2019 IEEE 35th International Conference on Data Engineering (ICDE)

ICDE 2019

Table of Contents

Message from the ICDE 2019 Chairs ..... **xxxiv**
Organizing Committee ..... **xxxvi**
Program Committee ..... **xxxviii**

Keynotes

AsterixDB Mid-Flight: A Case Study in Building Systems in Academia 1
Michael J. Carey (University of California, Irvine)

Data Management at Huawei: Recent Accomplishments and Future Challenges 13
Jianjun Chen (Huawei America Research), Yu Chen (Huawei America Research), Zhibiao Chen (Huawei Technologies Co., Ltd.), Ahmad Ghazal (Huawei America Research), Guoliang Li (Huawei Technologies Co., Ltd.), Sihao Li (Huawei Technologies Co., Ltd.), Weijie Ou (Huawei Technologies Co., Ltd.), Yang Sun (Huawei America Research), Mingyi Zhang (Huawei America Research), and Minqi Zhou (Huawei Technologies Co., Ltd.)

Building a Broad Knowledge Graph for Products 25
Xin Luna Dong (Amazon Inc)

Knowledge Graphs and Enterprise AI: The Promise of an Enabling Technology 26
Luigi Bellomarini (Banca d'Italia), Daniele Fakhoury (Universita Tor Vergata), Georg Gottlob (University of Oxford & TU Wien), and Emanuel Sallinger (University of Oxford & TU Wien)

Panel

Is There a Data Science and Engineering Brain Drain? If So, How Can We Rebalance Them? 38
Jian Pei (JD.com and Simon Fraser University)

Research Track

Research (1) - Graphs 1: Query Processing 1

Answering Why-Questions for Subgraph Queries in Multi-attributed Graphs 40
Qi Song (Washington State University), Mohammad Hossein Namaki (Washington State University), and Yinghui Wu (Washington State University; Pacific Northwest National Laboratory)

Enumerating k-Vertex Connected Components in Large Graphs 52
Dong Wen (University of Technology Sydney), Lu Qin (University of Technology Sydney), Ying Zhang (University of Technology Sydney), Lijun Chang (University of Sydney), and Ling Chen (University of Technology Sydney)

Index-Based Optimal Algorithm for Computing K-Cores in Large Uncertain Graphs 64
Bohua Yang (University of Technology Sydney), Dong Wen (University of Technology Sydney), Lu Qin (University of Technology Sydney), Ying Zhang (University of Technology Sydney), Lijun Chang (The University of Sydney), and Rong-Hua Li (Beijing Institute of Technology)

Computing a Near-Maximum Independent Set in Dynamic Graphs 76
Weiguo Zheng (Fudan University), Chengzhi Piao (The Chinese University of Hong Kong), Hong Cheng (The Chinese University of Hong Kong), and Jeffrey Xu Yu (The Chinese University of Hong Kong)

Contextual Community Search Over Large Social Networks 88
Lu Chen (Swinburne University of Technology), Chengfei Liu (Swinburne University of Technology), Kewen Liao (Charles Darwin University), Jianxin Li (Deakin University), and Rui Zhou (Swinburne University of Technology)

Research (2) - Distributed, Parallel and P2P Data Management

Rima: An RDMA-Accelerated Model-Parallelized Solution to Large-Scale Matrix Factorization 100
Jinkun Geng (Tsinghua University), Dan Li (Tsinghua University), and Shuai Wang (Tsinghua University)

Accelerating Partial Evaluation in Distributed SPARQL Query Evaluation 112
Peng Peng (Hunan University), Lei Zou (Peking University; Beijing Institute of Big Data Research; National Engineering Laboratory for Big Data Analysis Technology and Application (PKU), China), and Runyu Guan (Hunan University)

Blockplane: A Global-Scale Byzantizing Middleware 124
Faisal Nawab (University of California) and Mohammad Sadoghi (UC Davis)

BENU: Distributed Subgraph Enumeration with Backtracking-Based Framework 136
Zhaokang Wang (Nanjing University), Rong Gu (Nanjing University), Weiwei Hu (Nanjing University), Chunfeng Yuan (Nanjing University), and Yihua Huang (Nanjing University)

Efficient Synchronization of State-Based CRDTs 148
Vitor Enes (HASLab / INESC TEC and Universidade do Minho), Paulo Sérgio Almeida (HASLab / INESC TEC and Universidade do Minho), Carlos Baquero (HASLab / INESC TEC and Universidade do Minho), and João Leitão (NOVA LINCS, FCT and Universidade NOVA de Lisbon)

Research (3) - Data Cleaning

Learning Individual Models for Imputation 160
Aoqian Zhang (Tsinghua University), Shaoxu Song (Tsinghua University), Yu Sun (Tsinghua University), and Jianmin Wang (Tsinghua University)

CurrentClean: Spatio-Temporal Cleaning of Stale Data 172
Zheng Zheng (McMaster University), Mostafa Milani (McMaster University), and Fei Chiang (McMaster University)

Fine-Grained Provenance for Matching & ETL 184
Nan Zheng (University of Pennsylvania), Abdussalam Alawini (University of Illinois, Urbana-Champaign), and Zachary G. Ives (University of Pennsylvania)

Unsupervised String Transformation Learning for Entity Consolidation 196
Dong Deng (Rutgers University), Wenbo Tao (MIT), Ziawasch Abedjan (TU Berlin), Ahmed Elmagarmid (QCRI HBKU), Ihab F. Ilyas (University of Waterloo), Guoliang Li (Tsinghua University), Samuel Madden (MIT), Mourad Ouzzani (QCRI HBKU), Michael Stonebraker (MIT), and Nan Tang (QCRI HBKU)

A Semi-Supervised Framework of Clustering Selection for De-Duplication 208
Shrinu Kushagra (University of Waterloo), Hemant Saxena (University of Waterloo), Ihab F. Ilyas (University of Waterloo), and Shai Ben-David (University of Waterloo)

Research (4) - Graphs 2: Query Processing and Indexing

Scaling Up Subgraph Query Processing with Efficient Subgraph Matching 220
Shixuan Sun (Hong Kong University of Science and Technology) and Qiong Luo (Hong Kong University of Science and Technology)

Efficient Parallel Subgraph Enumeration on a Single Machine 232
Shixuan Sun (Hong Kong University of Science and Technology), Yulin Che (Hong Kong University of Science and Technology), Lipeng Wang (Hong Kong University of Science and Technology), and Qiong Luo (Hong Kong University of Science and Technology)

Fast Dual Simulation Processing of Graph Database Queries 244
Stephan Mennicke (Institute for Information Systems, TU Braunschweig), Jan-Christoph Kalo (Institute for Information Systems, TU Braunschweig), Denis Nagel (Institute for Information Systems, TU Braunschweig), Hermann Kroll (Institute for Information Systems, TU Braunschweig), and Wolf-Tilo Balke (Institute for Information Systems, TU Braunschweig)

Efficient and Incremental Clustering Algorithms on Star-Schema Heterogeneous Graphs 256
Lu Chen (Zhejiang University), Yunjun Gao (Zhejiang University), Yuanliang Zhang (Zhejiang University), Christian S. Jensen (Aalborg University), and Bolong Zheng (Aalborg University)

G*-Tree: An Efficient Spatial Index on Road Networks 268
Zijian Li (Hong Kong University of Science and Technology), Lei Chen (Hong Kong University of Science and Technology), and Yue Wang (Hong Kong University of Science and Technology)

Research (5) - Recommendation and Clustering

DBSVEC: Density-Based Clustering Using Support Vector Expansion 280
Zhen Wang (Tsinghua University), Rui Zhang (The University of Melbourne), Jianzhong Qi (The University of Melbourne), and Bo Yuan (Tsinghua University)

A Joint Context-Aware Embedding for Trip Recommendations 292
Jiayuan He (The University of Melbourne), Jianzhong Qi (The University of Melbourne), and Kotagiri Ramamohanarao (The University of Melbourne)

AIR: Attentional Intention-Aware Recommender Systems 304
Tong Chen (The University of Queensland), Hongzhi Yin (The University of Queensland), Hongxu Chen (The University of Queensland), Rui Yan (Peking University), Quoc Viet Hung Nguyen (Griffith University), and Xue Li (The University of Queensland and Dalian Neusoft University of Information)

No, That's Not My Feedback: TV Show Recommendation Using Watchable Interval 316
Kyung-Jae Cho (Hanyang University), Yeon-Chang Lee (Hanyang University), Kyungsik Han (Ajou University), Jaeho Choi (Naver Corporation), and Sang-Wook Kim (Hanyang University)

Adaptive Wavelet Clustering for Highly Noisy Data 328
Zengjian Chen (Huazhong University of Science and Technology), Jiayi Liu (University of Massachusetts Amherst), Yihe Deng (University of California, Los Angeles), Kun He (Huazhong University of Science and Technology), and John E. Hopcroft (Cornell University)

Research (6) - Strings, Texts, and Keyword Search

An Efficient Parallel Keyword Search Engine on Knowledge Graphs 338
Yueji Yang (National University of Singapore), Divykant Agrawal (University of California at Santa Barbara), H. V. Jagadish (University of Michigan of Ann Arbor), Anthony K. H. Tung (National University of Singapore), and Shuang Wu (National University of Singapore)

Towards Longitudinal Analytics on Social Media Data 350
Fan Xia (East China Normal University), Bin Yang (Aalborg University), Chengcheng Yu (Shanghai Polytechnic University), Weining Qian (East China Normal University), and Aoying Zhou (East China Normal University)

LCJoin: Set Containment Join via List Crosscutting 362
Dong Deng (Inception Institute of Artificial Intelligence), Chengcheng Yang (Inception Institute of Artificial Intelligence), Shuo Shang (Inception Institute of Artificial Intelligence), Fan Zhu (Inception Institute of Artificial Intelligence), Li Liu (Inception Institute of Artificial Intelligence), and Ling Shao (Inception Institute of Artificial Intelligence)

Bridging the Semantic Gap with SQL Query Logs in Natural Language Interfaces to Databases 374
Christopher Baik (University of Michigan), H. V. Jagadish (University of Michigan), and Yunyao Li (IBM Research - Almaden)

MF-Join: Efficient Fuzzy String Similarity Join with Multi-level Filtering 386
Jin Wang (UCLA), Chunbin Lin (Amazon AWS), and Carlo Zaniolo (UCLA)

Research (7) - Graphs 3: Social Networks

Finding Temporal Influential Users Over Evolving Social Networks 398
Shixun Huang (RMIT University), Zhifeng Bao (RMIT University), J. Shane Culpepper (RMIT University), and Bang Zhang (Damo Academy Alibaba)

Seed Selection and Social Coupon Allocation for Redemption Maximization in Online Social Networks 410
Tung-Chun Chang (Research Center for Information Technology Innovation, Academia Sinica, Taiwan), Yishuo Shi (Institute of Information Science, Academia Sinica, Taiwan), De-Nian Yang (Institute of Information Science, Academia Sinica, Taiwan), and Wen-Tsuen Chen (Department of Computer Science, National Tsing Hua University, Taiwan)

Keyword-Centric Community Search 422
Zhiwei Zhang (Hong Kong Baptist University), Xin Huang (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), Byron Choi (Hong Kong Baptist University), and Zechao Shang (The University of Chicago)

Cohesive Group Nearest Neighbor Queries Over Road-Social Networks 434
Fangda Guo (Northeastern University), Ye Yuan (Northeastern University), Guoren Wang (Beijing Institute of Technology), Lei Chen (Hong Kong University of Science and Technology), Xiang Lian (Kent State University), and Zimeng Wang (Northeastern University)

Maximizing Multifaceted Network Influence 446
Yuchen Li (Singapore Management University), Ju Fan (Renmin University of China), George Ovchinnikov (Skolkovo Institute of Science and Technology), and Panagiotis Karras (Aarhus University)

Research (8) - Approximation

GB-KMV: An Augmented KMV Sketch for Approximate Containment Similarity Search 458
Yang Yang (University of New South Wales), Ying Zhang (University of Technology Sydney), Wenjie Zhang (University of New South Wales), and Zengfeng Huang (Fudan University)

ARROW: Approximating Reachability Using Random Walks Over Web-Scale Graphs 470
Neha Sengupta (Indian Institute of Technology Delhi), Amitabha Bagchi (Indian Institute of Technology Delhi), Maya Ramanath (Indian Institute of Technology Delhi), and Srikanta Bedathur (Indian Institute of Technology Delhi)

Taster: Self-Tuning, Elastic and Online Approximate Query Processing 482
Matthaios Olma (EPFL), Odysseas Papapetrou (TU Eindhoven), Raja Appuswamy (EURECOM), and Anastasia Ailamaki (EPFL)

An Iterative Scheme for Leverage-Based Approximate Aggregation 494
Shanshan Han (Harbin Institute of Technology), Hongzhi Wang (Harbin Institute of Technology), Jialin Wan (Harbin Institute of Technology), and Jianzhong Li (Harbin Institute of Technology)

Deletion Propagation for Multiple Key Preserving Conjunctive Queries: Approximations and Complexity 506
Zhipeng Cai (Georgia State University), Dongjing Miao (Harbin Institute of Technology; Georgia State University), and Yingshu Li (Georgia State University)

Research (9) - Machine Learning and Data Science

Enumerating Minimal Weight Set Covers 518
Zahi Ajami (The Hebrew University of Jerusalem) and Sara Cohen (The Hebrew University of Jerusalem)

Constraints-Based Explanations of Classifications 530
Daniel Deutch (Tel Aviv University) and Nave Frost (Tel Aviv University)

KARL: Fast Kernel Aggregation Queries 542
Tsz Nam Chan (Hong Kong Polytechnic University, The University of Hong Kong), Man Lung Yiu (Hong Kong Polytechnic University), and Leong Hou U (University of Macau)

Assessing and Remedying Coverage for a Given Dataset 554
Abolfazl Asudeh (University of Michigan), Zhongjun Jin (University of Michigan), and H. V. Jagadish (University of Michigan)

Research (10) - Graphs 4: Mining 1

Social Influence-Based Group Representation Learning for Group Recommendation 566
Hongzhi Yin (The University of Queensland), Qinyong Wang (The University of Queensland), Kai Zheng (University of Electronic Science and Technology of China), Zhixu Li (Soochow University), Jiali Yang (Soochow University), and Xiaofang Zhou (The University of Queensland)

MIDAS: Finding the Right Web Sources to Fill Knowledge Gaps 578
Xiaolan Wang (University of Massachusetts Amherst), Xin Luna Dong (Amazon Inc.), Yang Li (Google Inc.), and Alexandra Meliou (University of Massachusetts Amherst)

Exploiting Centrality Information with Graph Convolutions for Network Representation Learning 590
Hongxu Chen (The University of Queensland), Hongzhi Yin (The University of Queensland), Tong Chen (The University of Queensland), Quoc Viet Hung Nguyen (Griffith University), Wen-Chih Peng (National Chiao Tung University), and Xue Li (The University of Queensland; Dalian Neusoft University of Information)

Route Recommendations on Road Networks for Arbitrary User Preference Functions 602
Pranali Yawalkar (Google) and Sayan Ranu (IIT Delhi)

NSCaching: Simple and Efficient Negative Sampling for Knowledge Graph Embedding 614
Yongqi Zhang (the Hong Kong University of Science and Technology), Quanming Yao (4Paradigm), Yingxia Shao (Beijing University of Posts and Telecommunications), and Lei Chen (the Hong Kong University of Science and Technology)

Research (11) - Database Privacy, Security, and Trust

ServeDB: Secure, Verifiable, and Efficient Range Queries on Outsourced Database 626
Songrui Wu (Tsinghua University), Qi Li (Tsinghua University), Guoliang Li (Tsinghua University), Dong Yuan (Tsinghua University), Xingliang Yuan (Monash University), and Cong Wang (City University of Hong Kong)

Collecting and Analyzing Multidimensional Data with Local Differential Privacy 638
Ning Wang (Ocean University of China), Xiaokui Xiao (National University of Singapore), Yin Yang (Hamad Bin Khalifa University), Jun Zhao (Nanyang Technological University), Siu Cheung Hui (Nanyang Technological University), Hyejin Shin (Samsung Electronics), Junbum Shin (Samsung Electronics), and Ge Yu (Northeastern University)

Partitioned Data Security on Outsourced Sensitive and Non-Sensitive Data 650
Sharad Mehrotra (University of California, Irvine), Shantanu Sharma (University of California, Irvine), Jeffrey Ullman (Stanford University), and Anurag Mishra (University of California, Irvine)

SecEQP: A Secure and Efficient Scheme for SkNN Query Problem Over Encrypted Geodata on Cloud 662
Xinyu Lei (Michigan State University), Alex X. Liu (Michigan State University), Rui Li (Dongguan University of Technology), and Guan-Hua Tu (Michigan State University)

Joins Over Encrypted Data with Fine Granular Security 674
Florian Hahn (SAP SE), Nicolas Loza (SAP SE), and Florian Kerschbaum (University of Waterloo)

Research (12) - Modern Hardware and In-Memory Database Systems

Column-Oriented Database Acceleration Using FPGAs 686
Satoru Watanabe (Hitachi, Ltd.), Kazuhisa Fujimoto (Hitachi, Ltd.), Yuji Saeki (Hitachi, Ltd.), Yoshifumi Fujikawa (Hitachi, Ltd.), and Hiroshi Yoshino (Hitachi ULSI Systems Co., Ltd.)

Hardware-Conscious Hash-Joins on GPUs 698
Panagiotis Sioulas (École polytechnique fédérale de Lausanne), Periklis Chrysogelos (École polytechnique fédérale de Lausanne), Manos Karpathiotakis (École polytechnique fédérale de Lausanne), Raja Appuswamy (École polytechnique fédérale de Lausanne), and Anastasia Ailamaki (École polytechnique fédérale de Lausanne)

TuFast: A Lightweight Parallelization Library for Graph Analytics 710
Zechao Shang (University of Chicago), Jeffrey Xu Yu (Chinese University of Hong Kong), and Zhiwei Zhang (Hong Kong Baptist University)

LDC: A Lower-Level Driven Compaction Method to Optimize SSD-Oriented Key-Value Stores 722
Yunpeng Chai (Renmin University of China), Yanfeng Chai (Renmin University of China), Xin Wang (Tianjin University), Haocheng Wei (Renmin University of China), Ning Bao (Renmin University of China), and Yushi Liang (Renmin University of China)

No False Negatives: Accepting All Useful Schedules in a Fast Serializable Many-Core System 734
Dominik Durner (Technical University of Munich) and Thomas Neumann (Technical University of Munich)

Research (13) - Graphs 5: Mining 2

Discovering Maximal Motif Cliques in Large Heterogeneous Information Networks 746
Jiafeng Hu (The University of Hong Kong), Reynold Cheng (The University of Hong Kong), Kevin Chen-Chuan Chang (University of Illinois at Urbana-Champaign), Aravind Sankar (University of Illinois at Urbana-Champaign), Yixiang Fang (The University of Hong Kong), and Brian Y.H. Lam (University of Cambridge)

REPT: A Streaming Algorithm of Approximating Global and Local Triangle Counts in Parallel 758
Pinghui Wang (Xi'an Jiaotong University), Peng Jia (Xi'an Jiaotong University, China), Yiyan Qi (Xi'an Jiaotong University, China), Yu Sun (Xi'an Jiaotong University, China), Jing Tao (Xi'an Jiaotong University), and Xiaohong Guan (Xi'an Jiaotong University)

Information Diffusion Prediction via Recurrent Cascades Convolution 770
Xueqin Chen (University of Electronic Science and Technology of China), Fan Zhou (University of Electronic Science and Technology of China), Kunpeng Zhang (University of Maryland, College Park), Goce Trajcevski (Iowa State University), Ting Zhong (University of Electronic Science and Technology of China), and Fengli Zhang (University of Electronic Science and Technology of China)

Finding Densest Lasting Subgraphs in Dynamic Graphs: A Stochastic Approach 782
Xuanming Liu (University of Massachusetts, Lowell), Tingjian Ge (University of Massachusetts, Lowell), and Yinghui Wu (Washington State University)

Multicapacity Facility Selection in Networks 794
Alvis Logins (Aarhus University), Panagiotis Karras (Aarhus University), and Christian S. Jensen (Aalborg University)

Research (14) - Query Processing, Indexing and Optimization

An MBR-Oriented Approach for Efficient Skyline Query Processing 806
Ji Zhang (Auburn University), Wenlu Wang (Auburn University), Xunfei Jiang (Earlham College), Wei-Shinn Ku (Auburn University), and Hua Lu (Aalborg University)

Dynamic Set kNN Self-Join 818
Daichi Amagata (Osaka University), Takahiro Hara (Osaka University), and Chuan Xiao (Nagoya University)

Packed Memory Arrays - Rewired 830
Dean De Leo (CWI) and Peter Boncz (CWI)

GEM^2-Tree: A Gas-Efficient Structure for Authenticated Range Queries in Blockchain 842
Ce Zhang (Hong Kong Baptist University), Cheng Xu (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), Yuzhe Tang (Syracuse University), and Byron Choi (Hong Kong Baptist University)

Effective Filters and Linear Time Verification for Tree Similarity Joins 854
Thomas Hütter (University of Salzburg), Mateusz Pawlik (University of Salzburg), Robert Löschinger (University of Salzburg), and Nikolaus Augsten (University of Salzburg)

Research (15) - Temporal and Spatial Data 1

KV-Match: A Subsequence Matching Approach Supporting Normalization and Time Warping 866
Jiaye Wu (Fudan University), Peng Wang (Fudan University), Ningting Pan (Fudan University), Chen Wang (Tsinghua University), Wei Wang (Fudan University), and Jianmin Wang (Tsinghua University)

Efficient Maximal Spatial Clique Enumeration 878
Chen Zhang (University of New South Wales, Australia), Ying Zhang (University of Technology Sydney, Australia), Wenjie Zhang (University of New South Wales, Australia), Lu Qin (University of Technology Sydney, Australia), and Jianye Yang (Ant Financial Services Group, Alibaba Group, China)

Cluster-Based Subscription Matching for Geo-Textual Data Streams 890
Lisi Chen (Inception Institute of Artificial Intelligence (IIAI), UAE), Shuo Shang (UESTC and IIAI), Kai Zheng (UESTC), and Panos Kalnis (KAUST)

Time-Dependent Hop Labeling on Road Network 902
Lei Li (The University of Queensland, Australia), Sibo Wang (The Chinese University of Hong Kong), and Xiaofang Zhou (The University of Queensland, Australia)

Weight-Constrained Route Planning Over Time-Dependent Graphs 914
Ye Yuan (Beijing Institute of Technology), Xiang Lian (Kent State University), Guoren Wang (Beijing Institute of Technology), Lei Chen (Hong Kong University of Science and Technology), Yuliang Ma (Northeastern University, China), and Yishu Wang (Northeastern University, China)

Research (16) - Graphs 6: Data Science

Skyline Queries Constrained by Multi-cost Transportation Networks 926
Qixu Gong (New Mexico State University), Huiping Cao (New Mexico State University), and Parth Nagarkar (New Mexico State University)

Online Social Media Recommendation Over Streams 938
Xiangmin Zhou (RMIT University, Melbourne, Australia), Dong Qin (RMIT University, Melbourne, Australia), Xiaolu Lu (RMIT University, Melbourne, Australia), Lei Chen (Hong Kong University of Science and Technology, Hong Kong, China), and Yanchun Zhang (Victoria University, Melbourne, Australia; CIAT, Guangzhou University, China)

Canonicalization of Open Knowledge Bases with Side Information from the Source Text 950
Xueling Lin (The Hong Kong University of Science and Technology) and Lei Chen (The Hong Kong University of Science and Technology)

Walking with Perception: Efficient Random Walk Sampling via Common Neighbor Awareness 962
Yongkun Li (University of Science and Technology of China), Zhiyong Wu (University of Science and Technology of China), Shuai Lin (University of Science and Technology of China), Hong Xie (Chongqing University), Min Lv (University of Science and Technology of China), Yinlong Xu (University of Science and Technology of China), and John C.S. Lui (The Chinese University of Hong Kong)

Research (17) - Search and Information Extraction

SimMeme: A Search Engine for Internet Memes 974
Tova Milo (Tel Aviv Univesity), Amit Somech (Tel Aviv Univesity), and Brit Youngmann (Tel Aviv Univesity)

A Hierarchical Framework for Top-k Location-Aware Error-Tolerant Keyword Search 986
Junye Yang (Tsinghua University), Yong Zhang (Tsinghua University), Xiaofang Zhou (The University of Queensland), Jin Wang (University of California), Huiqi Hu (East China Normal University), and Chunxiao Xing (Tsinghua University)

2ED: An Efficient Entity Extraction Algorithm Using Two-Level Edit-Distance 998
Zeyi Wen (National University of Singapore), Dong Deng (Rutgers University & Inception Institute of Artificial Intelligence), Rui Zhang (The University of Melbourne), and Ramamohanarao Kotagiri (The University of Melbourne)

Bridging Quantities in Tables and Text 1010
Yusra Ibrahim (Max Planck Institute for Informatics), Mirek Riedewald (Northeastern University), Gerhard Weikum (Max Planck Institute for Informatics), and Demetrios Zeinalipour-Yazti (University of Cyprus)

Research (18) - Temporal and Spatial Data 2

An Efficient Insertion Operator in Dynamic Ridesharing Services 1022
Yi Xu (Beihang University, China), Yongxin Tong (Beihang University, China), Yexuan Shi (Beihang University, China), Qian Tao (Beihang University, China), Ke Xu (Beihang University, China), and Wei Li (Beihang University, China)

Auction-Based Order Dispatch and Pricing in Ridesharing 1034
Libin Zheng (The Hong Kong University of Science and Technology), Peng Cheng (The Hong Kong University of Science and Technology), and Lei Chen (The Hong Kong University of Science and Technology)

When Geo-Text Meets Security: Privacy-Preserving Boolean Spatial Keyword Queries 1046
Ningning Cui (Northeastern University), Jianxin Li (Deakin University), Xiaochun Yang (Northeastern University), Bin Wang (Northeastern University), Mark Reynolds (University of Western Australia), and Yong Xiang (Deakin University)

Moving Object Linking Based on Historical Trace 1058
Fengmei Jin (Renmin University of China; The University of Queensland), Wen Hua (The University of Queensland), Jiajie Xu (Soochow University; Neusoft Corporation), and Xiaofang Zhou (The University of Queensland; Soochow University)

ImageProof: Enabling Authentication for Large-Scale Image Retrieval 1070
Shangwei Guo (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), Ce Zhang (Hong Kong Baptist University), Cheng Xu (Hong Kong Baptist University), and Tao Xiang (Chongqing University)

Research (19) - Graphs 7: Mining 3

Time Constrained Continuous Subgraph Search Over Streaming Graphs 1082
Youhuan Li (Center for Data Science, Peking University, China), Lei Zou (Institute of Computer Science and Technology, Peking University, China), M. Tamer Özsu (David R. Cheriton School of Computer Science), and Dongyan Zhao (Institute of Computer Science and Technology, Peking University, China)

Utilizing Dynamic Properties of Sharing Bits and Registers to Estimate User Cardinalities Over Time 1094
Pinghui Wang (MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, China; Shenzhen Research Institute of Xi'an Jiaotong University, Shenzhen, China), Peng Jia (MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, China), Xiangliang Zhang (King Abdullah University of Science and Technology, Thuwal, SA), Jing Tao (MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, China; Shenzhen Research Institute of Xi'an Jiaotong University, Shenzhen, China; Zhejiang Research Institute of Xi'an Jiaotong University, Hangzhou, China), Xiaohong Guan (Shenzhen Research Institute of Xi'an Jiaotong University, Shenzhen, China; MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, China; Department of Automation and NLIST Lab, Tsinghua University, Beijing, China), and Don Towsley (School of Computer Science, University of Massachusetts Amherst, MA, USA)

Tracking Influential Nodes in Time-Decaying Dynamic Interaction Networks 1106
Junzhou Zhao (King Abdullah University of Science and Technology), Shuo Shang (Inception Institute of Artificial Intelligence), Pinghui Wang (Xi'an Jiaotong University), John C.S. Lui (The Chinese University of Hong Kong), and Xiangliang Zhang (King Abdullah University of Science and Technology)

Fast and Accurate Graph Stream Summarization 1118
Xiangyang Gou (Peking University), Lei Zou (Peking University), Chenxingyu Zhao (Peking University), and Tong Yang (Peking University)

Mining Periodic Cliques in Temporal Networks 1130
Hongchao Qin (Northeastern University), Rong-Hua Li (Beijing Institute of Technology), Guoren Wang (Beijing Institute of Technology), Lu Qin (University of Technology Sydney), Yurong Cheng (Beijing Institute of Technology), and Ye Yuan (Northeastern University)

Research (20) - Data Integration, Probabilistic Databases

Coloring Embedder: A Memory Efficient Data Structure for Answering Multi-set Query 1142
Yang Tong (Peking University), Dongsheng Yang (Peking University), Jie Jiang (Peking University), Siang Gao (Peking University), Bin Cui (Peking University), Lei Shi (Beihang University), and Xiaoming Li (Peking University)

Mining Order-Preserving Submatrices Under Data Uncertainty: A Possible-World Approach 1154
Ji Cheng (The Hong Kong University of Science and Technology), Da Yan (The University of Alabama at Birmingham), Xiaotian Hao (The Hong Kong University of Science and Technology), and Wilfred Ng (The Hong Kong University of Science and Technology)

Multi-Dimensional Genomic Data Management for Region-Preserving Operations 1166
Olha Horlova (Politecnico di Milano), Abdulrahman Kaitoua (DFKI & TU-Berlin), Volker Markl (DFKI & TU-Berlin), and Stefano Ceri (Politecnico di Milano)

Improved Algorithms for Maximal Clique Search in Uncertain Networks 1178
Rong-Hua Li (Beijing Institute of Technology; National Engineering Laboratory for Big Data System Computing Technology), Qiangqiang Dai (Beijing Institute of Technology), Guoren Wang (Beijing Institute of Technology), Zhong Ming (Shenzhen University; National Engineering Laboratory for Big Data System Computing Technology), Lu Qin (University of Technology, Sydney), and Jeffrey Xu Yu (The Chinese University of Hong Kong)

Lazo: A Cardinality-Based Method for Coupled Estimation of Jaccard Similarity and Containment 1190
Raul Castro Fernandez (MIT), Jisoo Min (MIT), Demitri Nava (MIT), and Samuel Madden (MIT)

Research (21) - Indexing

TARDIS: Distributed Indexing Framework for Big Time Series Data 1202
Liang Zhang (Worcester Polytechnic Institute), Noura Alghamdi (Worcester Polytechnic Institute), Mohamed Y. Eltabakh (Worcester Polytechnic Institute), and Elke A. Rundensteiner (Worcester Polytechnic Institute)

Mostly Order Preserving Dictionaries 1214
Chunwei Liu (University of Chicago), McKade Umbenhower (University of Wyoming), Hao Jiang (University of Chicago), Pranav Subramaniam (University of Chicago), Jihong Ma (Alibaba Group), and Aaron J. Elmore (University of Chicago)

Multi-copy Cuckoo Hashing 1226
Dagang Li (Peking University Shenzhen Graduate School), Rong Du (Peking University), Ziheng Liu (Peking University), Tong Yang (Peking University), and Bin Cui (Peking University)

Efficient Scalable Multi-attribute Index Selection Using Recursive Strategies 1238
Rainer Schlosser (Hasso Plattner Institute), Jan Kossmann (Hasso Plattner Institute), and Martin Boissier (Hasso Plattner Institute)

To Index or Not to Index: Optimizing Exact Maximum Inner Product Search 1250
Firas Abuzaid (Stanford DAWN Project), Geet Sethi (Stanford DAWN Project), Peter Bailis (Stanford DAWN Project), and Matei Zaharia (Stanford DAWN Project)

Research (22) - Road Networks

Distributed In-memory Trajectory Similarity Search and Join on Road Network 1262
Haitao Yuan (Tsinghua University) and Guoliang Li (Tsinghua University)

Stochastic Weight Completion for Road Networks Using Graph Convolutional Networks 1274
Jilin Hu (Aalborg University), Chenjuan Guo (Aalborg University), Bin Yang (Aalborg University), and Christian S. Jensen (Aalborg University)

Identifying the Most Interactive Object in Spatial Databases 1286
Daichi Amagata (Osaka University) and Takahiro Hara (Osaka University)

An Efficient Framework for Correctness-Aware kNN Queries on Road Networks 1298
Dan He (University of Queensland), Sibo Wang (the Chinese University of Hong Kong), Xiaofang Zhou (University of Queensland), and Reynold Cheng (the University of Hong Kong)

MPR — A Partitioning-Replication Framework for Multi-Processing kNN Search on Road Networks 1310
Siqiang Luo (The University of Hong Kong), Ben Kao (The University of Hong Kong), Xiaowei Wu (University of Vienna), and Reynold Cheng (The University of Hong Kong)

Research (23) - Learning, Temporal and Spatial Data

AppUsage2Vec: Modeling Smartphone App Usage for Prediction 1322
Sha Zhao (Zhejiang University), Zhiling Luo (Zhejiang University), Ziwen Jiang (Zhejiang University), Haiyan Wang (Zhejiang University), Feng Xu (Zhejiang University), Shijian Li (Zhejiang University), Jianwei Yin (Zhejiang University), and Gang Pan (Zhejiang University)

iFair: Learning Individually Fair Data Representations for Algorithmic Decision Making 1334
Preethi Lahoti (Max Planck Institute for Informatics), Krishna P. Gummadi (Max Planck Institute for Software Systems), and Gerhard Weikum (Max Planck Institute for Informatics)

DBSCAN-MS: Distributed Density-Based Clustering in Metric Spaces 1346
Keyu Yang (Zhejiang University), Yunjun Gao (Zhejiang University), Rui Ma (Zhejiang University), Lu Chen (Aalborg University), Sai Wu (Zhejiang University), and Gang Chen (Zhejiang University)

Computing Trajectory Similarity in Linear Time: A Generic Seed-Guided Neural Metric Learning Approach 1358
Di Yao (Institute of Computing Technology, Chinese Academy of Sciences), Gao Cong (Nanyang Technological University, Singapore), Chao Zhang (University of Illinois at Urbana-Champaign), and Jingping Bi (Institute of Computing Technology, Chinese Academy of Sciences)

Bursty Event Detection Throughout Histories 1370
Debjyoti Paul (University of Utah), Yanqing Peng (University of Utah), and Feifei Li (University of Utah)

Research (24) - Knowledge Discovery

RUM: Network Representation Learning Using Motifs 1382
Yanlei Yu (Renmin University of China), Zhiwu Lu (Renmin University of China), Jiajun Liu (Renmin University of China), Guoping Zhao (Renmin University of China), and Ji-rong Wen (Renmin University of China)

Finding Significant Items in Data Streams 1394
Tong Yang (Peking University, China), Haowei Zhang (Peking University, China), Dongsheng Yang (Peking University, China), Yucheng Huang (Peking University, China), and Xiaoming Li (Peking University, China)

Knowledge-Aware Deep Dual Networks for Text-Based Mortality Prediction 1406
Ning Liu (Tsinghua University), Pan Lu (Tsinghua University), Wei Zhang (East China Normal University), and Jianyong Wang (Tsinghua University)

Robust High Dimensional Stream Classification with Novel Class Detection 1418
Zhuoyi Wang (University of Texas at Dallas), Zelun Kong (University of Texas at Dallas), Swarup Changra (University of Texas at Dallas), Hemeng Tao (University of Texas at Dallas), and Latifur Khan (University of Texas at Dallas)

Towards the Completion of a Domain-Specific Knowledge Base with Emerging Query Terms 1430
Sihang Jiang (Fudan University), Jiaqing Liang (Fudan University), Yanghua Xiao (Fudan University), Haihong Tang (Alibaba Group), Haikuan Huang (Alibaba Group), and Jun Tan (Alibaba Group)

Research (25) - Crowdsourcing and Learning

Cooperation-Aware Task Assignment in Spatial Crowdsourcing 1442
Peng Cheng (HKUST), Lei Chen (HKUST), and Jieping Ye (AI Labs, DiDi Chuxing)

Minimizing Maximum Delay of Task Assignment in Spatial Crowdsourcing 1454
Zhao Chen (HKUST), Peng Cheng (HKUST), Yuxiang Zeng (HKUST), and Lei Chen (HKUST)

Physical Representation-Based Predicate Optimization for a Visual Analytics Database 1466
Michael R. Anderson (University of Michigan), Michael Cafarella (University of Michigan), German Ros (Intel Labs), and Thomas F. Wenisch (University of Michigan)

Adaptive Dynamic Bipartite Graph Matching: A Reinforcement Learning Approach 1478
Yansheng Wang (Beihang University), Yongxin Tong (Beihang University), Cheng Long (Nanyang Technological University), Pan Xu (University of Maryland, College Park), Ke Xu (Beihang University), and Weifeng Lv (Beihang University)

Research (26) - Data Mining

Scalable Frequent Sequence Mining with Flexible Subsequence Constraints 1490
Alexander Renz-Wieland (Technische Universität Berlin), Matthias Bertsch (Universität Mannheim), and Rainer Gemulla (Universität Mannheim)

Adaptive Influence Blocking: Minimizing the Negative Spread by Observation-Based Policies 1502
Qihao Shi (Zhejiang University), Can Wang (Zhejiang University), Deshi Ye (Zhejiang University), Jiawei Chen (Zhejiang University), Yan Feng (Zhejiang University), and Chun Chen (Zhejiang University)

Fraction-Score: A New Support Measure for Co-location Pattern Mining 1514
Harry Kai-Ho Chan (The Hong Kong University of Science and Technology), Cheng Long (Nanyang Technological University), Da Yan (The University of Alabama at Birmingham), and Raymond Chi-Wing Wong (The Hong Kong University of Science and Technology)

Discovery and Ranking of Functional Dependencies 1526
Ziheng Wei (The University of Auckland) and Sebastian Link (The University of Auckland)

Adaptive Deep Reuse: Accelerating CNN Training on the Fly 1538
Lin Ning (North Carolina State University), Hui Guan (North Carolina State University), and Xipeng Shen (North Carolina State University)

Research Short Papers

Slice Finder: Automated Data Slicing for Model Validation 1550
Yeounoh Chung (Brown University), Tim Kraska (MIT), Neoklis Polyzotis (Google Research), Ki Hyun Tae (KAIST), and Steven Euijong Whang (KAIST)

Neural Multi-task Recommendation from Multi-behavior Data 1554
Chen Gao (Tsinghua University), Xiangnan He (University of Science and Technology of China), Dahua Gan (Tsinghua University), Xiangning Chen (Tsinghua University), Fuli Feng (National University of Singapore), Yong Li (Tsinghua University), Tat-Seng Chua (National University of Singapore), and Depeng Jin (Tsinghua University)

AUC-MF: Point of Interest Recommendation with AUC Maximization 1558
Peng Han (King Abdullah University of Science and Technology), Shuo Shang (Inception Institute of Artificial Intelligence), Aixin Sun (Nanyang Technological University), Peilin Zhao (Tencent AI Lab), Kai Zheng (University of Electronic Science and Technology of China), and Panos Kalnis (King Abdullah University of Science and Technology)

Efficient Batch One-Hop Personalized PageRanks 1562
Siqiang Luo (The University of Hong Kong), Xiaokui Xiao (National University of Singapore), Wenqing Lin (Tencent), and Ben Kao (The University of Hong Kong)

BB-Tree: A Main-Memory Index Structure for Multidimensional Range Queries 1566
Stefan Sprenger (Humboldt-University Berlin), Patrick Schäfer (Humboldt-University Berlin), and Ulf Leser (Humboldt-University Berlin)

Explaining Queries Over Web Tables to Non-experts 1570
Jonathan Berant (Tel Aviv University), Daniel Deutch (Tel Aviv University), Amir Globerson (Tel Aviv University), Tova Milo (Tel Aviv University), and Tomer Wolfson (Tel Aviv University)

Nonlinear Models Over Normalized Data 1574
Zhaoyue Cheng (University of Toronto) and Nick Koudas (University of Toronto)

Continuous Search on Dynamic Spatial Keyword Objects 1578
Yuyang Dong (University of Tsukuba), Hanxiong Chen (University of Tsukuba), and Hiroyuki Kitagawa (University of Tsukuba)

Top-K Frequent Term Queries on Streaming Data 1582
Sara Farazi (University of Alberta) and Davood Rafiei (University of Alberta)

Parallel and Distributed Processing of Reverse Top-k Queries 1586
Panagiotis Nikitopoulos (University of Piraeus), Georgios A. Sfyris (University of Piraeus), Akrivi Vlachou (University of Piraeus), Christos Doulkeridis (University of Piraeus), and Orestis Telelis (University of Piraeus)

Distributed Discovery of Functional Dependencies 1590
Hemant Saxena (University of Waterloo), Lukasz Golab (University of Waterloo), and Ihab F. Ilyas (University of Waterloo)

AID: An Adaptive Image Data Index for Interactive Multilevel Visualization 1594
Saheli Ghosh (University of California, Riverside), Ahmed Eldawy (University of California, Riverside), and Shipra Jais (University of California, Riverside)

Collecting Preference Rankings Under Local Differential Privacy 1598
Jianyu Yang (Beijing University of Posts and Telecommunications), Xiang Cheng (Beijing University of Posts and Telecommunications), Sen Su (Beijing University of Posts and Telecommunications), Rui Chen (Samsung Research America, Mountain View, USA), Qiyu Ren (Beijing University of Posts and Telecommunications), and Yuhan Liu (Beijing University of Posts and Telecommunications)

Muses: Distributed Data Migration System for Polystores 1602
Abdulrahman Kaitoua (DFKI), Tilmann Rabl (TU-Berlin), Asterios Katsifodimos (Delft University of Technology), and Volker Markl (TU-Berlin)

PriSTE: From Location Privacy to Spatiotemporal Event Privacy 1606
Yang Cao (Kyoto University), Yonghui Xiao (Google Inc.), Li Xiong (Emory University), and Liquan Bai (Emory University)

Continuous Range Queries Over Multi-attribute Trajectories 1610
jianqiu xu (Nanjing University of Aeronautics and Astronautics, China), Zhifeng Bao (RMIT University, Australia), and Hua Lu (Aalborg University, Denmark)

Insecurity and Hardness of Nearest Neighbor Queries Over Encrypted Data 1614
Rui Li (Dongguan University of Technology), Alex X. Liu (Dongguan University of Technology), Ying Liu (Hunan University), Huanle Xu (Dongguan University of Technology), and Huaqiang Yuan (Dongguan University of Technology)

Modeling Multidimensional User Preferences for Collaborative Filtering 1618
Farhan Khawar (The Hong Kong University of Science and Technology) and Nevin L. Zhang (The Hong Kong University of Science and Technology)

A Queueing-Theoretic Framework for Vehicle Dispatching in Dynamic Car-Hailing 1622
Peng Cheng (HKUST), Chao Feng (AI Labs, DiDi Chuxing), Lei Chen (HKUST), and Zheng Wang (AI Labs, DiDi Chuxing)

Maximizing the Utility in Location-Based Mobile Advertising 1626
Peng Cheng (HKUST), Xiang Lian (Kent State University), Lei Chen (HKUST), and Siyuan Liu (Pennsylvania State University)

Automated Grading of SQL Queries 1630
Bikash Chandra (IIT Bombay), Ananyo Banerjee (IIT Bombay), Udbhas Hazra (IIT Bombay), Mathew Joseph (IIT Bombay), and S. Sudarshan (IIT Bombay)

Parameter Discovery in Unsupervised Clustering 1634
Valentin Clement (Imeprial College London) and Thomas Heinis (Imeprial College London)

Interaction-Aware Arrangement for Event-Based Social Networks 1638
Feifei Kou (Beijing University of Posts and Telecommunications), Zimu Zhou (ETH Zurich), Hao Cheng (Beihang University), Junping Du (Beijing University of Posts and Telecommunications), Yexuan Shi (Beihang University), and Pan Xu (University of Maryland)

Optimizing Cross-Platform Data Movement 1642
Sebastian Kruse (HPI), Zoi Kaoudi (QCRI), Jorge-Arnulfo Quiane-Ruiz (QCRI), Sanjay Chawla (QCRI), Felix Naumann (QCRI), and Bertty Contreras-Rojas (QCRI)

Highly Efficient Pattern Mining Based on Transaction Decomposition 1646
Youcef Djenouri (NTNU, Trondheim, Norway), Jerry Chun-Wei Lin (HVL, Bergen, Norway), Kjetil Nørvåg (NTNU, Trondheim, Norway), and Heri Ramampiaro (HVL, Bergen, Norway)

Procrastination-Aware Scheduling: A Bipartite Graph Perspective 1650
Libin Wang (Beihang University), Yongxin Tong (Beihang University), Chunming Hu (Beihang University), Lei Chen (The Hong Kong University of Science and Technology), and Yiming Li (Beihang University)

Hankel Matrix Factorization for Tagged Time Series to Recover Missing Values During Blackouts 1654
Simeng Wu (State Key Laboratory for Novel Software Technology, Nanjing University), Liang Wang (State Key Laboratory for Novel Software Technology, Nanjing University), Tianheng Wu (State Key Laboratory for Novel Software Technology, Nanjing University), Xianping Tao (State Key Laboratory for Novel Software Technology, Nanjing University), and Jian Lu (State Key Laboratory for Novel Software Technology, Nanjing University)

PerRD: A System for Personalized Route Description 1658
Han Su (Zhejiang Gongshang University), Guanglin Cong (University of Electronic Science and Technology of China), Wei Chen (University of Electronic Science and Technology of China), Qinyuan Su (University of Electronic Science and Technology of China), Bolong Zheng (Huazhong University of Science and Technology), and Kai Zheng (University of Electronic Science and Technology of China)

Scalable Metric Similarity Join Using MapReduce 1662
Jiacheng Wu (Tsinghua University), Yong Zhang (Tsinghua University), Jin Wang (University of California, Los Angeles), Chunbin Lin (Amazon AWS), Yingjia Fu (University of California, San Diego), and Chunxiao Xing (Tsinghua University)

An Indexing Framework for Efficient Visual Exploratory Subgraph Search in Graph Databases 1666
Chaohui Wang (Nanyang Technological University), Miao Xie (Alibaba Group), Sourav S. Bhowmick (Nanyang Technological University), Byron Choi (Hong Kong Baptist University), Xiaokui Xiao (National University of Singapore), and Shuigeng Zhou (Fudan University)

I-LSH: I/O Efficient c-Approximate Nearest Neighbor Search in High-Dimensional Space 1670
Wanqi Liu (University of Technology Sydney), Hanchen Wang (University of Technology Sydney), Ying Zhang (University of Technology Sydney), Wei Wang (university of new south wales), and Lu Qin (University of Technology Sydney)

T-Sample: A Dual Reservoir-Based Sampling Method for Characterizing Large Graph Streams 1674
Lingling Zhang (Huazhong University of Science and Technology), Hong Jiang (Department of Computer Science and Engineering, University of Texas at Arlington, USA), Fang Wang (Huazhong University of Science and Technology), Dan Feng (Huazhong University of Science and Technology), and Yanwen Xie (Huazhong University of Science and Technology)

Real Time Principal Component Analysis 1678
Ranak Roy Chowdhury (Bangladesh University of Engineering and Technology (BUET)), Muhammad Abdullah Adnan (Bangladesh University of Engineering and Technology (BUET)), and Rajesh K. Gupta (University of California San Diego)

A Fast Sketch Method for Mining User Similarities Over Fully Dynamic Graph Streams 1682
Peng Jia (Xi'an Jiaotong University), Pinghui Wang (Xi'an Jiaotong University), Jing Tao (Xi'an Jiaotong University), and Xiaohong Guan (Xi'an Jiaotong University)

A Collaborative Framework for Similarity Enforcement in Synthetic Scaling of Relational Datasets 1686
Jiangwei W. Zhang (National University of Singapore) and Yong Chiang Tay (National University of Singapore)

Meta Diagram Based Active Social Networks Alignment 1690
Yuxiang Ren (IFM lab, Florida State University), Charu C. Aggarwal (IBM Research AI), and Jiawei Zhang (IFM lab, Florida State University)

Entity Integrity, Referential Integrity, and Query Optimization with Embedded Uniqueness Constraints..... 1694
Ziheng Wei (The University of Auckland), Uwe Leck (The University of Flensburg), and Sebastian Link (The University of Auckland)

Efficient Pattern Mining Based Cryptanalysis for Privacy-Preserving Record Linkage 1698
Anushka Vidanage (The Australian National University, Canberra, Australia), Thilina Ranbaduge (The Australian National University, Canberra, Australia), Peter Christen (The Australian National University, Canberra, Australia), and Rainer Schnell (University Duisburg-Essen, Germany)

ECOQUG: An Effective Ensemble Community Scoring Function 1702
Chunnan Wang (Harbin Institute of Technology), Hongzhi Wang (Harbin Institute of Technology), Chang Zhou (Harbin Institute of Technology), Jianzhong Li (Harbin Institute of Technology), and Hong Gao (Harbin Institute of Technology)

CN-Probase: A Data-Driven Approach for Large-Scale Chinese Taxonomy Construction 1706
Jindong Chen (Fudan University), Ao Wang (Fudan University), Jiangjie Chen (Fudan University), Yanghua Xiao (Fudan University), Zhendong Chu (Fudan University), Jingping Liu (Fudan University), Jiaqing Liang (Fudan University), and Wei Wang (Fudan University)

Understanding Data Science Lifecycle Provenance via Graph Segmentation and Summarization 1710
Hui Miao (University of Maryland) and Amol Deshpande (University of Maryland)

Efficient Partitioning and Query Processing of Spatio-Temporal Graphs with Trillion Edges 1714
Mengsu Ding (SKL Computer Architecture, ICT, CAS, University of Chinese Academy of Sciences) and Shimin Chen (SKL Computer Architecture, ICT, CAS, University of Chinese Academy of Sciences)

Top-K Entity Resolution with Adaptive Locality-Sensitive Hashing 1718
Vasilis Verroios (Stanford University) and Hector Garcia-Molina

Finding Average Regret Ratio Minimizing Set in Database 1722
Sepanta Zeighami (Hong Kong University of Science and Technology) and Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)

HyMJ: A Hybrid Structure-Aware Approach to Distributed Multi-way Join Query 1726
Guanghui Zhu (Nanjing University), Xiaoqi Wu (Nanjing University), Liangliang Yin (Nanjing University), Haogang Wang (Nanjing University), Rong Gu (Nanjing University), Chunfeng Yuan (Nanjing University), and Yihua Huang (Nanjing University)

Accelerate MaxBRkNN Search by kNN Estimation 1730
Xuefeng Chen (University of New South Wales, Australia), Xin Cao (University of New South Wales, Australia), Zhiqiang Xu (Baidu, China), Ying Zhang (University of Technology Sydney, Australia), Shuo Shang (Inception Institute of Artificial Intelligence, United Arab Emirates), and Wenjie Zhang (University of New South Wales, Australia)

Efficient Bottom-Up Discovery of Multi-scale Time Series Correlations Using Mutual Information 1734
Nguyen Ho (Aalborg University), Torben Bach Pedersen (Aalborg University), Mai Vu (Tufts University), Van Long Ho (Aalborg University), and Christophe A.N. Biscio (Aalborg University)

Fingerprinting Big Data: The Case of KNN Graph Construction 1738
Rachid Guerraoui (EPFL), Anne-Marie Kermarrec (EPFL, Mediego), Olivier Ruas (Inria, Univ Rennes, CNRS, IRISA), and François Taïani (Univ Rennes, Inria, CNRS, IRISA)

Outer and Anti Joins in Temporal-Probabilistic Databases 1742
Katerina Papaioannou (University of Zurich), Martin Theobald (University of Luxembourg), and Michael Boehlen (University of Zurich)

Workload-Driven Fragment Allocation for Partially Replicated Databases Using Linear Programming 1746
Stefan Halfpap (Hasso Plattner Institute, University of Potsdam, Germany) and Rainer Schlosser (Hasso Plattner Institute, University of Potsdam, Germany)

OSMAC: Optimizing Subgraph Matching Algorithms with Community Structure 1750
Yunkai Lou (Tsinghua University) and Chaokun Wang (Tsinghua University)

Workload-Aware Subgraph Query Caching and Processing in Large Graphs 1754
Yongjiang Liang (Florida State University) and Peixiang Zhao (Florida State University)

How I Learned to Stop Worrying and Love Re-optimization 1758
Matthew Perron (MIT CSAIL), Zeyuan Shang (MIT CSAIL), Tim Kraska (MIT CSAIL), and Michael Stonebraker (MIT CSAIL)

CRA: Enabling Data-Intensive Applications in Containerized Environments 1762
Ibrahim Sabek (University of Minnesota, USA), Badrish Chandramouli (Microsoft Research, USA), and Umar Farooq Minhas (Microsoft Research, USA)

Industry Track

Industry (1) - Data Systems Optimizations

Scalable Similarity Joins of Tokenized Strings 1766
Ahmed Metwally (LinkedIn) and Chun-Heng Huang (Google)

MLlib*: Fast Training of GLMs Using Spark MLlib 1778
Zhipeng Zhang (Peking University), Jiawei Jiang (Tencent Inc.), Wentao Wu (Microsoft Research, Redmond, USA), Ce Zhang (ETH Zurich, Switzerland), Lele Yu (Tencent Inc.), and Bin Cui (Peking University)

DirectLoad: A Fast Web-Scale Index System Across Large Regional Centers 1790
An Qin (Baidu Inc.), Mengbai Xiao (The Ohio State University), Jin Ma (Baidu Inc.), Dai Tan (Baidu Inc.), Rubao Lee (United Parallel Computing Corporation), and Xiaodong Zhang (The Ohio State University)

Presto: SQL on Everything 1802
Raghav Sethi (Facebook, Inc.), Martin Traverso (Facebook, Inc.), Dain Sundstrom (Facebook, Inc.), David Phillips (Facebook, Inc.), Wenlei Xie (Facebook, Inc.), Yutian Sun (Facebook, Inc.), Nezih Yegitbasi (Facebook, Inc.), Haozhun Jin (Facebook, Inc.), Eric Hwang (Facebook, Inc.), Nileema Shingte (Facebook, Inc.), and Christopher Berner (Facebook, Inc.)

Improving RDF Query Performance Using In-memory Virtual Columns in Oracle Database 1814
Eugene Inseok Chong (Oracle), Matthew Perry (Oracle), and Souripriya Das (Oracle)

Industry (2) - Data Analytics

SEBDB: Semantics Empowered BlockChain DataBase 1820
Yanchao Zhu (East China Normal University), Zhao Zhang (East China Normal University), Cheqing Jin (East China Normal University), Aoying Zhou (East China Normal University), and Ying Yan (Ant Financial)

Large Scale Traffic Signal Network Optimization - A Paradigm Shift Driven by Big Data 1832
Liang Yu (Alibaba Cloud), Jinqiang Yu (Alibaba Cloud), Maolei Zhang (Alibaba Cloud), Xin Zhang (Alibaba Cloud), Yuehu Liu (Alibaba Cloud), Hui Zhang (Alibaba Cloud), and Wanli Min (Alibaba Cloud)

Domain-Independent Automated Processing of Free-Form Text Data in Telecom 1841
Rajarshi Bhowmik (Rutgers University) and Ahmet Akyamac (Bell Labs, Nokia)

DRIVEN: a Framework for Efficient Data Retrieval and Clustering in Vehicular Networks 1850
Bastian Havers (Chalmers University), Romaric Duvignau (Chalmers University), Hannaneh Najdataei (Chalmers University), Vincenzo Gulisano (Chalmers University), Ashok Chaitanya Koppisetty (Volvo Cars), and Marina Papatriantafilou (Chalmers University)

Industry (3) - Driving Business Applications

Accurate Product Attribute Extraction on the Field 1862
Martin Rezk (Rakuten Inc), Laura Alonso Alemany (Universidad Nacional de Cordoba), Lasguido Nio (Rakuten Inc), and Ted Zhang (Rakuten Inc)

CATS: Cross-Platform E-Commerce Fraud Detection 1874
Haiqin Weng (Zhejiang University), Shouling Ji (Zhejiang University), Fuzheng Duan (Zhejiang University), Zhao Li (Alibaba Group), Jianhai Chen (Zhejiang University), Qinming He (Zhejiang University), and Ting Wang (Lehigh University)

Caladrius: A Performance Modelling Service for Distributed Stream Processing Systems 1886
Faria Kalim (University of Illinois at Urbana-Champaign), Thomas Cooper (Newcastle University, UK), Huijun Wu (Twitter, Inc.), Yao Li (Twitter, Inc.), Ning Wang (Twitter, Inc.), Neng Lu (Twitter, Inc.), Maosong Fu (Twitter, Inc.), Xiaoyao Qian (Twitter, Inc.), Hao Luo (Twitter, Inc.), Da Cheng (Twitter, Inc.), Yaliang Wang (Twitter, Inc.), Fred Dai (Twitter, Inc.), Mainak Ghosh (Twitter, Inc.), and Beinan Wang (Twitter, Inc.)

FAIR: Fraud Aware Impression Regulation System in Large-Scale Real-Time E-Commerce Search Platform . 1898
Zhao Li (Alibaba Group), Junshuai Song (Peking University), Shichang Hu (Alibaba Group), Shasha Ruan (Alibaba Group), Long Zhang (Alibaba Group), Zehong Hu (Alibaba Group), and Jun Gao (Peking University)

Micro-Browsing Models for Search Snippets 1904
Muhammad Asiful Islam (Google Inc), Ramakrishnan Srikant (Google Inc), and Sugato Basu (Google Inc)

Industry (4) - Learning from Data

Interpretable Multi-task Learning for Product Quality Prediction with Attention Mechanism 1910
Cheng-Han Yeh (National Chiao Tung University, Hsinchu, Taiwan), Yao-Chung Fan (National Chung Hsing University, Taichung, Taiwan), and Wen-Chih Peng (National Chiao Tung University, Hsinchu, Taiwan)

Learning Effective Embeddings From Crowdsourced Labels: An Educational Case Study 1922
Guowei Xu (TAL AI Lab), Wenbiao Ding (TAL AI Lab), Jiliang Tang (Michigan State University), Songfan Yang (TAL AI Lab), Gale Yan Huang (TAL AI Lab), and Zitao Liu (TAL AI Lab)

A Prescription Trend Analysis using Medical Insurance Claim Big Data 1928
Kazutoshi Umemoto (The University of Tokyo), Kazuo Goda (The University of Tokyo), Naohiro Mitsutake (Institute for Health Economics and Policy), and Masaru Kitsuregawa (The University of Tokyo)

Differential Data Quality Verification on Partitioned Data 1940
Sebastian Schelter (Amazon Research), Stefan Grafberger (Amazon Research), Philipp Schmidt (Amazon Research), Tammo Rukat (Amazon Research), Mario Kiessling (Amazon Research), Andrey Taptunov (Amazon Research), Felix Biessmann (Amazon Research), and Dustin Lange (Amazon Research)

Logan: A Distributed Online Log Parser 1946
Amey Agrawal (Qubole India), Rohit Karlupia (Qubole India), and Rajat Gupta (Qubole India)

Demo Track

Demo Session 1

WebPut: A Web-Aided Data Imputation System for the General Type of Missing String Attribute Values 1952
Shuangli Shan (Soochow University, China; Neusoft Corporation), Zhixu Li (Soochow University, China), Yang Li (Soochow University, China), Qiang Yang (King Abdullah University of Science and Technology, Saudi Arabia), Jia Zhu (South China Normal University), Mohamed Sharaf (The University of Queensland), and Xiaofang Zhou (The University of Queensland)

FGreat: Focused Graph Query Autocompletion 1956
Nathan Ng (Hong Kong Baptist University), Peipei Yi (Lenovo Machine Intelligence Center), Zhiwei Zhang (Hong Kong Baptist University), Byron Choi (Hong Kong Baptist University), Sourav S. Bhowmick (Nanyang Technological University), and Jianliang Xu (Hong Kong Baptist University)

Aucher: Multi-modal Queries on Live Audio Streams in Real-Time 1960
Zeyi Wen (National University of Singapore), Mingyu Liang (Shanghai Jiao Tong University), Bingsheng He (National University of Singapore), Zexin Xia (Shanghai Jiao Tong University), and Bo Li (Hong Kong University of Science and Technology)

SAC: A System for Big Data Lineage Tracking 1964
Mingjie Tang (Hortonworks), Saisai Shao (Hortonworks), Weiqing Yang (Hortonworks), Yanbo Liang (Hortonworks), Yongyang Yu (Facebook), Bikas Saha (Hortonworks), and Dongjoon Hyun (Hortonworks)

A Gossip-Based System for Fast Approximate Score Computation in Multinomial Bayesian Networks 1968
Arun Zachariah (University of Missouri-Kansas City), Praveen Rao (University of Missouri-Kansas City), Anas Katib (University of Missouri-Kansas City), Monica Senapati (University of Missouri-Kansas City), and Kobus Barnard (University of Arizona)

Faster, Higher, Stronger: Redesigning Spreadsheets for Scale 1972
Mangesh Bendre (University of Illinois at Urbana-Champaign), Tana Wattanawaroon (University of Illinois at Urbana-Champaign), Sajjadur Rahman (University of Illinois at Urbana-Champaign), Kelly Mack (University of Illinois at Urbana-Champaign), Yuyang Liu (University of Illinois at Urbana-Champaign), Shichu Zhu (University of Illinois at Urbana-Champaign), Yu Lu (University of Illinois at Urbana-Champaign), Ping-Jing Yang (University of Illinois at Urbana-Champaign), Xinyan Zhou (University of Illinois at Urbana-Champaign), Kevin Chen-Chuan Chang (University of Illinois at Urbana-Champaign), Karrie Karahalios (University of Illinois at Urbana-Champaign), and Aditya Parameswaran (University of Illinois at Urbana-Champaign)

RecovDB: Accurate and Efficient Missing Blocks Recovery for Large Time Series 1976
Ines Arous (University of Fribourg), Mourad Khayati (University of Fribourg), Philippe Cudré-Mauroux (University of Fribourg), Ying Zhang (MonetDB Solutions), Martin Kersten (MonetDB Solutions), and Svetlin Stalinlov (MonetDB Solutions)

AI Pro: Data Processing Framework for AI Models 1980
Richie Frost (University of Utah), Debjyoti Paul (University of Utah), and Feifei Li (University of Utah)

IVLG: Interactive Visualization of Large Graphs 1984
Maria Krommyda (NTUA), Verena Kantere (University of Ottawa), and Yannis Vassiliou (NTUA)

Just in Time: Personal Temporal Insights for Altering Model Decisions 1988
Naama Boer (Tel Aviv University), Daniel Deutch (Tel Aviv University), Nave Frost (Tel Aviv University), and Tova Milo (Tel Aviv University)

GeoSparkViz in Action: A Data System with Built-in Support for Geospatial Visualization 1992
Jia Yu (Arizona State University), Anique Tahir (Arizona State University), and Mohamed Sarwat (Arizona State University)

Hybrid.Poly: A Consolidated Interactive Analytical Polystore System 1996
Maksim Podkorytov (Florida State University) and Michael Gubanov (Florida State University)

Demo Session 2

EXPLAINER: Entity Resolution Explanations 2000
Amr Ebaid (Purdue University), Saravanan Thirumuruganathan (Qatar Computing Research Institute, HBKU), Walid G. Aref (Purdue University), Ahmed Elmagarmid (Qatar Computing Research Institute, HBKU), and Mourad Ouzzani (Qatar Computing Research Institute, HBKU)

CEP-Wizard: Automatic Deployment of Distributed Complex Event Processing 2004
Yooju Shin (KAIST), Susik Yoon (KAIST), Patara Trirat (KAIST), and Jae-Gil Lee (KAIST)

A Comparison of Allocation Algorithms for Partially Replicated Databases 2008
Stefan Halfpap (Hasso Plattner Institute, University of Potsdam, Germany) and Rainer Schlosser (Hasso Plattner Institute, University of Potsdam, Germany)

PePPer: Fine-Grained Personal Access Control via Peer Probing 2012
Yael Amsterdamer (Bar-Ilan University) and Osnat Drien (Bar-Ilan University)

COBRA: Compression Via Abstraction of Provenance for Hypothetical Reasoning 2016
Daniel Deutch (Tel Aviv University), Yuval Moskovitch (Tel Aviv University), and Noam Rinetzky (Tel Aviv University)

CogLearn: A Cognitive Graph-Oriented Online Learning System 2020
Yang Pian (Beijing Normal University), Yu Lu (Beijing Normal University), Penghe Chen (Beijing Normal University), and Qinglong Duan (Beijing Normal University)

GRIT: Consistent Distributed Transactions Across Polyglot Microservices with Multiple Databases 2024
Guogen Zhang (eBay Inc.), Kun Ren (eBay Inc.), Jung-Sang Ahn (eBay Inc.), and Sami Ben-Romdhane (eBay Inc.)

vABS: Towards Verifiable Attribute-Based Search Over Shared Cloud Data 2028
Yang Ji (Hong Kong Baptist University), Cheng Xu (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), and Haibo Hu (Hong Kong Polytechnic University)

An Environment-Aware Market Strategy for Data Allocation and Dynamic Migration in Cloud Database 2032
Tengjiao Wang (Peking University), Binyang Li (University of International Relations), Wei Chen (Peking University), Yuxiao Zhang (Peking University), Ying Han (Peking University), Jinzhong Niu (City University of New York), and Kam-fai Wong (The Chinese University of Hong Kong)

Vaite: A Visualization-Assisted Interactive Big Urban Trajectory Data Exploration System 2036
Chuang Yang (Southern University of Science and Technology; PCL Research Center of Networks and Communications, Peng Cheng Laboratory), Yilan Zhang (Sichuan University), Bo Tang (Southern University of Science and Technology; PCL Research Center of Networks and Communications, Peng Cheng Laboratory), and Min Zhu (Sichuan University)

SciDetector: Scientific Event Discovery by Tracking Variable Source Data Streaming 2040
Zhiqiang Duan (Renmin University of China), Chen Yang (Renmin University of China), Xiaofeng Meng (Renmin University of China), Yongjie Du (Renmin University of China), Jiaming Qiu (Tsinghua University), Xiaobin Ma (Tsinghua University), Zhihui Du (Tsinghua University), Xukang Zhang (Taiyuan University of Technology), Baoning Niu (Taiyuan University of Technology), and Chao Wu (National Astronomical Observatories, Chinese Academy of Sciences)

Demonstrating Spindra: A Geographic Knowledge Graph Management System 2044
Yuhan Sun (Arizona State University), Jia Yu (Arizona State University), and Mohamed Sarwat (Arizona State University)

Tutorials

Native Storage Techniques for Data Management 2048
Ilia Petrov (Reutlingen University), Andreas Koch (Technische Universität Darmstadt), Sergey Hardock (Technische Universität Darmstadt), Tobias Vincon (Reutlingen University), and Christian Riegger (Reutlingen University)

Crowdsourcing Database Systems: Overview and Challenges 2052
Chengliang Chai (Tsinghua University), Ju Fan (Renmin University), Guoliang Li (Tsinghua University), Jiannan Wang (Simon Fraser University), and Yudian Zheng (Twitter)

Telco Big Data Research and Open Problems 2056
Constantinos Costa (University of Pittsburgh) and Demetrios Zeinalipour-Yazti (University of Cyprus)

Geospatial Data Management in Apache Spark: A Tutorial 2060
Jia Yu (Arizona State University) and Mohamed Sarwat (Arizona State University)

Hierarchical Decomposition of Big Graphs 2064
Ying Zhang (University of Technology Sydney), Lu Qin (University of Technology Sydney), Fan Zhang (University of New South Wales), and Wenjie Zhang (University of New South Wales)

Cohesive Subgraph Computation Over Large Sparse Graphs 2068
Lijun Chang (The University of Sydney) and Lu Qin (University of Technology Sydney)

Robust Query Processing: Mission Possible 2072
Jayant R. Haritsa (Indian Institute of Science)

PhD Symposium

Automated Documentation of End-to-End Experiments in Data Science 2076
Sergey Redyuk (Technische Universität Berlin)

Explaining Results of Data-Driven Applications 2081
Nave Frost (Tel Aviv University)

Towards Explaining the Effects of Data Preprocessing on Machine Learning 2086
Carlos Vladimiro Gonzalez Zelaya (Newcastle University)

Don't Fear the REAPER: A Framework for Materializing and Reusing Deep-Learning Models 2091
Melanie B. Sigl (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Knowledge Representation for Emotion Intelligence 2096
Shuo Wang (Renmin University of China)

Disambiguation and Result Expansion in Keyword Search Over Relational Databases 2101
Niousha Hormozi (University of Athens)

Event Recommendation using Social Media 2106
Sreekanth Madisetty (Indian Institute of Technology Hyderabad)

TKDE Posters

Location Inference for Non-Geotagged Tweets in User Timelines [Extended Abstract] 2111
Pengfei Li (Zhejiang University, China), Hua Lu (Aalborg University), Nattiya Kanhabua (NTENT, Barcelona, Spain), Sha Zhao (Zhejiang University), and Gang Pan (Zhejiang University)

Efficient Parallel Skyline Query Processing for High-Dimensional Data 2113
Mingjie Tang (Ant Financial), Yongyang Yu (Purdue University), Walid G. Aref (Purdue University), Qutaibah M. Malluhi (Qatar University), and Mourad Ouzzani (Qatar Computing Research Institute)

On Generalizing Collective Spatial Keyword Queries (Extended Abstract) 2115
Harry Kai-Ho Chan (Hong Kong University of Science and Technology), Cheng Long (Nanyang Technological University), and Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)

A Novel Representation and Compression for Queries on Trajectories in Road Networks (Extended Abstract) 2117
Xiaochun Yang (Northeastern University, China), Bin Wang (Northeastern University, China), Kai Yang (Northeastern University, China), Chengfei Liu (Swinburne University of Techenology), and Baihua Zheng (Singapore Management University)

Efficient Multi-Class Probabilistic SVMs on GPUs 2119
Zeyi Wen (National University of Singapore), Jiashuai Shi (South China University of Technology), Bingsheng He (National University of Singapore), Jian Chen (South China University of Technology), and Yawen Chen (South China University of Technology)

C2Net: A Network-Efficient Approach to Collision Counting LSH Similarity Join(Extended Abstract) 2121
Hangyu Li (City University of Hong Kong), Sarana Nutanong (Vidyasirimedhi Institute of Science and Technology), Hong Xu (City University of Hong Kong), Chenyun Yu (City University of Hong Kong), and Foryu Ha (City University of Hong Kong)

LinkBlackHole*: Robust Overlapping Community Detection Using Link Embedding (Extended Abstract) 2123
Jungeun Kim (KAIST), Sungsu Lim (Chungnam National University), Jae-Gil Lee (KAIST), and Byung Suk Lee (University of Vermont)

Fusion OLAP: Fusing the Pros of MOLAP and ROLAP Together for In-memory OLAP (Extended Abstract) 2125
Yansong Zhang (Renmin University of China), Yu Zhang (National Satellite Meteorological Centre, Beijing, China), Shan Wang (Renmin University of China), and Jiaheng Lu (University of Helsinki)

In Search of Indoor Dense Regions: An Approach Using Indoor Positioning Data 2127
Huan Li (Zhejiang University), Hua Lu (Aalborg University), Lidan Shou (Zhejiang University), Gang Chen (Zhejiang University), and Ke Chen (Zhejiang University)

Optimizing Quality for Probabilistic Skyline Computation and Probabilistic Similarity Search (Extended Abstract) 2129
Xiaoye Miao (Zhejiang University), Yunjun Gao (Zhejiang University), Linlin Zhou (Zhejiang University), Wei Wang (University of New South Wales), and Qing Li (The Hong Kong Polytechnic University)

On Efficiently Answering Why-Not Range-Based Skyline Queries in Road Networks (Extended Abstract) 2131
Xiaoye Miao (Zhejiang University), Yunjun Gao (Zhejiang University), Su Guo (Zhejiang University), and Gang Chen (Zhejiang University)

SLADE: A Smart Large-Scale Task Decomposer in Crowdsourcing 2133
Yongxin Tong (Beihang University), Lei Chen (Hong Kong University of Science and Technology), Zimu Zhou (ETH Zurich), Hosagrahar Visvesvaraya Jagadish (University of Michigan), Lidan Shou (Zhejiang University), and Weifeng Lv (Beihang University)

XINA: Explainable Instance Alignment using Dominance Relationship (Extended Abstract) 2135
Jinyoung Yeo (SK T-Brain), Haeju Park (Yonsei University), Sanghoon Lee (NAVER Corporation), Eric Wonhee Lee (Emory University), and Seung-won Hwang (Yonsei University)

A Hardware-Accelerated Solution for Hierarchical Index-Based Merge-Join(Extended Abstract) 2137
Zimeng Zhou (City University of Hong Kong), Chenyun Yu (City University of Hong Kong), Sarana Nutanong (Vidyasirimedhi Institute of Science and Technology), Yufei Cui (City University of Hong Kong), Chenchen Fu (City University of Hong Kong), and Chun Jason Xue (City University of Hong Kong)

Finding Most Popular Indoor Semantic Locations Using Uncertain Mobility Data 2139
Huan Li (Zhejiang University), Hua Lu (Aalborg University), Lidan Shou (Zhejiang University), Gang Chen (Zhejiang University), and Ke Chen (Zhejiang University)

Uncertain Graph Sparsification (Extended Abstract) 2141
Panos Parchas (Amazon Web Services), Nikolaos Papailiou (NTUA), Dimitris Papadias (HKUST), and Francesco Bonchi (ISI Foundation & Eurecat)

Rule-Based Entity Resolution on Database with Hidden Temporal Information (Extended Abstract) 2143
Hongzhi Wang (Harbin Institute of Technology), Xiaoou Ding (Harbin Institute of Technology), Jianzhong Li (Harbin Institute of Technology), and Hong Gao (Harbin Institute of Technology)

BRIGHT - Drift-Aware Demand Predictions for Taxi Networks (Extended Abstract) 2145
Amal Saadallah (TU Dortmund), Luis Moreira-Matias (Kreditech Holding SSL, Hamburg, Germany), Ricardo Sousa (University of Porto, Portugal), Jihed Khiari (NEC Laboratories Europe, Germany), Erik Jenelius (KTH Royal Institute of Technology, Sweden), and Joao Gama (University of Porto, Portugal)

Order-Sensitive Imputation for Clustered Missing Values (Extended Abstract) 2147
Qian Ma (Dalian Maritime University, China), Yu Gu (Northeastern University, China), Wang-Chien Lee (The Pennsylvania State University), and Ge Yu (Northeastern University, China)

DeepDirect: Learning Directions of Social Ties with Edge-Based Network Embedding (Extended Abstract)... 2149
Chaokun Wang (Tsinghua University), Changping Wang (Tsinghua University), Zheng Wang (Tsinghua University), Xiaojun Ye (Tsinghua University), Jeffrey Xu Yu (Chinese University of Hong Kong), and Bin Wang (Tsinghua University)

A Utility-Optimized Framework for Personalized Private Histogram Estimation (Extended Abstract) 2151
Yiwen Nie (University of Science and Technology of China), Wei Yang (University of Science and Technology of China), Liusheng Huang (University of Science and Technology of China), Xike Xie (University of Science and Technology of China), Zhenhua Zhao (University of Science and Technology of China), and Shaowei Wang (University of Science and Technology of China)

Near-Accurate Multiset Reconciliation (Extended Abstract) 2153
Lailong Luo (National University of Defense Technology), Deke Guo (National University of Defence Technology), Xiang Zhao (National University of Defense Technology), Jie Wu (Temple University), Ori Rottenstreich (Orbs Research and Technion), and Xueshan Luo (National University of Defense Technology)

Answering Why-Not Group Spatial Keyword Queries (Extended Abstract) 2155
Bolong Zheng (Huazhong University of Science and Technology), Kai Zheng (University of Electronic Science and Technology of China), Christian S. Jensen (Aalborg University), Nguyen Quoc Viet Hung (Griffith University), Han Su (University of Electronic Science and Technology of China), Guohui Li (Huazhong University of Science and Technology), and Xiaofang Zhou (The University of Queensland)

Effective and Efficient Community Search Over Large Directed Graphs (Extended Abstract) 2157
Yixiang Fang (University of Hong Kong), Zhongran Wang (Harbin Institute of Technology), Reynold Cheng (The University of Hong Kong), Hongzhi Wang (Harbin Institute of Technology), and Jiafeng Hu (University of Hong Kong)

Exploring Communities in Large Profiled Graphs (Extended Abstract) 2159
Yankai Chen (The University of Hong Kong), Yixiang Fang (The University of Hong Kong), Reynold Cheng (The University of Hong Kong), Yun Li (Nanjing University), Xiaojun Chen (Shenzhen University), and Jie Zhang (Nanyang Technological University)

Index-Based Densest Clique Percolation Community Search in Networks (Extended Abstract) 2161
Long Yuan (Nanjing University of Science and Technology), Lu Qin (QCIS, University of Technology, Sydney), Wenjie Zhang (The University of New South Wales), Lijun Chang (The University of Sydney), and Jianye Yang (Central South University)

Author Index **2163**