

# 2018 IEEE 34th International Conference on Data Engineering **ICDE 2018**

## Table of Contents

|   |         |
|---|---------|
| Message from the ICDE 2018 Chairs ..... | xxxiii  |
| ICDE 2018 Organizing Committee .....    | xxxvi   |
| ICDE 2018 Program Committees .....      | xxxviii |
| ICDE 2018 Keynotes .....                | xlvi    |

## Keynotes

|   |    |
|---|----|
| Human Factors in Data Science .....   | 1  |
| <i>Sihem Amer-Yahia (CNRS, University Grenoble Alpes)</i>   |    |
| Actor-Oriented Database Systems .....   | 13 |
| <i>Philip A. Bernstein (Microsoft Research)</i>   |    |
| An NVM Carol: Visions of NVM Past, Present, and Future .....  | 15 |
| <i>Margo Seltzer (John A. Paulson School of Engineering and Applied Sciences), Virendra Marathe (Oracle Labs), and Steve Byan (Oracle Labs)</i> |    |
| My Top Ten Fears about the DBMS Field .....   | 24 |
| <i>Michael Stonebraker (Massachusetts Institute of Technology)</i>  |    |

## Research Track: Regular Papers

### Research Session 1: Data Integration, Metadata Management, and Interoperability

|   |    |
|---|----|
| GPH: Similarity Search in Hamming Space .....   | 29 |
| <i>Jianbin Qin (The University of New South Wales), Yaoshu Wang (The University of New South Wales), Chuan Xiao (Nagoya University), Wei Wang (The University of New South Wales), Xuemin Lin (The University of New South Wales), and Yoshiharu Ishikawa (Nagoya University)</i> |    |
| Extracting Syntactical Patterns from Databases .....  | 41 |
| <i>Andrew Ilyas (Massachusetts Institute of Technology), Joana M. F. da Trindade (Massachusetts Institute of Technology), Raul Castro Fernandez (Massachusetts Institute of Technology), and Samuel Madden (Massachusetts Institute of Technology)</i>                            |    |

|   |    |
|---|----|
| Schema-Agnostic Progressive Entity Resolution .....   | 53 |
| <i>Giovanni Simonini (Università di Modena e Reggio Emilia), George Papadakis (University of Athens), Themis Palpanas (Paris Descartes University), and Sonia Bergamaschi (Università di Modena e Reggio Emilia)</i>  |    |
| Ares: Automatic Disaggregation of Historical Data .....   | 65 |
| <i>Fan Yang (University of Pittsburgh), Hyun Ah Song (Carnegie Mellon University), Zongge Liu (Carnegie Mellon University), Christos Faloutsos (Carnegie Mellon University), Vladimir Zadorozhny (University of Pittsburgh), and Nicholas Sidiropoulos (University of Virginia)</i> |    |

## Research Session 2: Data Exploration and Visualization

|  |     |
|--|-----|
| Augmented Access for Querying and Exploring a Polystore .....  | 77  |
| <i>Antonio Maccioni (Collective[ij]) and Riccardo Torlone (Roma Tre University)</i>  |     |
| What-If Analysis with Conflicting Goals: Recommending Data Ranges for Exploration .....  | 89  |
| <i>Quoc Viet Hung Nguyen (Griffith University), Kai Zheng (University of Electronic Science and Technology of China), Matthias Weidlich (Humboldt-Universität zu Berlin), Bolong Zheng (Aalborg University), Hongzhi Yin (The University of Queensland), Thanh Tam Nguyen (Ecole Polytechnique Federale de Lausanne), and Bela Stantic (Griffith University)</i> |     |
| DeepEye: Towards Automatic Data Visualization .....  | 101 |
| <i>Yuyu Luo (Tsinghua University), Xuedi Qin (Tsinghua University), Nan Tang (Qatar Computing Research Institute), and Guoliang Li (Tsinghua University)</i>   |     |
| Towards a Holistic Integration of Spreadsheets with Databases: A Scalable Storage Engine for Presentational Data Management .....  | 113 |
| <i>Mangesh Bendre (University of Illinois at Urbana-Champaign), Vipul Venkataraman (University of Illinois at Urbana-Champaign), Xinyan Zhou (University of Illinois at Urbana-Champaign), Kevin Chang (University of Illinois at Urbana-Champaign), and Aditya Parameswaran (University of Illinois at Urbana-Champaign)</i>                                    |     |

## Research Session 3: Data Mining and Knowledge Discovery

|   |     |
|---|-----|
| Rule Sharing for Fraud Detection via Adaptation .....   | 125 |
| <i>Ariel Jarovsky (Tel Aviv University), Tova Milo (Tel Aviv University), Slava Novgorodov (Tel Aviv University), and Wang-Chiew Tan (Recruit Institute of Technology)</i>                              |     |
| DisTenC: A Distributed Algorithm for Scalable Tensor Completion on Spark .....  | 137 |
| <i>Hancheng Ge (Texas A&amp;M University), Kai Zhang (Temple University), Majid Alfifi (Texas A&amp;M University), Xia Hu (Texas A&amp;M University), and James Caverlee (Texas A&amp;M University)</i> |     |

|  |     |
|--|-----|
| A Generic Top-N Recommendation Framework for Trading-Off Accuracy, Novelty, and Coverage ....  | 149 |
| <i>Zainab Zolaktaf (University of British Columbia), Reza Babanezhad (University of British Columbia), and Rachel Pottinger (University of British Columbia)</i> |     |
| Predicting Named Entity Location Using Twitter .....   | 161 |
| <i>Wei Shen (Nankai University), Yanan Liu (Nankai University), and Jianyong Wang (Tsinghua University)</i>  |     |

## Research Session 4: Modern and Distributed Hardware

|   |     |
|---|-----|
| CUB, a Consensus Unit-Based Storage Scheme for Blockchain System .....  | 173 |
| <i>Zihuan Xu (Hong Kong University of Science and Technology), Siyuan Han (Hong Kong University of Science and Technology), and Lei Chen (Hong Kong University of Science and Technology)</i>     |     |
| LeanStore: In-Memory Data Management beyond Main Memory .....   | 185 |
| <i>Viktor Leis (Technische Universität München), Michael Haubenschild (Tableau Software), Alfons Kemper (Technische Universität München), and Thomas Neumann (Technische Universität München)</i> |     |
| Adaptive Execution of Compiled Queries .....  | 197 |
| <i>André Kohn (Technische Universität München), Viktor Leis (Technische Universität München), and Thomas Neumann (Technische Universität München)</i>   |     |
| Hybrid Data Layouts for Tiered HTAP Databases with Pareto-Optimal Data Placements .....   | 209 |
| <i>Martin Boissier (Hasso Plattner Institute), Rainer Schlosser (Hasso Plattner Institute), and Matthias Uflacker (Hasso Plattner Institute)</i>  |     |

## Research Session 5: Scalable Analytics, Graph, RDF, Web Data, and Social Networks

|   |     |
|---|-----|
| Mining Density Contrast Subgraphs .....   | 221 |
| <i>Yu Yang (Simon Fraser University), Lingyang Chu (Simon Fraser University), Yanyan Zhang (Simon Fraser University), Zhefeng Wang (Huawei Technologies), Jian Pei (Simon Fraser University &amp; JD.com), and Enhong Chen (University of Science and Technology of China)</i>  |     |
| Efficient Computing of Radius-Bounded k-Cores .....   | 233 |
| <i>Kai Wang (University of New South Wales), Xin Cao (University of New South Wales), Xuemin Lin (University of New South Wales), Wenjie Zhang (University of New South Wales), and Lu Qin (University of Technology Sydney)</i>  |     |
| Efficient Signed Clique Search in Signed Networks .....   | 245 |
| <i>Rong-Hua Li (Beijing Institute of Technology), Qiangqiang Dai (Shenzhen University), Lu Qin (University of Technology), Guoren Wang (Beijing Institute of Technology), Xiaokui Xiao (National University of Singapore), Jeffrey Xu Yu (The Chinese University of Hong Kong), and Shaojie Qiao (Chengdu University of Information Technology)</i> |     |

|  |     |
|--|-----|
| Scalable Hypergraph-Based Image Retrieval and Tagging System .....   | 257 |
| <i>Lu Chen (Zhejiang University), Yunjun Gao (Zhejiang University),<br/>Yuanliang Zhang (Zhejiang University), Sibao Wang (The University of<br/>Queensland), and Baihua Zheng (Singapore Management University)</i> |     |

## Research Session 6: Temporal, Spatial, Mobile, and Multimedia Data

|  |     |
|--|-----|
| Waterwheel: Realtime Indexing and Temporal Range Query Processing over Massive Data<br>Streams .....   | 269 |
| <i>Li Wang (Advanced Digital Sciences Center), Ruichu Cai (Guangdong<br/>University of Technology), Tom Z. J. Fu (Advanced Digital Sciences<br/>Center), Jiong He (Advanced Digital Sciences Center), Zijie Lu<br/>(Guangdong University of Technology), Marianne Winslett (University of<br/>Illinois Urbana-Champaign), and Zhenjie Zhang (Advanced Digital<br/>Sciences Center)</i> |     |
| Incremental Graph Pattern Based Node Matching .....  | 281 |
| <i>Guohao Sun (Macquarie University), Guanfeng Liu (Soochow University),<br/>Yan Wang (Macquarie University), Mehmet A. Orgun (Macquarie<br/>University), and Xiaofang Zhou (The University of Queensland)</i>   |     |
| On Spatial Pattern Matching .....  | 293 |
| <i>Yixiang Fang (The University of Hong Kong), Reynold Cheng (The<br/>University of Hong Kong), Gao Cong (Nanyang Technological University),<br/>Nikos Mamoulis (University of Ioannina), and Yun Li (Nanjing<br/>University)</i>  |     |
| FAST: Frequency-Aware Indexing for Spatio-Textual Data Streams .....   | 305 |
| <i>Ahmed R. Mahmood (Purdue University), Ahmed M. Aly (Google Inc.), and<br/>Walid G. Aref (Purdue University)</i>   |     |

## Research Session 7: Crowdsourcing

|  |     |
|--|-----|
| Latency-Oriented Task Completion via Spatial Crowdsourcing .....   | 317 |
| <i>Yuxiang Zeng (Hong Kong University of Science and Technology), Yongxin<br/>Tong (Beihang University), Lei Chen (Hong Kong University of Science<br/>and Technology), and Zimu Zhou (ETH Zurich)</i> |     |
| Realtime Traffic Speed Estimation with Sparse Crowdsourced Data .....  | 329 |
| <i>Zheng Liu (Hong Kong University of Science and Technology), Lei Chen<br/>(Beihang University), and Yongxin Tong (Beihang University)</i>  |     |
| Incentive-Based Entity Collection Using Crowdsourcing .....  | 341 |
| <i>Chengliang Chai (Tsinghua University), Ju Fan (Renmin University of<br/>China), and Guoliang Li (Tsinghua University)</i>   |     |
| Top-k Query Processing on Encrypted Databases with Strong Security Guarantees .....  | 353 |
| <i>Xianrui Meng (Amazon Web Services), Haohan Zhu (Facebook Inc.), and<br/>George Kollios (Boston University)</i>  |     |
| Task Relevance and Diversity as Worker Motivation in Crowdsourcing .....   | 365 |
| <i>Julien Pilourdault (Universite Grenoble Alpes), Sihem Amer-Yahia<br/>(CNRS, University Grenoble Alpes), Senjuti Basu Roy (NJIT), and<br/>Dongwon Lee (Penn State University)</i>                    |     |

## Research Session 8: Potpourri

|  |     |
|--|-----|
| Sensitive Task Assignments in Crowdsourcing Markets with Colluding Workers .....   | 377 |
| <i>Haipai Sun (Stevens Institute of Technology), Boxiang Dong (Montclair State University), Bo Zhang (Stevens Institute of Technology), Wendy Hui Wang (Stevens Institute of Technology), and Murat Kantarcioglu (The University of Texas at Dallas)</i>                   |     |
| RStore: A Distributed Multi-Version Document Store .....   | 389 |
| <i>Souvik Bhattacharjee (University of Maryland) and Amol Deshpande (University of Maryland)</i>   |     |
| Anna: A KVS for Any Scale .....  | 401 |
| <i>Chenggang Wu (UC Berkeley), Jose Faleiro (Yale University), Yihan Lin (Columbia University), and Joseph Hellerstein (UC Berkeley)</i>   |     |
| Discovering Mis-Categorized Entities .....   | 413 |
| <i>Shuang Hao (Tsinghua University), Nan Tang (QCRI), Guoliang Li (Tsinghua University), and Jianhua Feng (Tsinghua University)</i>  |     |
| Improving Spatial Data Processing by Clipping Minimum Bounding Boxes .....   | 425 |
| <i>Darius Sidlauskas (Ecole Polytechnique Federale de Lausanne), Sean Chester (Norwegian University of Science and Technology), Eleni Tzirita Zacharitou (Ecole Polytechnique Federale de Lausanne), and Anastasia Ailamaki (Ecole Polytechnique Federale de Lausanne)</i> |     |

## Research Session 9: Modern Hardware and In-Memory Database Systems

|   |     |
|---|-----|
| Accelerating Concurrent Workloads with CPU Cache Partitioning .....   | 437 |
| <i>Stefan Noll (TU Dortmund University), Jens Teubner (TU Dortmund University), Norman May (SAP SE), and Alexander Böhm (SAP SE)</i>  |     |
| LevelHeaded: A Unified Engine for Business Intelligence and Linear Algebra Querying .....   | 449 |
| <i>Christopher Aberger (Stanford University), Andrew Lamb (Stanford University), Kunle Olukotun (Stanford University), and Christopher Re (Stanford University)</i>   |     |
| Easy Lock-Free Indexing in Non-Volatile Memory .....  | 461 |
| <i>Tianzheng Wang (University of Toronto), Justin Levandoski (Microsoft Research), and Per-Ake Larson (University of Waterloo)</i>  |     |
| An Elastic Multi-Core Allocation Mechanism for Database Systems .....   | 473 |
| <i>Simone Dominico (Federal University of Paraná), Eduardo Cunha de Almeida (Federal University of Paraná), Jorge Augusto Meira (University of Luxembourg), and Marco Antonio Zanata Alves (Federal University of Paraná)</i> |     |

## Research Session 10: Data Mining, Knowledge Discovery, and Interaction

|   |     |
|---|-----|
| Adaptive Lightweight Regularization Tool for Complex Analytics .....  | 485 |
| <i>Zhaojing Luo (National University of Singapore), Shaofeng Cai (National University of Singapore), Jinyang Gao (National University of Singapore), Meihui Zhang (Beijing Institute of Technology), Kee Yuan Ngiam (National University Health System), Gang Chen (Zhejiang University), and Wang-Chien Lee (Pennsylvania State University)</i>  |     |
| Efficient Learning Interpretable Shapelets for Accurate Time Series Classification .....  | 497 |
| <i>Zicheng Fang (Fudan University), Peng Wang (Fudan University), and Wei Wang (Fudan University)</i>   |     |
| Learning Association Relationship and Accurate Geometric Structures for Multi-Type Relational Data .....  | 509 |
| <i>Khanh Luong (Queensland University of Technology) and Richi Nayak (Queensland University of Technology)</i>  |     |
| Generalized Dynamic Time Warping: Unleashing the Warping Power Hidden in Point-Wise Distances .....   | 521 |
| <i>Rodica Neamtu (Worcester Polytechnic Institute), Ramoza Ahsan (Worcester Polytechnic Institute), Elke A. Rundensteiner (Worcester Polytechnic Institute), Gabor Sarkozy (Worcester Polytechnic Institute), Eamonn Keogh (University of California Riverside), Hoang Anh Dau (University of California Riverside), Cuong Nguyen (Worcester Polytechnic Institute), and Charles Lovering (Worcester Polytechnic Institute)</i> |     |

## Research Session 11: Scalable Analytics, Graph, RDF, Web Data, and Social Networks

|  |     |
|--|-----|
| An Efficient Probabilistic Approach for Graph Similarity Search .....  | 533 |
| <i>Zijian Li (Hong Kong University of Science and Technology), Xun Jian (Hong Kong University of Science and Technology), Xiang Lian (Kent State University), and Lei Chen (Hong Kong University of Science and Technology)</i>        |     |
| Efficient SimRank Tracking in Dynamic Graphs .....   | 545 |
| <i>Yue Wang (Hong Kong University of Science and Technology), Xiang Lian (Kent State University), and Lei Chen (Hong Kong University of Science and Technology)</i>  |     |
| Efficiently Reinforcing Social Networks over User Engagement and Tie Strength .....  | 557 |
| <i>Fan Zhang (University of New South Wales), Ying Zhang (University of Technology Sydney), Lu Qin (University of Technology Sydney), Wenjie Zhang (University of New South Wales), and Xuemin Lin (University of New South Wales)</i> |     |
| Finding Top-k Optimal Sequenced Routes .....   | 569 |
| <i>Huiping Liu (East China Normal University), Cheqing Jin (East China Normal University), Bin Yang (Aalborg University), and Aoying Zhou (East China Normal University)</i>   |     |

## Research Session 12: Miscellaneous Query Processing

|  |     |
|--|-----|
| Interactive Inference of SPARQL Queries Using Provenance .....   | 581 |
| <i>Efrat Abramovitz (Tel Aviv University), Daniel Deutch (Tel Aviv University), and Amir Gilad (Tel Aviv University)</i>   |     |
| Parallel Reasoning of Graph Functional Dependencies .....  | 593 |
| <i>Wenfei Fan (University of Edinburgh), Xueli Liu (Harbin Institute of Technology), and Yingjie Cao (Beihang University)</i>  |     |
| Range Optimistic Concurrency Control for a Composite OLTP and Bulk Processing Workload .....   | 605 |
| <i>Jiahao Wang (East China Normal University), Peng Cai (East China Normal University), Jinwei Guo (East China Normal University), Weining Qian (East China Normal University), and Aoying Zhou (East China Normal University)</i>                       |     |
| Deep Representation Learning for Trajectory Similarity Computation .....   | 617 |
| <i>Xiucheng Li (Nanyang Technological University), Kaiqi Zhao (Nanyang Technological University), Gao Cong (Nanyang Technological University), Christian S. Jensen (Aalborg University), and Wei Wei (Huazhong University of Science and Technology)</i> |     |

## Research Session 13: Query Processing, Indexing, and Optimization

|  |     |
|--|-----|
| Continuous Proximity Detection via Predictive Safe Region Construction .....   | 629 |
| <i>Ying Xu (National University of Defense Technology), Dongxiang Zhang (UESTC), Meihui Zhang (Beijing Institute of Technology), Dongsheng Li (National University of Defense Technology), Xiaoling Wang (East China Normal University), and Heng Tao Shen (UESTC)</i> |     |
| A Comprehensive Performance Evaluation of Modern In-Memory Indices .....   | 641 |
| <i>Zhongle Xie (National University of Singapore), Qingchao Cai (National University of Singapore), Gang Chen (Zhejiang University), Rui Mao (Shenzhen University), and Meihui Zhang (Beijing Institute of Technology)</i>   |     |
| Skyline Diagram: Finding the Voronoi Counterpart for Skyline Queries .....   | 653 |
| <i>Jinfei Liu (Emory University &amp; Georgia Institute of Technology), Juncheng Yang (Emory University), Li Xiong (Emory University), Jian Pei (JD.com &amp; Simon Fraser University), and Jun Luo (Lenovo &amp; Chinese Academy of Sciences)</i>                     |     |
| Adaptive Adaptive Indexing .....   | 665 |
| <i>Felix Martin Schuhknecht (Saarland University), Jens Dittrich (Saarland University), and Laurent Linden (Saarland University)</i>   |     |

## Research Session 14: Potpourri

|  |     |
|--|-----|
| Impatience Is a Virtue: Revisiting Disorder in High-Performance Log Analytics .....  | 677 |
| <i>Badrish Chandramouli (Microsoft Research), Jonathan Goldstein (Microsoft Research), and Yinan Li (Microsoft Research)</i> |     |
| Cobra: A Framework for Cost-Based Rewriting of Database Applications .....   | 689 |
| <i>K. Venkatesh Emani (IIT Bombay) and S. Sudarshan (IIT Bombay)</i>   |     |

|  |     |
|--|-----|
| Diversified Coherent Core Search on Multi-Layer Graphs .....   | 701 |
| <i>Rong Zhu (Harbin Institute of Technology), Zhaonian Zou (Harbin Institute of Technology), and Jianzhong Li (Harbin Institute of Technology)</i>         |     |
| A Graph-Theoretic Fusion Framework for Unsupervised Entity Resolution .....  | 713 |
| <i>Dongxiang Zhang (UESTC), Long Guo (Peking University), Xiangnan He (NUS), Jie Shao (UESTC), Sai Wu (Zhejiang University), and Heng Tao Shen (UESTC)</i> |     |

## Research Session 15: Semistructured Data and Pub/Sub

|   |     |
|---|-----|
| Learning Graphical Models from a Distributed Stream .....   | 725 |
| <i>Yu Zhang (Iowa State University), Srikanta Tirthapura (Iowa State University), and Graham Cormode (University of Warwick)</i>  |     |
| Sharon: Shared Online Event Sequence Aggregation .....  | 737 |
| <i>Olga Poppe (Worcester Polytechnic Institute), Allison Rozet (Worcester Polytechnic Institute), Chuan Lei (IBM Research), Elke A. Rundensteiner (Worcester Polytechnic Institute), and David Maier (Portland State University)</i>  |     |
| Location-Aware Top-k Term Publish/Subscribe .....   | 749 |
| <i>Lisi Chen (University of Wollongong), Shuo Shang (King Abdullah University of Science and Technology), Zhiwei Zhang (Hong Kong Baptist University), Xin Cao (University of New South Wales), Christian S. Jensen (Aalborg University), and Panos Kalnis (King Abdullah University of Science and Technology)</i> |     |
| Flexible Aggregate Nearest Neighbor Queries in Road Networks .....  | 761 |
| <i>Bin Yao (Shanghai Jiao Tong University), Zhongpu Chen (Shanghai Jiao Tong University), Xiaofeng Gao (Shanghai Jiao Tong University), Shuo Shang (King Abdullah University of Science and Technology), Shuai Ma (Beihang University), and Minyi Guo (Shanghai Jiao Tong University)</i>                           |     |

## Research Session 16: Scalable Analytics, Graph, RDF, Web Data, and Social Networks

|   |     |
|---|-----|
| Rule-Based Graph Repairing: Semantic and Efficient Repairing Methods .....  | 773 |
| <i>Yurong Cheng (Beijing Institute of Technology), Lei Chen (Hong Kong University of Science and Technology), Ye Yuan (Northeastern University), and Guoren Wang (Beijing Institute of Technology)</i>  |     |
| Exacting Eccentricity for Small-World Networks .....  | 785 |
| <i>Wentao Li (University of Technology Sydney), Miao Qiao (Massey University), Lu Qin (University of Technology Sydney), Ying Zhang (University of Technology Sydney), Lijun Chang (The University of Sydney), and Xuemin Lin (University of New South Wales)</i> |     |
| Persistent Community Search in Temporal Networks .....  | 797 |
| <i>Rong-Hua Li (Beijing Institute of Technology), Jiao Su (The Chinese University of Hong Kong), Lu Qin (Centre for Artificial Intelligence), Jeffrey Xu Yu (Chinese University of Hong Kong), and Qiangqiang Dai (Shenzhen University)</i>                       |     |



|  |     |
|--|-----|
| Multi-Example Search in Rich Information Graphs .....  | 809 |
| <i>Matteo Lissandrini (University of Trento), Davide Mottin (Hasso Plattner Institute), Themis Palpanas (Paris Descartes University), and Yannis Velegrakis (University of Trento)</i> |     |

## Research Session 17: Database Privacy, Security, and Trust

|  |     |
|--|-----|
| PrivTrie: Effective Frequent Term Discovery under Local Differential Privacy .....   | 821 |
| <i>Ning Wang (Northeastern University), Xiaokui Xiao (National University of Singapore), Yin Yang (Hamad Bin Khalifa University), Ta Duy Hoang (Hanoi University of Science and Technology), Hyejin Shin (Samsung Electronics), Junbum Shin (Samsung Electronics), and Ge Yu (Northeastern University)</i> |     |
| Privacy-Preserving Online Task Assignment in Spatial Crowdsourcing with Untrusted Server .....   | 833 |
| <i>Hien To (University of Southern California), Cyrus Shahabi (University of Southern California), and Li Xiong (Emory University)</i>   |     |
| Constrained Private Mechanisms for Count Data .....  | 845 |
| <i>Graham Cormode (University of Warwick), Tejas Kulkarni (University of Warwick), and Divesh Srivastava (AT&amp;T Labs-Research)</i>  |     |
| A Differentially Private Index for Range Query Processing in Clouds .....  | 857 |
| <i>Cetin Sahin (University of California Santa Barbara), Tristan Allard (Universite Rennes), Reza Akbarinia (INRIA), Amr El Abbadi (University of California Santa Barbara), and Esther Pacitti (INRIA)</i>  |     |

## Research Session 18: Query Processing, Indexing, and Optimization

|  |     |
|--|-----|
| Efficient Computation of a Near-Maximum Independent Set over Evolving Graphs .....   | 869 |
| <i>Weiguo Zheng (The Chinese University of Hong Kong), Qichen Wang (Hong Kong University of Science and Technology), Jeffrey Xu Yu (The Chinese University of Hong Kong), Hong Cheng (The Chinese University of Hong Kong), and Lei Zou (Peking University)</i>  |     |
| A GPU Accelerated Update Efficient Index for kNN Queries in Road Networks .....  | 881 |
| <i>Chuanwen Li (Northeastern University), Yu Gu (Northeastern University), Jianzhong Qi (The University of Melbourne), Jiayuan He (The University of Melbourne), Qingxu Deng (Northeastern University), and Ge Yu (Northeastern University)</i>  |     |
| A Generic Inverted Index Framework for Similarity Search on the GPU .....  | 893 |
| <i>Jingbo Zhou (National University of Singapore), Qi Guo (National University of Singapore), H. V. Jagadish (University of Michigan), Lubos Krcaľ (National University of Singapore), Siyuan Liu (Nanyang Technological University), Wenhao Luan (National University of Singapore), Anthony K. H. Tung (National University of Singapore), Yueji Yang (National University of Singapore), and Yuxin Zheng (National University of Singapore)</i> |     |
| Finding Diverse Neighbors in High Dimensional Space .....  | 905 |
| <i>Qi Guo (National University of Singapore), Hosagrahar Visvesvaraya Jagadish (University of Michigan), Anthony Kum Hoe Tung (National University of Singapore), and Yu Xin Zheng (National University of Singapore)</i>  |     |

## Research Session 19: Scalable Analytics, Graph, RDF, Web Data, and Social Networks

|   |     |
|---|-----|
| Budget-Constrained Organization of Influential Social Events .....  | 917 |
| <i>Kai Han (University of Science and Technology of China), Yuntian He (University of Science and Technology of China), Xiaokui Xiao (National University of Singapore), Shaojie Tang (University of Texas at Dallas), Fei Gui (University of Science and Technology of China), Chaoting Xu (University of Science and Technology of China), and Jun Luo (Nanyang Technological University)</i> |     |
| Joint Event-Partner Recommendation in Event-Based Social Networks .....   | 929 |
| <i>Hongzhi Yin (The University of Queensland), Lei Zou (The University of Queensland), Quoc Viet Hung Nguyen (The University of Queensland), Zi Huang (The University of Queensland), and Xiaofang Zhou (The University of Queensland)</i>  |     |
| Inf2vec: Latent Representation Model for Social Influence Embedding .....   | 941 |
| <i>Shanshan Feng (Institute of High Performance Computing), Gao Cong (Nanyang Technological University), Arijit Khan (Nanyang Technological University), Xiucheng Li (Nanyang Technological University), Yong Liu (Institute of High Performance Computing), and Yeow Meng Chee (Nanyang Technological University)</i>  |     |
| Query Independent Scholarly Article Ranking .....   | 953 |
| <i>Shuai Ma (Beihang University), Chen Gong (Beihang University), Renjun Hu (Beihang University), Dongsheng Luo (Beihang University), Chunming Hu (Beihang University), and Jinpeng Huai (Beihang University)</i>   |     |

## Research Session 20: Data Science and Scientific Data

|   |     |
|---|-----|
| Why-Not Questions on Top-k Geo-Social Keyword Queries in Road Networks .....  | 965 |
| <i>Jingwen Zhao (Zhejiang University), Yunjun Gao (Zhejiang University), Gang Chen (Zhejiang University), and Rui Chen (Zhejiang University)</i>  |     |
| ArrayBridge: Interweaving Declarative Array Processing in SciDB with Imperative HDF5-Based Programs .....   | 977 |
| <i>Haoyuan Xing (The Ohio State University), Sofoklis Floratos (The Ohio State University), Spyros Blanas (The Ohio State University), Suren Byna (Lawrence Berkeley National Laboratory), M. Prabhat (Lawrence Berkeley National Laboratory), Kesheng Wu (Lawrence Berkeley National Laboratory), and Paul Brown (Paradigm)</i>                      |     |
| Seeping Semantics: Linking Datasets Using Word Embeddings for Data Discovery .....  | 989 |
| <i>Raul Castro Fernandez (Massachusetts Institute of Technology), Essam Mansour (QCRI), Abdulhakim A. Qahtan (QCRI), Ahmed Elmagarmid (QCRI), Ihab Ilyas (University of Waterloo), Samuel Madden (Massachusetts Institute of Technology), Mourad Ouzzani (QCRI), Michael Stonebraker (Massachusetts Institute of Technology), and Nan Tang (QCRI)</i> |     |

|  |      |
|--|------|
| Aurum: A Data Discovery System .....   | 1001 |
| <i>Raul Castro Fernandez (Massachusetts Institute of Technology),<br/>Ziawasch Abedjan (TU Berlin), Famien Koko (Massachusetts Institute of<br/>Technology), Gina Yuan (Massachusetts Institute of Technology), Samuel<br/>Madden (Massachusetts Institute of Technology), and Michael<br/>Stonebraker (Massachusetts Institute of Technology)</i> |      |

## Research Session 21: Query Processing, Indexing, and Optimization

|   |      |
|---|------|
| Generating Optimal Plans for Boolean Expressions .....  | 1013 |
| <i>Fisnik Kastrati (University of Mannheim) and Guido Moerkotte<br/>(University of Mannheim)</i>  |      |
| A Graph-Based Database Partitioning Method for Parallel OLAP Query Processing .....   | 1025 |
| <i>Yoon-Min Nam (DGIST), Min-Soo Kim (DGIST), and Donghyoung Han (DGIST)</i>  |      |
| A GPU-Accelerated Framework for Processing Trajectory Queries .....   | 1037 |
| <i>Bowen Zhang (Shanghai Jiao Tong University), Yanyan Shen (Shanghai<br/>Jiao Tong University), Yanmin Zhu (Shanghai Jiao Tong University), and<br/>Jiadi Yu (Shanghai Jiao Tong University)</i> |      |
| Reproducible Floating-Point Aggregation in RDBMSs .....   | 1049 |
| <i>Ingo Mueller (ETH Zurich), Andrea Arteaga (Federal Institute of<br/>Meteorology and Climatology MeteoSwiss), Torsten Hoefler (ETH Zurich),<br/>and Gustavo Alonso (ETH Zurich)</i>             |      |

## Research Session 22: Temporal, Spatial, Mobile, and Multimedia Data

|  |      |
|--|------|
| Price-and-Time-Aware Dynamic Ridesharing .....   | 1061 |
| <i>Lu Chen (Zhejiang University), Qilu Zhong (Zhejiang University),<br/>Xiaokui Xiao (National University of Singapore), Yunjun Gao (Zhejiang<br/>University), Pengfei Jin (Zhejiang University), and Christian S.<br/>Jensen (Aalborg University)</i> |      |
| Learning to Route with Sparse Trajectory Sets .....  | 1073 |
| <i>Chenjuan Guo (Aalborg University), Bin Yang (Aalborg University),<br/>Jilin Hu (Aalborg University), and Christian Jensen (Aalborg<br/>University)</i>  |      |
| Effective and Efficient User Account Linkage across Location Based Social Networks .....   | 1085 |
| <i>Wei Chen (Soochow University), Hongzhi Yin (The University of<br/>Queensland), Weiqing Wang (The University of Queensland), Lei Zhao<br/>(Soochow University), and Xiaofang Zhou (The University of Queensland)</i>                                 |      |
| CiNCT: Compression and Retrieval for Massive Vehicular Trajectories via Relative Movement<br>Labeling .....  | 1097 |
| <i>Satoshi Koide (Toyota Central R&amp;D Labs), Yukihiro Tadokoro (Toyota<br/>Central R&amp;D Labs), Chuan Xiao (Nagoya University), and Yoshiharu<br/>Ishikawa (Nagoya University)</i>  |      |

## Research Session 23: Data Mining and Knowledge Discovery

|   |      |
|---|------|
| Knowledge Base Enhancement via Data Facts and Crowdsourcing .....   | 1109 |
| <i>Linnan Jiang (Hong Kong University of Science and Technology), Lei Chen (Hong Kong University of Science and Technology), and Zhao Chen (Hong Kong University of Science and Technology)</i> |      |
| Scalable Tucker Factorization for Sparse Tensors - Algorithms and Discoveries .....   | 1120 |
| <i>Sejoon Oh (Seoul National University), Namyong Park (Carnegie Mellon University), Sael Lee (Seoul National University), and U Kang (Seoul National University)</i>                           |      |
| TPA: Fast, Scalable, and Accurate Method for Approximate Random Walk with Restart on Billion Scale Graphs .....   | 1132 |
| <i>Minji Yoon (Seoul National University), Jinhong Jung (Seoul National University), and U Kang (Seoul National University)</i>   |      |
| M2TD: Multi-Task Tensor Decomposition for Sparse Ensemble Simulations .....   | 1144 |
| <i>Xinsheng Li (Arizona State University), Kasim Selçuk Candan (Arizona State University), and Maria Luisa Sapino (University of Torino)</i>  |      |

## Research Session 24: Temporal and Probabilistic Data

|   |      |
|---|------|
| Enabling Quality Control for Entity Resolution: A Human and Machine Cooperation Framework .....   | 1156 |
| <i>Zhaoqiang Chen (Northwestern Polytechnical University), Qun Chen (Northwestern Polytechnical University), Fengfeng Fan (Northwestern Polytechnical University), Yanyan Wang (Northwestern Polytechnical University), Zhuo Wang (Northwestern Polytechnical University), Youcef Nafa (Northwestern Polytechnical University), Zhanhuai Li (Northwestern Polytechnical University), Hailong Liu (Northwestern Polytechnical University), and Wei Pan (Northwestern Polytechnical University)</i> |      |
| Robust Discovery of Positive and Negative Rules in Knowledge Bases .....  | 1168 |
| <i>Stefano Ortona (Meltwater), Venkata Vamsikrishna Meduri (Arizona State University), and Paolo Papotti (Eurecom)</i>  |      |
| Supporting Set Operations in Temporal-Probabilistic Databases .....   | 1180 |
| <i>Katerina Papaioannou (University of Zurich), Martin Theobald (University of Luxembourg), and Michael Boehlen (University of Zurich)</i>  |      |
| Efficient Probabilistic K-Core Computation on Uncertain Graphs .....  | 1192 |
| <i>You Peng (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), and Lu Qin (University of Technology Sydney)</i>   |      |

## Research Track: Short Papers

|  |      |
|--|------|
| Histogramming Privately Ever After: Differentially-Private Data-Dependent Error Bound Optimisation ..... | 1204 |
| <i>Maryam Fanaeepour (Duke University) and Benjamin I. P. Rubinstein (University of Melbourne)</i>       |      |

|  |      |
|--|------|
| Interactive Visual Data Exploration with Subjective Feedback: An Information-Theoretic Approach .....  | 1208 |
| <i>Kai Puolamäki (Aalto University), Emilia Oikarinen (Aalto University), Bo Kang (Ghent University), Jefrey Lijffijt (Ghent University), and Tijl De Bie (Ghent University)</i>   |      |
| Local Tensor Completion Based on Locality Sensitive Hashing .....  | 1212 |
| <i>Kun Xie (Hunan University), Yuxiang Chen (Hunan University), Xin Wang (State of New York University at Stony Brook), Gaogang Xie (Institute of Computing Technology), Jigang Wen (Institute of Computing Technology), and Dafang Zhang (Hunan University)</i> |      |
| Scalable De Novo Genome Assembly Using Pregel .....  | 1216 |
| <i>Da Yan (The University of Alabama at Birmingham), Hongzhi Chen (The Chinese University of Hong Kong), James Cheng (The Chinese University of Hong Kong), Zhenkun Cai (The Chinese University of Hong Kong), and Bin Shao (Microsoft Research Asia)</i>        |      |
| Fast k-Means Based on k-NN Graph .....   | 1220 |
| <i>Cheng-Hao Deng (Xiamen University) and Wan-Lei Zhao (Xiamen University)</i>   |      |
| Reuse-Centric K-Means Configuration .....  | 1224 |
| <i>Hui Guan (North Carolina State University), Yufei Ding (North Carolina State University), Xipeng Shen (North Carolina State University), and Hamid Krim (North Carolina State University)</i>   |      |
| AssureMR: Verifiable SQL Execution on MapReduce .....  | 1228 |
| <i>Bo Zhang (Stevens Institute of Technology), Boxiang Dong (Montclair State University), and Wendy Hui Wang (Stevens Institute of Technology)</i>   |      |
| Wadjet: Finding Outliers in Multiple Multi-Dimensional Heterogeneous Data Streams .....  | 1232 |
| <i>Shiblee Sadik (Amazon Web Services), Le Gruenwald (University of Oklahoma), and Eleazar Leal (University of Minnesota Duluth)</i>   |      |
| Onion Curve: A Space Filling Curve with Near-Optimal Clustering .....  | 1236 |
| <i>Pan Xu (University of Maryland), Cuong Nguyen (Iowa State University), and Srikantha Tirthapura (Iowa State University)</i>   |      |
| Scalable and Robust Set Similarity Join .....  | 1240 |
| <i>Tobias Christiani (IT University of Copenhagen), Rasmus Pagh (IT University of Copenhagen), and Johan Sivertsen (IT University of Copenhagen)</i>   |      |
| DIFusion: Fast Skip-Scan with Zero Space Overhead .....  | 1244 |
| <i>Wenjian Xu (The Hong Kong Polytechnic University), Eric Lo (Chinese University of Hong Kong), and Pengfei Zhang (Chinese University of Hong Kong)</i>   |      |
| Multi-Campaign Oriented Spatial Crowdsourcing .....  | 1248 |
| <i>Libin Zheng (Hong Kong University of Science and Technology) and Lei Chen (Hong Kong University of Science and Technology)</i>  |      |
| Parallelizing Multiple Pipelines of One Query in a Main Memory Database Cluster .....  | 1252 |
| <i>Zhuhe Fang (East China Normal University), Chuliang Weng (East China Normal University), Li Wang (Advanced Digital Sciences Center), and Aoying Zhou (East China Normal University)</i>   |      |

|   |      |
|---|------|
| Interactive Probabilistic Post-Mining of User-Preferred Spatial Co-Location Patterns .....  | 1256 |
| <i>Lizhen Wang (Yunnan University), Xuguang Bao (Yunnan University), and Longbing Cao (University of Technology Sydney)</i>   |      |
| Hebe: An Order-Oblivious and High-Performance Execution Scheme for Conjunctive Predicates ...   | 1260 |
| <i>Zeke Wang (ETH Zurich), Kai Zhang (Fudan University), Haihang Zhou (National University of Singapore), Xue Liu (Northeastern University), and Bingsheng He (National University of Singapore)</i>                                    |      |
| Distance-Based Data Mining over Encrypted Data .....  | 1264 |
| <i>Christine Tex (Karlsruhe Institute of Technology), Martin Schäler (Karlsruhe Institute of Technology), and Klemens Böhm (Karlsruhe Institute of Technology)</i>  |      |
| A Sliding-Window Framework for Representative Subset Selection .....  | 1268 |
| <i>Yanhao Wang (National University of Singapore), Yuchen Li (Singapore Management University), and Kian-Lee Tan (National University of Singapore)</i>   |      |
| Social Event Scheduling .....   | 1272 |
| <i>Nikos Bikakis (Athens University of Economics and Business), Vana Kalogeraki (Athens University of Economics and Business), and Dimitrios Gunopulos (University of Athens)</i>   |      |
| Query-Centric Failure Recovery for Distributed Stream Processing Engines .....  | 1276 |
| <i>Li Su (University of Copenhagen) and Yongluan Zhou (University of Copenhagen)</i>  |      |
| A Scalable Circular Pipeline Design for Multi-Way Stream Joins in Hardware .....  | 1280 |
| <i>Mohammadreza Najafi (Technical University of Munich), Mohammad Sadoghi (University of California), and Hans-Arno Jacobsen (Technical University of Munich)</i>   |      |
| Consistent Snapshot Algorithms for In-Memory Database Systems: Experiments and Analysis .....   | 1284 |
| <i>Liang Li (Northeastern University of China), Guoren Wang (Beijing Institute of Technology), Gang Wu (Northeastern University of China), and Ye Yuan (Northeastern University of China)</i>   |      |
| Performance of OLTP via Intelligent Scheduling .....  | 1288 |
| <i>Tieying Zhang (Alibaba Inc.), Anthony Tomasic (Carnegie Mellon University), Yangjun Sheng (Carnegie Mellon University), and Andrew Pavlo (Carnegie Mellon University)</i>  |      |
| SURGE: Continuous Detection of Bursty Regions over a Stream of Spatial Objects .....  | 1292 |
| <i>Kaiyu Feng (Nanyang Technological University), Tao Guo (Nanyang Technological University), Gao Cong (Nanyang Technological University), Sourav S. Bhowmick (Nanyang Technological University), and Shuai Ma (Beihang University)</i> |      |
| Simplifying Entity Resolution on Web Data with Schema-Agnostic, Non-Iterative Matching .....  | 1296 |
| <i>Vasilis Efthymiou (ICS-FORTH), George Papadakis (University of Athens), Kostas Stefanidis (University of Tampere), and Vassilis Christophides (INRIA-Paris &amp; University of Crete)</i>  |      |

|  |      |
|--|------|
| Scotty: Efficient Window Aggregation for Out-of-Order Stream Processing .....  | 1300 |
| <i>Jonas Traub (Technische Universität Berlin), Philipp Marian Grulich (DFKI GmbH), Alejandro Rodriguez Cuellar (Technische Universität Berlin), Sebastian Bress (DFKI GmbH), Asterios Katsifodimos (Delft University of Technology), Tilmann Rabl (Technische Universität Berlin), and Volker Markl (Technische Universität Berlin)</i> |      |
| Get Your Workload in Order: Game Theoretic Prioritization of Database Auditing .....   | 1304 |
| <i>Chao Yan (Vanderbilt University), Bo Li (University of California), Yevgeniy Vorobeychik (Vanderbilt University), Aron Laszka (University of Houston), Daniel Fabbri (Vanderbilt University), and Bradley Malin (Vanderbilt University)</i>   |      |
| Polygraph: A Plug-n-Play Framework to Quantify Anomalies .....   | 1308 |
| <i>Yazeed Alabdulkarim (University of Southern California), Marwan Almaymoni (University of Southern California), and Shahram Ghandeharizadeh (University of Southern California)</i>  |      |
| Labeled Graph Sketches .....   | 1312 |
| <i>Chunyao Song (Nankai University) and Tingjian Ge (University of Massachusetts)</i>  |      |
| T-Crowd: Effective Crowdsourcing for Tabular Data .....  | 1316 |
| <i>Caihua Shan (The University of Hong Kong), Nikos Mamoulis (University of Ioannina), Guoliang Li (Tsinghua University), Reynold Cheng (The University of Hong Kong), Zhipeng Huang (The University of Hong Kong), and Yudian Zheng (Twitter Inc.)</i>  |      |
| A Novel Framework for Constructing Partially Monotone Rule Ensembles .....   | 1320 |
| <i>Christopher Bartley (University of Western Australia), Wei Liu (University of Western Australia), and Mark Reynolds (University of Western Australia)</i>   |      |
| Querying Cohesive Subgraphs by Keywords .....  | 1324 |
| <i>Yuanyuan Zhu (Wuhan University), Qian Zhang (Wuhan University), Lu Qin (University of Technology Sydney), Lijun Chang (The University of Sydney), and Jeffrey Xu Yu (The Chinese University of Hong Kong)</i>   |      |
| On Efficiently Detecting Overlapping Communities over Distributed Dynamic Graphs .....   | 1328 |
| <i>Xun Jian (Hong Kong University of Science and Technology), Xiang Lian (Kent State University), and Lei Chen (Hong Kong University of Science and Technology)</i>  |      |
| Discovering Expert Drivers from Trajectories .....   | 1332 |
| <i>Jiabao Sun (Soochow University), Jiajie Xu (Soochow University), Rui Zhou (Swinburne University of Technology), Kai Zheng (University of Electronic Science and Technology of China), and Chengfei Liu (Swinburne University of Technology)</i>   |      |
| Sharing Uncertain Graphs Using Syntactic Private Graph Models .....  | 1336 |
| <i>Dongqing Xiao (Worcester Polytechnic Institute), Mohamed Y. Eltabakh (Worcester Polytechnic Institute), and Xiangnan Kong (Worcester Polytechnic Institute)</i>   |      |

|   |      |
|---|------|
| Holistic Influence Maximization for Targeted Advertisements in Spatial Social Networks .....  | 1340 |
| <i>Jianxin Li (University of Western Australia), Taotao Cai (University of Western Australia), Ajmal Mian (University of Western Australia), Rong-Hua Li (Beijing Institute of Technology), Timos Sellis (Swinburne University of Technology), and Jeffrey Xu Yu (The Chinese University of Hong Kong)</i>  |      |
| Match-Based Candidate Network Generation for Keyword Queries over Relational Databases .....  | 1344 |
| <i>Pericles Oliveira (Nokia), Altigran da Silva (Universidade Federal do Amazonas), Edleno de Moura (Universidade Federal do Amazonas), and Rosiane Rodrigues (Universidade Federal do Amazonas)</i>  |      |
| Detection and Delineation of Events and Sub-Events in Social Networks .....   | 1348 |
| <i>Antonia Saravanou (National and Kapodistrian University of Athens), Ioannis Katakis (University of Nicosia), George Valkanas (Detectica), and Dimitrios Gunopulos (National and Kapodistrian University of Athens)</i>   |      |
| Subjectively Interesting Subgroup Discovery on Real-Valued Targets .....  | 1352 |
| <i>Jefrey Lijffijt (Ghent University), Bo Kang (Ghent University), Wouter Duivesteyn (Eindhoven University of Technology), Kai Puolamaki (Aalto University), Emilia Oikarinen (Aalto University), and Tijl De Bie (Ghent University)</i>  |      |
| ULISSE: ULtra Compact Index for Variable-Length Similarity Search in Data Series .....  | 1356 |
| <i>Michele Linardi (Paris Descartes University) and Themis Palpanas (Paris Descartes University)</i>  |      |
| Approximate Geospatial Joins with Precision Guarantees .....  | 1360 |
| <i>Andreas Kipf (Technical University of Munich), Harald Lang (Technical University of Munich), Varun Pandey (Technical University of Munich), Raul Alexandru Persa (Technical University of Munich), Peter Boncz (Centrum Wiskunde and Informatica), Thomas Neumann (Technical University of Munich), and Alfons Kemper (Technical University of Munich)</i> |      |
| Ensemble Direct Density Ratio Estimation for Multistream Classification .....   | 1364 |
| <i>Swarup Chandra (University of Texas at Dallas), Ahsanul Haque (University of Texas at Dallas), Hemeng Tao (University of Texas at Dallas), Jie Liu (University of Texas at Dallas), Latifur Khan (University of Texas at Dallas), and Charu Aggarwal (IBM T. J. Watson Research Center)</i>  |      |
| GraphZ: Improving the Performance of Large-Scale Graph Analytics on Small-Scale Machines .....  | 1368 |
| <i>Zhixuan Zhou (University of Chicago) and Henry (Hank) Hoffmann (University of Chicago)</i>   |      |
| MGTag: a Multi-Dimensional Graph Labeling Scheme for Fast Reachability Queries .....  | 1372 |
| <i>Shuang Zhou (Huazhong University of Science and Technology), Pingpeng Yuan (Huazhong University of Science and Technology), Ling Liu (Georgia Institute of Technology), and Hai Jin (Huazhong University of Science and Technology)</i>  |      |
| On the Design of Adaptive and Speculative Concurrency Control in Distributed Databases .....  | 1376 |
| <i>Qian Lin (National University of Singapore), Gang Chen (Zhejiang University), and Meihui Zhang (Beijing Institute of Technology)</i>   |      |



|  |      |
|--|------|
| On Configuring a Hierarchy of Storage Media in the Age of NVM .....  | 1380 |
| <i>Shahram Ghandeharizadeh (University of Southern California), Sandy Irani (University of California), and Jenny Lam (San Jose State University)</i>  |      |
| Query and Resource Optimization: Bridging the Gap .....  | 1384 |
| <i>Lalitha Viswanathan (Microsoft), Alekh Jindal (Microsoft), and Konstantinos Karanasos (Microsoft)</i>   |      |
| Burstiness in Query Log: Web Search Analysis by Combining Global and Local Evidences .....   | 1388 |
| <i>Chen Zhang (Shandong University of Finance &amp; Economics), Sen Zhang (Shandong University of Finance &amp; Economics), Chen Lei (Hong Kong University of Science &amp; Technology), and Peiguang Lin (Shandong University of Finance &amp; Economics)</i> |      |

## Industry Track

### Industry Session 1: Predictive Modeling

|  |      |
|--|------|
| Did You Enjoy the Ride? Understanding Passenger Experience via Heterogeneous Network Embedding .....   | 1392 |
| <i>Carl Yang (University of Illinois), Chao Zhang (University of Illinois), Xuwen Chen (Didichuxing Inc.), Jieping Ye (Didichuxing Inc.), and Jiawei Han (University of Illinois)</i>  |      |
| Recommendations for All: Solving Thousands of Recommendation Problems Daily .....  | 1404 |
| <i>Bhargav Kanagal (Google) and Sandeep Tata (Google)</i>  |      |
| Predictive Maintenance in Aviation: Failure Prediction from Post-Flight Reports .....  | 1414 |
| <i>Panagiotis Korveis (Ecole Polytechnique), Stephane Besseau (Airbus), and Michalis Vazirgiannis (Ecole Polytechnique)</i>  |      |
| KNADIA: Enterprise KNowledge Assisted DIAlogue Systems Using Deep Learning .....   | 1423 |
| <i>Mahesh Singh (Tata Consultancy Services Ltd.), Puneet Agarwal (Tata Consultancy Services Ltd.), Ashish Chaudhary (Tata Consultancy Services Ltd.), Gautam Shroff (Tata Consultancy Services Ltd.), Prerna Khurana (Tata Consultancy Services Ltd.), Mayur Patidar (Tata Consultancy Services Ltd.), Vivek Bisht (Tata Consultancy Services Ltd.), Rachit Bansal (Tata Consultancy Services Ltd.), Prateek Sachan (Tata Consultancy Services Ltd.), and Rohit Kumar (Tata Consultancy Services Ltd.)</i> |      |
| Online E-Commerce Fraud: A Large-Scale Detection and Analysis .....  | 1435 |
| <i>Haiqin Weng (Zhejiang University), Zhao Li (Alibaba Group), Shouling Ji (Zhejiang University), Chen Chu (Alibaba Group), Haifeng Lu (Alibaba Group), Tianyu Du (Zhejiang University), and Qinming He (Zhejiang University)</i>  |      |

### Industry Session 2: Query Processing & Concurrency Control

|  |      |
|--|------|
| Accelerating Joins and Aggregations on the Oracle In-Memory Database .....   | 1441 |
| <i>Shasank Chavan (Oracle America), Albert Hopeman (Oracle America), Sangho Lee (Oracle America), Dennis Lui (Oracle America), Ajit Mylavarapu (Oracle America), and Ekrem Soylemez (Oracle America)</i> |      |

|  |      |
|--|------|
| Rethinking Concurrency Control for In-Memory OLAP DBMSs .....  | 1453 |
| <i>Pedro Pedreira (Facebook Inc.), Yinghai Lu (Facebook Inc.), Sergey Pershin (Facebook Inc.), Amit Dutta (Facebook Inc.), and Chris Croswite (Facebook Inc.)</i>  |      |
| SHC: Distributed Query Processing for Non-Relational Data Store .....  | 1465 |
| <i>Weiying Yang (Hortonworks), Mingjie Tang (Hortonworks), Yongyang Yu (Purdue University), Yanbo Liang (Hortonworks), and Bikas Saha (Hortonworks)</i>  |      |
| Partial Update: Efficient Materialized View Maintenance in a Distributed Graph Database .....  | 1477 |
| <i>SungJu Cho (LinkedIn Corp.), Roman Averbukh (LinkedIn Corp.), Yanwei Zhang (Georgia Institute of Technology), Andrew Carter (LinkedIn Corp.), and Jane Alam Jan (Uber)</i>                              |      |
| Efficiently Processing Temporal Queries on Hyperledger Fabric .....  | 1489 |
| <i>Himanshu Gupta (IBM Research), Sandeep Hans (IBM Research), Kushagra Aggarwal (IIT Kharagpur), Sameep Mehta (IBM Research), Bapi Chatterjee (IBM Research), and Praveen Jayachandran (IBM Research)</i> |      |

### Industry Session 3: Miscellaneous Topics

|  |      |
|--|------|
| RTSI: An Index Structure for Multi-Modal Real-Time Search on Live Audio Streaming Services.....  | 1495 |
| <i>Zeyi Wen (National University of Singapore), Xingyang Liu (Shanghai Jiao Tong University), Hongjian Cao (Shanghai Jiao Tong University), and Bingsheng He (National University of Singapore)</i>                |      |
| Benchmarking Distributed Stream Data Processing Systems .....  | 1507 |
| <i>Jeyhun Karimov (DFKI), Tilmann Rabl (TU-Berlin), Asterios Katsifodimos (Delft University of Technology), Roman Samarev (TU-Berlin), Henri Heiskanen (Rovio), and Volker Markl (TU-Berlin)</i>                   |      |
| Analysis of TPCx-IoT: The First Industry Standard Benchmark for IoT Gateway Systems .....  | 1519 |
| <i>Meikel Poess (Oracle), Raghunath Nambiar (Cisco), Karthik Kulkarni (Cisco), Chinmayi Narasimhadevara (Cisco), Tilmann Rabl (TU-Berlin), and Hans-Arno Jacobsen (TU-Munich)</i>                                  |      |
| PRIMA: An End-to-End Framework for Privacy at Scale .....  | 1531 |
| <i>Spiros Antonatos (IBM Research Ireland), Stefano Braghin (IBM Research Ireland), Naoise Holohan (IBM Research Ireland), Yiannis Gkoufas (IBM Research Ireland), and Pol Mac Aonghusa (IBM Research Ireland)</i> |      |
| Traffic-Aware Routing in Road Networks .....   | 1543 |
| <i>Daniel Delling (Apple), Dennis Schieferdecker (Apple), and Christian Sommer (Apple)</i>   |      |

## Demo Track

### Session 1: Data Analytics & Discovery

|  |      |
|--|------|
| An SQL Data Summarization Tool for the Acquisition of Meaningful Cardinality Constraints and Functional Dependencies .....   | 1549 |
| <i>Aniruddh Gandhi (Amrita University), Sven Hartmann (Clausthal University of Technology), Henning Koehler (Massey University), and Sebastian Link (The University of Auckland)</i>   |      |
| AutoMode: Relational Learning with Less Black Magic .....  | 1553 |
| <i>Jose Picado (Oregon State University), Sudhanshu Pathak (Oregon State University), Arash Termehchy (Oregon State University), and Alan Fern (Oregon State University)</i>   |      |
| Exploration of User Groups in VEXUS .....  | 1557 |
| <i>Sihem Amer-Yahia (CNRS, University Grenoble Alpes), Behrooz Omidvar-Tehrani (Universite Grenoble Alpes), João Comba (UFRGS Brazil), Viviane Moreira (UFRGS Brazil), and Fabian Colque Zegarra (UFRGS Brazil)</i>  |      |
| FASTOD: Bringing Order to Data .....   | 1561 |
| <i>Alexandar Mihaylov (University of Ontario), Parke Godfrey (York University), Lukasz Golab (University of Waterloo), Mehdi Kargar (University of Windsor), Divesh Srivastava (AT&amp;T Labs-Research), and Jaroslaw Szlichta (University of Ontario)</i>   |      |
| MeDIAR: Multi-Drug Adverse Reactions Analytics .....   | 1565 |
| <i>Xiao Qin (Worcester Polytechnic Institute), Tabassum Kakar (Worcester Polytechnic Institute), Susmitha Wunnava (Worcester Polytechnic Institute), Brian MacCarthy (Worcester Polytechnic Institute), Andrew Schade (Worcester Polytechnic Institute), Huy Quoc Tran (Worcester Polytechnic Institute), Brian Zylich (Worcester Polytechnic Institute), Elke Rundensteiner (Worcester Polytechnic Institute), Lane Harrison (Worcester Polytechnic Institute), Sanjay Sahoo (U.S. Food and Drug Administration), and Suranjan De (U.S. Food and Drug Administration)</i> |      |
| OCTOPUS: An Online Topic-Aware Influence Analysis System for Social Networks .....   | 1569 |
| <i>Ju Fan (Renmin University of China), Jiarong Qiu (Renmin University of China), Yuchen Li (Singapore Management University), Qingfei Meng (Renmin University of China), Dongxiang Zhang (University of Electronic Science and Technology of China), Guoliang Li (Tsinghua University), Kian-Lee Tan (National University of Singapore), and Xiaoyong Du (Renmin University of China)</i>   |      |
| RheemStudio: Cross-Platform Data Analytics Made Easy .....   | 1573 |
| <i>Ji Lucas (QCRI), Yasser Idris (QCRI), Bertty Contreras-Rojas (QCRI), Jorge-Arnulfo Quijane-Ruiz (QCRI), and Sanjay Chawla (QCRI)</i>  |      |
| SpaceKey: Exploring Patterns in Spatial Databases .....  | 1577 |
| <i>Yixiang Fang (University of Hong Kong), Reynold Cheng (University of Hong Kong), Jikun Wang (University of Hong Kong), -- Budiman (University of Hong Kong), Gao Cong (Nanyang Technological University), and Nikos Mamoulis (University of Ioannina)</i>   |      |

|  |      |
|--|------|
| Time-Aware Sub-Trajectory Clustering in Hermes@PostgreSQL .....  | 1581 |
| <i>Panagiotis Tampakis (University of Piraeus), Nikos Pelekis (University of Piraeus), Natalia Andrienko (Fraunhofer Institute IAIS), Gennady Andrienko (Fraunhofer Institute IAIS), Georg Fuchs (Fraunhofer Institute IAIS), and Yannis Theodoridis (University of Piraeus)</i> |      |
| VizCS: Online Searching and Visualizing Communities in Dynamic Graphs .....  | 1585 |
| <i>Yuli Jiang (Hong Kong Baptist University), Xin Huang (Hong Kong Baptist University), Hong Cheng (The Chinese University of Hong Kong), and Jeffrey Xu Yu (The Chinese University of Hong Kong)</i>  |      |

## Session 2: Data Cleaning, Fusion, & Generation

|   |      |
|---|------|
| A Framework to Integrate User Feedback for Rapid Conflict Resolution .....  | 1589 |
| <i>Romila Pradhan (Purdue University), Siarhei Bykau (Bloomberg L.P.), and Sunil Prabhakar (Purdue University)</i>  |      |
| Building Data Civilizer Pipelines with an Advanced Workflow Engine .....  | 1593 |
| <i>Essam Mansour (Qatar Computing Research Institute/HBKU), Dong Deng (MIT CSAIL), Raul Castro Fernandez (MIT CSAIL), Abdulhakim A. Qahtan (Qatar Computing Research Institute), Wenbo Tao (MIT CSAIL), Ziawasch Abedjan (TU Berlin), Ahmed Elmagarmid (Qatar Computing Research Institute), Ihab F. Ilyas (University of Waterloo), Samuel Madden (MIT CSAIL), Mourad Ouzzani (Qatar Computing Research Institute), Michael Stonebraker (MIT CSAIL), and Nan Tang (Qatar Computing Research Institute)</i> |      |
| Cleaning Your Wrong Google Scholar Entries .....  | 1597 |
| <i>Shuang Hao (Tsinghua University), Yi Xu (Tsinghua University), Nan Tang (QCRI), Guoliang Li (Tsinghua University), and Jianhua Feng (Tsinghua University)</i>  |      |
| CloudTP: A Cloud-Based Flexible Trajectory Preprocessing Framework .....  | 1601 |
| <i>Sijie Ruan (Xidian University), Ruiyuan Li (Xidian University), Jie Bao (Microsoft Research Asia), Tianfu He (Harbin Institution of Technology), and Yu Zheng (Xidian University)</i>  |      |
| Curating Variational Data in Application Development .....  | 1605 |
| <i>Uta Störl (Darmstadt University of Applied Sciences), Daniel Müller (Darmstadt University of Applied Sciences), Alexander Tekleab (Darmstadt University of Applied Sciences), Stephane Tolale (Darmstadt University of Applied Sciences), Julian Stenzel (Darmstadt University of Applied Sciences), Meike Klettke (University of Rostock), and Stefanie Scherzinger (OTH Regensburg)</i>  |      |
| FAHES: Detecting Disguised Missing Values .....   | 1609 |
| <i>Abdulhakim Qahtan (QCRI), Ahmed Elmagarmid (QCRI), Mourad Ouzzani (QCRI), and Nan Tang (QCRI)</i>  |      |
| LUSTRE: An Interactive System for Entity Structured Representation and Variant Generation .....   | 1613 |
| <i>Kun Qian (IBM Research Almaden), Nikita Bhutani (University of Michigan), Yunyao Li (IBM Research Almaden), H. V. Jagadish (University of Michigan), and Mauricio Hernandez (IBM Research Almaden)</i>   |      |

|  |      |
|--|------|
| TableView: A Visual Interface for Generating Preview Tables of Entity Graphs .....   | 1617 |
| <i>Sona Hasani (University of Texas at Arlington), Ning Yan (Huawei U.S. R&amp;D Center), and Chengkai Li (University of Texas at Arlington)</i>       |      |
| Test Data Generation for Database Applications .....   | 1621 |
| <i>Pooja Agrawal (IIT Bombay), Bikash Chandra (IIT Bombay), K. Venkatesh Emani (IIT Bombay), Neha Garg (IIT Bombay), and S. Sudarshan (IIT Bombay)</i> |      |

### Session 3: Data-Centric Applications, Platforms, & Techniques

|   |      |
|---|------|
| Continuous Maintenance of Range Sum Heat Maps .....   | 1625 |
| <i>Jianzhong Qi (The University of Melbourne), Vivek Kumar (The University of Melbourne), Rui Zhang (The University of Melbourne), Egemen Tanin (The University of Melbourne), Goce Trajcevski (Iowa State University), and Peter Scheuermann (Northwestern University)</i>   |      |
| CrowdOTA: An Online Task Assignment System in Crowdsourcing .....   | 1629 |
| <i>Xiang Yu (Tsinghua University), Guoliang Li (Tsinghua University), Yudian Zheng (Twitter), Yan Huang (TAL), Songfan Zhang (TAL), and Fei Chen (TAL)</i>  |      |
| CrowdSheet: Instant Implementation and Out-of-Hand Execution of Complex Crowdsourcing .....   | 1633 |
| <i>Rikuya Suzuki (University of Tsukuba), Tetsuo Sakaguchi (University of Tsukuba), Masaki Matsubara (University of Tsukuba), Hiroyuki Kitagawa (University of Tsukuba), and Atsuyuki Morishima (University of Tsukuba)</i>   |      |
| Dione: A Framework for Automatic Profiling and Tuning Big Data Applications .....   | 1637 |
| <i>Nikos Zacheilas (Athens University of Economics and Business), Stathis Maroulis (Athens University of Economics and Business), Thanasis Priovolos (Athens University of Economics and Business), Vana Kalogeraki (Athens University of Economics and Business), and Dimitrios Gunopulos (University of Athens)</i> |      |
| Elton: A Cloud Resource Scaling-Out Manager for NoSQL Databases .....   | 1641 |
| <i>Athanasios Naskos (Aristotle University of Thessaloniki), Anastasios Gounaris (Aristotle University of Thessaloniki), and Ioannis Konstantinou (National Technical University of Athens)</i>   |      |
| iZone: Efficient Influence Zone Evaluation over Geo-Textual Data .....  | 1645 |
| <i>Qing Liu (Hong Kong Baptist University), Zijin Feng (Hong Kong Baptist University), Xike Xie (University of Science and Technology of China), Jianliang Xu (Hong Kong Baptist University), Xin Lin (East China Normal University), and Christian S. Jensen (Aalborg University)</i>                                |      |
| Jupiter: A Blockchain Platform for Mobile Devices .....   | 1649 |
| <i>Siyuan Han (Hong Kong University of Science and Technology), Zihuan Xu (Hong Kong University of Science and Technology), and Lei Chen (Hong Kong University of Science and Technology)</i>   |      |

|  |      |
|--|------|
| QR2: A Third-Party Query Reranking Service over Web Databases .....  | 1653 |
| <i>Yeshwanth Durairaj Gunasekaran (University of Texas at Arlington),<br/>Abolfazl Asudeh (University of Michigan), Sona Hasani (University of<br/>Texas at Arlington), Nan Zhang (Pennsylvania State University), Ali<br/>Jaoua (Qatar University), and Gautam Das (University of Texas at<br/>Arlington)</i> |      |
| Rainbow: Adaptive Layout Optimization for Wide Tables .....  | 1657 |
| <i>Haoqiong Bian (Renmin University of China), Youxian Tao (Renmin<br/>University of China), Guodong Jin (Renmin University of China), Yueguo<br/>Chen (Renmin University of China), Xiongpai Qin (Renmin University of<br/>China), and Xiaoyong Du (Renmin University of China)</i>                           |      |

## Lightning Talks

|   |      |
|---|------|
| Caching Data Stores: High Performance at Low Cost .....   | 1661 |
| <i>David Lomet (Microsoft Research)</i>   |      |
| In for a Surprise When Migrating NoSQL Data .....   | 1662 |
| <i>Uta Störl (Darmstadt University of Applied Sciences), Alexander<br/>Tekleab (Darmstadt University of Applied Sciences), Meike Klettke<br/>(University of Rostock), and Stefanie Scherzinger (OTH Regensburg)</i> |      |
| Teaching In-Memory Database Systems the Detection of Hardware Errors .....  | 1663 |
| <i>Wolfgang Lehner (TU Dresden), Dirk Habich (TU Dresden), and Till<br/>Kolditz (TU Dresden)</i>  |      |
| Artifact Evaluation: FAD or Real News? .....  | 1664 |
| <i>Bruce R. Childers (University of Pittsburgh) and Panos K. Chrysanthis<br/>(University of Pittsburgh)</i>   |      |
| Data-Less Big Data Analytics (Towards Intelligent Data Analytics Systems) .....   | 1666 |
| <i>Peter Triantafillou (University of Warwick)</i>  |      |
| The Data Exploration Game .....   | 1668 |
| <i>Ben McCamish (Oregon State University), Arash Termehchy (Oregon State<br/>University), and Behrouz Touri (University of California San Diego)</i>  |      |
| Decomposing and Re-Composing Lightweight Compression Schemes - And Why It Matters .....   | 1670 |
| <i>Eyal Rozenberg (CWI Amsterdam)</i>   |      |
| Love at First Sight: MonetDB/TensorFlow .....   | 1672 |
| <i>Ying Zhang (MonetDB Solutions), Richard Koopmanschap (MonetDB<br/>Solutions), and Martin Kersten (MonetDB Solutions)</i>   |      |
| Human-Database Interaction: A Holistic Approach .....   | 1673 |
| <i>Mohamed Sarwat (Arizona State University) and Venkata Vamsikrishna<br/>Meduri (Arizona State University)</i>   |      |
| There Is No Dichotomy between Effectiveness and Efficiency in Keyword Search over<br>Databases .....  | 1675 |
| <i>Vahid Ghadakchi (Oregon Sate University) and Arash Termehchy (Oregon<br/>State University)</i>   |      |
| Data Series Management: Fulfilling the Need for Big Sequence Analytics .....  | 1677 |
| <i>Kostas Zoumpatianos (Harvard University) and Themis Palpanas (Paris<br/>Descartes University)</i>  |      |

|  |      |
|--|------|
| Towards an Industry Standard for Benchmarking Artificial Intelligence Systems .....    | 1679 |
| <i>Raghunath Nambiar (Cisco)</i>   |      |
| You Can Check Others' Work More Quickly Than Doing It Yourself .....                   | 1681 |
| <i>Graham Cormode (University of Warwick) and Chris Hickey (University of Warwick)</i> |      |
| Reasoning in the Presence of NULLs .....   | 1682 |
| <i>Thomas Neumann (Technische Universität München)</i>                                 |      |

## PhD Symposium

### Session 1

|   |      |
|---|------|
| Designing an Adaptive VM That Combines Vectorized and JIT Execution on Heterogeneous Hardware ..... | 1684 |
| <i>Tim Gubner (Centrum Wiskunde &amp; Informatica)</i>  |      |

### Session 2

|   |      |
|---|------|
| Node Selection in Large Networks .....  | 1689 |
| <i>Alvis Logins (Aalborg University)</i>  |      |
| ZIP-Code Classification Using Spatial and Crowdsourced Data .....                       | 1694 |
| <i>Tunaggina Subrina Khan (George Mason University)</i>                                 |      |
| Complex Event Processing under Constrained Resources by State-Based Load Shedding ..... | 1699 |
| <i>Bo Zhao (Humboldt-Universität zu Berlin)</i>   |      |

### Session 3

|  |      |
|--|------|
| Auditing DBMSes through Forensic Analysis .....                          | 1704 |
| <i>James Wagner (DePaul University)</i>                                  |      |
| Towards Not Re-Inventing the Wheel: Managing Data Management Tools ..... | 1709 |
| <i>Kathryn Dahlgren (University of California)</i>                       |      |
| Internet of Metaproteomics .....   | 1714 |
| <i>Roman Zoun (University of Magdeburg)</i>                              |      |

## Tutorials

|   |      |
|---|------|
| Sorting in Space and Words .....  | 1719 |
| <i>Hanan Samet (University of Maryland at College Park)</i>   |      |
| Cross-Platform Data Processing: Use Cases and Challenges .....  | 1723 |
| <i>Zoi Kaoudi (Qatar Computing Research Institute) and Jorge-Arnulfo Quiane-Ruiz (Qatar Computing Research Institute)</i> |      |

|  |      |
|--|------|
| Online Temporal Analysis of Complex Systems Using IoT Data Sensing .....   | 1727 |
| <i>Avigdor Gal (Technion Israel Institute of Technology), Arik Senderovich (Technion Israel Institute of Technology), and Matthias Weidlich (Humboldt-Universität zu Berlin)</i> |      |
| Data Security and Privacy for Outsourced Data in the Cloud .....   | 1731 |
| <i>Cetin Sahin (University of California) and Amr El Abbadi (University of California)</i>   |      |
| Machine Learning to Data Management: A Round Trip .....  | 1735 |
| <i>Berti-Equille Laure (Aix-Marseille Université), Bonifati Angela (Université Claude Bernard Lyon 1), and Milo Tova (Tel Aviv University)</i>                                   |      |
| Blockchains and Databases: A New Era in Distributed Computing .....  | 1739 |
| <i>C. Mohan (IBM Almaden Research Center)</i>  |      |

## TKDE Posters

|   |      |
|---|------|
| Searching Trajectories by Regions of Interest .....   | 1741 |
| <i>Shuo Shang (KAUST), Lisi Chen (Wollongong University), Christian S. Jensen (Aalborg University), Ji-Rong Wen (Renmin University of China), and Panos Kalnis (KAUST)</i>  |      |
| Sparse Poisson Latent Block Model for Document Clustering (Extended Abstract) .....   | 1743 |
| <i>Melissa Ailem (INRIA and University of Southern California), Francois Role (Paris Descartes University), and Mohamed Nadif (Paris Descartes University)</i>  |      |
| Feature Grouping-Based Outlier Detection upon Streaming Trajectories (Extended Abstract) .....  | 1745 |
| <i>Jiali Mao (East China Normal University), Tao Wang (East China Normal University), Cheqing Jin (East China Normal University), and Aoying Zhou (East China Normal University)</i>  |      |
| Mining Top-k Co-Occurrence Patterns across Multiple Streams (Extended Abstract) .....   | 1747 |
| <i>Daichi Amagata (Osaka University) and Takahiro Hara (Osaka University)</i>   |      |
| Computing Crowd Consensus with Partial Agreement .....  | 1749 |
| <i>Quoc Viet Hung Nguyen (Griffith University), Huu Viet Huynh (Deakin University), Thanh Tam Nguyen (Ecole Polytechnique Federale de Lausanne), Matthias Weidlich (Humboldt-Universität zu Berlin), Hongzhi Yin (The University of Queensland), and Xiaofang Zhou (The University of Queensland)</i> |      |
| Cleaning Antipatterns in an SQL Query Log .....   | 1751 |
| <i>Natalia Arzamasova (Karlsruhe Institute of Technology), Martin Schäler (Karlsruhe Institute of Technology), and Klemens Böhm (Karlsruhe Institute of Technology)</i>   |      |
| Fast Parallel Path Concatenation for Graph Extraction .....   | 1753 |
| <i>Yingxia Shao (Peking University), Kai Lei (Peking University), Lei Chen (HKUST), Zi Huang (The University of Queensland), Bin Cui (Peking University), Zhongyi Liu (Alibaba Group), Yunhai Tong (Peking University), and Jin Xu (Peking University)</i>  |      |



|   |      |
|---|------|
| Human-Powered Data Cleaning for Probabilistic Reachability Queries on Uncertain Graphs .....  | 1755 |
| <i>Xin Lin (Shanghai Key Laboratory of Multidimensional Information Processing), Yun Peng (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), and Byron Choi (Hong Kong Baptist University)</i>  |      |
| Reducing Uncertainty of Probabilistic Top-k Ranking via Pairwise Crowdsourcing .....  | 1757 |
| <i>Xin Lin (Shanghai Key Laboratory of Multidimensional Information Processing), Jianliang Xu (Shanghai Key Laboratory of Multidimensional Information Processing), Haibo Hu (Shanghai Key Laboratory of Multidimensional Information Processing), and Fan Zhe (Shanghai Key Laboratory of Multidimensional Information Processing)</i> |      |
| KDE-Track: An Efficient Dynamic Density Estimator for Data Streams (Extended Abstract) .....  | 1759 |
| <i>Abdulhakim Qahtan (Qatar Computing Research Institute), Suojin Wang (Texas A&amp;M University), and Xiangliang Zhang (KAUST)</i>   |      |
| Finding Top-k Shortest Paths with Diversity .....   | 1761 |
| <i>Huiping Liu (East China Normal University), Cheqing Jin (East China Normal University), Bin Yang (Aalborg University), and Aoying Zhou (East China Normal University)</i>  |      |
| Linking Fine-Grained Locations in User Comments (Extended Abstract) .....   | 1763 |
| <i>Jialong Han (Tencent AI Lab), Aixin Sun (Nanyang Technological University), Gao Cong (Nanyang Technological University), Wayne Xin Zhao (Renmin University of China), Zongcheng Ji (Nanyang Technological University), and Minh C. Phan (Nanyang Technological University)</i>   |      |
| A Multi-Objective Optimization Approach for Question Routing in Community Question Answering Services (Extended Abstract) .....   | 1765 |
| <i>Xiang Cheng (State Key Laboratory of Networking and Switching Technology), Shuguang Zhu (State Key Laboratory of Networking and Switching Technology), Sen Su (State Key Laboratory of Networking and Switching Technology), and Gang Chen (State Key Laboratory of Networking and Switching Technology)</i>                         |      |
| Incremental Frequent Subgraph Mining on Large Evolving Graphs .....   | 1767 |
| <i>Ehab Abdelhamid (Imperial College London), Mustafa Caim (IBM), Mohammad Sadoghi (University of California), Bishwaranjan Bhattacharjee (IBM), Yuan-Chi Chang (IBM), and Panos Kalnis (King Abdullah University of Science and Technology)</i>  |      |
| Efficient Computation of G-Skyline Groups (Extended Abstract) .....   | 1769 |
| <i>Changping Wang (Tsinghua University), Chaokun Wang (Tsinghua University), Gaoyang Guo (Tsinghua University), Xiaojun Ye (Tsinghua University), and Philip S. Yu (University of Illinois at Chicago)</i>  |      |
| Efficient Top-k Dominating Computation on Massive Data (Extended Abstract) .....  | 1771 |
| <i>Xixian Han (Harbin Institute of Technology), Jianzhong Li (Harbin Institute of Technology), and Hong Gao (Harbin Institute of Technology)</i>  |      |

|   |      |
|---|------|
| MOSS-5: A Fast Method of Approximating Counts of 5-Node Graphlets in Large Graphs<br>(Extended Abstract) .....  | 1773 |
| <i>Pinghui Wang (Xi'an Jiaotong University), Junzhou Zhao (King Abdullah University of Science and Technology), Xiangliang Zhang (King Abdullah University of Science and Technology), Zhenguo Li (Huawei Noah's Ark Lab), Jiefeng Cheng (Tencent), John C.S. Lui (The Chinese University of Hong Kong), Don Towsley (University of Massachusetts Amherst), and Xiaohong Guan (Xi'an Jiaotong University)</i> |      |
| Geo-Social Influence Spanning Maximization .....  | 1775 |
| <i>Jianxin Li (University of Western Australia), Timos Sellis (Swinburne University of Technology), J. Shane Culpepper (RMIT), Zhenying He (Fudan University), Chengfei Liu (Swinburne University of Technology), and Junhu Wang (Griffith University)</i>  |      |
| Redundancy Reduction for Prevalent Co-Location Patterns .....   | 1777 |
| <i>Lizhen Wang (Yunnan University), Xuguang Bao (Yunnan University), and lihua Zhou (Yunnan University)</i>   |      |
| LS-Join: Local Similarity Join on String Collections (Extended Abstract) .....  | 1779 |
| <i>Jiaying Wang (Shenyang Jianzhu University), Xiaochun Yang (Northeastern University), Bin Wang (Northeastern University), and Chengfei Liu (Swinburne University of Technology)</i>   |      |
| Time-Aware Boolean Spatial Keyword Queries (Extended Abstract) .....  | 1781 |
| <i>Gang Chen (Zhejiang University), Jingwen Zhao (Zhejiang University), Yunjun Gao (Zhejiang University), Lei Chen (Hong Kong University of Science and Technology), and Rui Chen (Zhejiang University)</i>   |      |
| Efficient Clue-Based Route Search on Road Networks (Extended Abstract) .....  | 1783 |
| <i>Bolong Zheng (Sun Yat-sen University), Han Su (Big Data Research Center), Wen Hua (University of Queensland), Kai Zheng (Big Data Research Center), Xiaofang Zhou (University of Queensland), and Guohui Li (Huazhong University of Science and Technology)</i>  |      |
| Reverse k Nearest Neighbor Search over Trajectories (Extended Abstract) .....   | 1785 |
| <i>Sheng Wang (RMIT University), Zhifeng Bao (RMIT University), J. Shane Culpepper (RMIT University), Timos Sellis (Swinburne University of Technology), and Gao Cong (Nanyang Technological University)</i>  |      |
| Metric Similarity Joins Using MapReduce (Extended Abstract) .....   | 1787 |
| <i>Gang Chen (Zhejiang University), Keyu Yang (Zhejiang University), Lu Chen (Zhejiang University), Yunjun Gao (Zhejiang University), Baihua Zheng (Singapore Management University), and Chun Chen (Zhejiang University)</i>   |      |
| Community Deception - Or: How to Stop Fearing Community Detection Algorithms (Extended Abstract) .....  | 1789 |
| <i>Valeria Fionda (DeMACS) and Giuseppe Pirrò (Institute for High Performance Computing and Networking)</i>   |      |
| Deep Learning of Graphs with Ngram Convolutional Neural Networks (Extended Abstract) .....  | 1791 |
| <i>Zhiling Luo (Zhejiang University), Ling Liu (Georgia Institute of Technology), Jianwei Yin (Zhejiang University), Ying Li (Zhejiang University), and Zhaohui Wu (Zhejiang University)</i>  |      |

|  |      |
|--|------|
| An Empirical Evaluation of Techniques for Ranking Semantic Associations (Extended Abstract) .....  | 1793 |
| <i>Gong Cheng (Nanjing University), Fei Shao (Nanjing University), and Yuzhong Qu (Nanjing University)</i>   |      |
| An Efficient Ride-Sharing Framework for Maximizing Shared Routes .....   | 1795 |
| <i>Na Ta (Renmin University of China), Guoliang Li (Tsinghua University), Tianyu Zhao (Tsinghua University), Jianhua Feng (Tsinghua University), Hanchao Ma (Washington State University), and Zhiguo Gong (Macau University)</i>  |      |
| Authenticating Aggregate Queries over Set-Valued Data with Confidentiality (Extended Abstract) .....   | 1797 |
| <i>Cheng Xu (Hong Kong Baptist University), Qian Chen (Hong Kong Baptist University), Haibo Hu (Hong Kong Polytechnic University), Jianliang Xu (Hong Kong Baptist University), and Xiaojun Hei (Huazhong University of Science and Technology)</i>  |      |
| FORWARD: A Model for FOREcasting Reservoir WAtER Dynamics Using Spatial Bayesian Network (SpaBN) (Extended Abstract) .....   | 1799 |
| <i>Monidipa Das (Indian Institute of Technology Kharagpur), Soumya K. Ghosh (Indian Institute of Technology Kharagpur), Pramesh Gupta (Indian Institute of Technology Kharagpur), Vemuri M. Chowdary (Indian Space Research Organization), Ravoori Nagaraja (Indian Space Research Organization), and Vinay Kumar Dadhwal (Indian Space Research Organization)</i> |      |
| Efficient Information Flow Maximization in Probabilistic Graphs (Extended Abstract) .....  | 1801 |
| <i>Christian Frey (Ludwig-Maximilians-Universität München), Andreas Züfle (George Mason University), Tobias Emrich (Harman International), and Matthias Renz (George Mason University)</i>   |      |
| Continuous Top-k Monitoring on Document Streams (Extended Abstract) .....  | 1803 |
| <i>Leong Hou U (University of Macau), Junjie Zhang (University of Macau), Kyriakos Mouratidis (Singapore Management University), and Ye Li (University of Macau)</i>   |      |
| Finding Related Forum Posts through Content Similarity over Intention-Based Segmentation (Extended Abstract) .....   | 1805 |
| <i>Dimitra Papadimitriou (University of Trento), Georgia Koutrika (Athena Research Center), Yannis Velegrakis (University of Trento), and John Mylopoulos (University of Trento)</i>   |      |
| Keyword Search on Temporal Graphs .....  | 1807 |
| <i>Ziyang Liu (Facebook Inc.), Chong Wang (New Jersey Institute of Technology), and Yi Chen (New Jersey Institute of Technology)</i>   |      |
| Game-Theoretic Cross Social Media Analytic: How Do Yelp Ratings Affect Deal Selection on Groupon? (Extended Abstract) .....  | 1809 |
| <i>Chih-Yu Wang (Academia Sinica), Yan Chen (University of Electronic Science and Technology of China), and K.J. Ray Liu (University of Maryland)</i>  |      |

|  |             |
|--|-------------|
| Efficient Retrieval of Bounded-Cost Informative Routes .....   | 1811        |
| <i>Wengen Li (The Hong Kong Polytechnic University), Jiannong Cao (The Hong Kong Polytechnic University), Jihong Guan (Tongji University), Man Lung Yiu (The Hong Kong Polytechnic University), and Shuigeng Zhou (Fudan University)</i>   |             |
| Dynamic Facet Ordering for Faceted Product Search Engines (Extended Abstract) .....  | 1813        |
| <i>Damir Vandic (Erasmus University Rotterdam), Steven S. Aanen (Grible.co), Flavius Frasinca (Erasmus University Rotterdam), and Uzay Kaymak (Eindhoven University of Technology)</i>   |             |
| Answering Natural Language Questions by Subgraph Matching over Knowledge Graphs (Extended Abstract) .....  | 1815        |
| <i>Sen Hu (Peking University), Lei Zou (Peking University), Jeffery Xu Yu (The Chinese University of Hong Kong), Haixun Wang (Google Research), and Dongyan Zhao (Peking University)</i>   |             |
| Efficient Parameter Estimation for Information Retrieval Using Black-Box Optimization (Extended Abstract) .....  | 1817        |
| <i>Alberto Costa (ETH Zurich), Emanuele Di Buccio (Universita' di Padova), Massimo Melucci (University of Padova), and Giacomo Nannicini (IBM Research)</i>  |             |
| SAP: Improving Continuous Top-K Queries over Streaming Data .....  | 1819        |
| <i>Rui Zhu (Shenyang Aerospace University), Bin Wang (Northeastern University-Liaoning), Xiaochun Yang (Northeastern University-Liaoning), Baihua Zheng (Singapore Management University), and Guoren Wang (Northeastern University-Liaoning)</i>  |             |
| Efficient and Scalable Integrity Verification of Data and Query Results for Graph Databases .....  | 1821        |
| <i>Muhammad U. Arshad (Purdue University), Ashish Kundu (IBM T. J. Watson Research Center), Elisa Bertino (Purdue University), Arif Ghafoor (Purdue University), and Chinmay Kundu (IEEE Member)</i>   |             |
| Workload Management in Database Management System: A Taxonomy (Extended Abstract) .....  | 1823        |
| <i>Mingyi Zhang (Huawei America Research), Patrick Martin (Queen's University), Wendy Powley (Queen's University), and Jianjun Chen (Huawei America Research)</i>  |             |
| Clarifying Trust in Social Internet of Things (Extended Abstract) .....  | 1825        |
| <i>Zhiting Lin (Anhui University) and Liang Dong (Baylor University)</i>   |             |
| Enhancing Binary Classification by Modeling Uncertain Boundary in Three-Way Decisions (Extended Abstract) .....  | 1827        |
| <i>Yuefeng Li (Queensland University of Technology), Libiao Zhang (Queensland University of Technology), Yue Xu (Queensland University of Technology), Yiyu Yao (University of Regina), Raymond Y.K. Lau (City University of Hong Kong), and Yutong Wu (Queensland University of Technology)</i> |             |
| <b>Author Index .....</b>  | <b>1829</b> |