a unified argument for the selection of the two methods. Experimental results on both real and synthetic data show that the proposed method could repair the data effectively and efficiently.

#### Combining Quantitative and Logical Data Cleaning

Nataliya Prokoshyna (University of Toronto); Jaroslaw Szlichta (University of Ontario Institute of Technology); Fei Chiang (McMaster University); Renée Miller (University of Toronto); Divesh Srivastava (AT&T Labs Research)

Quantitative data cleaning relies on the use of statistical methods to identify and repair data quality problems while logical data cleaning tackles the same problems using various forms of logical reasoning over declarative dependencies. Each of these approaches has its strengths: the logical approach is able to capture subtle data quality problems using sophisticated dependencies, while the quantitative approach excels at ensuring that the repaired data has desired statistical properties. We propose a novel framework within which these two approaches can be used synergistically to combine their respective strengths. We instantiate our framework using (i) metric functional dependencies, a type of dependency that generalizes functional dependencies (FDs) to identify inconsistencies in domains where only large differences in metric data are considered to be a data quality problem, and (ii) repairs that modify the inconsistent data so as to minimize statistical distortion, measured using the Earth Mover's Distance. We show that the problem of computing a statistical distortion minimal repair is NP-hard. Given this complexity, we present an efficient algorithm for finding a minimal repair that has a small statistical distortion using EMD computation over semantically related attributes. To identify semantically related attributes, we present a sound and complete axiomatization and an efficient algorithm for testing implication of metric FDs. While the complexity of inference for some other FD extensions is co-NP complete, we show that the inference problem for metric FDs remains linear, as in traditional FDs. We prove that every instance that can be generated by our repair algorithm is set-minimal (with no unnecessary changes).

Our experimental evaluation demonstrates that our techniques obtain a considerably lower statistical distortion than existing repair techniques, while achieving similar levels of efficiency.

### **Temporal Rules Discovery for Web Data Cleaning**

Ziawasch Abedjan (MIT); Cuneys Akcora (Qatar Computing Research Institute); Mourad Ouzzani (Qatar Computing Research Institute); Paolo Papotti (Qatar Computing Research Institute); Michael Stonebraker (MIT)

Declarative rules, such as functional dependencies, are widely used for cleaning data. Several systems take them as input for detecting errors and computing a “clean” version of the data. To support domain experts, in specifying these rules, several tools have been proposed to profile the data and mine rules. However, existing discovery techniques have traditionally ignored the time dimension. Recurrent events, such as persons reported in locations, have a duration in which they are valid, and this duration should be part of the rules or the cleaning process would simply fail. In this work, we study the rule discovery problem for temporal web data. Such a discovery process is challenging because of the nature of web data; extracted facts are (i) sparse over time, (ii) reported with delays, and (iii) often reported with errors over the values because of inaccurate sources or non robust extractors. We handle these challenges with a new discovery approach that is more robust to noise. Our solution uses machine learning methods, such as association measures and outlier detection, for the discovery of the rules, together with an aggressive repair of the data in the mining step itself. Our experimental evaluation over real-world data from Recorded Future, an intelligence company that monitors over 700K Web sources, shows that temporal rules improve the quality of the data with an increase of the average precision in the cleaning process from 0.37 to 0.84, and a 40% relative increase in the average F-measure.

## **Research 6: Graph Processing - 1**

### **Session Chair: Bin Cui (Peking University)**

#### **I/O Efficient ECC Graph Decomposition via Graph Reduction**

Long Yuan (The University of New South Wales); Lu Qin (University of Technology, Sydney); Xuemin Lin (The University of New South Wales); Lijun Chang (The University of New South Wales); Wenjie Zhang (The University of New South Wales)

The problem of computing  $k$ -edge connected components ( $k$ -ECCs) of a graph  $G$  for a specific  $k$  is a fundamental graph problem and has been investigated recently. In this paper, we study the problem of ECC decomposition, which computes the  $k$ -ECCs of a graph  $G$  for all  $k$  values. ECC decomposition can be widely applied in a variety of applications such as graph-topology analysis, community detection, Steiner component search, and graph visualization. A straightforward solution for ECC decomposition is to apply the existing  $k$ -ECC computation algorithm to compute the  $k$ -ECCs for all  $k$  values. However, this solution is not applicable to large graphs for two challenging reasons. First, all existing  $k$ -ECC computation algorithms are highly memory intensive due to the complex data structures used in the algorithms. Second, the number of possible  $k$  values can be very large, resulting in a high computational cost when each  $k$  value is independently considered. In this paper, we address the above challenges, and study I/O efficient ECC decomposition via graph reduction. We introduce two elegant graph reduction operators which aim to reduce the size of the graph loaded in memory while preserving the connectivity information of a certain set of edges to be computed for a specific  $k$ . We also propose three novel I/O efficient algorithms, Bottom-Up, Top-Down, and Hybrid, that explore the  $k$  values in different orders to reduce the redundant computations between different  $k$  values. We analyze the I/O and memory costs for all proposed algorithms. In our experiments, we evaluate our algorithms using seven real large datasets with various graph properties, one of which contains 1.95 billion edges. The experimental results show that our proposed algorithms are scalable and efficient.

#### **LEOPARD: Lightweight Edge-Oriented Partitioning and Replication for Dynamic Graphs**

Jiewen Huang (Yale University); Daniel Abadi (Yale University)

This paper introduces a dynamic graph partitioning algorithm, designed for large, constantly changing graphs. We propose a partitioning framework that adjusts on the fly as the graph structure changes. We also introduce a replication algorithm that is tightly integrated with the partitioning algorithm, which further reduces the number of edges cut by the partitioning algorithm. Even though the proposed approach is handicapped by only taking into consideration local parts of the graph when reassigning vertices, extensive evaluation shows that the proposed approach maintains a quality partitioning over time, which is comparable at any point in time to performing a full partitioning from scratch using a state-the-art static graph partitioning algorithm such as METIS. Furthermore, when vertex replication is turned on, edge-cut can improve by an order of magnitude.

#### **Towards Maximum Independent Sets on Massive Graphs**

Yu Liu (Renmin University of China); Jiaheng Lu (University of Helsinki); Hua Yang (Renmin University of China); Xiaokui Xiao (Nanyang Technological University); Zhewei Wei (Renmin University of China)

Maximum independent set (MIS) is a fundamental problem in graph theory and it has important applications in many areas such as social network analysis, graphical information systems and coding theory. The problem is NP-hard, and there has been numerous studies on its approximate solutions. While successful to a certain degree, the existing methods require memory space at least linear in the size of the input graph. This has become a serious concern in view of the massive volume of today’s fast-growing graphs. In this paper, we study the MIS problem under the semi-external setting, which assumes that the main memory can accommodate all vertices of the graph but not all edges. We present a greedy algorithm and a general vertex-swap framework, which swaps vertices

to incrementally increase the size of independent sets. Our solutions require only few sequential scans of graphs on the diskfile, thus enabling in-memory computation without costly random disk accesses. Experiments on large-scale datasets show that our solutions are able to compute a large independent set for a massive graph with 59 million vertices and 151 million edges using a commodity machine, with a memory cost of 469MB and a time cost of three minutes, while yielding an approximation ratio that is around 99% of the theoretical optimum.

### **Weaver: A High-Performance, Transactional Graph Database Based on Refinable Timestamps**

Ayush Dubey (Cornell University); Greg Hill (Stanford University); Robert Escriva (Cornell University); Emin Sifer (Cornell University)

Graph databases have become an increasingly common infrastructure component. Yet existing systems either operate on offline snapshots, provide weak consistency guarantees, or use expensive concurrency control techniques that limit performance. In this paper, we introduce a new distributed graph database, called Weaver, which enables efficient, transactional graph analyses as well as strictly serializable ACID transactions on dynamic graphs. The key insight that allows Weaver to combine strict serializability with horizontal scalability and high performance is a novel request ordering mechanism called refinable timestamps. This technique couples coarse-grained vector timestamps with a fine-grained timeline oracle to pay the overhead of strong consistency only when needed. Experiments show that Weaver enables a Bitcoin blockchain explorer that is 8x faster than Blockchain.info, and achieves 12x higher throughput than the Titan graph database on social network workloads and 4x lower latency than GraphLab on offline graph traversal workloads.

## **Research 7: Query Execution - 1**

### **Session Chair: Anastasia Ailamaki (EPFL)**

#### **Parallel Evaluation of Multi-Semi-Joins**

Jonny Daenen (Hasselt University); Frank Neven (Hasselt University); Tony Tan (National Taiwan University); Stijn Vansummeren (Université Libre de Bruxelles)

While services such as Amazon AWS make computing power abundantly available, adding more computing nodes can incur high costs in, for instance, pay-as-you-go plans while not always significantly improving the net running time (aka wall-clock time) of queries. In this work, we provide algorithms for parallel evaluation of SGF queries in MapReduce that optimize total time, while retaining low net time. Not only can SGF queries specify all semi-join reducers, but also more expressive queries involving disjunction and negation. Since SGF queries can be seen as Boolean combinations of (potentially nested) semi-joins, we introduce a novel multi-semi-join (MSJ) MapReduce operator that enables the evaluation of a set of semi-joins in one job. We use this operator to obtain parallel query plans for SGF queries that outvalue sequential plans w.r.t. net time and provide additional optimizations aimed at minimizing total time without severely affecting net time. Even though the latter optimizations are NP-hard, we present effective greedy algorithms. Our experiments, conducted using our own implementation Gumbo on top of Hadoop, confirm the usefulness of parallel query plans, and the effectiveness and scalability of our optimizations, all with a significant improvement over Pig and Hive.

#### **Lightning Fast and Space Efficient Inequality Joins**

Zuhair Khayyat (King Abdullah University of Science and Technology); William Lucia (Qatar Computing Research Institute); Meghna Singh (Qatar Computing Research Institute); Mourad Ouzzani (Qatar Computing Research Institute); Paolo Papotti (Qatar Computing Research Institute); Jorge-Arnulfo Quijane-Ruiz (Qatar Computing Research Institute); Nan Tang (Qatar Computing Research Institute); Panos Kalnis (King Abdullah University of Science and Technology)

Inequality joins, which join relational tables on inequality conditions, are used in various applications. While there have been a wide range of optimization methods for joins in database systems, from algorithms such as sort-merge join and band join, to various indices such as B+-tree, R-tree and Bitmap, inequality joins have received little attention and queries containing such joins are usually very slow. In this paper, we introduce fast inequality join algorithms. We put columns to be joined in sorted arrays and we use permutation arrays to encode positions of tuples in one sorted array w.r.t. the other sorted array. In contrast to sort-merge join, we use space efficient bit-arrays that enable optimizations, such as Bloom filter indices, for fast computation of the join results. We have implemented a centralized version of these algorithms on top of PostgreSQL, and a distributed version on top of Spark SQL. We have compared against well known optimization techniques for inequality joins and show that our solution is more scalable and several orders of magnitude faster.

#### **MQJoin: Efficient Shared Execution of Main-Memory Joins**

Darko Makreshanski (ETH Zurich); Georgios Giannikis (Oracle Labs); Gustavo Alonso (ETH Zurich); Donald Kossman (Microsoft Research)

Database architectures typically process queries one-at-a-time, executing concurrent queries in independent execution contexts. Often, such a design leads to unpredictable performance and poor scalability. One approach to circumvent the problem is to take advantage of sharing opportunities across concurrently running queries. In this paper we propose Many-Query Join (MQJoin), a novel method for sharing the execution of a join that can efficiently deal with hundreds of concurrent queries. This is achieved by

minimizing redundant work and making efficient use of main-memory bandwidth and multi-core architectures. Compared to existing proposals, MQJoin is able to efficiently handle larger workloads regardless of the schema by exploiting more sharing opportunities. We also compared MQJoin to two commercial main-memory column-store databases. For a TPC-H based workload, we show that MQJoin provides 2-5x higher throughput with significantly more stable response times.

### **Incremental Computation of Common Windowed Holistic Aggregates**

Richard Wesley (Tableau Software); Fei Xu (Tableau Software)

Windowed aggregates are a SQL 2003 feature for computing aggregates in moving windows. Common examples include cumulative sums, local maxima and moving quantiles. With the advent over the last few years of easy-to-use data analytics tools, these functions are becoming widely used by more and more analysts, but some aggregates (such as local maxima) are much easier to compute than others (such as moving quantiles). Nevertheless, aggregates that are more difficult to compute, like quantile and mode (or “most frequent”) provide more appropriate statistical summaries in the common situation when a distribution is not Gaussian and are an essential part of a data analysis toolkit. Recent work has described highly efficient windowed implementations of the most common aggregate function categories, including distributive aggregates such as cumulative sums and algebraic aggregates such as moving averages. But little has been published on either the implementation or the performance of the more complex holistic windowed aggregates such as moving quantiles. This paper provides the first in-depth study of how to efficiently implement the three most common holistic windowed aggregates (count distinct, mode and quantile) by reusing the aggregate state between consecutive frames. Our measurements show that these incremental algorithms generally achieve improvements of about 10+ over naive implementations, and that they can effectively detect when to reset the internal state during extreme frame variation.

## **Research 8: Data Security and Privacy**

### **Session Chair: Lei Chen (Hong Kong University of Science and Technology)**

#### **Oblivious RAM: A Dissection and Experimental Evaluation**

Zhao Chang (University of Utah); Dong Xie (University of Utah); Feifei Li (University of Utah)

Many companies choose the cloud as their data and IT infrastructure platform. The remote access of the data brings the issue of trust. Despite the use of strong encryption schemes, adversaries can still learn valuable information regarding encrypted data by observing the data access patterns. To that end, one can hide the access patterns, which may leak sensitive information, using Oblivious RAMs (ORAMs). Numerous works have proposed different ORAM constructions, but they have never been thoroughly compared against and tested on large databases. There are also no open source implementation of these schemes. These limitations make it difficult for researchers and practitioners to choose and adopt a suitable ORAM for their applications. To address this issue, we provide a thorough study over several practical ORAM constructions, and implement them under the same library. We perform extensive experiments to provide insights into their performance characteristics with respect to efficiency, scalability, and communication cost.

#### **Design of Policy-Aware Differentially Private Algorithms**

Samuel Haney (Duke University); Ashwin Machanavajjhala (Duke University); Bolin Ding (Microsoft Research)

The problem of designing error optimal differentially private algorithms is well studied. Recent work applying differential privacy to real world settings have used variants of differential privacy that appropriately modify the notion of neighboring databases. The problem of designing error optimal algorithms for such variants of differential privacy is open. In this paper, we show a novel transformational equivalence result that can turn the problem of query answering under differential privacy with a modified notion of neighbors to one of query answering under standard differential privacy, for a large class of neighbor definitions. We utilize the Blowfish privacy framework that generalizes differential privacy. Blowfish uses a policy graph to instantiate different notions of neighboring databases. We show that the error incurred when answering a workload  $W$  on a database  $x$  under a Blowfish policy graph  $G$  is identical to the error required to answer a transformed workload  $f_G(W)$  on database  $g_G(x)$  under standard differential privacy, where  $f_G$  and  $g_G$  are linear transformations based on  $G$ . Using this result, we develop error efficient algorithms for releasing histograms and multidimensional range queries under different Blowfish policies. We believe the tools we develop will be useful for finding mechanisms to answer many other classes of queries with low error under other policy graphs.

#### **Behavior Query Discovery in System-Generated Temporal Graphs**

Bo Zong (University of California at Santa Barbara); Xusheng Xiao (NEC Labs America, Inc.); Zhichun Li (NEC Labs America, Inc.); Zhenyu Wu (NEC Labs America, Inc.); Zhiyun Qian (University of California at Riverside); Xifeng Yan (University of California at Santa Barbara); Ambuj Singh (University of California at Santa Barbara); Guofei Jiang (NEC Labs America, Inc.)

Computer system monitoring generates huge amounts of logs that record the interaction of system entities. How to query such data to better understand system behaviors and identify potential system risks and malicious behaviors becomes a challenging task for system administrators due to the dynamics and heterogeneity of the data. System monitoring data are essentially heterogeneous temporal graphs with nodes being system entities and edges being their interactions over time. Given the complexity of such graphs, it becomes time-consuming for system administrators to manually formulate useful queries in order to examine abnormal



activities, attacks, and vulnerabilities in computer systems. In this work, we investigate how to query temporal graphs and treat query formulation as a discriminative temporal graph pattern mining problem. We introduce TGMIner to mine discriminative patterns from system logs, and these patterns can be taken as templates for building more complex queries. TGMIner leverages temporal information in graphs to prune graph patterns that share similar growth trend without compromising pattern quality. Experimental results on real system data show that TGMIner is 6-32 times faster than baseline methods. The discovered patterns were verified by system experts; they achieved high precision (97%) and recall (91%).

## Research 9: Ranking Queries

**Session Chair: Renée Miller (University of Toronto)**

### Query Reranking As A Service

Abolfazl Asudeh (UT Arlington); Nan Zhang (George Washington University); Gautam Das (UT Arlington)

The ranked retrieval model has rapidly become the de facto way for search query processing in client-server databases, especially those on the web. Despite of the extensive efforts in the database community on designing better ranking functions/mechanisms, many such databases in practice still fail to address the diverse and sometimes contradicting preferences of users on tuple ranking, perhaps (at least partially) due to the lack of expertise and/or motivation for the database owner to design truly effective ranking functions. This paper takes a different route on addressing the issue by defining a novel (query reranking problem), i.e., we aim to design a third-party service that uses nothing but the public search interface of a client-server database to enable the on-the-fly processing of queries with any user-specified ranking functions (with or without selection conditions), no matter if the ranking function is supported by the database or not. We analyze the worst-case complexity of the problem and introduce a number of ideas, e.g., on-the-fly indexing, domination detection and virtual tuple pruning, to reduce the average-case cost of the query reranking algorithm. We also present extensive experimental results on real-world datasets, in both offline and live online systems, that demonstrate the effectiveness of our proposed techniques.

### Finding Pareto Optimal Groups: Group-based Skyline

Jinfei Liu (Emory University); Li Xiong (Emory University); Jian Pei (Simon Fraser University); Jun Luo (Lenovo); Haoyu Zhang (Emory University)

Skyline computation, aiming at identifying a set of skyline points that are not dominated by any other point, is particularly useful for multi-criteria data analysis and decision making. Traditional skyline computation, however, is inadequate to answer queries that need to analyze not only individual points but also groups of points. To address this gap, we generalize the original skyline definition to the novel group-based skyline (G-Skyline), which represents Pareto optimal groups that are not dominated by other groups. In order to compute G-Skyline groups consisting of  $k$  points efficiently, we present a novel structure that represents the points in a directed skyline graph and captures the dominance relationships among the points based on the first  $k$  skyline layers. We propose efficient algorithms to compute the first  $k$  skyline layers. We then present two heuristic algorithms to efficiently compute the G-Skyline groups: the point-wise algorithm and the unit group-wise algorithm, using various pruning strategies. The experimental results on the real NBA dataset and the synthetic datasets show that G-Skyline is interesting and useful, and our algorithms are efficient and scalable.

### k-Regret Queries with Nonlinear Utilities

Taylor Kessler Faulkner (Denison University); Will Brackenkbury (Denison University); Ashwin Lal (Denison University)

In exploring representative databases, a primary issue has been finding accurate models of user preferences. Given this, our work generalizes the method of regret minimization as proposed by Nanongka et al. to include nonlinear utility functions. Regret minimization is an approach for selecting  $k$  representative points from a database such that every user's ideal point in the entire database is similar to one of the  $k$  points. This approach combines benefits of the methods top- $k$  and skyline; it controls the size of the output but does not require knowledge of users' preferences. Prior work with  $k$ -regret queries assumes users' preferences to be modeled by linear utility functions. In this paper, we derive upper and lower bounds for non-linear utility functions, as these functions can better fit occurrences such as diminishing marginal returns, propensity for risk, and substitutability of preferences. To model these phenomena, we analyze a broad subset of convex, concave, and constant elasticity of substitution functions. We also run simulations on real and synthetic data to prove the efficacy of our bounds in practice.

### Discovering the Skyline of Web Databases

Abolfazl Asudeh (UT Arlington); Saravanan Thirumuruganathan (UT Arlington); Nan Zhang (George Washington University); Gautam Das (UT Arlington)

Many web databases are "hidden" behind proprietary search interfaces that enforce the top- $k$  output constraint, i.e., each query returns at most  $k$  of all matching tuples, preferentially selected and returned according to a proprietary ranking function. In this paper, we initiate research into the novel problem of skyline discovery over top- $k$  hidden web databases. Since skyline tuples provide critical insights into the database and include the top-ranked tuple for every possible ranking function following the monotonic order of attribute values, skyline discovery from a hidden web database can enable a wide variety of innovative third-party applications over one or multiple web databases. Our research in the paper shows that the critical factor affecting the cost of skyline discovery is the

type of search interface controls provided by the website. As such, we develop efficient algorithms for three most popular types, i.e., one-ended range, free range and point predicates, and then combine them to support web databases that feature a mixture of these types. Rigorous theoretical analysis and extensive real-world online and offline experiments demonstrate the effectiveness of our proposed techniques and their superiority over baseline solutions.

## Research 10: Query Optimization - 1

### Session Chair: S Sudarshan (IIT Bombay)

#### Faster Plan Generation through Consideration of Functional Dependencies and Keys

Marius Eich (University of Mannheim); Pit Fender (Oracle Labs); Guido Moerkotte (University of Mannheim)

It has been a recognized fact for many years that query execution can benefit from pushing group-by operators down in the operator tree and applying them before a join. This so-called eager aggregation reduces the size(s) of the join argument(s), making join evaluation faster. Lately, the idea enjoyed a revival when it was applied to outer joins for the first time and incorporated in a state-of-the-art plan generator. However, this recent approach is highly dependent on the use of heuristics because of the exponential growth of the search space that goes along with eager aggregation. Finding an optimal solution for larger queries calls for effective optimality preserving pruning mechanisms to reduce the search space size as far as possible. By a more thorough investigation of functional dependencies and keys, we provide a set of new pruning criteria and evaluate their effectiveness with respect to the runtime and memory consumption of the resulting plan generator.

#### Exploiting Soft and Hard Correlations in Big Data Query Optimization

Hai Liu (Worcester Polytechnic Institute); Dongqing Xiao (Worcester Polytechnic Institute); Pankaj Didwania (Worcester Polytechnic Institute); Mohamed Eltabakh (Worcester Polytechnic Institute)

Big data infrastructures are increasingly supporting datasets that are relatively structured. These datasets are full of correlations among their attributes, which if managed in systematic ways would enable optimization opportunities that otherwise will be missed. Unlike relational databases in which discovering and exploiting the correlations in query optimization have been extensively studied, in big data infrastructures, such important data properties and their utilization have been mostly abandoned. The key reason is that domain experts may know many correlations but with a degree of uncertainty. Since the data is big, it is very challenging to validate such correlations, judge their worthiness, and put strategies for utilizing them in query optimization. Existing techniques for exploiting soft correlations in RDBMSs are heavily tailored towards optimizing factors inherent in relational databases, e.g., predicate selectivity and random I/O accesses of secondary indexes, which are issues not applicable to big data infrastructures, e.g., Hadoop. In this paper, we propose the EXORD system to fill in this gap by exploiting the data's correlations in big data query optimization. EXORD supports two types of correlations; hard and soft correlations. We introduce a new three-phase approach for (1) Validating and judging the worthiness of soft correlations, (2) Selecting and preparing the soft correlations for deployment by specially handling the violating data records, and (3) Deploying and exploiting the correlations in query optimization. We propose a novel cost-benefit model for adaptively selecting the most beneficial soft correlations w.r.t a given query workload while minimizing the introduced overhead. We show the complexity of this problem (NP-Hard), and propose a heuristic to efficiently solve it in a polynomial time. EXORD can be integrated with various state-of-art big data query optimization techniques, e.g., indexing and partitioning. EXORD prototype is implemented as an extension to the Hive engine on top of Hadoop. The experimental evaluation using real-world and synthetic datasets shows the potential of EXORD in achieving more than 10x speedup while introducing minimal storage overheads.

#### How Good Are Query Optimizers, Really?

Viktor Leis (Technische Universität München); Andrey Gubichev (Technische Universität München); Atanas Mirchev (Technische Universität München); Peter Boncz (CWI); Alfons Kemper (Technische Universität München); Thomas Neumann (Technische Universität München)

Finding a good join order is crucial for query performance. In this paper, we introduce the Join Order Benchmark (JOB) and experimentally revisit the main components in the classic query optimizer architecture using a complex, real-world data set and realistic multi-join queries. We investigate the quality of industrial-strength cardinality estimators and find that all estimators routinely produce large errors. We further show that while estimates are essential for finding a good join order, query performance is unsatisfactory if the query engine relies too heavily on these estimates. Using another set of experiments that measure the impact of the cost model, we find that it has much less influence on query performance than the cardinality estimates. Finally, we investigate plan enumeration techniques comparing exhaustive dynamic programming with heuristic algorithms and find that exhaustive enumeration improves performance despite the sub-optimal cardinality estimates.

#### Teaching an RDBMS about ontological constraints

Damian Bursztyjn (INRIA & LIX); François Goasdoué (University of Rennes 1 and INRIA); Ioana Manolescu (INRIA & LIX)

In the presence of an ontology, query answers must reflect not only data explicitly present in the database, but also implicit data, which holds due to the ontology, even though it is not present in the database. A large and useful set of ontologies enjoys FOL reducibility of query answering, that is: answering a query  $q$  can be reduced to evaluating a certain first-order logic (FOL) formula (obtained from

the query and ontology) against only the explicit facts. We present a novel query optimization framework for ontology-based data access settings enjoying FOL reducibility. Our framework is based on searching within a set of alternative equivalent FOL queries, i.e., FOL reformulations, one with minimal evaluation cost when evaluated through a relational database system. We apply this framework to the DL-LiteR Description Logic underpinning the W3C's OWL2 QL ontology language, and demonstrate through experiments its performance benefits when two leading SQL systems, one open-source and one commercial, are used for evaluating the FOL query reformulations.

## Research 11: Spatial Data and Queries - 1

**Session Chair: Amr El Abbadi (University of California at Santa Barbara)**

### **Processing and Optimizing Main Memory Spatial-Keyword Queries**

Taesung Lee (Yonsei University); Jin-Woo Park (POSTECH); Sanghoon Lee (POSTECH); Seung-won Hwang (Yonsei University); Sameh Elnikety (Microsoft Research); Yuxiong He (Microsoft Research)

Important cloud services rely on spatial-keyword queries, containing a spatial predicate and arbitrary boolean keyword queries. In particular, we study the processing of such queries in main memory to support short response times. In contrast, current state-of-the-art spatial-keyword indexes and relational engines are designed for different assumptions. Rather than building a new spatial-keyword index, we employ a cost-based optimizer to process these queries using a spatial index and a keyword index. We address several technical challenges to achieve this goal. We introduce three operators as the building blocks to construct plans for main memory query processing. We then develop a cost model for the operators and query plans. We introduce five optimization techniques that efficiently reduce the search space and produce a query plan with low cost. The optimization techniques are computationally efficient, and they identify a query plan with a formal approximation guarantee under the common independence assumption. Furthermore, we extend the framework to exploit interesting orders. We implement the query optimizer to empirically validate our proposed approach using real-life datasets. The evaluation shows that the optimizations provide significant reduction in the average and tail latency of query processing: 7- to 11-fold reduction over using a single index in terms of 99th percentile response time. In addition, this approach outperforms existing spatial-keyword indexes, and DBMS query optimizers for both average and high-percentile response times.

### **Spatial Online Sampling and Aggregation**

Lu Wang (Hong Kong University of Science and Technology); Robert Christensen (University of Utah); Feifei Li (University of Utah); Ke Yi (Hong Kong University of Science and Technology)

The massive adoption of smart phones and other mobile devices has generated humongous amount of spatial and spatio-temporal data. The importance of spatial analytics and aggregation is ever-increasing. An important challenge is to support interactive exploration over such data. However, spatial analytics and aggregation using all data points that satisfy a query condition is expensive, especially over large data sets, and could not meet the needs of interactive exploration. To that end, we present novel indexing structures that support spatial online sampling and aggregation on large spatial and spatio-temporal data sets. In spatial online sampling, random samples from the set of spatial (or spatio-temporal) points that satisfy a query condition are generated incrementally in an online fashion. With more and more samples, various spatial analytics and aggregations can be performed in an online, interactive fashion, with estimators that have better accuracy over time. Our design works well for both memory-based and disk-resident data sets, and scales well towards different query and sample sizes. More importantly, our structures are dynamic, hence, they are able to deal with insertions and deletions efficiently. Extensive experiments on large real data sets demonstrate the improvements achieved by our indexing structures compared to other baseline methods.

### **Online Minimum Matching in Real-Time Spatial Data: Experiments and Analysis**

Yongxin Tong (Beihang University); Jiaying She (Hong Kong University of Science and Technology); Bolin Din (Microsoft Research); Lei Chen (Hong Kong University of Science and Technology); Tianyu Wo (Beihang University); Ke Xu (Beihang University)

Recently, with the development of mobile Internet and smartphones, the online minimum bipartite matching in real time spatial data (OMBM) problem becomes popular. Specifically, given a set of service providers with specific locations and a set of users who dynamically appear one by one, the OMBM problem is to find a maximum-cardinality matching with minimum total distance following that once a user appears, s/he must be immediately matched to an unmatched service provider, which cannot be revoked, before subsequent users arrive. To address this problem, existing studies mainly focus on analyzing the worst-case competitive ratios of the proposed online algorithms, but study on the performance of the algorithms in practice is absent. In this paper, we present a comprehensive experimental comparison of the representative algorithms of the OMBM problem. Particularly, we observe a surprising result that the simple and efficient greedy algorithm, which has been considered as the worst due to its exponential worst-case competitive ratio, is significantly more effective than other algorithms. We investigate the results and further show that the competitive ratio of the worst case of the greedy algorithm is actually just a constant, 3.195, in the average-case analysis. We try to clarify a 24-year misunderstanding towards the greedy algorithm and justify that the greedy algorithm is not bad at all. Finally, we provide a uniform implementation for all the algorithms of the OMBM problem and clarify their strengths and weaknesses, which can guide practitioners to select appropriate algorithms for various scenarios.

## **AQWA: Adaptive Query-Workload-Aware Partitioning of Big Spatial Data**

Ahmed M. Aly (Purdue University); Ahmed R. Mahmood (Purdue University); Mohamed S. Hassan (Purdue University); Walid G. Aref (Purdue University); Mourad Ouzzani (Qatar Computing Research Institute); Hazem Elmeleegy (Turn Inc.); Thamiir Qadah (Purdue University)

The unprecedented spread of location-aware devices has resulted in a plethora of location-based services in which huge amounts of spatial data need to be efficiently processed by large-scale computing clusters. Existing cluster-based systems for processing spatial data employ static data-partitioning structures that cannot adapt to data changes, and that are insensitive to the query workload. Hence, these systems are incapable of consistently providing good performance. To close this gap, we present AQWA, an adaptive and query-workload-aware mechanism for partitioning large-scale spatial data. AQWA does not assume prior knowledge of the data distribution or the query workload. Instead, as data is consumed and queries are processed, the data partitions are incrementally updated. With extensive experiments using real spatial data from Twitter, and various workloads of range and k-nearest-neighbor queries, we demonstrate that AQWA can achieve an order of magnitude enhancement in query performance compared to the state-of-the-art systems.

## **Research 12: Distributed and Cloud Systems - 2**

### **Session Chair: Krithi Ramamritham (IIT Bombay)**

#### **High-Speed Query Processing over High-Speed Networks**

Wolf Roediger (Technische Universität München); Tobias Muehlbauer (Technische Universität München); Alfons Kemper (Technische Universität München); Thomas Neumann (Technische Universität München)

Modern database clusters entail two levels of networks: connecting CPUs and NUMA regions inside a single server in the small and multiple servers in the large. The huge performance gap between these two types of networks used to slow down distributed query processing to such an extent that a cluster of machines actually performed worse than a single many-core server. The increased main-memory capacity of the cluster remained the sole benefit of such a scale-out. The economic viability of high-speed interconnects such as InfiniBand has narrowed this performance gap considerably. However, InfiniBand's higher network bandwidth alone does not improve query performance as expected when the distributed query engine is left unchanged. The scalability of distributed query processing is impaired by TCP overheads, switch contention due to uncoordinated communication, and load imbalances resulting from the inflexibility of the classic exchange operator model. This paper presents the blueprint for a distributed query engine that addresses these problems by considering both levels of networks holistically. It consists of two parts: First, hybrid parallelism that distinguishes local and distributed parallelism for better scalability in both the number of cores as well as servers. Second, a novel communication multiplexer tailored for analytical database workloads using remote direct memory access (RDMA) and low-latency network scheduling for high-speed communication with almost no CPU overhead. An extensive evaluation within the HyPer database system using the TPC-H benchmark shows that our holistic approach indeed enables high-speed query processing over high-speed networks.

#### **The End of Slow Networks: It's Time for a Redesign**

Carsten Binnig (Brown University); Andrew Crotty (Brown University); Alex Galakatos (Brown University); Tim Kraska (Brown University); Erfan Zamanian (Brown University)

The next generation of high-performance networks with remote direct memory access (RDMA) capabilities requires a fundamental rethinking of the design of distributed in-memory DBMSs. These systems are commonly built under the assumption that the network is the primary bottleneck and should be avoided at all costs, but this assumption no longer holds. For instance, with InfiniBand FDR 4x, the bandwidth available to transfer data across the network is in the same ballpark as the bandwidth of one memory channel. Moreover, RDMA transfer latencies continue to rapidly improve as well. In this paper, we first argue that traditional distributed DBMS architectures cannot take full advantage of high-performance networks and suggest a new architecture to address this problem. Then, we discuss initial results from a prototype implementation of our proposed architecture for OLTP and OLAP, showing remarkable performance improvements over existing designs.

#### **Tempo: Robust and Self-Tuning Resource Management in Multi-tenant Parallel Databases**

Zilong Tan (Duke University); Shivnath Babu (Duke University)

Multi-tenant database systems have a component called the Resource Manager, or RM that is responsible for allocating resources to tenants. RMs today do not provide direct support for performance objectives such as: "Average job response time of tenant A must be less than two minutes", or "No more than 5% of tenant B's jobs can miss the deadline of 1 hour." Thus, DBAs have to tinker with the RM's low-level configuration settings to meet such objectives. We propose a framework called Tempo that brings simplicity, self-tuning, and robustness to existing RMs. Tempo provides a simple interface for DBAs to specify performance objectives declaratively, and optimizes the RM configuration settings to meet these objectives. Tempo has a solid theoretical foundation which gives key robustness guarantees. We report experiments done on Tempo using production traces of data-processing workloads from companies such as Facebook and Cloudera. These experiments demonstrate significant improvements in meeting desired performance objectives over RM configuration settings specified by human experts.

## Measuring and Optimizing Distributed Array Programs

Mingxing Zhang (Tsinghua University); Yongwei Wu (Tsinghua University); Kang Chen (Tsinghua University); Teng Ma (Tsinghua University); Weimin Zheng (Tsinghua University)

Nowadays, there is a rising trend of building array-based distributed computing frameworks. By taking advantage of highly optimized array/matrix operations, these frameworks have achieved significant improvement on performance for many important algorithms. However, due to the absence of a comprehensive optimizer, most of these frameworks execute each primitive in an isolated manner and in the exact order defined by programmers, which implies a huge space for optimization. In this paper, we propose a novel array-based programming model, named as KASEN, which distinguishes itself from models in the existing literature by defining a strict computation and communication model. This model makes it easy to analyze programs' behavior and measure their performance, with which we design a corresponding optimizer that can automatically apply many high-level optimizations to the original programs written by programmers. According to our evaluation, KASEN is sufficient for implementing many important machine learning and graph algorithms, and it can be extended very easily. More importantly, the optimizer of KASEN can achieve significant reduction on memory read/write, buffer allocation and network traffic. The speedup achieved by our optimizer is up to 5.82X for real-world and synthetic datasets.

## Research 13: Query Execution - 2

### Session Chair: Wolfgang Lehner (Technische Universität Dresden)

#### A Seven-Dimensional Analysis of Hashing Methods and its Implications on Query Processing

Stefan Richter (Saarland University); Victor Alvarez (TU Braunschweig); Jens Dittrich (Saarland University)

Hashing is a solved problem. It allows us to get constant time access for lookups. Hashing is also simple. It is safe to use an arbitrary method as a black box and expect good performance, and optimizations to hashing can only improve it by a negligible delta. Why are all of the previous statements plain wrong? That is what this paper is about. In this paper we thoroughly study hashing for integer keys and carefully analyze the most common hashing methods in a five-dimensional requirements space: (1) data-distribution, (2) load factor, (3) dataset size, (4) read/write-ratio, and (5) un/successful-ratio. Each point in that design space may potentially suggest a different hashing scheme, and additionally also a different hash function. We show that a right or wrong decision in picking the right hashing scheme and hash function combination may lead to significant difference in performance. To substantiate this claim, we carefully analyze two additional dimensions: (6) five representative hashing schemes (which includes an improved variant of Robin Hood hashing), (7) four important classes of hash functions widely used today. That is, we consider 20 different combinations in total. Finally, we also provide a glimpse about the effect of table memory layout and the use of SIMD instructions. Our study clearly indicates that picking the right combination may have considerable impact on insert and lookup performance, as well as memory footprint. A major conclusion of our work is that hashing should be considered a white box before blindly using it in applications, such as query processing. Finally, we also provide a strong guideline about when to use which hashing method.

#### Query-Aware Locality-Sensitive Hashing for Approximate Nearest Neighbor Search

Qiang Huang (Sun Yat-sen University); Jianlin Feng (Sun Yat-sen University); Yikai Zhang (Sun Yat-sen University); Qiong Fang (South China University of Technology); Wilfred Ng (Hong Kong University of Science and Technology)

Locality-Sensitive Hashing (LSH) and its variants are the well-known indexing schemes for the problem of  $c$ -Approximate Nearest Neighbor ( $c$ -ANN) search in high-dimensional Euclidean space. Traditionally, LSH functions are constructed in a query-oblivious manner in the sense that buckets are partitioned before any query arrives. However, objects closer to a query may be partitioned into different buckets, which is undesirable. Due to the use of query-oblivious bucket partition, the state-of-the-art LSH schemes for external memory, namely C2LSH and LSB-Forest, only work with approximation ratio of integer  $c \geq 2$ . In this paper, we introduce a novel concept of query-aware bucket partition which uses a given query as the "anchor" for bucket partition. Accordingly, a query-aware LSH function is a random projection coupled with query-aware bucket partition, which removes random shift required by traditional query-oblivious LSH functions. Notably, query-aware bucket partition can be easily implemented so that query performance is guaranteed. We propose a novel query-aware LSH scheme named QALSH for  $c$ -ANN search over external memory. Our theoretical studies show that QALSH enjoys guarantee on query quality. The use of query-aware LSH function enables QALSH to work with any approximation ratio  $c > 1$ . Extensive experiments on four real datasets show that QALSH outperforms C2LSH and LSB-Forest. Specifically, by using a ratio  $c < 2$ , QALSH can achieve much better query quality.

#### Neighbor-Sensitive Hashing

Yongjoo Park (University of Michigan); Michael Cafarella (University of Michigan); Barzan Mozafari (University of Michigan)

Approximate kNN ( $k$ -nearest neighbor) techniques using binary hash functions are among the most commonly used approaches for overcoming the prohibitive cost of performing exact kNN queries. However, the success of these techniques largely depends on their hash functions' ability to distinguish kNN items; that is, the kNN items retrieved based on data items' hashcodes, should include as many true kNN items as possible. A widely-adopted principle for this process is to ensure that similar items are assigned to the same hashcode so that the items with the hashcodes similar to a query's hashcode are likely to be true neighbors. In this work, we abandon this heavily-utilized principle and pursue the opposite direction for generating more effective hash functions for kNN tasks. That is, we

aim to increase the distance between similar items in the hashcode space, instead of reducing it. Our contribution begins by providing theoretical analysis on why this revolutionary and seemingly counter-intuitive approach leads to a more accurate identification of kNN items. Our analysis is followed by a proposal for a hashing algorithm that embeds this novel principle. Our empirical studies confirm that a hashing algorithm based on this counter-intuitive idea significantly improves the efficiency and accuracy of state-of-the-art techniques.

### **LSH Ensemble: Internet Scale Domain Search**

Erkang Zhu (University of Toronto); Fatemeh Nargesian (University of Toronto); Ken Pu (UOIT); Renée Miller (University of Toronto)

We study (the problem of domain search) where a domain is a set of distinct values from an unspecified universe. We use Jaccard set containment score, defined as  $\frac{|Q \cap X|}{|Q|}$ , as the measure of relevance of a domain  $X$  to a query domain  $Q$ . Our choice of Jaccard set containment over Jaccard similarity as a measure of relevance makes our work particularly suitable for searching Open Data and data on the web, as Jaccard similarity is known to have poor performance over sets with large differences in their domain sizes. We demonstrate that the domains found in several real-life Open Data and web data repositories show a power-law distribution over their domain sizes. We present a new index structure, Locality Sensitive Hashing (LSH) Ensemble, that solves the domain search problem using set containment at Internet scale. Our index structure and search algorithm cope with the data volume and skew by means of data sketches using Minwise Hashing and domain partitioning. Our index structure does not assume a prescribed set of data values. We construct a cost model that describes the accuracy of LSH Ensemble with any given partitioning. This allows us to formulate the data partitioning for LSH Ensemble as an optimization problem. We prove that there exists an (optimal) partitioning for any data distribution. Furthermore, for datasets following a power-law distribution, as observed in Open Data and Web data corpora, we show that the optimal partitioning can be approximated using equi-depth, making it particularly efficient to use in practice. We evaluate our algorithm using real data (Canadian Open Data and WDC Web Tables) containing over 262 million domains. The experiments demonstrate that our index consistently outperforms other leading alternatives in accuracy and performance. The improvements are most dramatic for data with large skew in the domain sizes. Even at 262 million domains, our index sustains query performance with under 3 seconds response time.

## **Research 14: Graph Processing - 2**

### **Session Chair: Kyuseok Shim (Seoul National University)**

#### **G-SQL: Fast Query Processing via Graph Exploration**

Hongbin Ma (eBay); Bin Shao (Microsoft Research Asia); Yanghua Xiao (Fudan University); Liang Jeff Chen (Microsoft Research Asia); Haixun Wang (Facebook)

A lot of real-life data is of graph nature. However, it is not until recently that business begins to exploit data's connectedness for business insights. On the other hand, RDBMSs are a mature technology for data management, but they are not for graph processing. Take graph traversal, a common graph operation for example, it heavily relies on a graph primitive that accesses a given node's neighborhood. We need to join tables following foreign keys to access the nodes in the neighborhood if an RDBMS is used to manage graph data. Graph exploration is a fundamental building block of many graph algorithms. But this simple operation is costly due to large volumes of I/O caused by the massive amount of table joins. In this paper, we present G-SQL, our first effort towards the integration of a RDBMS and a native graph processing engine (Trinity). G-SQL leverages fast graph exploration provided by Trinity's memory cloud to answer multi-way join queries, and it uses RDBMSs to provide mature data management functionalities, including reliable data storage, additional data access methods, etc. Specifically, G-SQL is a SQL dialect augmented with graph exploration functionalities and it dispatches query tasks to its underlying RDBMS and the Trinity graph engine. The G-SQL runtime coordinates the two query processors, and uses a unified cost model to ensure that the entire query is processed efficiently. Experimental results show that our approach greatly expands capabilities of RDBMSs and delivers exceptional performance for SQL-graph hybrid queries.

#### **Fully Dynamic Betweenness Centrality Maintenance on Massive Networks**

Takanori Hayashi (The University of Tokyo); Takuya Akiba (National Institute of Informatics); Yuichi Yoshida (National Institute of Informatics)

Measuring the relative importance of each vertex in a network is one of the most fundamental building blocks in network analysis. Among several importance measures, betweenness centrality, in particular, plays key roles in many real applications. Considerable effort has been made for developing algorithms for static settings. However, real networks today are highly dynamic and are evolving rapidly, and scalable dynamic methods that can instantly reflect graph changes into centrality values are required. In this paper, we present the first fully dynamic method for managing betweenness centrality of all vertices in a large dynamic network. Its main data structure is the weighted hyperedge representation of shortest paths called hypergraph sketch. We carefully design dynamic update procedure with theoretical accuracy guarantee. To accelerate updates, we further propose two auxiliary data structures called two-ball index and special-purpose reachability index. Experimental results using real networks demonstrate its high scalability and efficiency. In particular, it can reflect a graph change in less than a millisecond on average for a large-scale web graph with 106M vertices and 3.7B edges, which is several orders of magnitude larger than the limits of previous dynamic methods.

### **Parallel Local Graph Clustering**

Julian Shun (University of California at Berkeley); Farbod Roosta-Khorasani (University of California at Berkeley); Kimon Fountoulakis (University of California at Berkeley); Michael Mahoney (University of California at Berkeley)

Graph clustering has many important applications in computing, but due to growing sizes of graph, even traditionally fast clustering methods such as spectral partitioning can be computationally expensive for real-world graphs of interest. Motivated partly by this, so-called local algorithms for graph clustering have received significant interest due to the fact that they can find good clusters in a graph with work proportional to the size of the cluster rather than that of the entire graph. This feature has proven to be crucial in making such graph clustering and many of its downstream applications efficient in practice. While local clustering algorithms are already faster than traditional algorithms that touch the entire graph, there is an opportunity to make them even more efficient via parallelization. Existing local clustering algorithms are sequential. In this paper, we show how to parallelize many of these algorithms in the shared-memory multicore setting, and we analyze the parallel complexity of these algorithms. We present comprehensive experiments on large-scale graphs showing that the parallel algorithms achieve good parallel speedups on a modern multicore machine, thus significantly speeding up the analysis of local graph clusters in the very large-scale setting.

### **KCore Decomposition for Large Networks on a Single PC**

Wissam Khaoiid (University of Victoria); Marina Barsky (University of Victoria); Venkatesh Srinivasan (University of Victoria); Alex Thomo (University of Victoria)

Studying the topology of a network is critical to inferring underlying dynamics such as tolerance to failure, group behavior and spreading patterns.  $k$ -core decomposition is a well-established metric which partitions a graph into layers from external to more central vertices. In this paper we aim to explore whether  $k$ -core decomposition of large networks can be computed using a consumer-grade PC. We feature implementations of the “vertex-centric” distributed protocol introduced by Montresor, De Pellegrini and Miorandi on GraphChi and Webgraph. Also, we present an accurate implementation of the Batagelj and Zaversnik algorithm for  $k$ -core decomposition in Webgraph. With our implementations, we show that we can efficiently handle networks of billions of edges using a single consumer-level machine within reasonable time and can produce excellent approximations in only a fraction of the execution time. To the best of our knowledge, our biggest graphs are considerably larger than the graphs considered in the literature. Next, we present an optimized implementation of an external-memory algorithm (EMCore) by Cheng, Ke, Chu, and Ozsu. We show that this algorithm also performs well for large datasets, however, it cannot predict whether a given memory budget is sufficient for a new dataset. We present a thorough analysis of all algorithms concluding that it is viable to compute  $k$ -core decomposition for large networks in a consumer-grade PC.

## **Research 15: Data Cleaning - 2**

### **Session Chair: Panos Chrysanthis (University of Pittsburgh)**

#### **Messing Up with BART: Error Generation for Evaluating Data-Cleaning Algorithms**

Patricia Arocena (University of Toronto); Boris Glavic (Illinois Institute of Technology); Giansalvatore Mecca (Università di Basilicata); Renée Miller (University of Toronto); Paolo Papotti (Qatar Computing Research Institute); Donatello Santoro (Università di Basilicata)

We study the problem of introducing errors into clean databases for the purpose of benchmarking data-cleaning algorithms. Our goal is to provide users with the highest possible level of control over the error-generation process, and at the same time develop solutions that scale to large databases. We show in the paper that the error-generation problem is surprisingly challenging, and in fact, NP-complete. To provide a scalable solution, we develop a correct and efficient greedy algorithm that sacrifices completeness, but succeeds under very reasonable assumptions. To scale to millions of tuples, the algorithm relies on several non-trivial optimizations, including a new symmetry property of data quality constraints. The trade-off between control and scalability is the main technical contribution of the paper.

#### **Detecting Data Errors: Where are we and what needs to be done?**

Ziawasch Abedjan (MIT); Xu Chu (University of Waterloo); Dong Deng (Tsinghua University); Raul Castro Fernandez (MIT); Ihab Ilyas (University of Waterloo); Mourad Ouzzani (Qatar Computing Research Institute); Paolo Papotti (Arizona State University); Michael Stonebraker (MIT); Nan Tang (Qatar Computing Research Institute)

Data cleaning has played a critical role in ensuring data quality for enterprise applications. Naturally, there has been extensive research in this area, and many data cleaning algorithms have been translated into tools to detect and to possibly repair certain classes of errors such as outliers, duplicates, missing values, and violations of integrity constraints. Since different types of errors may coexist in the same dataset, we often need to run more than one kind of tool. In this paper, we investigate two pragmatic questions: (1) are these tools robust enough to capture most errors in real-world data sets? and (2) what is the best strategy to holistically run multiple tools to optimize the detection effort? To answer these two questions, we obtained multiple data cleaning tools that utilize a variety of error detection techniques. We also collected five real-world data sets, for which we could obtain both the raw data and the ground truth on existing errors. In this paper, we report our experimental findings on the errors detected by the tools we tested. First, we show that the coverage of each tool is well below 100%. Second, we show that the order in which multiple tools are run makes a big difference. Hence, we propose a holistic multi-tool strategy that orders

the invocations of the available tools to maximize their benefit, while minimizing human effort in verifying results. Third, since this holistic approach still does not lead to acceptable error coverage, we discuss two simple strategies that have the potential to improve the situation, namely domain specific tools and data enrichment. We close this paper by reasoning about the errors that are not detectable by any of the tools we tested.

### **ActiveClean: Interactive Data Cleaning for Statistical Modeling**

Sanjay Krishnan (University of California at Berkeley); Jiannan Wang (Simon Fraser University); Eugene Wu (Columbia University); Michael Franklin (University of California at Berkeley); Ken Goldberg (University of California at Berkeley)

Analysts often clean dirty data iteratively—cleaning some data, executing the analysis, and then, cleaning more data based on the results. We explore the iterative cleaning process in the context of statistical model training, which is an increasingly popular form of data analytics. The challenge is that training high-dimensional statistical models on partially clean data lacks guarantees, and the iterative cleaning process may actually diverge. We propose ActiveClean which allows for progressive and iterative cleaning in statistical modeling problems while preserving convergence guarantees. ActiveClean supports a popular class of models called convex loss models (e.g., linear regression and SVMs), and also prioritizes cleaning those records likely to affect the results. We evaluate ActiveClean on four real-world datasets UCI Adult, UCI EEG, MNIST, and Dollars for Docs with both real and synthetic errors. Our results suggest that our proposed optimizations can improve model accuracy by up-to 2.5x for the same amount of data cleaned. Furthermore, for a fixed cleaning budget and on all real dirty datasets, ActiveClean returns more accurate models than uniform sampling and Active Learning.

### **META: An Efficient Matching-Based Method for Error-Tolerant Autocompletion**

Dong Deng (Tsinghua University); Guoliang Li (Tsinghua University); He Wen (Tsinghua University); H. V. Jagadish (University of Michigan); Feng Jianhua (Tsinghua University)

Autocompletion has been widely adopted in many computing systems because it can instantly provide users with results as users type in queries. Since the typing task is tedious and prone to error, especially on mobile devices, a recent trend is to tolerate errors in autocompletion. Existing error-tolerant autocompletion methods build a trie to index the data, utilize the trie index to compute the trie nodes that are similar to the query, called active nodes, and identify the leaf descendants of active nodes as the results. However these methods have two limitations. First, they involve many redundant computations to identify the active nodes. Second, they do not support top-k queries. To address these problems, we propose a matching-based framework, which computes the answers based on matching characters between queries and data. We design a compact tree index to maintain active nodes in order to avoid the redundant computations. We devise an incremental method to efficiently answer top-k queries. Experimental results on real datasets show that our method outperforms state-of-the-art approaches by 1-2 orders of magnitude.

## **Research 16: Data and Query Models - 1**

### **Session Chair: Yanlei Diao (University of Massachusetts Amherst)**

#### **SlimShot: In-Database Probabilistic Inference for Knowledge Bases**

Eric Gribkoff (University of Washington); Dan Suciu (University of Washington)

Increasingly large Knowledge Bases are being created, by crawling the Web or other corpora of documents, and by extracting facts and relations using machine learning techniques. To manage the uncertainty in the data, these KBs rely on probabilistic engines based on Markov Logic Networks (MLN), for which probabilistic inference remains a major challenge. Today's state of the art systems use variants of MCMC, which have no theoretical error guarantees, and, as we show, suffer from poor performance in practice. In this paper we describe SlimShot (Scalable Lifted Inference and Monte Carlo Sampling Hybrid Optimization Technique), a probabilistic inference engine for knowledge bases. SlimShot converts the MLN to a tuple-independent probabilistic database, then uses a simple Monte Carlo-based inference, with three key enhancements: (1) it combines sampling with safe query evaluation, (2) it estimates a conditional probability by jointly computing the numerator and denominator, and (3) it adjusts the proposal distribution based on the sample cardinality. In combination, these three techniques allow us to give formal error guarantees, and we demonstrate empirically that SlimShot outperforms today's state of the art probabilistic inference engines used in knowledge bases.

#### **Scalable Package Queries in Relational Database Systems**

Matteo Brucato (University of Massachusetts Amherst); Juan Beltran (New York University Abu Dhabi); Azza Abouzied (New York University Abu Dhabi); Alexandra Meliou (University of Massachusetts Amherst)

Traditional database queries follow a simple model: they define constraints that each tuple in the result must satisfy. This model is computationally efficient, as the database system can evaluate the query conditions on each tuple individually. However, many practical, real-world problems require a collection of result tuples to satisfy constraints collectively, rather than individually. In this paper, we present package queries, a new query model that extends traditional database queries to handle complex constraints and preferences over answer sets. We develop a full-fledged package query system, implemented on top of a traditional database engine. Our work makes several contributions. First, we design PaQL, a SQL-based query language that supports the declarative specification of package queries. We prove that PaQL is as least as expressive as integer



linear programming, and therefore, evaluation of package queries is in general NP-hard. Second, we present a fundamental evaluation strategy that combines the capabilities of databases and constraint optimization solvers to derive solutions to package queries. The core of our approach is a set of translation rules that transform a package query to an integer linear program. Third, we introduce an offline data partitioning strategy allowing query evaluation to scale to large data sizes. Fourth, we introduce SketchRefine, a scalable algorithm for package evaluation, with strong approximation guarantees  $(1 \pm \epsilon)^6$ -factor approximation). Finally, we present extensive experiments over real-world and benchmark data. The results demonstrate that SketchRefine is effective at deriving high-quality package results, and achieves runtime performance that is an order of magnitude faster than directly using ILP solvers over large datasets.

### Explaining Query Answers with Explanation-Ready Databases

Sudeepa Roy (Duke University); Laurel Orr (University of Washington); Dan Suciu (University of Washington)

With the increased generation and availability of big data in different domains, there is an imminent requirement for data analysis tools that are able to 'explain' the trends and anomalies obtained from this data to a range of users with different backgrounds. Wu-Madden (PVLDB 2013) and Roy-Suciu (SIGMOD 2014) recently proposed solutions that can explain interesting or unexpected answers to simple aggregate queries in terms of predicates on attributes. In this paper, we propose a generic framework that can support much richer, insightful explanations by preparing the database offline, so that top explanations can be found interactively at query time. The main idea in such 'explanation-ready databases' is to pre-compute the effects of potential explanations (called 'interventions'), and efficiently re-evaluate the original query taking into accounts these effects. We formalize this notion and define an 'explanation-query' that can evaluate all possible explanations simultaneously without having to run an iterative process, develop algorithms and optimizations, and evaluate our approach with experiments on real data.

### A framework for annotating CSV-like data

Marcelo Arenas (PUC Chile); Francisco Maturana (PUC Chile); Cristian Riveros (PUC Chile); Domagoj Vrgoc (PUC Chile)

In this paper, we propose a simple and expressive framework for adding metadata to CSV documents and their noisy variants. The framework is based on annotating parts of the document that can be later used to read, query, or exchange the data. The core of our framework is a language based on extended regular expressions that are used for selecting data. These expressions are then combined using a set of rules in order to annotate the data. We study the computational complexity of implementing our framework and present an efficient evaluation algorithm that runs in time proportional to its output and linear in its input. As a proof of concept, we test an implementation of our framework against a large number of real world datasets and show that it can be efficiently used in practice.

## Research 17: Scalable Analytics - 1

### Session Chair: Matthias Boehm (IBM Research - Almaden)

#### A General-Purpose Query-Centric Framework for Querying Big Graphs

Da Yan (The Chinese University of Hong Kong); James Cheng (The Chinese University of Hong Kong); Tamer Özsu (University of Waterloo); Fan Yang (The Chinese University of Hong Kong); Yi Lu (The Chinese University of Hong Kong); John C.S. Lui (The Chinese University of Hong Kong); Qizhen Zhang (The Chinese University of Hong Kong); Wilfred Ng (The Hong Kong University of Science and Technology)

Pioneered by Google's Pregel, many distributed systems have been developed for large-scale graph analytics. These systems expose the user-friendly "think like a vertex" programming interface to users, and exhibit good horizontal scalability. However, these systems are designed for tasks where the majority of graph vertices participate in computation, but are not suitable for processing light-workload graph queries where only a small fraction of vertices need to be accessed. The programming paradigm adopted by these systems can seriously under-utilize the resources in a cluster for graph query processing. In this work, we develop a new open-source system, called Quegel, for querying big graphs, which treats queries as first-class citizens in the design of its computing model. Users only need to specify the Pregel-like algorithm for a generic query, and Quegel processes light-workload graph queries on demand using a novel superstep-sharing execution model to effectively utilize the cluster resources. Quegel further provides a convenient interface for constructing graph indexes, which significantly improve query performance but are not supported by existing graph-parallel systems. Our experiments verified that Quegel is highly efficient in answering various types of graph queries and is up to orders of magnitude faster than existing systems.

#### Distance-based Outlier Detection in Data Streams

Luan Tran (University of Southern California); Liyue Fan (University of Southern California); Cyrus Shahabi (University of Southern California)

Continuous outlier detection in data streams has important applications in fraud detection, network security, and public health. The arrival and departure of data objects in a streaming manner impose new challenges for outlier detection algorithms, especially in time and space efficiency. In the past decade, several studies have been performed to address the problem of distance-based outlier detection in data streams (DODDS), which adapts an unsupervised definition and does not have any distributional assumption on data values. Our work is motivated by the lack of comparative evaluation among the state-of-the-art algorithms using the same dataset selections and on the same platform. We systematically evaluate five most recent exact algorithms for DODDS under various stream

settings and outlier rates. Our extensive results show that in most settings, the MCO algorithm provides superior performance among all algorithms, including the most recent algorithm `Thresh_LEAP`.

### **SEEDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics**

Manasi Vartak (MIT); Sajjadur Rahman (University of Illinois at Urbana-Champaign); Samuel Madden (MIT); Aditya Parameswaran (University of Illinois at Urbana-Champaign); Neoklis Polyzotis (Google)

Data analysts often build visualizations as the first step in their analytical workflow. However, when working with high-dimensional datasets, identifying visualizations that show relevant or desired trends in data can be laborious. We propose SEEDB, a visualization recommendation engine to facilitate fast visual analysis: given a subset of data to be studied, SEEDB intelligently explores the space of visualizations, evaluates promising visualizations for trends, and recommends those it deems most “useful” or “inter-esting”. The two major obstacles in recommending interesting visualizations are (a) scale: evaluating a large number of candidate visualizations while responding within interactive time scales, and (b) utility: identifying an appropriate metric for assessing interest-ingness of visualizations. For the former, SEEDB introduces pruning optimizations to quickly identify high-utility visualizations and sharing optimizations to maximize sharing of computation across visualizations. For the latter, as a first step, we adopt a deviation-based metric for visualization utility, while indicating how we may be able to generalize it to other factors influencing utility. We implement SEEDB as a middleware layer that can run on top of any DBMS. Our experiments show that our framework can identify interesting visualizations with high accuracy. Our optimizations lead to multiple orders of magnitude speedup on relational row and column stores and provide recommendations at interactive time scales. Finally, we demonstrate via a user study the effectiveness of our deviation-based utility metric and the value of recommendations in supporting visual analytics

### **Explaining Outputs in Modern Data Analytics**

Zaheer Chothia (ETH Zurich); John Liagouris (ETH Zurich); Frank McSherry (ETH Zurich); Timothy Roscoe (ETH Zurich)

We report on the design and implementation of a general framework for interactively explaining the outputs of modern data-parallel computations, including iterative data analytics. To produce explanations, existing works adopt a naive backward tracing approach which runs into known issues; naive backward tracing may identify: (i) too much information that is difficult to process, and (ii) not enough information to reproduce the output, which hinders the logical debugging of the program. The contribution of this work is twofold. First, we provide methods to effectively reduce the size of explanations based on the first occurrence of a record in an iterative computation. Second, we provide a general method for identifying explanations that are sufficient to reproduce the target output in arbitrary computations - a problem for which no viable solution existed until now. We implement our approach on differential dataflow, a modern high-throughput, low-latency dataflow platform. We add a small (but extensible) set of rules to explain each of its data-parallel operators, and we implement these rules as differential dataflow operators themselves. This choice allows our implementation to inherit the performance characteristics of differential dataflow, and results in a system that efficiently computes and updates explanatory inputs even as the inputs of the reference computation change. We evaluate our system with various analytic tasks on real datasets, and we show that it produces concise explanations in tens of milliseconds, while remaining faster - up to two orders of magnitude - than even the best implementations that do not support explanations.

## **Research 18: Database Hardware-Software Codesign**

### **Session Chair: Ryan Stutsman (University of Utah)**

#### **YourSQL: A High-Performance Database System Leveraging In-Storage Computing**

Insoon Jo (Samsung Electronics Co.); Duck-Ho Bae (Samsung Electronics Co.); Andre Yoon (Samsung Electronics Co.); Jeong-Uk Kang (Samsung Electronics Co.); Sangyeun Cho (Samsung Electronics Co.); Daniel Lee (Samsung Electronics Co.); Jaeheon Jeong (Samsung Electronics Co.)

This paper presents YourSQL, a database system that accelerates data-intensive queries with the help of additional in-storage computing capabilities. YourSQL realizes very early filtering of data by offloading data scanning of a query to user-programmable solid-state drives. We implement our system on a recent branch of MariaDB (a variant of MySQL). In order to quantify the performance gains of YourSQL, we evaluate SQL queries with varying complexities. Our result shows that YourSQL reduces the execution time of the whole TPC-H queries by 3.6x, compared to a vanilla system. Moreover, the average speed-up of the five TPC-H queries with the largest performance gains reaches over 15x. Thanks to this significant reduction of execution time, we observe sizable energy savings. Our study demonstrates that the YourSQL approach, combining the power of early filtering with end-to-end datapath optimization, can accelerate large-scale analytic queries with lower energy consumption.

#### **Cheap Data Analytics using Cold Storage Devices**

Renata Borovica-Gajic (EPFL); Raja Appuswamy (EPFL); Anastasia Ailamaki (EPFL)

Enterprise databases use storage tiering to lower capital and operational expenses. In such a setting, data waterfalls from an SSD based high-performance tier when it is “hot” (frequently accessed) to a disk-based capacity tier and finally to a tape-based archival tier when “cold” (rarely accessed). To address the unprecedented growth in the amount of cold data, hardware vendors introduced new devices named Cold Storage Devices (CSD) explicitly targeted at cold data workloads. With access latencies in tens of seconds and cost/GB as low as 0.01\$/GB/month, CSD provide a middle ground between the low-latency (ms), high-cost, HDD-based capacity tier, and

high-latency (min to h), low-cost, tape-based, archival tier. Driven by the price/performance aspects of CSD, this paper makes a case for using CSD as a replacement for both capacity and archival tiers of enterprise databases. Although CSD offer major cost savings, we show that current database systems can suffer from severe performance drop when CSD are used as a replacement for HDD due to the mismatch between design assumptions made by the query execution engine and actual storage characteristics of the CSD. We then build a CSD-driven query execution framework, called Skipper, that modifies both the database execution engine and CSD scheduling algorithms to be aware of each other. Using results from our implementation of the architecture based on PostgreSQL and OpenStack Swift, we show that Skipper is capable of completely masking the high latency overheads of CSD, thereby opening up CSD for wider adoption as a storage tier for cheap data analytics over cold data.

### **An Experimental Evaluation of Datacenter Workloads On Low-Power Embedded Micro Servers**

Yiran Zhao (University of Illinois at Urbana-Champaign); Shen Li (University of Illinois at Urbana-Champaign); Shaohan Hu (University of Illinois at Urbana-Champaign); Hongwei Wang (University of Illinois at Urbana-Champaign); Shuochao Yao (University of Illinois at Urbana-Champaign); Huajie Shao (University of Illinois at Urbana-Champaign); Tarek Abdelzaher (University of Illinois at Urbana-Champaign)

This paper presents a comprehensive evaluation of an ultra-low power cluster, built upon the Intel Edison based micro servers. The improved performance and high energy efficiency of micro servers have driven both academia and industry to explore the possibility of replacing conventional brawny servers with a larger swarm of embedded micro servers. Existing attempts mostly focus on mobile-class micro servers, whose capacities are similar to mobile phones. We, on the other hand, target on sensor-class micro servers, which are originally intended for uses in wearable technologies, sensor networks, and Internet-of-Things. Although sensor-class micro servers have much less capacity, they are touted for minimal power consumption (<1 Watt), which opens new possibilities of achieving higher energy efficiency in datacenter workloads. Our systematic evaluation of the Edison cluster and comparisons to conventional brawny clusters involve careful workload choosing and laborious parameter tuning, which ensures maximum server utilization and thus fair comparisons. Results show that the Edison cluster achieves up to 3.5 times improvement on work-done-per-joule for web service applications and data-intensive MapReduce jobs. In terms of scalability, the Edison cluster scales linearly on the throughput of web service workloads, and also shows satisfactory scalability for MapReduce workloads despite coordination overhead.

## **Research 19: Query Optimization - 2**

### **Session Chair: Johann Christoph-Freytag (Humboldt University)**

#### **Optimization of Conjunctive Predicates for Main Memory Column Stores**

Fisnik Kastrati (University of Mannheim); Guido Moerkotte (University of Mannheim)

Optimization of queries with conjunctive predicates for main memory databases remains a challenging task. The traditional way of optimizing this class of queries relies on predicate ordering based on selectivities or ranks. However, the optimization of queries with conjunctive predicates is a much more challenging task, requiring a holistic approach in view of (1) an accurate cost model that is aware of CPU architectural characteristics such as branch (mis)prediction, (2) a storage layer, allowing for a streamlined query execution, (3) a common subexpression elimination technique, minimizing column access costs, and (4) an optimization algorithm able to pick the optimal plan even in presence of a small (bounded) estimation error. In this work, we embrace the holistic approach, and show its superiority experimentally. Current approaches typically base their optimization algorithms on at least one of two assumptions: (1) the predicate selectivities are assumed to be independent, (2) the predicate costs are assumed to be constant. Our approach is not based on these assumptions, as they in general do not hold.

#### **Parallelizing Query Optimization on Shared-Nothing Architectures**

Immanuel Trummer (EPFL); Christoph Koch (EPFL)

Data processing systems offer an ever increasing degree of parallelism on the levels of cores, CPUs, and processing nodes. Query optimization must exploit high degrees of parallelism in order not to gradually become the bottleneck of query evaluation. We show how to parallelize query optimization at a massive scale. We present algorithms for parallel query optimization in left-deep and bushy plan spaces. At optimization start, we divide the plan space for a given query into partitions of equal size that are explored in parallel by worker nodes. At the end of optimization, each worker returns the optimal plan in its partition to the master which determines the globally optimal plan from the partition-optimal plans. No synchronization or data exchange is required during the actual optimization phase. The amount of data sent over the network, at the start and at the end of optimization, as well as the complexity of serial steps within our algorithms increase only linearly in the number of workers and in the query size. The time and space complexity of optimization within one partition decreases uniformly in the number of workers. We parallelize single- and multi-objective query optimization over a cluster with 100 nodes in our experiments, using more than 250 concurrent worker threads (Spark executors). Despite high network latency and task assignment overheads, parallelization yields speedups of up to one order of magnitude for large queries whose optimization takes minutes on a single node.

#### **Multiple Query Optimization on the D-Wave 2X Adiabatic Quantum Computer**

Immanuel Trummer (EPFL); Christoph Koch (EPFL)

The D-Wave adiabatic quantum annealer solves hard combinatorial optimization problems leveraging quantum physics. The newest version features over 1000 qubits and was released in August 2015. We were given access to one of the first exemplars, currently hosted at NASA Ames research center in California, to explore the potential for hard optimization problems that arise in the context of databases. In this paper, we tackle the problem of multiple query optimization (MQO). We show how a MQO problem instance can be transformed into a mathematical formula that complies with the restrictive input format accepted by the quantum annealer. This formula is translated into ferromagnetic weights on and between qubits such that the configuration minimizing the input formula can be found via a process called adiabatic quantum annealing. We analyze the asymptotic growth rate of the number of required qubits in the MQO problem dimensions as the number of qubits is currently the main factor restricting applicability. We experimentally compare the performance of the quantum annealer against other MQO algorithms executed on a traditional computer. While the problem sizes that can be treated with 1000 qubits are still limited, we already find a narrow class of problem instances where the quantum annealer is three orders of magnitude faster than other approaches.

## Research 20: Spatial Data and Queries - 2

**Session Chair: Walid G. Aref (Purdue University)**

### **k-Nearest Neighbors on Road Networks: A Journey in Experimentation and In-Memory Implementation**

Tenindra Abeywickrama (Monash University); Muhammad Cheema (Monash University); David Taniar (Monash University)

A  $k$  nearest neighbor (kNN) query on road networks retrieves the  $k$  closest points of interest (POIs) by their network distances from a given location. Today, in the era of ubiquitous mobile computing, this is a highly pertinent query. While Euclidean distance has been used as a heuristic to search for the closest POIs by their road network distance, its efficacy has not been thoroughly investigated. The most recent methods have shown significant improvement in query performance. Earlier studies, which proposed disk-based indexes, were compared to the current state-of-the-art in main memory. However, recent studies have shown that main memory comparisons can be challenging and require careful adaptation. This paper presents an extensive experimental investigation in main memory to settle these and several other issues. We use efficient and fair memory-resident implementations of each method to reproduce past experiments and conduct additional comparisons for several overlooked evaluations. Notably we revisit a previously discarded technique (IER) showing that, through a simple improvement, it is often the best performing technique.

### **SKYPE: Top-k Spatial-keyword Publish/Subscribe Over Sliding Window**

Xiang Wang (University of New South Wales); Ying Zhang (University of Technology Sydney); Wenjie Zhang (University of New South Wales); Xuemin Lin (University of New South Wales); Zengfeng Huang (University of New South Wales)

As the prevalence of social media and GPS-enabled devices, a massive amount of geo-textual data has been generated in a stream fashion, leading to a variety of applications such as location-based recommendation and information dissemination. In this paper, we investigate a novel real-time top- $k$  monitoring problem over sliding window of streaming data; that is, we continuously maintain the top- $k$  most relevant geo-textual messages (e.g., geo-tagged tweets) for a large number of spatial-keyword subscriptions (e.g., registered users interested in local events) simultaneously. To provide the most recent information under controllable memory cost, sliding window model is employed on the streaming geo-textual data. To the best of our knowledge, this is the first work to study top- $k$  spatial-keyword publish/subscribe over sliding window. A novel system, called Skype (Top- $k$  Spatial-keyword Publish/Subscribe), is proposed in this paper. In Skype, to continuously maintain top- $k$  results for massive subscriptions, we devise a novel indexing structure upon subscriptions such that each incoming message can be immediately delivered on its arrival. Moreover, to reduce the expensive top- $k$  re-evaluation cost triggered by message expiration, we develop a novel cost-based  $k$ -skyband technique to reduce the number of re-evaluations in a cost-effective way. Extensive experiments verify the great efficiency and effectiveness of our proposed techniques.

### **Maximizing Bichromatic Reverse Spatial and Textual $k$ Nearest Neighbor Queries**

Farhana Choudhury (RMIT University); Shane Culpepper (RMIT University); Timos Sellis (RMIT University); Xin Cao (Queen's University, Belfast)

The problem of maximizing bichromatic reverse  $k$  nearest neighbor queries (BRkNN) has been extensively studied in spatial databases. In this work, we present a related query for spatial-textual databases that finds an optimal location, and a set of keywords that maximizes the size of bichromatic reverse spatial textual  $k$  nearest neighbors (MaxBRSTkNN). Such a query has many practical applications including social media advertisements where a limited number of relevant advertisements are displayed to each user. The problem is to find the location and the text contents to include in an advertisement so that it will be displayed to the maximum number of users. The increasing availability of spatial-textual collections allows us to answer these queries for both spatial proximity and textual similarity. This paper is the first to consider the MaxBRSTkNN query. We show that the problem is NP-hard and present both approximate and exact solutions.

### **Cache locality is not enough: High-Performance Nearest Neighbor Search with Product Quantization Fast Scan**

Fabien André (Technicolor); Anne-Marie Kermarrec (INRIA); Nicolas Le Scouarnec (Technicolor)

Nearest Neighbor (NN) search in high dimension is an important feature in many applications (e.g., image retrieval, multimedia databases). Product Quantization (PQ) is a widely used solution which offers high performance, i.e. low response time, while preserving a high accuracy. PQ represents high-dimensional vectors (e.g., images descriptors) by compact codes. Hence, very large databases can be stored in memory, allowing NN queries without resorting to slow I/O operations. PQ computes distances to neighbors using cache-resident lookup tables, thus its performance remains limited by (i) the many cache accesses that the algorithm requires, and (ii) the inability to leverage SIMD instructions available on modern CPUs. In this paper, we advocate that cache locality is not sufficient for efficiency. To address these issues, we design a novel algorithm, PQ Fast Scan, that transforms the cache-resident lookup tables into small tables, sized to fit SIMD registers. This transformation allows (i) in-register lookups in place of cache accesses and (ii) an efficient SIMD implementation. PQ Fast Scan has the exact same accuracy as PQ, while having 4 to 5 times lower response time (e.g., for 25 million vectors, scan time is reduced from 74ms to 13ms).

## Research 21: Distributed and Cloud Systems - 3

**Session Chair: Mike Carey (University of California at Irvine)**

### **WiSeDB: A Learning-based Workload Management Advisor for Cloud Databases**

Ryan Marcus (Brandeis University); Olga Papaemmanouil (Brandeis University)

Workload management for cloud databases deals with the tasks of resource provisioning, query placement, and query scheduling in a manner that meets the application's performance goals while minimizing the cost of using cloud resources. Existing solutions have approached these three challenges in isolation while aiming to optimize a single performance metric. In this paper, we introduce WiSeDB, a learning-based framework for generating holistic workload management solutions customized to application-defined performance goals and workload characteristics. Our approach relies on supervised learning to train cost-effective decision tree models for guiding query placement, scheduling, and resource provisioning decisions. Applications can use these models for both batch and online scheduling of incoming workloads. A unique feature of our system is that it can adapt its offline model to stricter/looser performance goals with minimal re-training. This allows us to present to the application alternative workload management strategies that address the typical performance vs. cost trade-off of cloud services. Experimental results show that our approach has very low training overhead while offering low cost strategies for a variety of performance metrics and workload characteristics.

### **Husky: Towards a More Efficient and Expressive Distributed Computing Framework**

Fan Yang (The Chinese University of Hong Kong); Jinfeng Li (The Chinese University of Hong Kong); James Cheng (The Chinese University of Hong Kong)

Finding efficient, expressive and yet intuitive programming models for data-parallel computing system is an important and open problem. Systems like Hadoop and Spark have been widely adopted for massive data processing, as coarse-grained primitives like map and reduce are succinct and easy to master. However, sometimes over-simplified API hinders programmers from more fine-grained control and designing more efficient algorithms. Developers may have to resort to sophisticated domain-specific languages (DSLs), or even low-level layers like MPI, but this raises development cost--learning many mutually exclusive systems prolongs the development schedule, and the use of low-level tools may result in bug-prone programming. This motivated us to start the Husky open-source project, which is an attempt to strike a better balance between high performance and low development cost. Husky is developed mainly for in-memory large scale data mining, and also serves as a general research platform for designing efficient distributed algorithms. We show that many existing frameworks can be easily implemented and bridged together inside Husky, and Husky is able to achieve similar or even better performance compared with domain-specific systems.

### **Titian: Data Provenance Support in Spark**

Matteo Interlandi (University of California at Los Angeles); Kshitij Shah (University of California at Los Angeles); Sai Tetali (University of California at Los Angeles); Muhammad Gulzar (University of California at Los Angeles); Seunghyun Yoo (University of California at Los Angeles); Miryung Kim (University of California at Los Angeles); Todd Millstein (University of California at Los Angeles); Tyson Condie (University of California at Los Angeles)

Debugging data processing logic in Data-Intensive Scalable Computing (DISC) systems is a difficult and time consuming effort. Today's DISC systems offer very little tooling for debugging programs, and as a result programmers spend countless hours collecting evidence (e.g., from log files) and performing trial and error debugging. To aid this effort, we built Titian, a library that enables data provenance-- tracking data through transformations--in Apache Spark. Data scientists using the Titian Spark extension will be able to quickly identify the input data at the root cause of a potential bug or outlier result. Titian is built directly into the Spark platform and offers data provenance support at interactive speeds--orders-of-magnitude faster than alternative solutions--while minimally impacting Spark job performance; observed overheads for capturing data lineage rarely exceed 30% above the baseline job execution time.

## Research 22: Entity Matching - 1

**Session Chair: Ioana Manolescu (INRIA)**

### **Online Entity Resolution Using an Oracle**

Donatella Firmani (University of Rome "Tor Vergata"); Barna Saha (University of Massachusetts Amherst); Divesh Srivastava (AT&T Labs Research)

Entity resolution (ER) is the task of identifying all records in a database that refer to the same underlying entity. This is an expensive task, and can take a significant amount of money and time; the end-user may want to take decisions during the process, rather than waiting for the task to be completed. We formalize an online version of the entity resolution task, and use an oracle which correctly labels matching and non-matching pairs through queries. In this setting, we design algorithms that seek to maximize progressive recall, and develop a novel analysis framework for prior proposals on entity resolution with an oracle, beyond their worst case guarantees. Finally, we provide both theoretical and experimental analysis of the proposed algorithms.

### **BLAST: A Loosely Schema-aware Meta-blocking Approach for Entity Resolution**

Giovanni Simonini (Università di Modena e Reggio Emilia); Sonia Bergamaschi (Università di Modena e Reggio Emilia); H. V. Jagadish (University of Michigan)

To identify records that refer to the same entity is a fundamental step for data integration. Blocking techniques are typically employed to reduce the complexity of this task by partitioning records into blocks and limiting the comparison to those records co-occurring in them. Generally, to deal with highly heterogeneous and noisy data (e.g. semi-structured data of the Web), these techniques rely on redundancy to reduce the chance of missing matches. Meta-blocking is the task of restructuring blocks generated by redundancy-based blocking techniques, removing superfluous comparisons. Existing meta-blocking approaches rely exclusively on schema-agnostic features. In this paper, we demonstrate how "loose" schema information (i.e., statistics collected directly from the data) can be exploited to enhance the quality of the blocks in a holistic loosely schema-aware (meta-)blocking approach. We call it Blast (Blocking with Loosely-Aware Schema Techniques). We show how Blast can automatically extract this loose information by adopting a LSH-based step for efficiently scaling to large datasets. We experimentally demonstrate, on five real world datasets, how Blast outperforms the state-of-the-art unsupervised meta-blocking approaches, and, in many cases, also the supervised one.

### **QuERy: A Framework for Integrating Entity Resolution with Query Processing**

Hotham Altwajry (University of California at Irvine); Sharad Mehrotra (University of California at Irvine); Dmitri Kalashnikov (AT&T Labs Research)

This paper explores an analysis-aware data cleaning architecture for a large class of SPJ SQL queries. In particular, we propose QuERy, a novel framework for integrating entity resolution (ER) with query processing. The aim of QuERy is to correctly and efficiently answer complex queries issued on top of dirty data. The comprehensive empirical evaluation of the proposed solution demonstrates its significant advantage in terms of efficiency over the traditional techniques for the given problem settings.

### **Comparative Analysis of Approximate Blocking Techniques for Entity Resolution**

George Papadakis (University of Athens); Jonathan Svirsky (Technion); Avigdor Gal (Technion); Themis Palpanas (Paris Descartes University)

Entity Resolution is a core task for merging data collections. Due to its quadratic complexity, it typically scales to large volumes of data through blocking: similar entities are clustered into blocks and pair-wise comparisons are executed only between co-occurring entities, at the cost of some missed matches. There are numerous blocking methods, and the aim of this work is to offer a comprehensive empirical survey, extending the dimensions of comparison beyond what is commonly available in the literature. We consider 17 state-of-the-art blocking methods and use 6 popular real datasets to examine the robustness of their internal configurations and their relative balance between effectiveness and time efficiency. We also investigate their scalability over a corpus of 7 established synthetic datasets that range from 10,000 to 2 million entities.

## **Research 23: Query Execution - 3**

### **Session Chair: Peter Boncz (CWI)**

#### **An Empirical Evaluation of Set Similarity Join Techniques**

Willi Mann (University of Salzburg); Nikolaus Augsten (University of Salzburg); Panagiotis Bours (Aarhus University)

Set similarity joins compute all pairs of similar sets from two collections of sets. We conduct extensive experiments on seven state-of-the-art algorithms for set similarity joins. These algorithms adopt a filter-verification approach. Our analysis shows that verification has not received enough attention in previous works. In practice, efficient verification inspects only a small, constant number of set elements and is faster than some of the more sophisticated filter techniques. Although we can identify three winners, we find that most algorithms show very similar performance. The key technique is the prefix filter, and AllPairs, the first algorithm adopting this technique is still a relevant competitor. We repeat experiments from previous work and discuss diverging results. All our claims are supported by a detailed analysis of the factors that determine the overall runtime.

#### **Streaming Similarity Self-Join**

Gianmarco De Francisci Morales (Qatar Computing Research Institute); Aristides Gionis (Aalto University)

We introduce and study the problem of computing the similarity self-join in a streaming context (SSSJ), where the input is an unbounded stream of items arriving continuously. The goal is to find all pairs of items in the stream whose similarity is greater than a given threshold. The simplest formulation of the problem requires unbounded memory, and thus, it is intractable. To make the problem feasible, we introduce the notion of time-dependent similarity: the similarity of two items decreases with the difference in their arrival time. By leveraging the properties of this time-dependent similarity function, we design two algorithmic frameworks to solve the SSSJ problem. The first one, MiniBatch (MB), uses existing index-based filtering techniques for the static version of the problem, and combines them in a pipeline. The second framework, Streaming (STR), adds time filtering to the existing indexes, and integrates new time-based bounds deeply in the working of the algorithms. We also introduce a new indexing technique (L2), which is based on an existing state-of-the-art indexing technique (L2AP), but is optimized for the streaming case. Extensive experiments show that the STR algorithm, when instantiated with the L2 index, is the most scalable option across a wide array of datasets and parameters.

### **An Efficient Partition Based Method for Exact Set Similarity Joins**

Dong Deng (Tsinghua University); Guoliang Li (Tsinghua University); He Wen (Tsinghua University); Jianhua Feng (Tsinghua University)

We study the exact set similarity join problem, which, given two collections of sets, finds out all the similar set pairs from the collections. Existing methods generally utilize the prefix filter based framework. They generate a prefix for each set and prune all the pairs whose prefixes are disjoint. However the pruning power is limited, because if two dissimilar sets share a common element in their prefixes, they cannot be pruned. To address this problem, we propose a partition-based framework. We design a partition scheme to partition the sets into several subsets and guarantee that two sets are similar only if they share a common subset. To improve the pruning power, we propose a mixture of the subsets and their 1-deletion neighborhoods (the subset of a set by eliminating one element). As there are multiple allocation strategies to generate the mixture, we evaluate different allocations and design a dynamic-programming algorithm to select the optimal one. However the time complexity of generating the optimal one is  $O(s^3)$  for a set with size  $s$ . To speed up the allocation selection, we develop a greedy algorithm with an approximation ratio of 2. To further reduce the complexity, we design an adaptive grouping mechanism, and the two techniques can reduce the complexity to  $O(s \log s)$ . Experimental results on three real-world datasets show our method achieves high performance and outperforms state-of-the-art methods by 2-5 times.

### **Fast Queries Over Heterogeneous Data Through Engine Customization**

Manos Karpathiotakis (EPFL); Ioannis Alagiannis (EPFL); Anastasia Ailamaki (EPFL)

Industry and academia are continuously becoming more data-driven and data-intensive, relying on the analysis of a wide variety of heterogeneous datasets to gain insights. The different data models and formats pose a significant challenge on performing analysis over a combination of diverse datasets. Serving all queries using a single, general-purpose query engine is slow. On the other hand, using a specialized engine for each heterogeneous dataset increases complexity: queries touching a combination of datasets require an integration layer over the different engines. This paper presents a system design that natively supports heterogeneous data formats and also minimizes query execution times. For multi-format support, the design uses an expressive query algebra which enables operations over various data models. For minimal execution times, it uses a code generation mechanism to mimic the system and storage most appropriate to answer a query fast. We validate our design by building Proteus, a query engine which natively supports queries over CSV, JSON, and relational binary data, and which specializes itself to each query, dataset, and workload via code generation. Proteus outperforms state-of-the-art open-source and commercial systems on both synthetic and real-world workloads without being tied to a single data model or format, all while exposing users to a single query interface.

## **Research 24: Social Networks and Crowdsourcing**

### **Session Chair: Sihem Amer-Yahia (CNRS)**

#### **Crowdsourced Top-k Algorithms: An Experimental Evaluation**

Xiaohang Zhang (Tsinghua University); Guoliang Li (Tsinghua University); Jianhua Feng (Tsinghua University)

Crowdsourced top-k computation has attracted significant attention recently, thanks to emerging crowdsourcing platforms, e.g., Amazon Mechanical Turk and CrowdFlower. Crowdsourced top-k algorithms ask the crowd to compare the objects and infer the top-k objects based on the crowdsourced comparison results. The crowd may return incorrect answers, but traditional top-k algorithms cannot tolerate the errors from the crowd. To address this problem, the database and machine-learning communities have independently studied the crowdsourced top-k problem. The database community proposes the heuristic-based solutions while the machine-learning community proposes the learning-based methods (e.g., maximum likelihood estimation). However, these two types of techniques have not been compared systematically under the same experimental framework. Thus it is rather difficult for a practitioner to decide which algorithm should be adopted. Furthermore, the experimental evaluation of existing studies has several weaknesses. Some methods assume the crowd returns high-quality results and some algorithms are only tested on simulated experiments. To alleviate these limitations, in this paper we present a comprehensive comparison of crowdsourced top-k algorithms. Using various synthetic and real datasets, we evaluate each algorithm in terms of result quality and efficiency on real crowdsourcing

platforms. We reveal the characteristics of different techniques and provide guidelines on selecting appropriate algorithms for various scenarios.

### **CLAMShell: Speeding up Crowds for Low-latency Data Labeling**

Daniel Haas (University of California at Berkeley); Jiannan Wang (Simon Fraser University); Eugene Wu (Columbia University); Michael Franklin (University of California at Berkeley)

Data labeling is a necessary but often slow process that impedes the development of interactive systems for modern data analysis. Despite rising demand for manual data labeling, there is a surprising lack of work addressing its high and unpredictable latency. In this paper, we introduce `\sys`, a system that speeds up crowds in order to achieve consistently low-latency data labeling. We offer a taxonomy of the sources of labeling latency and study several large crowd-sourced labeling deployments to understand their empirical latency profiles. Driven by these insights, we comprehensively tackle each source of latency, both by developing novel techniques such as straggler mitigation and pool maintenance and by optimizing existing methods such as crowd retainer pools and active learning. We evaluate `\sys` in simulation and on live workers on Amazon's Mechanical Turk, demonstrating that our techniques can provide an order of magnitude speedup and variance reduction over existing crowdsourced labeling strategies.

### **Dynamic Influence Analysis in Evolving Networks**

Naoto Ohsaka (The University of Tokyo); Takuya Akiba (National Institute of Informatics); Yuichi Yoshida (National Institute of Informatics); Ken-ichi Kawarabayashi (National Institute of Informatics)

We propose the first real-time fully-dynamic index data structure designed for influence analysis on evolving networks. With this aim, we carefully redesign the data structure of the state-of-the-art sketching method introduced by Borgs et al., and construct corresponding update algorithms. Using this index, we present algorithms for two kinds of queries, influence estimation and influence maximization, which are strongly motivated by practical applications, such as viral marketing. We provide a thorough theoretical analysis, which guarantees the non-degeneracy of the solution accuracy after an arbitrary number of updates. Furthermore, we introduce a reachability-tree-based technique and a skipping method, which greatly reduce the time consumption required for edge/vertex deletions and vertex additions, respectively, and counter-based random number generators, which improve the space efficiency. Experimental evaluations using real dynamic networks with tens of millions of edges demonstrate the efficiency, scalability, and accuracy of our proposed indexing scheme. Specifically, it can reflect a graph modification within a time of several orders of magnitude smaller than that required to reconstruct an index from scratch, estimate the influence spread of a vertex set accurately within a millisecond, and select highly influential vertices at least ten times faster than state-of-the-art static algorithms.

### **From Competition to Complementarity: Comparative Influence Diffusion and Maximization**

Wei Lu (University of British Columbia); Wei Chen (Microsoft Research); Laks Lakshmanan (University of British Columbia)

Influence maximization is a well-studied problem that asks for a small set of influential users from a social network, such that by targeting them as early adopters, the expected total adoption through influence cascades over the network is maximized. However, almost all prior work focuses on cascades of a single propagating entity or purely-competitive entities. In this work, we propose the Comparative Independent Cascade (Com-IC) model that covers the full spectrum of entity interactions from competition to complementarity. In Com-IC, users' adoption decisions depend not only on edge-level information propagation, but also on a node-level automaton whose behavior is governed by a set of model parameters, enabling our model to capture not only competition, but also complementarity, to any possible degree. We study two natural optimization problems, Self Influence Maximization and Complementary Influence Maximization, in a novel setting with complementary entities. Both problems are NP-hard, and we devise efficient and effective approximation algorithms via non-trivial techniques based on reverse-reachable sets and a novel "sandwich approximation" strategy. The applicability of both techniques extends beyond our model and problems. Our experiments show that the proposed algorithms consistently outperform intuitive baselines in four real-world social networks, often by a significant margin. In addition, we learn model parameters from real user action logs.

## **Research 25: Graph Processing - 3**

### **Session Chair: Mohamed Mokbel (University of Minnesota)**

### **Data-driven Visual Graph Query Interface Construction and Maintenance: Challenges and Opportunities**

Sourav Bhowmick (Nanyang Technological University); Byron Choi (Hong Kong Baptist University); Curtis Dyreson (Utah State University)

Visual query interfaces make it easy for scientists and other non-expert users to query a data collection. Heretofore, visual query interfaces have been statically-constructed, independent of the data. In this paper we outline a vision of a different kind of interface, one that is built (in part) from the data. In our data-driven approach, the visual interface is dynamically constructed and maintained. A data-driven approach has many benefits such as reducing the cost in constructing and maintaining an interface, superior support for query formulation, and increased portability of the interface. We focus on graph databases, but our approach is applicable to several other kinds of databases such as JSON and XML.



### **The shortest path is not always a straight line**

Vasiliki Kalavri (KTH Royal Institute of Technology); Tiago Simas (Telefonica Research); Dionysios Logothetis (Facebook)

In this paper, we leverage the concept of the metric backbone to improve the efficiency of large-scale graph analytics. The metric backbone is the minimum subgraph that preserves the shortest paths of a weighted graph. We use the metric backbone in place of the original graph to compute various graph metrics exactly or with good approximation. By computing on a smaller graph, we improve the performance of graph analytics applications on two different systems, a batch graph processing system and a graph database. Further, we provide an algorithm for the computation of the metric backbone on large graphs. While one can compute the metric backbone by solving the all-pairs-shortest-paths problem, this approach incurs prohibitive time and space complexity for big graphs. Instead, we propose a heuristic that makes computing the metric backbone practical even for large graphs. Additionally, we analyze several real datasets of different sizes and domains and we show that we can approximate the metric backbone by removing only first-order semi-metric edges; edges for which a shorter two-hop path exists. We provide a distributed implementation of our algorithm and apply it in large scale scenarios. We evaluate our algorithm using a variety of real graphs, including a Facebook social network subgraph of 50 billion edges. We measure the impact of using the metric backbone on runtime performance in two graph management systems. We achieve query speedups of up to 6.7x in the Neo4j commercial graph database and job speedups of up to 6x in the Giraph graph processing system.

### **New Lower and Upper Bounds for Shortest Distance Queries on Terrains**

Manohar Kaul (Technische Universität Berlin); Raymond Chi-Wing Wong (The Hong Kong University of Science and Technology); Christian Jensen (Aalborg University)

The increasing availability of massive and accurate laser data enables the processing of spatial queries on terrains. As shortest-path computation, an integral element of query processing, is inherently expensive on terrains, a key approach to enabling efficient query processing is to reduce the need for exact shortest-path computation in query processing. We develop new lower and upper bounds on terrain shortest distances that are provably tighter than any existing bounds. Unlike existing bounds, the new bounds do not rely on the quality of the triangulation. We show how use of the new bounds speeds up query processing by reducing the need for exact distance computations. Speedups of nearly an order of magnitude are demonstrated empirically for well-known spatial queries.

### **On Measuring the Lattice of Commonalities Among Several Linked Datasets**

Michalis Mountantonakis (FORTH); Yannis Tzitzikas (FORTH)

A big number of datasets has been published according to the principles of Linked Data and this number keeps increasing. Although the ultimate objective is linking and integration, it is not currently evident how connected the current LOD cloud is. The classical measurements and visualizations of the LOD cloud count and show only the number of links that exist between pairs of datasets. Measurements (and indexes) that involve more than two datasets are not available although they would be useful in several tasks, e.g. (a) for obtaining complete information about one particular URI (or set of URIs) enriched (b) for aiding dataset discovery and dataset selection, (c) for assessing the connectivity between any set of datasets for quality checking and for monitoring their evolution over time, (d) for constructing visualizations that provide more informative overviews. Since it would be prohibitively expensive to perform all these measurements in a naive way, in this paper we introduce various indexes (and their construction algorithms) that can speedup such tasks. In brief, we introduce (i) a namespace-based prefix index, (ii) a sameAs catalog for computing the symmetric and transitive closure of the owl:sameAs relationships encountered in the datasets, (iii) a semantics-aware element index (that exploits the aforementioned indexes), and (iv) finally two lattice-based incremental algorithms for speeding up the computation of the intersection URIs of any set of datasets. We discuss the speedup obtained by the introduced indexes and algorithms through comparative results and finally we report measurements about connectivity of the LOD cloud that have never been carried out in the past.

## **Research 26: Data and Query Models - 2**

### **Session Chair: Tim Kraska (Brown University)**

#### **Interactive Browsing and Navigation in Relational Databases**

Minsuk Kahng (Georgia Institute of Technology); Shamkant Navathe (Georgia Institute of Technology); John Stasko (Georgia Institute of Technology); Duen Horng Chau (Georgia Institute of Technology)

Although database researchers have devoted considerable effort to helping database users formulate queries, many users still find it challenging to specify queries that involve joining tables. To help users construct join queries for exploring relational databases, we propose Etable, a novel presentation data model that provides users with a presentation-level interactive view. This view compactly presents one-to-many and many-to-many relationships within a single enriched table by allowing a cell to contain a set of entity references. Users can directly interact with the table and the entity references to incrementally construct complex queries. To enable users to explore data on a conceptual entity-relationship level, we also introduce a graph-based model, called typed graph model that provides an abstraction of relational databases. In a user study, participants performed a range of database querying tasks faster with Etable than with a commercial graphical query builder. Subjective feedback about Etable was also positive. All participants found that Etable was easier to learn and helpful for exploring databases.

## **ATHENA: An Ontology-Driven System for Natural Language Querying over Relational Data Stores**

Diptikalyan Saha (IBM Research - India); Avriilia Floratou (IBM Research - Almaden); Karthik Sankaranarayanan (IBM Research - India); Umar Farooq Minhas (IBM Research - Almaden); Ashish Mittal (IBM Research - India); Fatma Ozcan (IBM Research - Almaden)

In this paper, we present ATHENA, an ontology-driven system for natural language querying of complex relational databases. Natural language interfaces to databases enable users easy access to data, without the need to learn a complex query language, such as SQL. ATHENA uses domain specific ontologies, which describe the semantic entities, and their relationships in a domain. We propose a unique two-stage approach, where the input natural language query (NLQ) is first translated into an intermediate query language over the ontology, called OQL, and subsequently translated into SQL. Our two-stage approach allows us to decouple the physical layout of the data in the relational store from the semantics of the query, providing physical independence. Moreover, ontologies provide richer semantic information, such as inheritance and membership relations, that are lost in a relational schema. By reasoning over the ontologies, our NLQ engine is able to accurately capture the user intent. We study the effectiveness of our approach using three different workloads on top of geographical (GEO), academic (MAS) and financial (FIN) data. ATHENA achieves 100% precision on the GEO and MAS workloads, and 99% precision over the FIN workload which operates on a complex financial ontology. Moreover, ATHENA attains 87.2%, 88.3%, and 88.9% recall on the GEO, MAS, and FIN workloads, respectively.

## **BlinkFill: Semi-supervised Programming By Example for Syntactic String Transformations**

Rishabh Singh (Microsoft Research)

The recent Programming By Example (PBE) techniques such as FlashFill have shown great promise for enabling end-users to perform data transformation tasks using input-output examples. Since examples are inherently an under-specification, there are typically a large number of hypotheses conforming to the examples, and the PBE techniques suffer from scalability issues for finding the intended program amongst the large space. We present a semi-supervised learning technique to significantly reduce this ambiguity by using the logical information present in the input data to guide the synthesis algorithm. We develop a data structure InputDataGraph to succinctly represent a large set of logical patterns that are shared across the input data, and use this graph to efficiently learn substring expressions in a new PBE system BlinkFill. We evaluate BlinkFill on 207 real-world benchmarks and show that BlinkFill is significantly faster (on average 41x) and requires fewer input-output examples (1.27 vs 1.53) to learn the desired transformations in comparison to FlashFill.

## **Query from Examples: An Iterative, Data-Driven Approach to Query Construction**

Hao Li (National University of Singapore); Chee-Yong Chan (National University of Singapore); David Maier (Portland State University)

In this paper, we propose a new approach, called Query from Examples (QFE), to help non-expert database users construct SQL queries. Our approach, which is designed for users who might be unfamiliar with SQL, only requires that the user is able to determine whether a given output table is the result of his or her intended query on a given input database. To kick-start the construction of a target query  $Q$ , the user first provides a pair of inputs: a sample database  $D$  and an output table  $R$  which is the result of  $Q$  on  $D$ . As there will be many candidate queries that transform  $D$  to  $R$ , QFE winnows this collection by presenting the user with new database-result pairs that distinguish these candidates. Unlike previous approaches that use synthetic data for such pairs, QFE strives to make these distinguishing pairs as close to the original  $(D, R)$  pair as possible. By doing so, it seeks to minimize the effort needed by a user to determine if a new database-result pair is consistent with his or her desired query. We demonstrate the effectiveness and efficiency of our approach using real datasets from SQLShare, a cloud-based platform designed to help scientists utilize RDBMS technology for data analysis.

## **Research 27: Data and Query Models - 3**

### **Session Chair: Sudeepa Roy (Duke University)**

#### **Exploiting Equality Generating Dependencies in Checking Chase Termination**

Marco Calautti (University of Calabria); Sergio Greco (University of Calabria); Cristian Molinaro (University of Calabria); Irina Trubitsyna (University of Calabria)

The chase is a well-known algorithm with a wide range of applications in data exchange, data cleaning, data integration, query optimization, and ontological reasoning. Since the chase evaluation might not terminate and it is undecidable whether it terminates, the problem of defining (decidable) sufficient conditions ensuring termination has received a great deal of interest in recent years. In this regard, several termination criteria have been proposed. One of the main weaknesses of current approaches is the limited analysis they perform on equality generating dependencies (EGDs). In this paper, we propose sufficient conditions ensuring that a set of dependencies has at least one terminating chase sequence. We propose novel criteria which are able to perform a more accurate analysis of EGDs. Specifically, we propose a new stratification criterion and an adornment algorithm. The latter can both be used as a termination criterion and be combined with current techniques to make them more effective, in that strictly more sets of dependencies are identified. Our techniques identify sets of dependencies that are not recognized by any of the current criteria.

## **The Complexity of Resilience and Responsibility for Self-Join-Free Conjunctive Queries**

Cibele Freire (University of Massachusetts Amherst); Wolfgang Gatterbauer (Carnegie Mellon University); Neil Immerman (University of Massachusetts Amherst); Alexandra Meliou (University of Massachusetts Amherst)

Several research thrusts in the area of data management have focused on understanding how changes in the data affect the output of a view or standing query. Example applications are explaining query results, propagating updates through views, and anonymizing datasets. An important aspect of this analysis is the problem of deleting a minimum number of tuples from the input tables to make a given Boolean query false, which we refer to as “the resilience of a query”. In this paper, we study the complexity of resilience for self-join-free conjunctive queries with arbitrary functional dependencies. The cornerstone of our work is the novel concept of triads, a simple structural property of a query that leads to the several dichotomy results we show in this paper. The concepts of triads and resilience bridge the connections between the problems of deletion propagation and causal responsibility, and allow us to substantially advance the known complexity results in these topics. Specifically, we show a dichotomy for the complexity of resilience, which identifies previously unknown tractable families for deletion propagation with source side-effects, and we extend this result to account for functional dependencies. Further, we identify a mistake in a previous dichotomy for causal responsibility, and offer a revised characterization based purely on the structural form of the query (presence or absence of triads). Finally, we extend the dichotomy for causal responsibility in two ways: (a) we account for functional dependencies in the input tables, and (b) we compute responsibility for sets of tuples specified via wildcards.

## **A Shifting Bloom Filter Framework for Set Queries**

Tong Yang (Peking University); Alex X. Liu (Michigan State University); Muhammad Shahzad (North Carolina State University); Yuankun Zhong (Nanjing University); Qiaobin Fu (Boston University); Zi Li (Nanjing University); Gaogang Xie (ICT, CAS); Xiaoming Li (Peking University)

Set queries are fundamental operations in computer systems and applications. This paper addresses the fundamental problem of designing a probabilistic data structure that can quickly process set queries using a small amount of memory. We propose a Shifting Bloom Filter (ShBF) framework for representing and querying sets. We demonstrate the effectiveness of ShBF using three types of popular set queries: membership, association, and multiplicity queries. The key novelty of ShBF is on encoding the auxiliary information of a set element in a location offset. In contrast, prior BF based set data structures allocate additional memory to store auxiliary information. To evaluate ShBF in comparison with prior art, we conducted experiments using real-world network traces. Results show that ShBF significantly advances the state-of-the-art on all three types of set queries.

## **Decibel: The Relational Dataset Branching System**

Michael Maddox (MIT); David Goehring (MIT); Aaron Elmore (University of Chicago); Samuel Madden (MIT); Aditya Parameswaran (University of Illinois at Urbana-Champaign); Amol Deshpande (University of Maryland)

As scientific endeavors and data analysis becomes increasingly collaborative, there is a need for data management systems that natively support the versioning or branching of datasets to enable concurrent analysis, cleaning, integration, manipulation, or curation of data across teams of individuals. Common practice for sharing and collaborating on datasets involves creating or storing multiple copies of the dataset, one for each stage of analysis, with no provenance information tracking the relationships between these datasets. This results not only in wasted storage, but also makes it challenging to track and integrate modifications made by different users to the same dataset. In this paper, we introduce the Relational Dataset Branching System, Decibel, a new relational storage system with built-in version control designed to address these shortcomings. We present our initial design for Decibel and provide a thorough evaluation of three versioned storage engine designs that focus on efficient query processing with minimal storage overhead. We also develop an exhaustive benchmark to enable the rigorous testing of these and future versioned storage engine designs.

## **Research 28: Entity Matching - 2**

### **Session Chair: Themis Palpanas (Paris Descartes University)**

#### **Magellan: Toward Building Entity Matching Management Systems**

Pradap Konda (University of Wisconsin Madison); Sanjib Das (University of Wisconsin Madison); Paul Suganthan G. C. (University of Wisconsin Madison); Anhai Doan (University of Wisconsin Madison); Adel Ardanal (University of Wisconsin Madison); Jeff Ballard (University of Wisconsin Madison); Han Li (University of Wisconsin Madison); Fatemah Panahi (University of Wisconsin Madison); Haojun Zhang (University of Wisconsin Madison); Jeffrey Naughton (University of Wisconsin Madison); Shishir Prasad (Instacart); Ganesh Krishnan (WalmartLabs); Rohit Deep (WalmartLabs); Vijay Raghavendra (WalmartLabs)

Entity matching (EM) has been an important data integration problem and will become even more important in the age of Big Data. However, most EM works have focused only on developing matching algorithms. Going forward we argue that significantly more attention should be devoted to building EM systems. We discuss four major limitations of current EM systems, then describe Magellan, a new kind of EM system. Compared to current EM systems, Magellan is novel in several important aspects. (1) It provides a how-to guide that tells users what to do in each EM scenario, step by step. (2) It provides tools to help users do these steps. The tools seek to cover the entire EM pipeline, not just matching and blocking as current EM systems do. (3) Magellan is built on top of the data analysis and Big Data stacks in Python, allowing it to borrow a rich set of capabilities in data cleaning, IE, visualization, learning, etc. (4) In contrast

to current “closed world” EM systems, Magellan is a new kind of system that we call “open world” systems, but building such systems raises significant challenges. We present extensive experiments with 44 students and many real users at various organizations that show the promise of Magellan.

### **Distributed Data Deduplication**

Xu Chu (University of Waterloo); Ihab Ilyas (University of Waterloo); Paraschos Koutris (University of Wisconsin Madison)

Data deduplication refers to the process of identifying tuples in a relation that refer to the same real world entity. The complexity of the problem is inherently quadratic with respect to the number of tuples, since a similarity value must be computed for every pair of tuples. In order to avoid comparing tuple pairs that are obviously non-duplicates, matching algorithms use blocking techniques that divide the tuples into blocks and compare only tuples within the same block. However, even with the use of blocking, data deduplication remains a costly problem for large datasets. In this paper, we show how to further speed up data deduplication by leveraging parallelism in a shared-nothing computing environment. Our main contribution is a distribution strategy, called `\disdedup`, that minimizes the maximum workload across all worker nodes and provides strong theoretical guarantees. We demonstrate the effectiveness of our proposed strategy by performing extensive experiments on both synthetic datasets with varying block size distributions, as well as real world datasets.

### **The iBench Integration Metadata Generator**

Patricia Arocena (University of Toronto); Boris Glavic (Illinois Institute of Technology); Radu Ciucanu (University of Oxford); René Miller (University of Toronto)

Given the maturity of the data integration field it is surprising that rigorous empirical evaluations of research ideas are so scarce. We identify a major roadblock for empirical work the lack of comprehensive metadata generators that can be used to create benchmarks for different integration tasks. This makes it difficult to compare integration solutions, understand their generality, and understand their performance. We present iBench, the first metadata generator that can be used to evaluate a wide-range of integration tasks (data exchange, mapping creation, mapping composition, schema evolution, among many others). iBench permits control over the size and characteristics of the metadata it generates (schemas, constraints, and mappings). Our evaluation demonstrates that iBench can efficiently generate very large, complex, yet realistic scenarios with different characteristics. We also present an evaluation of three mapping creation systems using iBench and show that the intricate control that iBench provides over metadata scenarios can reveal new and important empirical insights. iBench is an open-source, extensible tool that we are providing to the community. We believe it will raise the bar for empirical evaluation and comparison of data integration systems.

### **Schema-agnostic vs Schema-based Configurations for Blocking Methods on Homogeneous Data**

George Papadakis (University of Athens); George Alexiou (IMIS, “Athena” Research Center); George Papastefanatos (IMIS, “Athena” Research Center); Georgia Koutrika (Hewlett-Packard Labs)

Entity Resolution constitutes a core task for data integration that, due to its quadratic complexity, typically scales to large datasets through blocking methods. These can be configured in two ways. The schema-based configuration relies on schema information in order to select signatures of high distinctiveness and low noise, while the schema-agnostic one treats every token from all attribute values as a signature. The latter approach has significant potential, as it requires no fine-tuning by human experts and it applies to heterogeneous data. Yet, there is no systematic study on its relative performance with respect to the schema-based configuration. This work covers this gap by comparing analytically the two configurations in terms of effectiveness, time efficiency and scalability. We apply them to 9 established blocking methods and to 11 benchmarks of structured data. We provide valuable insights into the internal functionality of the blocking methods with the help of a novel taxonomy. Our studies reveal that the schema-agnostic configuration offers unsupervised and robust definition of blocking keys under versatile settings, trading a higher computational cost for a consistently higher recall than the schema-based one. It also enables the use of state-of-the-art blocking methods without schema knowledge.

## **Research 29: Community Search and Mining**

### **Session Chair: Tamer Öszu (University of Waterloo)**

#### **Effective Community Search for Large Attributed Graphs**

Yixiang Fang (The University of Hong Kong); Reynold C.K. Cheng (The University of Hong Kong); Siqiang Luo (The University of Hong Kong); Jiafeng Hu (The University of Hong Kong)

Given a graph  $G$  and a vertex  $q \in G$ , the `\{it community search\}` query returns a subgraph of  $G$  that contains vertices related to  $q$ . Communities, which are prevalent in `\{it attributed graphs\}` such as social networks and knowledge bases, can be used in emerging applications such as product advertisement and setting up of social events. In this paper, we investigate the `\{it attributed community query\}` (or `\{it ACD\}`), which returns an `\{it attributed community\}` (AC) for an `\{it attributed graph\}`. The AC is a subgraph of  $G$ , which satisfies both `\{it structure cohesiveness\}` (i.e., its vertices are tightly connected) and `\{it keyword cohesiveness\}` (i.e., its vertices share common keywords). The AC enables a better understanding of how and why a community is formed (e.g., members of an AC

have a common interest in music, because they all have the same keyword "music"). An AC can be "personalized"; for example, an ACQ user may specify that an AC returned should be related to the keywords like "sports". To enable efficient AC search, we develop the CL-tree structure and three algorithms based on it. We evaluate our solutions on four large graphs, namely Flickr, DBLP, Tencent, and DBpedia. Our results show that ACs are more effective and efficient than existing community retrieval approaches. Moreover, an AC contains more precise and personalized information than existing community search and detection methods.

### Approximate Closest Community Search in Networks

Xin Huang (University of British Columbia); Laks Lakshmanan (University of British Columbia); Jeffrey Xu Yu (The Chinese University of Hong Kong); Hong Cheng (The Chinese University of Hong Kong)

Recently, there has been significant interest in the study of the community search problem in social and information networks: given one or more query nodes, find densely connected communities containing the query nodes. However, most existing studies do not address the "free rider" issue, that is, nodes far away from query nodes and irrelevant to them are included in the detected community. Some state-of-the-art models have attempted to address this issue, but not only are their formulated problems NP-hard, they do not admit any approximations without restrictive assumptions, which may not always hold in practice. In this paper, given an undirected graph  $G$  and a set of query nodes  $Q$ , we study community search using the  $k$ -truss based community model. We formulate our problem of finding a closest truss community (CTC), as finding a connected  $k$ -truss subgraph with the largest  $k$  that contains  $Q$ , and has the minimum diameter among such subgraphs. We prove this problem is NP-hard. Furthermore, it is NP-hard to approximate the problem within a factor  $(2-\epsilon)$ , for any  $\epsilon > 0$ . However, we develop a greedy algorithmic framework, which first finds a CTC  $k$ -truss graph with the largest  $k$  containing  $Q$ , and then iteratively removes the furthest nodes from  $Q$ , from the graph. The method achieves 2-approximation to the optimal solution. To further improve the efficiency, we make use of a compact truss index and develop efficient algorithms for  $k$ -truss identification and maintenance as nodes get eliminated. In addition, using bulk deletion optimization and local exploration strategies, we propose two more efficient algorithms. One of them trades some approximation quality for efficiency while the other is a very efficient heuristic. Extensive experiments on 6 real-world networks show the effectiveness and efficiency of our community model and search algorithms.

### Ego-net Community Mining Applied to Friend Suggestion

Alessandro Epasto (Brown University); Silvio Lattanzi (Google); Vahab Mirrokni (Google); Ismail Sebe (Google); Ahmed Taei (Google); Sunita Verma (Google)

In this paper, we present a study of the community structure of ego-networks --- the graphs representing the connections among the neighbors of a node --- for several online social networks. Toward this goal, we design a new technique to efficiently build and cluster all the ego-nets of a graph in parallel (note that even just building the ego-nets efficiently is challenging on large networks). Our experimental findings are quite compelling: at a microscopic level it is easy to detect high quality communities. Leveraging on this fact we, then, develop new features for friend suggestion based on co-occurrences of two nodes in different ego-nets' communities. Our new features can be computed efficiently on very large scale graphs by just analyzing the neighborhood of each node. Furthermore, we prove formally on a stylized model, and by experimental analysis that this new similarity measure outperforms the classic local features employed for friend suggestions.

### Tracking the Conductance of Rapidly Evolving Topic-Subgraphs

Sainyam Gahotra (Xerox Research Centre India); Amitabha Bagchi (IIT Delhi); Srikanta Bedathur (IBM Research - India); Maya Ramanath (IIT Delhi); Vidit Jain (American Express Big Data Labs)

Monitoring the formation and evolution of communities in large online social networks such as Twitter is an important problem that has generated considerable interest in both industry and academia. Fundamentally, the problem can be cast as studying evolving sub-graphs (each subgraph corresponding to a topical community) on an underlying social graph - with users as nodes and the connection between them as edges. A key metric of interest in this setting is tracking the changes to the conductance of subgraphs induced by edge activations. This metric quantifies how well or poorly connected a subgraph is to the rest of the graph relative to its internal connections. Conductance has been demonstrated to be of great use in many applications, such as identifying bursty topics, tracking the spread of rumors, and so on. However, tracking this simple metric presents a considerable scalability challenge - the underlying social network is large, the number of communities that are active at any moment is large, the rate at which these communities evolve is high, and moreover, we need to track conductance in real-time. We address these challenges in this paper. We propose an in-memory approximation called BloomGraphs to store and update these (possibly overlapping) evolving subgraphs. As the name suggests, we use Bloom filters to represent an approximation of the underlying graph. This representation is compact and computationally efficient to maintain in the presence of updates. This is especially important when we need to simultaneously maintain thousands of evolving subgraphs. BloomGraphs are used in computing and tracking conductance of these subgraphs as edge-activations arrive. BloomGraphs have several desirable properties in the context of this application, including a small memory footprint and efficient updateability. We also demonstrate mathematically that the error incurred in computing conductance is one-sided and that in the case of evolving subgraphs the change in approximate conductance has the same sign as the change in exact conductance in most cases. We validate the effectiveness of BloomGraphs through extensive experimentation on large Twitter graphs and other social networks.

**Session Chair: Jonathan Goldstein (Microsoft Research)****Streaming Anomaly Detection Using Randomized Matrix Sketching**

Hao Huang (General Electric Global Research); Shiva Kasiviswanathan (Samsung Research America)

Data is continuously being generated from sources such as machines, network traffic, application logs, etc. Timely and accurate detection of anomalies in massive data streams has important applications such as in preventing machine failures, intrusion detection, and dynamic load balancing. In this paper, we introduce a novel (unsupervised) anomaly detection framework which can be used to detect anomalies in a streaming fashion by making only one pass over the data while utilizing limited storage. We adapt ideas from matrix sketching to maintain, in a streaming model, a set of few orthogonal vectors that form a good approximate basis for all the observed data. Using this constructed orthogonal basis, anomalies in new incoming data are detected based on a simple reconstruction error test. We theoretically prove that our algorithm compares favorably with an offline approach based on expensive global singular value decomposition (SVD) updates. Additionally, we apply ideas from randomized low-rank matrix approximations to further speedup the algorithm. The experimental results show the effectiveness and efficiency of our approach over other popular scalable anomaly detection approaches.

**Index-Assisted Hierarchical Computations in Main-Memory RDBMS**

Robert Brunel (Technische Universität München); Norman May (SAP SE); Alfons Kemper (Technische Universität München)

We address the problem of expressing and evaluating computations on hierarchies represented as database tables. Engine support for such computations is very limited today, and so they are usually outsourced into stored procedures or client code. Recently, data model and SQL language extensions were proposed to conveniently represent and work with hierarchies. On that basis we introduce a concept of structural grouping to relational algebra, provide concise syntax to express a class of useful computations, and discuss algorithms to evaluate them efficiently by exploiting available indexing schemes. This extends the versatility of RDBMS towards a great many use cases dealing with hierarchical data.

**DEXTER: Large-Scale Discovery and Extraction of Product Specifications on the Web**

Disheng Qiu (Università degli Studi Roma Tre); Luciano Barbosa (IBM Research - Brazil); Xin Luna Dong (Google); Yanyan Shen (National University of Singapore); Divesh Srivastava (AT&amp;T Labs - Research)

The web is a rich resource of structured data. There has been an increasing interest in using web structured data for many applications such as data integration, web search and question answering. In this paper, we present DEXTER, a system to find product sites on the web, and detect and extract product specifications from them. Since product specifications exist in multiple product sites, our focused crawler relies on search queries and backlinks to discover product sites. To perform the detection, and handle the high diversity of specifications in terms of content, size and format, our system uses supervised learning to classify HTML fragments (e.g., tables and lists) present in web pages as specifications or not. To perform large-scale extraction of the attribute-value pairs from the HTML fragments identified by the specification detector, DEXTER adopts two lightweight strategies: a domain-independent and unsupervised wrapper method, which relies on the observation that these HTML fragments have very similar structure; and a combination of this strategy with a previous approach, which infers extraction patterns by annotations generated by automatic but noisy annotators. The results show that our crawler strategy to locate product specification pages is effective: (1) it discovered 146M product specification pages from 3,005 sites and 9 different categories; (2) the specification detector obtains high values of F-measure (close to 0.9) over a heterogeneous set of product specifications; and (3) our efficient wrapper methods for attribute-value extraction get very high values of precision (0.92) and recall (0.95) and obtain better results than a state-of-the-art, supervised rule-based wrapper.

**Walking in the Cloud: Parallel SimRank at Scale**

Zhenguo Li (Huawei); Yixiang Fang (The University of Hong Kong); Qin Liu (The Chinese University of Hong Kong); Jiefeng Cheng (Huawei); Reynold Cheng (University of Hong Kong); John Lui (The Chinese University of Hong Kong)

Despite its popularity, SimRank is computationally costly, in both time and space. In particular, its recursive nature poses a great challenge in using modern distributed computing power, and also prevents querying similarities individually. Existing solutions suffer greatly from these practical issues. In this paper, we break such dependency for maximum efficiency possible. Our method consists of offline and online phases. In offline phase, a length- $n$  indexing vector is derived by solving a linear system in parallel. At online query time, the similarities are computed instantly from the index vector. Throughout, the Monte Carlo method is used to maximally reduce time and space. Our algorithm, called CloudWalker, is highly parallelizable, with only linear time and space. Remarkably, it responds to both single-pair and single-source queries in constant time. CloudWalker is orders of magnitude more efficient and scalable than existing solutions for large-scale problems. Implemented on Spark with 10 machines and tested on the web-scale clue-web graph with 1 billion nodes and 43 billion edges, it takes 110 hours for offline indexing, 64 seconds for a single-pair query, and 188 seconds for a single-source query. To the best of our knowledge, our work is the first to report results on clue-web, which is 10x larger than the largest graph ever reported for SimRank computation.

# Industrial Track Paper Abstracts

---

## Industrial 1: Distributed Data Analytics Platforms - 1

### Session Chair: Gustavo Alonso (ETH Zurich)

#### Data Infrastructure at Flipkart (Invited talk)

Sharad Agarwal (Flipkart)

Data infrastructure lies at the core of Flipkart technology stack. The e-commerce business not only deals with huge traffic but the kind of data it deals with, is widely varied too. Whether it is user behavioural or transactional data of millions of customers or catalog with millions of products that being offered by thousands of sellers, or logistics network that covers the entire nation, the challenge to effectively derive value and build data enabled products lies in all aspects of technology stack.

Data infrastructure at Flipkart is a hosted central platform as a service that enables realtime systemic decision making and business intelligence across all product lines at Flipkart.

In this session, I will talk about the technical challenges in building this highly scalable and self-serve data platform that manages more than petabyte of data with thousand of realtime business critical data streams. I will also talk about the key technologies and platform abstractions that enable self serve realtime ETL and querying capabilities on Hadoop as well as on traditional data warehouses in seamless fashion.

#### Database System Support of Simulation Data

Hermano Lustosa (LNCC); Fabio Porto (LNCC); Pablo Blanco (LNCC); Patrick Valduriez (INRIA)

Supported by increasingly efficient HPC infra-structure, numerical simulations are rapidly expanding to fields such as oil and gas, medicine and meteorology. As simulations become more precise and cover longer periods of time, they may produce files with terabytes of data that need to be efficiently analyzed. In this paper, we investigate techniques for managing such data using an array DBMS. We take advantage of multidimensional arrays that nicely models the dimensions and variables used in numerical simulations. However, a naive approach to map simulation data files may lead to sparse arrays, impacting query response time, in particular, when the simulation uses irregular meshes to model its physical domain. We propose efficient techniques to map coordinate values in numerical simulations to evenly distributed cells in array chunks with the use of equi-depth histograms and space-filling curves. We implemented our techniques in SciDB and, through experiments over real-world data, compared them with two other approaches: row-store and column-store DBMS. The results indicate that multidimensional arrays and column-stores are much faster than a traditional row-store system for queries over a larger amount of simulation data. They also help identifying the scenarios where array DBMSs are most efficient, and those where they are outperformed by column-stores.

#### Hybrid Row-Column Partitioning in Teradata

Mohammed Al-Kateb (Teradata Labs); Paul Sinclair (Teradata Labs); Grace Au (Teradata Labs); Carrie Ballinger (Teradata Labs)

Data partitioning is an indispensable ingredient of database systems due to the performance improvement it can bring to any given mixed workload. Data can be partitioned horizontally or vertically. While some commercial proprietary and open source database systems have one flavor or mixed flavors of these partitioning forms, Teradata Database offers a unique hybrid row-column store solution that seamlessly combines both of these partitioning schemes. The key feature of this hybrid solution is that either row, column, or combined partitions are all stored and handled in the same way internally by the underlying file system storage layer. In this paper, we present the main characteristics and explain the implementation approach of Teradata's row-column store. We also discuss query optimization techniques applicable specifically to partitioned tables. Furthermore, we present a performance study that demonstrates how different partitioning options impact the performance of various queries.

#### Consistent Regions: Guaranteed Tuple Processing in IBM Streams

Gabriela Jacques da Silva (IBM Research - T.J. Watson); Fang Zheng (IBM Research - T.J. Watson); Daniel Debrunner (IBM); Kun-Lung Wu (IBM Research - T.J. Watson); Victor Dogaru (IBM); Eric Johnson (IBM); Michael Spicer (IBM); Ahmet Erdem Sariyuce (Sandia National Labs)

Guaranteed tuple processing has become critically important for many streaming applications. This paper describes how we enabled IBM Streams, an enterprise-grade stream processing system, to provide data processing guarantees. Our solution goes from language-level abstractions to a runtime protocol. As a result, with a couple of simple annotations at the source code level, IBM Streams developers can define consistent regions, allowing any subgraph of their streaming application to achieve guaranteed tuple processing. At runtime, a consistent region periodically executes a variation of the Chandy-Lamport snapshot algorithm to establish a consistent global state for that region. The coupling of consistent states with data replay enables guaranteed tuple processing.

**Session Chair: Justin Levandoski (Microsoft Research)****Big Data Analysis as a Service on Public Clouds (Invited talk)**

Rajat Venkatesh (Qubole)

Public clouds pose challenges (reliability, IO latency etc) as well as provide opportunities (elasticity, pricing models) for large volume data processing. In this talk, I will present the lessons learnt and features (auto-scaling, hybrid clusters with spot nodes, caching etc) implemented in Qubole's data service to provide an efficient (cost & performance) and a reliable data processing service in the top public clouds.

QDS allows users to run batch or ETL, interactive and streaming workloads. The technologies used for each of these workloads are different. Apache Hive, Apache Hadoop M/R and Apache Pig are popular for batch workloads. Presto & SparkSQL are used for interactive analytic queries. Apache Spark is used for machine learning. Spark Streaming is used for streaming workloads. Apache HBase is the K-V store. As such the design and architecture of each of these technologies is very different. We've focussed on each of these technologies separately and augmented them to run better on the cloud as well as a service.

This talk will iterate through each of these technologies, list areas where the open source distribution has gaps and how QDS has closed these gaps. Since these technologies are representative in their class, the features implemented in QDS is broadly applicable. I will also provide real-world metrics for performance gains and cost benefits where applicable.

**dmapply: A functional primitive to express distributed machine learning algorithms in R**

Edward Ma (Hewlett-Packard Enterprise Vertica); Vishrut Gupta (Hewlett-Packard Enterprise Vertica); Meichun Hsu (Hewlett-Packard Enterprise Vertica); Indrajit Roy (Hewlett Packard Labs)

Due to R's popularity as a data-mining tool, many distributed systems expose an R-based API to users who need to build a distributed application in R. As a result, data scientists have to learn to use different interfaces such as R Hadoop, SparkR, Revolution R's ScaleR, and HPE's Distributed R. Unfortunately, these interfaces are custom, non-standard, and difficult to learn. Not surprisingly, R applications written in one framework do not work in another, and each backend infrastructure has spent redundant effort in implementing distributed machine learning algorithms. Working with the members of R-core, we have created ddR (Distributed Data structures in R), a unified system that works across different distributed frameworks. In ddR, we introduce a novel programming primitive called dmapply that executes functions on distributed data structures. The dmapply primitive encapsulates different computation patterns: from function and data broadcast to pair-wise communication. We show that dmapply is powerful enough to express algorithms that fit the statistical query model, which includes many popular machine learning algorithms, as well as applications written in MapReduce. We have integrated ddR with many backends, such as R's single-node parallel framework, multi-node SNOW framework, Spark, and HPE Distributed R, with few or no modifications to any of these systems. We have also implemented multiple machine learning algorithms which are not only portable across different distributed systems, but also have performance comparable to the "native" implementations on the backends. We believe that ddR will standardize distributed computing in R, just like the SQL interface has standardized how relational data is manipulated.

**SystemML: Declarative Machine Learning on Spark**

Matthias Boehm (IBM Research - Almaden); Michael Dusenberry (IBM Spark Technology Center); Deron Eriksson (IBM Spark Technology Center); Alexandre Evfimievski (IBM Research - Almaden); Faraz Makari Manshadi (IBM Research - Almaden); Niketan Pansare (IBM Research - Almaden); Berthold Reinwald (IBM Research - Almaden); Frederick Reiss (IBM Research - Almaden); Prithviraj Sen (IBM Research - Almaden); Arvind Surve (IBM Spark Technology Center); Shirish Tatikonda (IBM Research - Almaden)

The rising need for custom machine learning (ML) algorithms and the growing data sizes that require the exploitation of distributed, data-parallel frameworks such as MapReduce or Spark, pose significant productivity challenges to data scientists. Apache SystemML addresses these challenges through declarative ML by (1) increasing the productivity of data scientists as they are able to express custom algorithms in a familiar domain-specific language covering linear algebra primitives and statistical functions, and (2) transparently running these ML algorithms on distributed, data-parallel frameworks by applying cost-based compilation techniques to generate efficient, low-level execution plans with in-memory single-node and large-scale distributed operations. This paper describes SystemML on Apache Spark, end to end, including insights into various optimizer and runtime techniques as well as performance characteristics. We also share lessons learned from porting SystemML to Spark and declarative ML in general. Finally, SystemML is open-source, which allows the database community to leverage it as a testbed for further research.

**Not for the Timid: On the Impact of Aggressive Over-booking in the Cloud**

Willis Lang (Microsoft); Karthik Ramachandra (Microsoft); David DeWitt (Microsoft); Shize Xu (Microsoft); Qun Guo (Microsoft); Ajay Kalhan (Microsoft); Peter Carlin (Microsoft)

To lower hosting costs and service prices, database-as-a-service (DBaaS) providers strive to maximize cluster utilization without negatively affecting their users' service experience. Some of the most effective approaches for increasing service efficiency result in the over-booking of the cluster with user databases. For instance, one approach is to reclaim cluster capacity from a database when



it is idle, temporarily re-using the capacity for some other purpose, and over-booking the cluster's resources. Such approaches are largely driven by policies that determine when it is prudent to temporarily reclaim capacity from an idle database. In this paper, we examine policies that inherently tune the system's idle sensitivity. Increased sensitivity to idleness leads to aggressive over-booking while the converse leads to conservative reclamation and lower utilization levels. Aggressive over-booking also incurs a "reserve" capacity cost (for when we suddenly "owe" capacity to previously idle databases.) We answer these key questions in this paper: (1) how to find a "good" resource reclamation policy for a given DBaaS cluster of users; and (2) how to forecast the needed near-term reserve capacity. To help us answer these questions, we used production user activity traces from Azure SQL DB and built models of an over-booking mechanism. We show that choosing the right policy can substantially boost the efficiency of the service, facilitating lower service prices via lower amortized infrastructure costs.

## Industrial 3: Data Engine Architectures - 1

### Session Chair: Wolfgang Lehner (Technische Universität Dresden)

#### Hardware-Software Co-design for Data Management (Invited talk)

Weimei Gong (Oracle); Shasank Chavan (Oracle)

The Oracle database is the world's leading data management product. Soon after the acquisition of Sun Microsystems in 2010, engineers from Oracle and Sun started design on a new category of microprocessor that processes data several times faster, many times more efficiently, and qualitatively safer. This goal cannot be achieved by running software unchanged -- we needed to design new hardware and modify our software to utilize it. The first step in this transformation was released as the SPARC M7 processor last year. The approach we took, and are taking, exploits the following ideas:

Big is better: Large scale computing systems give access to huge amounts of data without the costs of moving data between systems. This allows for larger tables, bigger sorts, fatter graphs, and more cloud tenants sharing the the same resource pool on SPARC systems that scale linearly in cost and performance from 8 to 512 cores, 64 to 4096 threads. Secure is better: Cache-line level memory access checking allows our instrumented memory allocators to manage memory at production speed while detecting bugs and reporting attacks in real time. Information Density is better: With hardware designed for scanning n-gram compressed, bit packed, dictionary and run-length encoded columnar data at full memory bandwidth, we make maximal use of every bit stored and every cache line transferred over the memory channels with no impact on performance. Fast is better: With hardware support for running database operators on specialized streaming processors, we drive the memory channels at maximum rate, freeing up power and cores for running user computations on the result of these operators. Connected is better: Integrating EDR InfiniBand on-chip and on-board with low-latency, high-throughput, one-sided networking. Portable is better: By supporting platform independent acceleration APIs inside the database we support a wide variety of acceleration techniques and give applications and query planners the information to make the best use of the available hardware. Integrated is better: By supporting and accelerating multiple storage types (In-memory, NVMe, Exadata, NFS, HDFS, Fibre Channel), data formats (row major, column major, graph, JSON, spatial, MIME, Hive), algorithms, query languages, network protocols, and hardware platforms in a single product, we optimize, share resources, increase usability and reduce the cost and the cognitive load in acquiring, storing, securing and understanding data.

In this talk, I will describe the experience that drives our microprocessor acceleration priorities, the constraints and joys of the hardware-software co-design process, the HW features that resulted, and how software engineers have exploiting these features in ways we expected and ways we didn't. The industry has accelerated linear algebra and computer graphics, and in this talk we'll see how we apply it to data processing but with changes to match the lower compute density of the problem space.

#### Nitro: A Fast, Scalable In-Memory Storage Engine for NoSQL Global Secondary Index

Sarath Lakshman (Couchbase Inc.); Sriram Melkote (Couchbase Inc.); John Liang (Couchbase Inc.); Ravi Mayuram (Couchbase Inc.)

We present Nitro, a high-performance in-memory key-value storage engine used in Couchbase 4.5 Global Secondary Indexes. The Nitro storage engine is well suited for the recent hardware trends like large amounts of memory and many CPU cores. The storage engine leverages latch-free data structures and tries to achieve linear scalability for the index read-write operations. The Nitro storage engine offers concurrent readers and writers, lightweight database snapshots, stable scan, backup and recovery operations. We integrated Nitro into the Couchbase Global Secondary Indexes (GSI) and observed significant improvement in performance compared to our disk oriented storage engine configured with the same amount of memory for buffer cache. On a 32 core machine, we observed an end-to-end GSI server insertion throughput of 1,650,000 entries/sec and index update throughput of 822,000 entries/sec. A single instance of Nitro data structure running on a 40 core machine achieved a peak insertion throughput of 4 million index entries/sec and entry lookup throughput of 10 million lookups/sec.

#### Aerospike: Architecture of a Real-Time Operational DBMS

V. Srinivasan (Aerospike Inc.); Brian Bulkowski (Aerospike Inc.); Wei-Ling Chu (Aerospike Inc.); Sunil Sayyaparaju (Aerospike Inc.); Andrew Gooding (Aerospike Inc.); Rajkumar Iyer (Aerospike Inc.); Ashish Shinde (Aerospike Inc.); Thomas Lopatic (Aerospike Inc.)

In this paper, we describe the solutions developed to address key technical challenges encountered while building a distributed database system that can smoothly handle demanding real-time workloads and provide a high level of fault tolerance.

Specifically, we describe schemes for the efficient clustering and data partitioning for the automatic scale out of processing across multiple nodes and for optimizing the usage of CPUs, DRAM, SSDs and networks to efficiently scale up performance on one node. The techniques described here were used to develop Aerospike (formerly Citrusleaf), a high performance distributed database system built to handle the needs of today's interactive online services. Most real-time decision systems that use Aerospike require very high scale and need to make decisions within a strict SLA by reading from, and writing to, a database containing billions of data items at a rate of millions of operations per second with sub-millisecond latency. For over five years, Aerospike has been continuously used in over a hundred successful production deployments, as many enterprises have discovered that it can substantially enhance their user experience.

### **Comdb2 - Bloomberg's Highly Available Relational Database System**

Alex Scotti (Bloomberg LP); Mark Hannum (Bloomberg LP); Michael Ponomarenko (Bloomberg LP); Dorin Hogeia (Bloomberg LP); Akshat Sikarwar (Bloomberg LP); Mohit Khullar (Bloomberg LP); Adi Zaimi (Bloomberg LP); James Leddy (Bloomberg LP); Fabio Angius (Bloomberg LP); Rivers Zhang (Bloomberg LP); Lingzhi Deng (Bloomberg LP)

Comdb2 is a distributed database designed for geographical replication and high availability. In contrast with the latest trends in this field Comdb2 offers full transactional support, a standard relational model, and the expressivity of SQL. Moreover, the system allows for rich stored procedures using a dialect of Lua. Comdb2 implements a serializable system in which reads from any node always return current values. Comdb2 provides transparent High Availability through built in service discovery and sophisticated retry logic embedded in the standard API. In addition to the relational data model, Comdb2 implements queues for publisher-to-subscriber message delivery. Queues can be combined with table triggers for time-consistent log distribution, two functionalities that are commonly needed in modern OLTP. In this paper we give an overview of our last twelve years of work. We focus on the design choices that have made Comdb2 the primary database solution within our company, Bloomberg LP (BLP).

## **Industrial 4: Data Engine Architectures - 2**

### **Session Chair: Spyros Blanas (The Ohio State University)**

#### **Kodiak: Leveraging Materialized Views for Very Low-Latency Analytics Over High-Dimensional Web-Scale Data**

Shaosu Liu (Turn Inc.); Bin Song (Turn Inc.); Sriharsha Gangam (Turn Inc.); Lawrence Lo (Turn Inc.); Khaled Elmeleegy (Turn Inc.)

Turn's online advertising campaigns produce petabytes of data. This data is composed of trillions of events, e.g. impressions, clicks, etc., spanning multiple years. In addition to a timestamp, each event includes hundreds of fields describing the user's attributes, campaign's attributes, attributes of where the ad was served, etc. Advertisers need advanced analytics to monitor their running campaigns' performance, as well as to optimize future campaigns. This involves slicing and dicing the data over tens of dimensions over arbitrary time ranges. Many of these queries need to power the web portal to provide reports and dashboards. For an interactive response time, they have to have tens of milliseconds latency. At Turn's scale of operations, no existing system was able to deliver this performance in a cost effective manner. Kodiak, a distributed analytical data platform for web-scale high-dimensional data, was built to serve this need. It relies on precomputations to materialize thousands of views to serve these advanced queries. These views are partitioned and replicated across Kodiak's storage nodes for scalability and reliability. They are system maintained as new events arrive. At query time, the system auto-selects the most suitable view to serve each query. Kodiak has been used in production for over a year. It hosts 2490 views for over three petabytes of raw data serving over 200K queries daily. It has median and 99% query latencies of 8 ms and 252 ms respectively. Our experiments show that its query latency is 3 orders of magnitude faster than leading big data platforms on head-to-head comparisons using Turn's query workload. Moreover, Kodiak uses 4 orders of magnitude less resources to run the same workload.

#### **Accelerating Analytics with Dynamic In-Memory Expressions**

Aurosish Mishra (Oracle); Shasank Chavan (Oracle); Allison Holloway (Oracle); Tirthankar Lahiri (Oracle); Zhen Hua Liu (Oracle); Sunil Chakkappen (Oracle); Dennis Lui (Oracle); Vinita Subramanian (Oracle); Ramesh Kumar (Oracle); Maria Colgan (Oracle); Jesse Kamp (Oracle); Niloy Mukherjee (LinkedIn); Vineet Marwah (Oracle)

Oracle Database In-Memory (DBIM) accelerates analytic workload performance by orders of magnitude through an in-memory columnar format utilizing techniques such as SIMD vector processing, in-memory storage indexes, and optimized predicate evaluation and aggregation. With Oracle Database 12.2, Database In-Memory is further enhanced to accelerate analytic processing through a novel lightweight mechanism known as Dynamic In-Memory Expressions (DIMEs). The DIME mechanism automatically detects frequently occurring expressions in a query workload, and then creates highly optimized, transactionally consistent, in-memory columnar representations of these expression results. At runtime, queries can directly access these DIMEs, thus avoiding costly expression evaluations. Furthermore, all the optimizations introduced in DBIM can apply directly to DIMEs. Since DIMEs are purely in-memory structures, no changes are required to the underlying tables. We show that DIMEs can reduce query elapsed times by several orders of magnitude without the need for costly pre-computed structures such as computed columns or materialized views or cubes.

## **TrafficDB: HERE's High Performance Shared-Memory Data Store**

Ricardo Fernandes (HERE Global B.V.); Piotr Zaczkowski (HERE Global B.V.); Bernd Göttinger (HERE Global B.V.); Conor Ettinoffe (HERE Global B.V.); Anis Moussa (HERE Global B.V.)

HERE's traffic-aware services enable route planning and traffic visualisation on web, mobile and connected car applications. These services process thousands of requests per second and require efficient ways to access the information needed to provide a timely response to end-users. The characteristics of road traffic information and these traffic-aware services require storage solutions with specific performance features. A route planning application utilising traffic congestion information to calculate the optimal route from an origin to a destination might hit a database with millions of queries per second. However, existing storage solutions are not prepared to handle such volumes of concurrent read operations, as well as to provide the desired vertical scalability. This paper presents TrafficDB, a shared-memory data store, designed to provide high rates of read operations, enabling applications to directly access the data from memory. Our evaluation demonstrates that TrafficDB handles millions of read operations and provides near-linear scalability on multi-core machines, where additional processes can be spawned to increase the systems' throughput without a noticeable impact on the latency of querying the data store. The paper concludes with a description of how TrafficDB improved the performance of our traffic-aware services running in production.

## **The MemSQL Query Optimizer: A modern optimizer for real-time analytics in a distributed database**

Jack Chen (MemSQL Inc.); Samir Jindel (MemSQL Inc.); Robert Walzer (MemSQL Inc.); Rajkumar Sen (MemSQL Inc.); Nika Jimshelishvili (MemSQL Inc.); Michael Andrews (MemSQL Inc.)

Real-time analytics on massive datasets has become a very common need in many enterprises. These applications require not only rapid data ingest, but also quick answers to analytical queries operating on the latest data. MemSQL is a distributed SQL database designed to exploit memory-optimized, scale-out architecture to enable real-time transactional and analytical workloads which are fast, highly concurrent, and extremely scalable. Many analytical queries in MemSQL's customer workloads are complex queries involving joins, aggregations, sub-queries, etc. over star and snowflake schemas, often ad-hoc or produced interactively by business intelligence tools. These queries often require latencies of seconds or less, and therefore require the optimizer to not only produce a high quality distributed execution plan, but also produce it fast enough so that optimization time does not become a bottleneck. In this paper, we describe the architecture of the MemSQL Query Optimizer and the design choices and innovations which enable it quickly produce highly efficient execution plans for complex distributed queries. We discuss how query rewrite decisions oblivious of distribution cost can lead to poor distributed execution plans, and argue that to choose high-quality plans in a distributed database, the optimizer needs to be distribution-aware in choosing join plans, applying query rewrites, and costing plans. We discuss methods to make join enumeration faster and more effective, such as a rewrite-based approach to exploit bushy joins in queries involving multiple star schemas without sacrificing optimization time. We demonstrate the effectiveness of the MemSQL optimizer over queries from the TPC-H benchmark and a real customer workload.

## **Industrial 5: Graph Systems and Analytics**

### **Session Chair: Prasad Deshpande (Consultant)**

### **LDBC Graphalytics: A Benchmark for Large-Scale Graph Analysis on Parallel and Distributed Platforms**

Alexandru Iosup (Delft University of Technology); Tim Hegeman (Delft University of Technology); Wing Lung Ngai (Delft University of Technology); Stijn Heldens (Delft University of Technology); Arnau Prat-Pérez (Universitat Politècnica de Catalunya); Thomas Manhardt (Oracle Labs); Hassan Chafi (Oracle Labs); Mihai Capotă (Intel Labs); Narayanan Sundaram (Intel Labs); Michael Anderson (Intel Labs); Ilie Gabriel Tanase (IBM Research); Yinglong Xia (Huawei Research America); Lifeng Nai (Georgia Tech); Peter Boncz (CWI)

In this paper we introduce LDBC Graphalytics, a new industrial-grade benchmark for graph analysis platforms. It consists of six deterministic algorithms, standard datasets, synthetic dataset generators, and reference output, that enable the objective comparison of graph analysis platforms. Its test harness produces deep metrics that quantify multiple kinds of system scalability, such as horizontal/vertical and weak/strong, and of robustness, such as failures and performance variability. The benchmark comes with open-source software for generating data and monitoring performance. We describe and analyze six implementations of the benchmark (three from the community, three from industry), providing insights into the strengths and weaknesses of the platforms. Key to our contribution, vendors perform the tuning and benchmarking of their platforms.

### **GraphJet: Real-Time Content Recommendations at Twitter**

Aneesh Sharma (Twitter Inc.); Jerry Jiang (Twitter Inc.); Praveen Bommannavar (Twitter Inc.); Brian Larson (Twitter Inc.); Jimmy Lin (University of Waterloo)

This paper presents GraphJet, a new graph-based system for generating content recommendations at Twitter. As motivation, we trace the evolution of our formulation and approach to the graph recommendation problem, embodied in successive generations of systems. Two trends can be identified: supplementing batch with real-time processing and a broadening of the scope of recommendations

from users to content. Both of these trends come together in GraphJet, an in-memory graph processing engine that maintains a real-time bipartite interaction graph between users and tweets. The storage engine implements a simple API, but one that is sufficiently expressive to support a range of recommendation algorithms based on random walks that we have refined over the years. Similar to Cassovary, a previous graph recommendation engine developed at Twitter, GraphJet assumes that the entire graph can be held in memory on a single server. The system organizes the interaction graph into temporally-partitioned index segments that hold adjacency lists. GraphJet is able to support rapid ingestion of edges while concurrently serving lookup queries through a combination of compact edge encoding and a dynamic memory allocation scheme that exploits power-law characteristics of the graph. Each GraphJet server ingests up to one million graph edges per second, and in steady state, computes up to 500 recommendations per second, which translates into several million edge read operations per second.

### **Using Domain-Specific Languages for Analytic Graph Databases**

Martin Sevenich (Oracle Labs); Sungpack Hong (Oracle Labs); Oskar Van Rest (Oracle Labs); Zhe Wu (Oracle); Jayanta Banerjee (Oracle); Hassan Chafi (Oracle Labs)

Recently graph has been drawing a lot of attention both as a natural data model that captures fine-grained relationships between data entities and as a tool for powerful data analysis that considers such relationships. In this paper, we present a new graph database system that integrates a robust graph storage with an efficient graph analytics engine. Primarily, our system adopts two domain-specific languages (DSLs), one for describing graph analysis algorithms and the other for graph pattern matching queries. Compared to the API-based approaches in conventional graph processing systems, the DSL-based approach provides users with more flexible and intuitive ways of expressing algorithms and queries. Moreover, the DSL-based approach has significant performance benefits as well, (1) by skipping (remote) API invocation overhead and (2) by applying high-level optimization from the compiler.

### **Cubrick: Indexing Millions of Records per Second for Interactive Analytics**

Pedro Pedreira (Facebook Inc); Chris Croswhite (Facebook Inc.); Luis Bona (Federal University of Parana)

This paper describes the architecture and design of Cubrick, a distributed multidimensional in-memory database suited for interactive analytics over highly dynamic datasets. Cubrick has a strictly multidimensional data model composed of cubes, dimensions and metrics, supporting sub-second OLAP operations such as slice and dice, roll-up and drill-down over terabytes of data. All data stored in Cubrick is range partitioned by every dimension and stored within containers called bricks in an unordered and sparse fashion, providing high data ingestion rates and indexed access through any combination of dimensions. In this paper, we describe details about Cubrick's internal data structures, distributed model, query execution engine and a few details about the current implementation. Finally, we present results from a thorough experimental evaluation that leveraged datasets and queries collected from a few internal Cubrick deployments at Facebook.

# Demonstration Track Paper Abstracts

---

## Demo 1A/1B: Data Engines and Analytics

### **Vita: A Versatile Toolkit for Generating Indoor Mobility Data for Real-World Buildings**

Huan Li (Zhejiang University); Hua Lu (Aalborg University); Xin Chen (Zhejiang University); Gang Chen (Zhejiang University); Ke Chen (Zhejiang University); Lidan Shou (Zhejiang University)

We demonstrate a generic, user-configurable toolkit for generating different types of indoor mobility data for real-world buildings. Our prototype generates the desired data in a three-layer pipeline. The Infrastructure Layer accepts industry-standard digital building information (DBI) files to generate the host indoor environment, allowing users to configure the generation of a variety of positioning devices, such as Wi-Fi, Bluetooth, RFID, etc. The Moving Object Layer offers the functionality of defining objects or trajectories, with configurable indoor moving patterns, distribution models, and sampling frequencies. The Positioning Layer generates synthetic signal strength measurements known as raw RSSI measurements according to the positioning device data and trajectory data generated at layered layers. It also generates different types of indoor positioning data through the customization of all typical indoor positioning methods on the raw RSSI data.

### **JexLog: A Sonar for the Abyss**

Tobias Scheuer (SAP SE); Norman May (SAP SE); Alexander Böhm (SAP SE); Daniel Scheibli (SAP SE)

Today's hardware architectures provide an ever-increasing number of CPU cores that can be used for running concurrent operations. A big challenge is to ensure that these operations are properly synchronized, and make efficient use of the available resources. Fellow database researchers have appropriately described this problem as staring into the abyss of complexity, where reasoning about the interplay of jobs on a thousand cores becomes extremely challenging. In this demonstration, we show how a new tool, JexLog, can help to visually analyze concurrent jobs in system software and how it is used to optimize for modern hardware.

### **SI2P: A Restaurant Recommendation System Using Preference Queries over Incomplete Information**

Xiaoye Miao (Zhejiang University); Yunjun Gao (Zhejiang University); Gang Chen (Zhejiang University); Huiyong Cui (Zhejiang University); Chong Guo (Zhejiang University); Weida Pan (Zhejiang University)

The incomplete data is universal in many real-life applications due to data integration, the limitation of devices, etc. In this demonstration, we present SI2P, a restaurant recommendation System with Preference queries on Incomplete Information. SI2P is capable of friendly recommending desirable restaurants based on preference queries that take the incomplete ratings information into consideration. It adopts the browser-server model, and incorporates three functionality modules including friendly and convenient query submission, flexible and useful result explanation, timely and incremental dataset interaction. SI2P provides the server side based on an extended PostgreSQL database that integrates two types of preference queries, namely, skyline and top-k dominating queries over incomplete data. It also offers the browser-based interface for the users to interact with the system. Using a real restaurant dataset from TripAdvisor, we demonstrate SI2P can recommend and explore the restaurants in a friendly way.

### **Mixed-instance querying: a lightweight integration architecture for data journalism**

Raphaël Bonaque (INRIA); Tien Cao (INRIA); Bogdan Cautis (Université Paris-Sud); François Goasdoué (Université de Rennes 1); Javier Letellier (INRIA); Ioana Manolescu (INRIA); Oscar Mendoza (INRIA); Swen Ribeiro (Université Paris-Sud); Xavier Tannier (Université Paris-Sud); Michaël Thomazo (INRIA)

As the world's affairs get increasingly more digital, timely production and consumption of news require to efficiently and quickly exploiting heterogeneous data sources. Discussions with journalists revealed that content management tools currently at their disposal fall very short of expectations. We demonstrate TATOINE, a lightweight data integration prototype, which allows to quickly set up integration queries across (very) heterogeneous data sources, capitalizing on the many data links (joins) available in this application domain. Our demonstration is based on scenarios we currently study in collaboration with Le Monde, France's major newspaper.

### **Precision Performance Surgery for PostgreSQL – LLVM-based Expression Compilation, Just in Time**

Dennis Butterstein (University of Tuebingen); Torsten Grust (Universität Tübingen)

We demonstrate how the compilation of SQL expressions into machine code leads to significant query runtime improvements in PostgreSQL 9. Our primary goal is to connect recent research in query code generation with one of the most widely deployed database engines. The approach calls on LLVM to translate arithmetic and filter expressions into native x86 instructions just before SQL query execution begins. We deliberately follow a non-invasive design that does not turn PostgreSQL on its head: interpreted and compiled expression evaluation coexist and both are used to execute the same query. We will bring an enhanced version of PostgreSQL that

exhibits notable runtime savings and provides visual insight into exactly where and how execution plans can benefit from SQL expression compilation.

### **Partial Marking for Automated Grading of SQL Queries**

Bikash Chandra (IIT Bombay); Mathew Joseph (IIT Bombay); Bharath Radhakrishnan (IIT Bombay); Shreevidhya Acharya (IIT Bombay); S. Sudarshan (IIT Bombay)

The XData system, currently being developed at IIT Bombay, provides an automated and interactive platform for grading student SQL queries, as well as for learning SQL. Prior work on the XData system focused on generating query specific test cases to catch common errors in queries. These test cases are used to check whether the student queries are correct or not. For grading student assignments, it is usually not sufficient to just check if a query is correct: if the query is incorrect, partial marks may need to be given, depending on how close the query is to being correct. In this paper, we extend the XData system by adding features that enable awarding of partial marks to incorrect student queries. Our system is able to go beyond numerous syntactic features when comparing a student query with a correct query. These features of our grading system allow the grading of SQL queries to be fully automated, and scalable to even large class sizes such as those of MOOCs.

### **Squall: Scalable Real-time Analytics**

Aleksandar Vitorovic (EPFL); Mohammed Elseidy (EPFL); Khayyam Guliyev (EPFL); Khue Vu Minh (EPFL); Daniel Espino Timón (EPFL); Mohammad Dashti (EPFL); Yannis Klonatos (EPFL); Christoph Koch (EPFL)

Squall is a scalable online query engine that runs complex analytics in a cluster using skew-resilient, adaptive operators. Squall builds on state-of-the-art partitioning schemes and local algorithms, including some of our own. This paper presents the overview of Squall, including some novel join operators. The paper also presents lessons learned over the five years of working on this system, and outlines the plan for the proposed system demonstration.

### **LocationSpark: A Distributed In-Memory Data Management System for Big Spatial Data**

Mingjie Tang (Purdue University); Yongyang Yu (Purdue University); Outaibah Malluhi (Qatar University); Mourad Ouzzani (Qatar Computing Research Institute); Walid Aref (Purdue University)

We present LocationSpark, a spatial data processing system built on top of Apache Spark, a widely used distributed data processing system. LocationSpark offers a rich set of spatial query operators, e.g., range search, Sk\$NN, spatio-textual operation, spatial-join, and Sk\$NN-join. To achieve high performance, LocationSpark employs various spatial indexes for in-memory data, and guarantees that immutable spatial indexes have low overhead with fault tolerance. In addition, we build two new layers over Spark, namely a query scheduler and a query executor. The query scheduler is responsible for mitigating skew in spatial queries, while the query executor selects the best plan based on the indexes and the nature of the spatial queries. Furthermore, to avoid unnecessary network communication overhead when processing overlapped spatial data, we embed an efficient spatial Bloom filter into LocationSpark's indexes. Finally, LocationSpark tracks frequently accessed spatial data, and dynamically flushes less frequently accessed data into disk. We evaluate our system on real workloads and demonstrate that it achieves an order of magnitude performance gain over a baseline framework.

### **Amoeba: A Shape changing Storage System for Big Data**

Anil Shanbhag (MIT); Samuel Madden (MIT); Alekh Jindal (Microsoft); Yi Lu (MIT)

Data partitioning significantly improves the query performance in distributed database systems. A large number of techniques have been proposed to efficiently partition a dataset for a given query workload. However, many modern analytic applications involve ad-hoc or exploratory analysis where users do not have a representative query workload upfront. Furthermore, workloads change over time as businesses evolve or as analysts gain better understanding of their data. Static workload-based data partitioning techniques are therefore not suitable for such settings. In this paper, we describe the demonstration of Amoeba, a distributed storage system which uses adaptive multi-attribute data partitioning to efficiently support ad-hoc as well as recurring queries. Amoeba applies a robust partitioning algorithm such that ad-hoc queries on all attributes have similar performance gains. Thereafter, Amoeba adaptively repartitions the data based on the observed query sequence, i.e., the system improves over time. All along Amoeba offers both adaptivity (i.e., adjustments according to workload changes) as well as robustness (i.e., avoiding performance spikes due to workload changes). We propose to demonstrate Amoeba on scenarios from an internet-of-things startup that tracks user driving patterns. We invite the audience to interactively fire fast ad-hoc queries, observe multi-dimensional adaptivity, and play with a robust/reactive knob in Amoeba. The web front end displays the layout changes, runtime costs, and compares it to Spark with both default and workload-aware partitioning.

### **F: Regression Models over Factorized Views**

Dan Olteanu (University of Oxford); Maximilian Schleich (University of Oxford);

We demonstrate F, a system for building regression models over database views. At its core lies the observation that the computation and representation of materialized views, and in particular of joins, entail non-trivial redundancy that is not necessary for the efficient computation of aggregates used for building regression models. F avoids this redundancy by factorizing data and computation and can outperform the state-of-the-art systems MADlib, R, and Python StatsModels by orders of magnitude on real-world datasets. We illustrate how to incrementally build regression models over factorized views using both an in-

memory implementation of F and its SQL encoding. We also showcase the effective use of F for model selection: F decouples the data-dependent computation step from the data-independent convergence of model parameters and only performs once the former to explore the entire model space.

### **Magellan: Toward Building Entity Matching Management Systems over Data Science Stacks**

Pradap Konda (University of Wisconsin Madison); Sanjib Das (University of Wisconsin Madison); Paul Suganthan G. C. (University of Wisconsin Madison); Anhai Doan (University of Wisconsin Madison); Adel Ardalani (University of Wisconsin Madison); Jeffrey Ballard (University of Wisconsin Madison); Han Li (University of Wisconsin Madison); Fatemah Panahi (University of Wisconsin Madison); Haojun Zhang (University of Wisconsin Madison); Jeffrey Naughton (University of Wisconsin Madison); Shishir Prasad (University of Wisconsin Madison); Ganesh Krishnan (WalmartLabs); Rohit Deep (WalmartLabs); Vijay Raghavendra (WalmartLabs)

Entity matching (EM) has been an important problem in data integration, and will become even more important in the age of Big Data. However, most EM works have focused only on developing matching algorithms. Going forward we argue that far more efforts should be devoted to building EM systems. Current EM systems suffer from four major limitations that prevent them from being used widely in practice. We present Magellan, a new kind of EM system. Magellan is novel in several important aspects. (1) It provides a howto guide that tells users what to do in each EM scenario, step by step. (2) It provides tools to help users do these steps. The tools seek to cover the entire EM pipeline, not just matching and blocking as current EM systems do. (3) It is built on top of the data science stacks in Python, allowing it to borrow a rich set of capabilities in data cleaning, IE, visualization, learning, etc. (4) In contrast to current “closed-world” EM systems, Magellan is a new kind of system called “open-world” systems, which must learn to “play well” with other systems in the data science stacks.

### **Schema Independent and Scalable Relational Learning by Castor**

Jose Picado (Oregon State University); Parisa Ataei (Oregon State University); Arash Termehchy (Oregon State University); Alan Fern (Oregon State University)

Learning novel relations from relational databases is an important problem with many applications in database systems and machine learning. Relational learning algorithms leverage the properties of the database schema to find the definition of the target relation in terms of the existing relations in the database. However, the same data set may be represented under different schemas for various reasons, such as efficiency and data quality. Unfortunately, current relational learning algorithms tend to vary quite substantially over the choice of schema, which complicates their off-the-shelf application. We demonstrate Castor, a relational learning system that efficiently learns the same definitions over common schema variations. The results of Castor are more accurate than well-known learning systems over large data.

### **AD-WIRE: Add-on for Web Item Reviewing System**

Rajeshkumar Kannapalli (University of Texas at Arlington); Azade Nazi (University of Texas at Arlington); Mahashweta Das (Hewlett Packard Labs); Gautam Das (University of Texas at Arlington)

Over the past few decades as purchasing options moved online, the widespread use and popularity of online review sites has simultaneously increased. In spite of the fact that a huge extent of buying choices today are driven by numeric scores (e.g., rating a product), detailed reviews play an important role for activities like purchasing an expensive DSLR camera. Since writing a detailed review for an item is usually time-consuming, the number of reviews available in the Web is far from many. In this paper, we build a system AD-WIRE that given a user and an item, our system identifies the top-k meaningful tags to help her review the item easily. AD-WIRE allows a user to compose her review by quickly selecting from among the set of returned tags or writes her own review. AD-WIRE also visualizes the dependency of the tags to different aspects of an item so a user can make an informed decision quickly. The system can be used for different type of the products. The current demonstration is built to explore review writing process for the mobile phones.

## **Demo 2A/2B: Interactive and Exploratory Systems**

### **ArchimedesOne: Query Processing over Probabilistic Knowledge Bases**

Xiaofeng Zhou (University of Florida); Yang Chen (University of Florida); Daisy Zhe Wang (University of Florida)

Knowledge bases are becoming increasingly important in structuring and representing information from the web. Meanwhile, web-scale information poses significant scalability and quality challenges to knowledge base systems. To address these challenges, we develop a probabilistic knowledge base system, ArchimedesOne, by scaling up the knowledge expansion and statistical inference algorithms. We design a web interface for users to query and update large knowledge bases. In this paper, we demonstrate the ArchimedesOne system to showcase its efficient query and inference engines. The demonstration serves two purposes: 1) to provide an interface for users to interact with ArchimedesOne through load, search, and update queries; and 2) to validate our approaches of knowledge expansion by applying inference rules in batches using relational operations and query-driven inference to focus computation on the query facts. We compare ArchimedesOne with state-of-the-art approaches using two large automatically extracted knowledge bases: NELL-sports with 4.5 million facts and Reverb with 15 million facts.

### **Rudolf: Interactive Rule Refinement System for Fraud Detection**

Tova Milo (Tel-Aviv University); Slava Novgorodov (Tel-Aviv University); Wang-Chiew Tan (University of California at Santa Cruz)

Credit card frauds are unauthorized transactions that are made or attempted by a person or an organization that is not authorized by the card holders. In addition to machine learning-based techniques, credit card companies often employ domain experts to manually specify rules that exploit domain knowledge for improving the detection process. Over time, however, as new (fraudulent and legitimate) transaction arrive, these rules need to be updated and refined to capture the evolving (fraud and legitimate) activity patterns. The goal of the Rudolf system that is demonstrated here is to guide and assist domain experts in this challenging task. Rudolf automatically determines a best set of candidate adaptations to existing rules to capture all fraudulent transactions and, respectively, omit all legitimate transactions. The proposed modifications can then be further refined by domain experts based on their domain knowledge, and the process can be repeated until the experts are satisfied with the resulting rules. Our experimental results on real-life datasets demonstrate the effectiveness and efficiency of our approach. We showcase Rudolf with two demonstration scenarios: detecting credit card frauds and network attacks. Our demonstration will engage the VLDB audience by allowing them to play the role of a security expert, a credit card fraudster, or a network attacker.

### **Ziggy: Characterizing Query Results for Data Explorers**

Thibault Sellam (CWI); Martin Kersten (CWI)

Data exploration has received much attention during the last few years. The aim is to learn interesting new facts from a possibly unfamiliar data set. Typically, explorers operate by trial and error: they write a query, inspect the results and refine their specifications accordingly. In this demo proposal, we present Ziggy, a system to help them understand their query results. Ziggy's aim is to complement an existing exploration system. It assumes that users already have a query in mind, but they do not know what is interesting about it. To assist them, it detects characteristic views, that is, small sets of columns on which the tuples in the results are different from those in the rest of the database. Thanks to these views, our explorers can understand why their selection is unique and make more informed exploration decisions.

### **Blaeu: Mapping and Navigating Large Tables with Cluster Analysis**

Thibault Sellam (CWI); Robin Cijvat (MonetDB Solutions); Richard Koopmanschap (MonetDB Solutions); Martin Kersten (CWI)

Blaeu is an interactive database exploration tool. Its aim is to guide casual users through large data tables, ultimately triggering insights and serendipity. To do so, it relies on a double cluster analysis mechanism. It clusters the data vertically: it detects themes, groups of mutually dependent columns that highlight one aspect of the data. Then it clusters the data horizontally. For each theme, it produces a data map, an interactive visualization of the clusters in the table. The data maps summarize the data. They provide a visual synopsis of the clusters, as well as facilities to inspect their content and annotate them. But they also let the users navigate further. Our explorers can change the active set of columns or drill down into the clusters to refine their selection. Our prototype is fully operational, ready to deliver insights from complex databases.

### **A Demonstration of VisDPT: Visual Exploration of Differentially Private Trajectories**

Xi He (Duke University); Nisarg Raval (Duke University); Ashwin Machanavajhala (Duke University)

The release of detailed taxi trips has motivated numerous useful studies, but has also triggered multiple privacy attacks on individuals' trips. Despite these attacks, no tools are available for systematically analyzing the privacy risk of released trajectory data. While, recent studies have proposed mechanisms to publish synthetic mobility data with provable privacy guarantees, the questions on - 1) how to explain the theoretical privacy guarantee to non-privacy experts; and 2) how well private data preserves the properties of ground truth, remain unclear. To address these issues, we propose a system - VisDPT that provides rich visualization of sensitive information in trajectory databases and helps data curators understand the impact on utility due to privacy preserving mechanisms. We believe VisDPT will enable data curators to take informed decisions while publishing sanitized data.

### **YASK: A Why-Not Question Answering Engine for Spatial Keyword Query Services**

Lei Chen (Hong Kong Baptist University); Jianliang Xu (Hong Kong Baptist University); Christian Jensen (Aalborg University); Yafei Li (Hong Kong Baptist University)

With the proliferation of the mobile use of the web, spatial keyword query (SKQ) services are gaining in importance. However, state-of-the-art SKQ systems do not provide systematic functionality that allows users to ask why some known object is unexpectedly missing from a query result and do not provide an explanation for such missing objects. In this demonstration, we present a system called YASK, a why-not question answering engine for Spatial Keyword query services, that is capable of answering why-not questions posed in response to answers to spatial keyword top-k queries. Two explanation and query refinement models, namely preference adjustment and keyword adaption, are implemented in YASK. The system provides users not only with the reasons why desired objects are missing from query results, but provides also relevant refined queries that revive the expected but missing objects. This demonstration gives attendees hands-on experience with YASK through a map-based GUI interface in which attendees can issue spatial keyword queries, pose why-not questions, and visualize the results.

### **Exploring Databases via Reverse Engineering Ranking Queries with PALEO**

Kiril Panev (TU Kaiserslautern); Sebastian Michel (TU Kaiserslautern); Evica Milchevski (TU Kaiserslautern); Koninika Pal (TU Kaiserslautern)

A novel approach to explore databases using ranked lists is demonstrated. Working with ranked lists, capturing the relative



performance of entities, is a very intuitive and widely applicable concept. Users can post lists of entities for which explanatory SQL queries and full result lists are returned. By refining the input, the results, or the queries, user can interactively explore the database content. The demonstrated system is centered around our PALEO framework for reverse engineering OLAP-style database queries and novel work on mining interesting categorical attributes.

### **ExRank: An Exploratory Ranking Interface**

Ramon Besspinoyong (National University of Singapore); Wei Chen (Zhejiang University); H. V. Jagadish (University of Michigan); Yuxin Ma (Zhejiang University)

Even with simple everyday tasks like online shopping or choosing a restaurant, users are easily overwhelmed with the large number of choices available today, each with a large number of inter-related attributes. We present ExRank, an interactive interface for exploring data that helps users understand the relationship between attribute values and find interesting items in the dataset. Based on a kNN graph and a PageRank algorithm, ExRank suggests which attributes the user should look at, and how expressed choices in particular attributes affect the distribution of values in other attributes for candidate objects. It solves the problem of empty result by showing similar items and when there are too many results, it ranks the data for the user. This demo consists of 1) the description of the software architecture and the user interface 2) the logic and reason behind our solution and 3) a list of demonstration scenarios for showing to the audience.

### **NLProv: Natural Language Provenance**

Daniel Deutch (Tel Aviv University); Nave Frost (Tel Aviv University); Amir Gilad (Tel Aviv University)

We propose to present NLProv: an end-to-end Natural Language (NL) interface for database queries. Previous work has focused on interfaces for specifying NL questions, which are then compiled into queries in a formal language (e.g. SQL). We build upon this work, but focus on presenting a detailed form of the answers in Natural Language. The answers that we present are importantly based on the provenance of tuples in the query result, detailing not only which are the results but also their explanations. We develop a novel method for transforming provenance information to NL, by leveraging the original NL query structure. Furthermore, since provenance information is typically large, we present two solutions for its effective presentation as NL text: one that is based on provenance factorization with novel desiderata relevant to the NL case, and one that is based on summarization.

### **Towards Personalized Maps: Mining User Preferences from Geo-textual Data**

Kaiqi Zhao (Nanyang Technological University); Yiding Liu (Nanyang Technological University); Quan Yuan (Nanyang Technological University); Lisi Chen (Hong Kong Baptist University); Zhida Chen (Nanyang Technological University); Gao Cong (Nanyang Technological University)

Rich geo-textual data is available online and the data keeps increasing at a high speed. We propose two user behavior models to learn several types of user preferences from geo-textual data, and a prototype system on top of the user preference models for mining and search geo-textual data (called PreMiner) to support personalized maps. Different from existing recommender systems and data analysis systems, PreMiner highly personalizes user experience on maps and supports several applications, including user mobility & interests mining, opinion mining in regions, user recommendation, point-of-interest recommendation, and querying and subscribing on geo-textual data.

### **A System for Region Search and Exploration**

Kaiyu Feng (Nanyang Technological University); Kaiqi Zhao (Nanyang Technological University); Yiding Liu (Nanyang Technological University); Gao Cong (Nanyang Technological University)

With the increasing popularity of mobile devices and location based services, massive amount of geo-textual data (e.g., geo-tagged tweets) is being generated everyday. Compared with traditional spatial data, the textual dimension of geo-textual data greatly enriches the data. Meanwhile, the spatial dimension of geo-textual data also adds a semantically rich new aspect to textual data. The large volume, together with its rich semantics, calls for the need for data exploration. First, it has many applications to retrieve a region for exploration that satisfies user-specified conditions (e.g., the size and shape of the region) while maximizing some other conditions (e.g., the relevance to the query keywords of the objects in the region). Second, it is useful to mine and explore the topics of the geo-textual data within a (specified or retrieved) region and perhaps a timespan. This demonstration proposal presents the main ideas of our system, the Region Search and Exploration System (RSE), for efficiently supporting region search and exploration, and our demonstration plan.

### **SigmaKB: Multiple Probabilistic Knowledge Base Fusion**

Miguel Rodriguez (University of Florida); Sean Goldberg (University of Florida); Daisy Zhe Wang (University of Florida)

The interest in integrating web-scale knowledge bases (KBs) has intensified in the last several years. Research has focused on knowledge base completion between two KBs with complementary information, lacking any notion of uncertainty or method of handling conflicting information. We present SigmaKB, a knowledge base system that utilizes Consensus Maximization Fusion and user feedback to integrate and improve the query results of a total of 71 KBs. This paper presents the architecture and demonstration details.

### **Collaborative Crowdsourcing with Crowd4U**

Kosetsu Ikeda (University of Tsukuba); Atsuyuki Morishima (University of Tsukuba); Habibur Rahman (UT Arlington); Senjuti Basu Roy (NJIT); Saravanan Thirumuruganathan (QCRI, HBKU); Sihem Amer-Yahia (CNRS LIG); Gautam Das (UT Arlington)

Collaborative crowdsourcing is an emerging paradigm where a set of workers, often with diverse and complementary skills, form groups and work together to complete complex tasks. While crowdsourcing has been used successfully in many applications, collaboration is essential for achieving a high quality outcome for a number of emerging applications such as text translation, citizen journalism and surveillance tasks. However, no crowdsourcing platform today enables the end-to-end deployment of collaborative tasks. We demonstrate Crowd4U, a volunteer-based system that enables the deployment of diverse crowdsourcing tasks with complex data-flows, in a declarative manner. In addition to treating workers and tasks as rich entities, Crowd4U also provides an easy-to-use form-based task UI. Crowd4U implements worker-to-task assignment algorithms that are appropriate for each kind of task. Once workers are assigned to tasks, appropriate worker collaboration schemes are enforced in order to enable effective result coordination.

### **Demo 3A/3B: Graph and Semistructured Data**

#### **GARUDA: A System for Large-Scale Mining of Statistically Significant Connected Subgraphs**

Satyajit Bhadange (IIT Kanpur); Akhil Arora (Xerox Research Centre India); Arnab Bhattacharya (IIT Kanpur)

Unraveling “interesting” subgraphs corresponding to disease/crime hotspots or characterizing habitation shift patterns is an important graph mining task. With the availability and growth of large-scale real-world graphs, mining for such subgraphs has become the need of the hour for graph miners as well as non-technical end-users. In this demo, we present GARUDA, a system capable of mining large-scale graphs for statistically significant subgraphs in a scalable manner, and provide: (1) a detailed description of the various features and user-friendly GUI of GARUDA; (2) a brief description of the system architecture; and (3) a demonstration scenario for the audience. The demonstration showcases one real graph mining task as well as its ability to scale to large real graphs, portraying speedups of upto 8-10 times over the state-of-the-art MSCS algorithm.

#### **Generating Flexible Workloads for Graph Databases**

Guillaume Bagan (CNRS LIRIS); Angela Bonifati (University of Lyon 1 & CNRS LIRIS); Radu Ciucanu (University of Oxford); George Fletcher (TU Eindhoven); Aurélien Lemay (University of Lille 3 & INRIA); Nicky Advokaat (TU Eindhoven)

Graph data management tools are nowadays evolving at a great pace. Key drivers of progress in the design and study of data intensive systems are solutions for synthetic generation of data and workloads, for use in empirical studies. Current graph generators, however, provide limited or no support for workload generation or are limited to fixed use-cases. Towards addressing these limitations, we demonstrate gMark, the first domain- and query language-independent framework for synthetic graph and query workload generation. Its novel features are: (i) fine-grained control of graph instance and query workload generation via expressive user-defined schemas; (ii) the support of expressive graph query languages, including recursion among other features; and, (iii) selectivity estimation of the generated queries. During the demonstration, we will showcase the highly tunable generation of graphs and queries through various user-defined schemas and targeted selectivities, and the variety of supported practical graph query languages. We will also show a performance comparison of four state-of-the-art graph database engines, which helps us understand their current strengths and desirable future extensions.

#### **Graph databases in the browser: using LevelGraph to explore New Delhi**

Antonio Maccioni (Roma Tre University); Matteo Collina (NearForm)

The pervasiveness of graphs on the Web is growing; however, the difficulty of managing complex graph structures curbs the development of web-oriented applications that embed network data. The open source project, LevelGraph, aims to overcome the obstacles that web developers face with graph data management. LevelGraph is an easy-to-use graph database layer for web applications. To demonstrate various capabilities of the system, we developed a web-based application that utilizes a graph database of a tourist network in New Delhi. The application allows users to move around the city while LevelGraph executes graph queries on the underlying database. In this demonstration, we show how LevelGraph's features facilitate development and maintenance of web applications that embed graph data.

#### **Sapphire: Querying RDF Data Made Simple**

Ahmed El-Roby (University of Waterloo); Khaled Ammar (University of Waterloo); Ashraf Aboulnaga (Qatar Computing Research Institute); Jimmy Lin (University of Waterloo)

There is currently a large amount of publicly accessible structured data available as RDF data sets. For example, the Linked Open Data (LOD) cloud now consists of thousands of RDF data sets with over 30 billion triples, and the number and size of the data sets is continuously growing. Many of the data sets in the LOD cloud provide public SPARQL endpoints to allow issuing queries over them. These endpoints enable users to retrieve data using precise and highly expressive SPARQL queries. However, in order to do so, the user must have sufficient knowledge about the data sets that she wishes to query, that is, the structure of data, the vocabulary used within the data set, the exact values of literals, their data types, etc. Thus, while SPARQL is powerful, it is not easy to use. An alternative to SPARQL that does not require as much prior knowledge of the data is some form of keyword search over the structured data. Keyword

search queries are user-friendly and easy to use, but inherently ambiguous in describing structured queries. In this demonstration, we introduce Sapphire, a system for querying RDF data that strikes a middle ground between ambiguous keyword search and difficult-to-use SPARQL. Sapphire does not replace keyword search or SPARQL, but utilizes both where they are most effective. Sapphire helps the user construct expressive SPARQL queries that represents her information needs, while not requiring her to have detailed knowledge about the queried data sets. These queries are then executed over public SPARQL endpoints from the LOD cloud. Sapphire guides the user in the query writing process by showing suggestions of query terms based on the queried data, and by recommending changes to the query based on a predictive user model.

### **December: A Declarative Tool for Crowd Member Selection**

Yael Amsterdamer (Bar Ilan University); Tova Milo (Tel-Aviv University); Amit Somech (Tel-Aviv University); Brit Youngmann (Tel-Aviv University)

Adequate crowd selection is an important factor in the success of crowdsourcing platforms, increasing the quality and relevance of crowd answers and their performance in different tasks. The optimal crowd selection can greatly vary depending on properties of the crowd and of the task. To this end, we present December, a declarative platform with novel capabilities for flexible crowd selection. December supports the personalized selection of crowd members via a dedicated query language Member-QL. This language enables specifying and combining common crowd selection criteria such as properties of the crowd member's profile and history, similarity between profiles in specific aspects and relevance of the member to a given task. This holistic, customizable approach differs from previous work that has mostly focused on dedicated algorithms for crowd selection in specific settings. To allow for efficient query execution, we implement in December novel algorithms based on our generic, semantically-aware definitions of crowd member similarity and expertise. We demonstrate the effectiveness of December and Member-QL by using the VLDB community as the crowd members and allowing conference participants to use the system for choosing among these members for different purposes and in different contexts.

### **AutoG: A Visual Query Autocompletion Framework for Graph Databases**

Peipei Yi (Hong Kong Baptist University); Byron Choi (Hong Kong Baptist University); Sourav Bhowmick (Nanyang Technological University); Jianliang Xu (Hong Kong Baptist University)

Composing queries is evidently a tedious task. This is particularly true of graph queries as they are typically complex and prone to errors, compounded by the fact that graph schemas can be missing or too loose to be helpful for query formulation. Despite the great success of query formulation aids, in particular, automatic query completion, graph query autocompletion has received much less attention from the data management community. In this demonstration, we present a novel interactive visual subgraph query autocompletion framework called AutoG which alleviates the potentially painstaking task of graph query formulation. Specifically, given a visual query fragment  $q$  formulated by a user, AutoG returns ranked query suggestions  $Q'$  as output at interactive time. Users may choose a query from  $Q'$  and iteratively apply AutoG to compose their queries. We demonstrate various features of AutoG and its superior ability to generate high quality suggestions to aid visual subgraph query formulation.

### **Exploratory Querying of Extended Knowledge Graphs**

Mohamed Yahya (Max-Planck Institute for Informatics, Germany); Klaus Berberich (Max-Planck Institute for Informatics, Germany); Maya Ramanath (IIT Delhi); Gerhard Weikum (Max-Planck Institute for Informatics, Germany)

Knowledge graphs (KGs) are important assets for search, analytics, and recommendations. However, querying a KG to explore entities and discover facts is difficult and tedious, even for users with skills in SPARQL. First, users are not familiar with the structure and labels of entities, classes and relations. Second, KGs are bound to be incomplete, as they capture only major facts about entities and their relationships and miss out on many of the subtler aspects. We demonstrate TriniT, a system that facilitates exploratory querying of large KGs, by addressing these issues of "vocabulary" mismatch and KG incompleteness. TriniT supports query relaxation rules that are invoked to allow for relevant answers which are not found otherwise. The incompleteness issue is addressed by extending a KG with additional text-style token triples obtained by running Open IE on Web and text sources. The query language, relaxation methods, and answer ranking are extended appropriately. The demo shows automatic query relaxation and has support for interactively adding user-customized relaxations. In both situations, the demo provides answer explanations and offers additional query suggestions.

### **SPARQLByE: Querying RDF data by example**

Gonzalo Diaz (University of Oxford); Marcelo Arenas (PUC Chile); Michael Benedikt (University of Oxford)

Semantic Web technologies such as RDF and its query language, SPARQL, offer the possibility of opening up the use of public datasets to a great variety of ordinary users. However, a key obstacle to the consumption of open data is the unfamiliarity of users with the structure of data, as well as their unfamiliarity with SPARQL. To deal with these issues, we introduce a system for querying RDF data by example. At its core is a technique for reverse-engineering SPARQL queries by example. We demonstrate how reverse engineering along with other techniques, such as query relaxation, enables our system, SPARQLByE, to guide users who are unfamiliar with both the dataset and with SPARQL to the desired query and result set.

### **Graph-based Exploration of Non-graph Datasets**

Udayan Khurana (IBM Research - T.J. Watson); Srinivasan Parthasarathy (IBM Research - T.J. Watson); Deepak Turaga (IBM Research - T.J. Watson)

Graphs or networks provide a powerful abstraction to view and analyze relationships among different entities present in a dataset. However, much of the data of interest to analysts and data scientists resides in non-graph forms such as relational databases, JSON, XML, CSV and text. The effort and skill required in identifying and extracting the relevant graph representation from data is often prohibitive and limits a wider adoption of graph-based analysis of non-graph data. In this paper, we demonstrate our system called GraphViewer, for accelerated graph-based exploration and analysis. It automatically discovers relevant graphs implicit within a given non-graph dataset using a set of novel rule-based and data-driven techniques, and optimizes their extraction and storage. It computes several node and graph level metrics and detects anomalous entities in data. Finally, it summarizes the results to support interpretation by a human analyst. While the system automates the computationally intensive aspects of the process, it is engineered to leverage human domain expertise and instincts to fine tune the data exploration process.

### **Rogas: A Declarative Framework for Network Analytics**

Minjian Liu (The Australian National University); Qing Wang (The Australian National University)

Network analytics has become increasingly popular in recent years. Various graph systems have been developed for analysing networks, while network data is still largely stored and managed in relational database systems in the first place. As two separate systems are often used to manage and analyse network data, it not only increases the difficulty for users to learn and maintain these different systems simultaneously, but also impedes performing more sophisticated analysis on relational and topological properties of network data. Aiming to tackle these issues, we present Rogas in this paper, which is a declarative framework that allows the user to formulate analysis queries naturally without thinking about the tedious implementation details of graph algorithms and query processing.

### **Large-scale Complex Analytics on Semi-structured Datasets using AsterixDB and Spark**

Wail Alkawaileet (CCEE at KACST and MIT); Sattam Alsubaiee (CCEE at KACST and MIT); Michael Carey (University of California at Irvine); Till Westmann (Couchbase Inc.); Yingyi Bu (Couchbase Inc.)

Large quantities of raw data are being generated by many different sources in different formats. Private and public sectors alike acclaim the valuable information and insights that can be mined from such data to better understand the dynamics of everyday life, such as traffic, worldwide logistics, and social behavior. For this reason, storing, managing, and analyzing "Big Data" at scale is getting a tremendous amount of attention, both in academia and industry. In this paper, we demonstrate the power of a parallel connection that we have built between Apache Spark and Apache AsterixDB (Incubating) to enable complex analytics such as machine learning and graph analysis on data drawn from large semi-structured data collections. The integration of these two systems allows researchers and data scientists to leverage AsterixDB capabilities, including fast ingestion and indexing of semi-structured data and efficient answering of geo-spatial and fuzzy text queries. Complex data analytics can then be performed on the resulting AsterixDB query output in order to obtain additional insights by leveraging the power of Spark's machine learning and graph libraries.





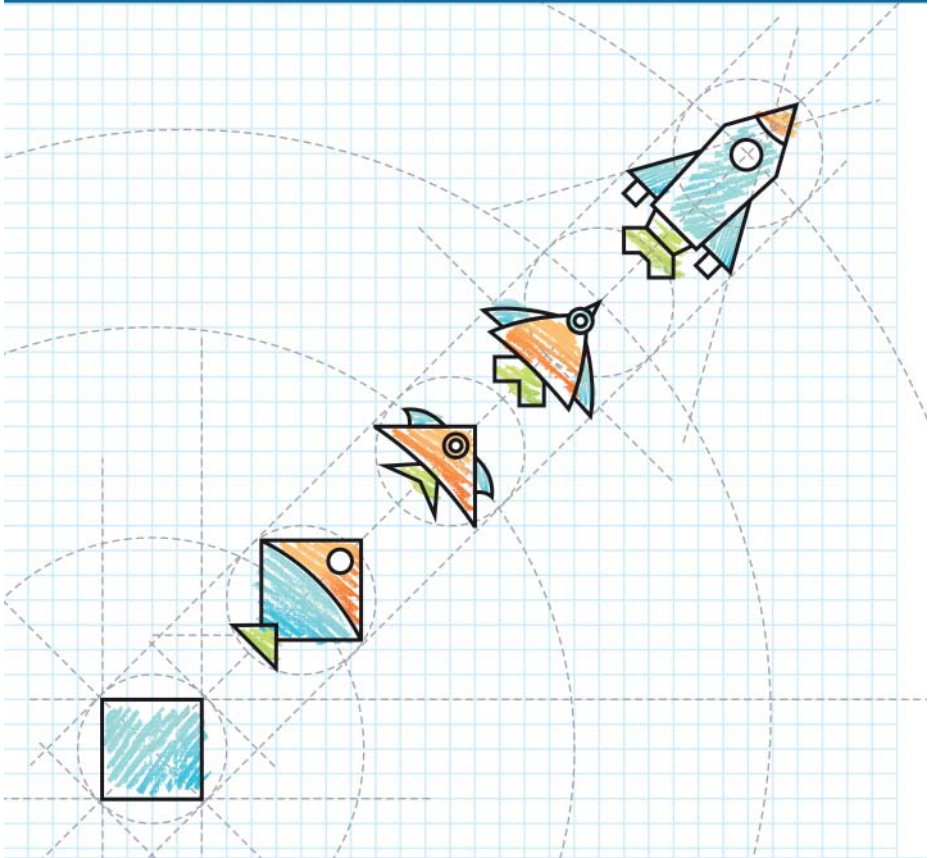
# Sponsors

# Continuous Transformation by Design



PERSISTENT

Join the team that obsessively focuses on the "how" of digital transformation, the ideas of API-centericity, design and data-driven applications for building software-driven businesses.



[www.persistent.com](http://www.persistent.com)



# THE DATABASE, REDEFINED

## ANNOUNCING COUCHBASE SERVER 4.5

The World's Most Powerful  
and Complete Database.

Scalability of NoSQL. Flexibility of JSON. Power of SQL.







# your future made with IBM

## IBM Research is hiring

We live in a moment of remarkable change and opportunity. Data and technology are transforming industries and societies, ushering in a new era of Cognitive Computing. IBM Research is a leader in this worldwide transformation, building on a long history of innovation.

For more than seven decades, IBM Research has defined the future of technology. Our scientists, among them six Nobel Laureates and six Turing Award winners, have produced ten U.S. National Medals of Technology and five U.S. National Medals of Science. Along the way we helped put a man on the moon, defeated Kasparov at chess, and built a *Jeopardy!*<sup>®</sup> champion named Watson.

At IBM, you can achieve what others think is impossible. And in doing so, you'll play a significant role in shaping the future. Join us and discover what you can make of this moment.

What will you make with IBM?  
[ibm.com/jobs](http://ibm.com/jobs)



Infosys®

Infosys Information Platform (IIP)  
is now available on **Amazon Web  
Services (AWS)** Marketplace



Run your **analytics** for just a few dollars per hour



Set up **Infosys Information Platform (IIP) cluster** in just an hour



Get **insights** real time without losing business continuity

[www.infosys.com/iip](http://www.infosys.com/iip)

# SQL Server 2016

Use your skills now to work with SQL Server's new features: on-premises, in the cloud, and in between

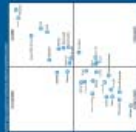


## Ready to leap ahead

SQL Server 2016 has prominent, enterprise-ready features. You are ready to embrace with SQL Server on Microsoft Azure, and with Azure SQL Database. Start benefiting today from this new world of SQL Server.

### Leader

SQL Server 2016 is the only database that has been named a Leader in the Gartner Magic Quadrant for Data Warehouse Management Systems, Q3 2015.



### Cloud first

SQL Server 2016 is designed to be cloud first. It's the only database that can run in the cloud, on-premises, or in a hybrid environment. It's the only database that can run in the cloud, on-premises, or in a hybrid environment.

### SQL Server on Azure virtual machines

SQL Server 2016 is designed to be cloud first. It's the only database that can run in the cloud, on-premises, or in a hybrid environment. It's the only database that can run in the cloud, on-premises, or in a hybrid environment.

### Azure SQL Database

SQL Server 2016 is designed to be cloud first. It's the only database that can run in the cloud, on-premises, or in a hybrid environment. It's the only database that can run in the cloud, on-premises, or in a hybrid environment.

Search for SQL Server 2016 on [msdn.microsoft.com](http://msdn.microsoft.com)

© 2015 Microsoft Corporation. All rights reserved. | Contact Us | Feedback | Home | <http://www.microsoft.com>

## Hybrid + hyperscale

SQL Server 2016 and Microsoft Azure work together to provide new ways to store and analyze data.



## High availability + disaster recovery

Reduce your business failure or unexpected interruption by reducing risks to secondary locations. Streamline backup from a previous time to recover from an event.



## Real-time + historical analytics

Analyze data and historical data in real time. Perform operational analysis with increased efficiency and performance of the OLTP operations. Query data from any point in time.



## Security

SQL Server 2016 offers a full suite of tools for ensuring your data is secure and your system is protected from threats.



## Advanced analytics

SQL Server 2016 integrates the language and scalable parallel processing directly into the database engine. This allows you to analyze data in real time, without the need for moving it, and then machine models to form data to the cloud for an advanced analytics to their applications.



## Beyond relational

With the world data types in all forms. Simplify big data using PolyBase + Azure HDInsight.



## Enterprise information management

SQL Server 2016 offers a full suite of tools for ensuring your data is secure and your system is protected from threats.



## Business intelligence

SQL Server 2016 helps your business insights faster and more easily through improved data integration and analysis. It's the only database that can run in the cloud, on-premises, or in a hybrid environment.

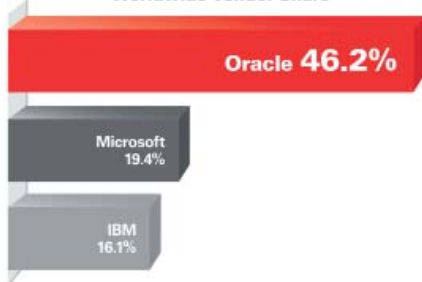


**STILL**

**#1**

# Database

Worldwide Vendor Share



## Oracle Database

Trusted by 310,000 Customers Worldwide

**ORACLE®**

[oracle.com/database](http://oracle.com/database)  
or call 1.800.ORACLE.1



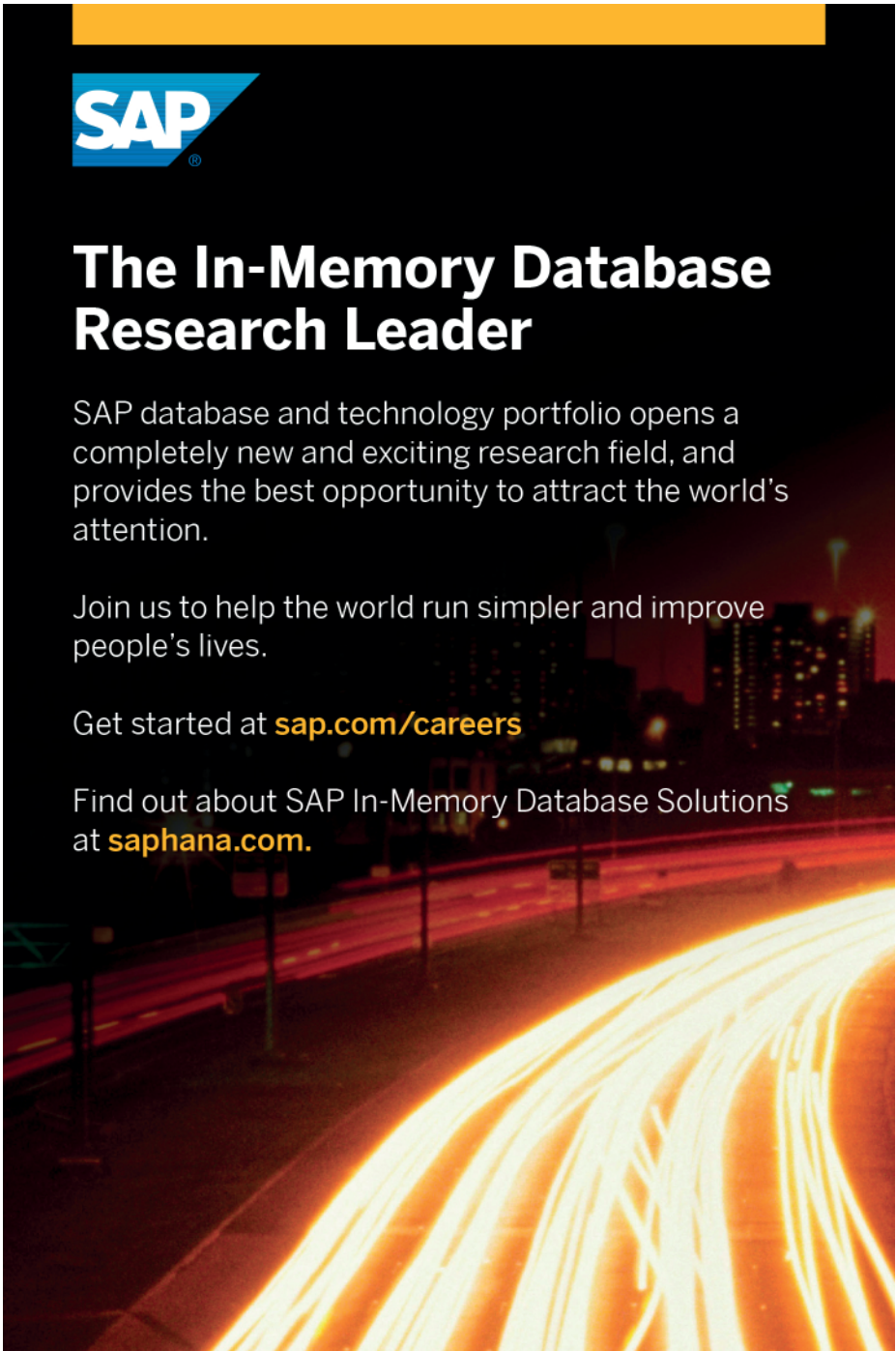
# The In-Memory Database Research Leader

SAP database and technology portfolio opens a completely new and exciting research field, and provides the best opportunity to attract the world's attention.

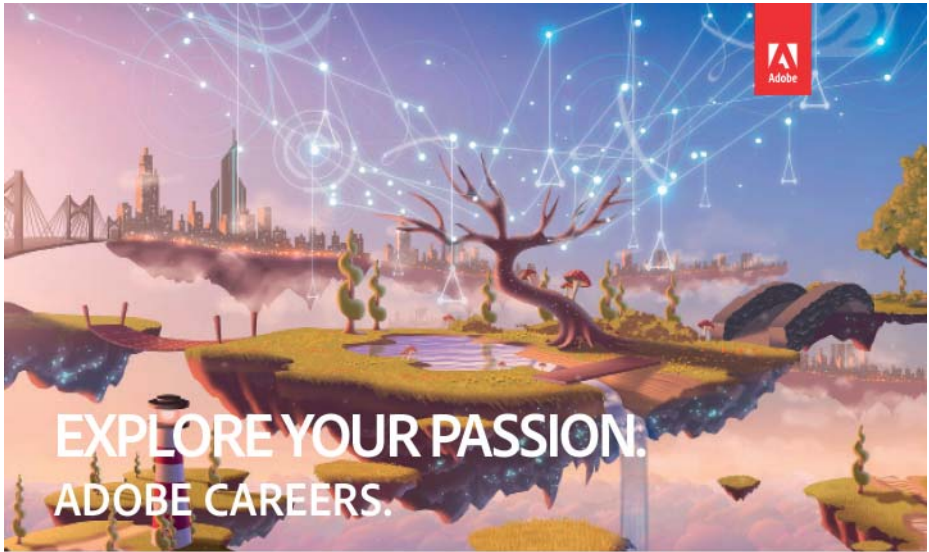
Join us to help the world run simpler and improve people's lives.

Get started at [sap.com/careers](http://sap.com/careers)

Find out about SAP In-Memory Database Solutions at [saphana.com](http://saphana.com).







# EXPLORE YOUR PASSION. ADOBE CAREERS.

Adobe's Big Data Experience Lab has exciting opportunities for PhD researchers in Bangalore & San Jose. Explore below Research Opportunities at Big Data Experience Lab

## **Machine Learning & Data Mining**

## **Big Data Systems and Algorithms**

## **Information & Data Visualization**

## **NLP & Information Retrieval**

We're always looking for individuals who share our values: genuine, exceptional, innovative and involved.

### **Stay connected**

- Subscribe to our Adobe Life blog at <http://blogs.adobe.com/adobelife>
- Follow Adobe on LinkedIn
- Subscribe to our AdobeCareers YouTube channel

### **How to apply**

- Learn more about career opportunities with Adobe's Big Data Experience Lab and apply online at <http://linkd.in/1AJtFIS>
- Connect with our team via email at [atli-hiring@adobe.com](mailto:atli-hiring@adobe.com)

Adobe Systems Incorporated  
345 Park Avenue  
San Jose, CA 95110-2704 USA  
[www.adobe.com](http://www.adobe.com)

Adobe and the Adobe logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.  
© 2016 Adobe Systems Incorporated. All rights reserved.  
91053267 03/16

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're 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".



Clone out our code  
at [github.com/bloomberg](https://github.com/bloomberg)

[techatbloomberg.com](https://techatbloomberg.com)

**Bloomberg**





## Research at Facebook

Giving people the power to share and connect requires constant innovation. At Facebook, research permeates everything we do. We believe the most interesting research questions are derived from real world problems. Working on cutting edge research with a practical focus, we push product boundaries every day. At the same time, we are publishing papers, giving talks, attending and hosting conferences, and collaborating with the academic community.

### We work to solve real-world problems that impact millions of people in areas such as:

- Artificial Intelligence
- Compilers
- Computer Vision
- Data Mining
- Databases
- Distributed Systems
- Economics and Computation
- Human-Computer Interaction
- Machine Learning
- Natural Language Processing
- Network
- Operating Systems
- Programming Languages
- Security / Privacy
- Social Computing
- Speech Technologies



[research.facebook.com/programs](https://research.facebook.com/programs)

Graduate Fellowship  
Research Collaborations  
Conference Sponsorships  
Internships, Full Time Positions, Post Docs  
Faculty Sabbaticals, Invited Talks, and Visiting Scholars

[research.facebook.com](https://research.facebook.com)





**Hewlett Packard**  
Enterprise

# Accelerate our ascent... and yours

Want to work on a data analytics platform that responds so quickly that incredulous prospects believe it must be broken? A platform that returns query results in seconds, instead of hours, days, or even months?

We're **HPE Vertica**, a Boston-based and rapidly growing Big Data analytics platform team. Trusted by hundreds of leading data-driven organizations like AT&T, Guess, and Twitter, HPE Vertica is the most advanced SQL database analytics platform built from the very first line of code to address the most demanding Big Data analytics initiatives.

Join our talented team of engineers solve interesting problems in distributed systems, fault tolerance, optimization, and machine learning. It's our job to help customers gain insights from petabytes of data easily and quickly. If you want to work with top-notch engineers in tackling the most important Big Data challenges in the industry, stop by our booth or visit

**[www.vertica.com/about/careers/](http://www.vertica.com/about/careers/)**

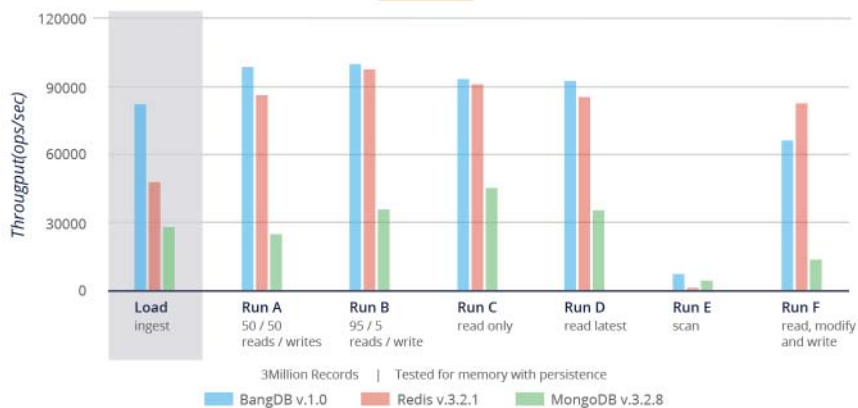
# Why use different servers for distribution, storage, messaging, analytics, & orchestration?

When IQLECT, a next gen real-time data analytics platform gives you the power to Ingest, Process, Store, Visualize, Act, Correlate and Notify - All-in-box scale out converged solution that simplifies, accelerates and humanize the big data analysis.

## Inhouse Technology stack overview



## YCSB benchmark results: BangDB vs MongoDB vs Redis



IQLECT is built on a proprietary NoSQL persistent DB known as BangDB. The goal of BangDB is to be a fast, reliable, robust, scalable and easy to use data store for various data management services required by applications.

MongoDB is a popular open-source database, and Redis is a widely used in-memory data store. It is interesting and exciting to note that BangDB outperforms both MongoDB and Redis, across a range of different loads.





Our mission is simple:

Connect the world's professionals to  
make them more productive and successful

Learn more at:

[linkedin.com/careers](https://linkedin.com/careers)

[data.linkedin.com](https://data.linkedin.com)



# MAKE DIGITAL



# REAL



**\$715 M**  
Annual Revenue



**16,000+**  
Mindtree Minds



**343**  
Active Clients



**42** Offices Across  
the Globe



Banking,  
Financial Services  
& Insurance



CPG, Retail &  
Manufacturing



Technology  
& Media



Travel,  
Transportation  
& Hospitality



E-commerce



Marketing



Data Analytics



Cloud



Mobility

Mindtree [NSE: MINDTREE] delivers digital transformation and technology services from ideation to execution, enabling Global 2000 clients to outperform the competition. "Born digital," Mindtree takes an agile, collaborative approach to creating customized solutions across the digital value chain. At the same time, our deep expertise in infrastructure and applications management helps optimize your IT into a strategic asset. Whether you need to differentiate your company, reinvent business functions or accelerate revenue growth, we can get you there. Visit [www.mindtree.com](http://www.mindtree.com) to learn more.



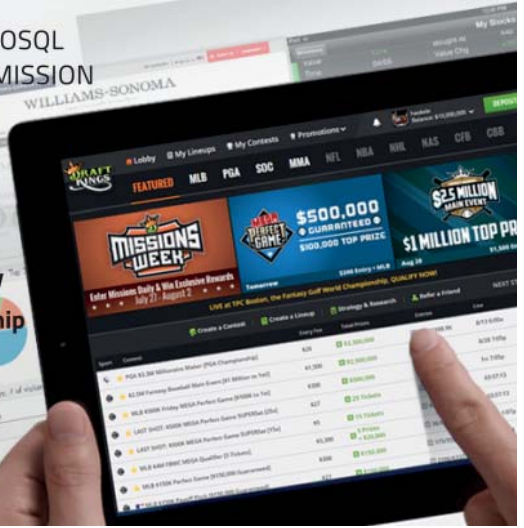
## AEROSPIKE

THE HIGH PERFORMANCE, NOSQL DATABASE FOR REAL-TIME, MISSION CRITICAL APPLICATIONS.

AEROSPIKE

- \* Predictable performance
- \* Highest up time & availability
- \* Lowest Total Cost of Ownership

[www.aerospike.com](http://www.aerospike.com)



# APIs fuel digital disruption & cloud



#DigitalKnowHow  
by **apigee**  
[apigee.com](http://apigee.com)

# Deloitte.

## About Deloitte



Deloitte drives progress. We advance the aims of our clients and their stakeholders, striving to make them leaders wherever they choose to compete. We focus on making a tangible positive difference by combining strategy with action and delivering measurable impact. We form unique collaborations to find smarter insights, innovative solutions and entrepreneurial ways to move ahead. We invest in outstanding people of diverse talents and backgrounds and empower them to achieve more than they could elsewhere. We believe that when our clients succeed, and when society succeeds, so do we.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see [www.deloitte.com/about](http://www.deloitte.com/about) for a more detailed description of DTTL and its member firms.

Deloitte globally provides audit, tax, consulting, and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and high-quality service to clients, delivering the insights they need to address their most complex business challenges. Deloitte has in the region of 210,000 professionals, all committed to becoming the standard of excellence.

In India, Deloitte member firms are spread across 13 locations with nearly 32,000 professionals who take pride in their ability to deliver to clients the right combination of local insight and international expertise.

About Deloitte Touche Tohmatsu India LLP

DTTL member firm, Deloitte Touche Tohmatsu India Private Limited (U74140MH1995PTC093339), a private company limited by shares, was converted into Deloitte Touche Tohmatsu India LLP (LLP Identification No. AAE-8458), a limited liability partnership, with effect from October 1, 2015

(C) 2016 Deloitte Touche Tohmatsu India LLP

# BRINGING ARTIFICIAL INTELLIGENCE TO LIFE



Intel® Architecture-based platforms + advanced math libraries analyze massive data sets to build smarter cities, power intelligent cars, and deliver personalized medicine...and that's just the beginning.

Learn more: [Intel.com/machinelearning](http://Intel.com/machinelearning)

Join us: [Intel.com/jobs](http://Intel.com/jobs)

Copyright © 2016 Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.  
\* Other names and brands may be claimed as the property of others.

# MAPR<sup>®</sup>

The Only

## Converged Data Platform

Real Time.  
Reliable.  
Results.

[mapr.com](http://mapr.com)

Hadoop +  
Spark +  
Streaming +  
Database +  
Storage



 MEMSQL

SHIFT AS QUICKLY  
AS TRENDS DO

The Database Platform  
For Real-Time Analytics

[memsql.com](http://memsql.com)





## At Tableau, innovation is key.

We constantly push the boundaries of Visual Analytics. Tableau is expanding its engineering and R&D power to continue to delight customers with the best analytics platform ever built. You can be part of the team that delivers this breakthroughs.

### We are currently hiring:

- Software Engineers in Development
- Software Engineers in Test
- Research Engineers

Visit [careers.tableau.com](http://careers.tableau.com) for all open positions.

### Haven't had a chance to check out Tableau?

Students & Professors download your free trial today at [www.tableau.com/academic](http://www.tableau.com/academic).



Transforming your business while you're busy running it.  
There is a certain way.



IT Services  
Business Solutions  
Consulting  
[www.tcs.com](http://www.tcs.com)

**TATA CONSULTANCY SERVICES**  
Experience certainty.





# Notes

---

# Notes

---

# Notes

---

# Notes

---

# Quick Dial Numbers

---

## Event Helpline

India Attitude ..... +91 98230 98230  
+91 98230 98231

## Taxi Operators

Easy Cab ..... +91 11 43434343  
Mega Cab ..... +91 11 41414141  
Meru Cab ..... +91 11 44224422

## Travel Assistance

Line 1 ..... +91 12 44409376  
Line 2 ..... +91 12 44409363  
Line 3 ..... +91 8527797855

## Leela Ambience Gurgaon Hotel & Residences

Line 1 ..... +91 12 44771234  
Line 2 ..... +91 12 44771235

## Hospitals

Paras Hospital ..... +91 12 44585555  
Privat Hospital ..... +91 12 44324300

## Police Station

DLF Phase 1 ..... +91 12 42567223  
DLF Phase 2 ..... +91 12 42566387

## ORGANIZING SPONSOR



PERSISTENT

## PLATINUM SPONSORS



Microsoft



ORACLE®



Couchbase

## GOLD SPONSORS



Hewlett Packard  
Enterprise



IQLECT



Baidu 百度



Adobe

facebook

LinkedIn

Bloomberg



Mindtree

Welcome to possible

## SILVER SPONSORS



TATA  
CONSULTANCY  
SERVICES



apigee



Deloitte

## BRONZE SPONSORS

TPC

ampol



EXHIBITORS

DIGITAL INDIA EVENT SPONSORS

INTERNET SPONSOR



Springer



acm  
India



de Asra  
Entrepreneur's Agency to Success

