

2021 IEEE 37th International Conference on Data Engineering (ICDE) **ICDE 2021**

Table of Contents

Message from the Chairs	xliv
Organizing Committee	xlvii
Program Committee	xliv

Research Papers

Data Integration and Cleaning 1

Profiles of Schema Evolution in Free Open Source Software Projects	1
<i>Panos Vassiliadis (Univ. Ioannina)</i>	
CleanML: A Study for Evaluating the Impact of Data Cleaning on ML Classification Tasks	13
<i>Peng Li (Georgia Institute of Technology), Xi Rao (ETH Zurich), Jennifer Blase (Georgia Institute of Technology), Yue Zhang (Georgia Institute of Technology), Xu Chu (Georgia Institute of Technology), and Ce Zhang (ETH Zurich)</i>	
Approximate Order Dependency Discovery	25
<i>Yifeng Jin (Fudan University, China; Shanghai Key Laboratory of Data Science), Zijing Tan (Fudan University, China; Shanghai Key Laboratory of Data Science), Weijun Zeng (Fudan University, China; Shanghai Key Laboratory of Data Science), and Shuai Ma (Beihang University, China; Beijing Advanced Innovation Center for Big Data and Brain Computing)</i>	
DBSCOUT: A Density-Based Method for Scalable Outlier Detection in Very Large Datasets	37
<i>Matteo Corain (Politecnico di Torino, Italy), Paolo Garza (Politecnico di Torino, Italy), and Abolfazl Asudeh (University of Illinois at Chicago, USA)</i>	
Bootstrapping Information Extraction via Conceptualization	49
<i>Jiaqing Liang (Fudan University, China; Fudan & Aishu Cognitive Intelligence Joint Research Center, China), Suo Feng (Fudan University, China), Chenhao Xie (Fudan University, China), Yanghua Xiao (Fudan University, China; Fudan & Aishu Cognitive Intelligence Joint Research Center, China), Jindong Chen (Fudan University, China), and Seung-Won Hwang (Seoul National University, South Korea)</i>	

Capturing Semantics for Imputation with Pre-Trained Language Models	61
<i>Yinan Mei (Tsinghua University), Shaoxu Song (Tsinghua University), Chenguang Fang (Tsinghua University), Haifeng Yang (HUAWEI Cloud BU), Jingyun Fang (HUAWEI Cloud BU), and Jiang Long (HUAWEI Cloud BU)</i>	

Graph Data Management 1

Manipulating Black-Box Networks for Centrality Promotion	73
<i>Wentao Li (University of Technology Sydney, Australia), Min Gao (Chongqing University, China), Fan Wu (Chongqing University, China), Wenge Rong (Beihang University, China), Junhao Wen (Chongqing University, China), and Lu Qin (University of Technology Sydney, Australia)</i>	
Efficient and Effective Community Search on Large-Scale Bipartite Graphs	85
<i>Kai Wang (University of New South Wales), Wenjie Zhang (University of New South Wales), Xuemin Lin (University of New South Wales), Ying Zhang (University of Technology Sydney), Lu Qin (University of Technology Sydney), and Yuting Zhang (University of New South Wales)</i>	
Efficient Community Search with Size Constraint	97
<i>Boge Liu (Guangzhou University; University of New South Wales), Fan Zhang (Guangzhou University), Wenjie Zhang (University of New South Wales), Xuemin Lin (University of New South Wales), and Ying Zhang (University of Technology Sydney)</i>	
Multi-attributed Community Search in Road-Social Networks	109
<i>Fangda Guo (Northeastern University, China), Ye Yuan (Beijing Institute of Technology, China), Guoren Wang (Beijing Institute of Technology, China), Xiangguo Zhao (Northeastern University, China), and Hao Sun (Northeastern University, China)</i>	
Peer Learning Through Targeted Dynamic Groups Formation	121
<i>Dong Wei (CS, New Jersey Institute of Technology), Yiannis Koutis (CS, New Jersey Institute of Technology), and Senjuti Basu Roy (CS, New Jersey Institute of Technology)</i>	
Efficient 2-Hop Labeling Maintenance in Dynamic Small-World Networks	133
<i>Mengxuan Zhang (University of Queensland, Australia), Lei Li (University of Queensland, Australia), Wen Hua (University of Queensland, Australia), and Xiaofang Zhou (University of Queensland, Australia)</i>	

Data Privacy

Differentially Private Publication of Multi-party Sequential Data	145
<i>Peng Tang (Shandong University, China), Rui Chen (Harbin Engineering University, China), Sen Su (Beijing University of Posts and Telecommunications, China), Shanqing Guo (Shandong University, China), Lei Ju (Shandong University, China), and Gaoyuan Liu (Shandong University, China)</i>	

Secure Dynamic Skyline Queries Using Result Materialization	157
<i>Sepanta Zeighami (University of Southern California), Gabriel Ghinita (University of Massachusetts Boston), and Cyrus Shahabi (University of Southern California)</i>	
P3GM: Private High-Dimensional Data Release via Privacy Preserving Phased Generative Model..	169
<i>Shun Takagi (Kyoto University), Tsubasa Takahashi (LINE Corporation), Yang Cao (Kyoto University), and Masatoshi Yoshikawa (Kyoto University)</i>	
Feature Inference Attack on Model Predictions in Vertical Federated Learning	181
<i>Xinjian Luo (National University of Singapore, Singapore), Yuncheng Wu (National University of Singapore, Singapore), Xiaokui Xiao (National University of Singapore, Singapore), and Beng Chin Ooi (National University of Singapore, Singapore)</i>	
Enabling Efficient Cyber Threat Hunting with Cyber Threat Intelligence	193
<i>Peng Gao (University of California), Fei Shao (Case Western Reserve University), Xiaoyuan Liu (University of California), Xusheng Xiao (Case Western Reserve University), Zheng Qin (Nanjing University), Fengyuan Xu (Nanjing University), Prateek Mittal (Princeton University), Sanjeev R. Kulkarni (Princeton University), and Dawn Song (University of California)</i>	
Twine: An Embedded Trusted Runtime for WebAssembly	205
<i>Jämes Ménétrety (University of Neuchâtel, Switzerland), Marcelo Pasin (University of Neuchâtel, Switzerland), Pascal Felber (University of Neuchâtel, Switzerland), and Valerio Schiavoni (University of Neuchâtel, Switzerland)</i>	

Crowdsourcing

Modeling Citywide Crowd Flows using Attentive Convolutional LSTM	217
<i>Chi Harold Liu (Beijing Institute of Technology, China), Chengzhe Piao (Beijing Institute of Technology, China), Xiaoxin Ma (Beijing Institute of Technology, China), Ye Yuan (Beijing Institute of Technology, China), Jian Tang (AI Labs, DiDi Chuxing, China), Guoren Wang (Beijing Institute of Technology, China), and Kin K. Leung (Imperial College, U.K.)</i>	
A Privacy-Enhanced and Personalized Safe Route Planner with Crowdsourced Data and Computation	229
<i>Fariha Tabassum Islam (Bangladesh University of Engineering and Technology Dhaka, Bangladesh), Tanzima Hashem (Bangladesh University of Engineering and Technology Dhaka, Bangladesh), and Rifat Shahriyar (Bangladesh University of Engineering and Technology Dhaka, Bangladesh)</i>	
Coalition-Based Task Assignment in Spatial Crowdsourcing	241
<i>Yan Zhao (Aalborg University, Denmark), Jiannan Guo (China Mobile Cloud Centre, China), Xuanhao Chen (University of Electronic Science and Technology of China, China), Jianye Hao (Tianjin University, China), Xiaofang Zhou (The Hong Kong University of Science and Technology, China), and Kai Zheng (University of Electronic Science and Technology of China, China)</i>	

Crowdsensing Data Trading Based on Combinatorial Multi-armed Bandit and Stackelberg Game ..	253
<i>Baoyi An (University of Science and Technology of China, China), Mingjun Xiao (University of Science and Technology of China, China), An Liu (Soochow University, China), Xike Xie (University of Science and Technology of China, China), and Xiaofang Zhou (The Hong Kong University of Science and Technology, Hong Kong)</i>	
Fairness-Aware Task Assignment in Spatial Crowdsourcing: Game-Theoretic Approaches	265
<i>Yan Zhao (Aalborg University, Denmark), Kai Zheng (University of Electronic Science and Technology of China, China), Jiannan Guo (China Mobile Cloud Centre, China), Bin Yang (Aalborg University, China), Torben Bach Pedersen (Aalborg University, China), and Christian S. Jensen (Aalborg University, China)</i>	
A Human-in-the-Loop Approach to Social Behavioral Targeting	277
<i>Jingru Yang (Renmin University of China), Xiaoman Zhao (Renmin University of China), Ju Fan (Renmin University of China), Gong Chen (Tencent Inc.), Chong Peng (Tencent Inc.), Sheng Yao (Tencent Inc.), and Xiaoyong Du (Renmin University of China)</i>	
CrowdRL: An End-to-End Reinforcement Learning Framework for Data Labelling	289
<i>Kaiyu Li (Tsinghua University, China), Guoliang Li (Tsinghua University, China), Yong Wang (Tsinghua University, China), Yan Huang (TAL Education Group, China), Zitao Liu (TAL Education Group, China), and Zhongqin Wu (TAL Education Group, China)</i>	

Spatial and Temporal Data Management 1

Rebuilding City-Wide Traffic Origin Destination from Road Speed Data	301
<i>Guanjie Zheng (The Pennsylvania State University, USA), Chang Liu (Shanghai Jiao Tong University, China), Hua Wei (The Pennsylvania State University, USA), Chacha Chen (The Pennsylvania State University, USA), and Zhenhui Li (The Pennsylvania State University, USA)</i>	
Constrained Route Planning over Large Multi-modal Time-Dependent Networks	313
<i>Yishu Wang (Northeastern University, China), Ye Yuan (Beijing Institute of Technology, China), Hao Wang (Nanjing University of Information Science & Technology, China), Xiangmin Zhou (RMIT University, Australia), Congcong Mu (Northeastern University, China), and Guoren Wang (Beijing Institute of Technology, China)</i>	
Online Route Planning over Time-Dependent Road Networks	325
<i>Di Chen (Northeastern University, China), Ye Yuan (Beijing Institute of Technology), Wenjin Du (Beijing Institute of Technology), Yurong Cheng (Beijing Institute of Technology), and Guoren Wang (Beijing Institute of Technology)</i>	
Dynamic Hub Labeling for Road Networks	336
<i>Mengxuan Zhang (University of Queensland, Australia), Lei Li (University of Queensland, Australia), Wen Hua (University of Queensland, Australia), Rui Mao (Shenzhen University, China), Pingfu Chao (University of Queensland, Australia), and Xiaofang Zhou (The Hong Kong University of Science and Technology)</i>	

An Effective Joint Prediction Model for Travel Demands and Traffic Flows	348
<i>Haitao Yuan (Tsinghua University, China), Guoliang Li (Tsinghua University, China), Zhifeng Bao (RMIT University, Australia), and Ling Feng (Tsinghua University, China)</i>	
A Learning-Based Method for Computing Shortest Path Distances on Road Networks	360
<i>Shuai Huang (Tsinghua University, China), Yong Wang (Tsinghua University, China), Tianyu Zhao (Tsinghua University, China), and Guoliang Li (Tsinghua University, China)</i>	

Distributed Data Management 1

Efficient Federated-Learning Model Debugging	372
<i>Anran Li (University of Science and Technology of China, China), Lan Zhang (University of Science and Technology of China, China), Junhao Wang (University of Science and Technology of China, China), Juntao Tan (University of Science and Technology of China, China), Feng Han (University of Science and Technology of China, China), Yaxuan Qin (University of Science and Technology of China, China), Nikolaos M. Freris (University of Science and Technology of China, China), and Xiang-Yang Li (University of Science and Technology of China, China)</i>	
Communication-Efficient Decentralized Machine Learning over Heterogeneous Networks	384
<i>Pan Zhou (University of Electronic Science and Technology of China, China), Qian Lin (National University of Singapore), Dumitrel Loghin (National University of Singapore), Beng Chin Ooi (National University of Singapore), Yuncheng Wu (National University of Singapore), and Hongfang Yu (University of Electronic Science and Technology of China, China)</i>	
Spark-Based Cloud Data Analytics using Multi-objective Optimization	396
<i>Fei Song (Ecole Polytechnique, France), Khaled Zaouk (Ecole Polytechnique, France), Chenghao Lyu (University of Massachusetts, USA), Arnab Sinha (Ecole Polytechnique, France), Qi Fan (Ecole Polytechnique, France), Yanlei Diao (Ecole Polytechnique, France; University of Massachusetts, USA), and Prashant Shenoy (University of Massachusetts, USA)</i>	
WedgeChain: A Trusted Edge-Cloud Store with Asynchronous (Lazy) Trust	408
<i>Faisal Nawab (University of California, Santa Cruz, United States)</i>	
CooLSM: Distributed and Cooperative Indexing Across Edge and Cloud Machines	420
<i>Natasha Mittal (University of California, Santa Cruz) and Faisal Nawab (University of California, Santa Cruz)</i>	
Interactive Analytic DBMSs: Breaching the Scalability Wall	432
<i>Pedro Pedreira (Facebook Inc.), Amit Dutta (Facebook Inc.), Sergey Pershin (Facebook Inc.), Lin Liu (Facebook Inc.), Sushant Shringarpure (Facebook Inc.), Jialiang Tan (Facebook Inc.), Brian Landers (Facebook Inc.), Ge Gao (Facebook Inc.), and Karen Pieper (Facebook Inc.)</i>	

Data Integration and Data Science

Relational Header Discovery using Similarity Search in a Table Corpus	444
<i>Hazar Harmouch (Hasso Plattner Institute, University of Potsdam), Thorsten Papenbrock (Hasso Plattner Institute, University of Potsdam), and Felix Naumann (Hasso Plattner Institute, University of Potsdam)</i>	
Efficient Joinable Table Discovery in Data Lakes: A High-Dimensional Similarity-Based Approach	456
<i>Yuyang Dong (NEC Corporation, Japan), Kunihiro Takeoka (NEC Corporation, Japan), Chuan Xiao (Osaka University and Nagoya University, Japan), and Masafumi Oyamada (NEC Corporation, Japan)</i>	
Valentine: Evaluating Matching Techniques for Dataset Discovery	468
<i>Christos Koutras (Delft University of Technology), George Siachamis (Delft University of Technology; ING Bank Netherlands), Andra Ionescu (Delft University of Technology), Kyriakos Psarakis (Delft University of Technology), Jerry Brons (ING Bank Netherlands), Marios Fragkoulis (Delft University of Technology), Christoph Lofi (Delft University of Technology), Angela Bonifati (Lyon 1 University), and Asterios Katsifodimos (Delft University of Technology)</i>	
Odess: Speeding up Resemblance Detection for Redundancy Elimination by Fast Content-Defined Sampling	480
<i>Xiangyu Zou (Harbin Institute of Technology), Cai Deng (Harbin Institute of Technology), Wen Xia (Harbin Institute of Technology; State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences), Philip Shilane (Dell Technologies), Haoliang Tan (Harbin Institute of Technology), Haijun Zhang (Harbin Institute of Technology), and Xuan Wang (Harbin Institute of Technology)</i>	
Latent Low-Rank Graph Learning for Multimodal Clustering	492
<i>Guo Zhong (University of Macau, China) and Chi-Man Pun (University of Macau, China)</i>	
Hate is the New Infodemic: A Topic-Aware Modeling of Hate Speech Diffusion on Twitter	504
<i>Sarah Masud (IIIT-Delhi, India), Subhabrata Dutta (Jadavpur University, India), Sakshi Makkar (IIIT-Delhi, India), Chhavi Jain (IIIT-Delhi, India), Vikram Goyal (IIIT-Delhi, India), Amitava Das (Wipro AI, India), and Tanmoy Chakraborty (IIIT-Delhi, India)</i>	

Graph Data Management 2

UniNet: Scalable Network Representation Learning with Metropolis-Hastings Sampling	516
<i>Xingyu Yao (BUPT), Yingxia Shao (BUPT), Bin Cui (Peking University), and Lei Chen (HKUST)</i>	
Towards Efficient Motif-Based Graph Partitioning: An Adaptive Sampling Approach	528
<i>Shixun Huang (RMIT University), Yuchen Li (Singapore Management University), Zhifeng Bao (RMIT University), and Zhao Li (Alibaba Group)</i>	

LineageBA: A Fast, Exact and Scalable Graph Generation for the Barabási-Albert Model	540
<i>Himchan Park (Korea Advanced Institute of Science and Technology, Korea) and Min-Soo Kim (Korea Advanced Institute of Science and Technology, Korea)</i>	
Search to Aggregate Neighborhood for Graph Neural Network	552
<i>Huan Zhao (4Paradigm Inc.), Quanming Yao (4Paradigm Inc.; Tsinghua University), and Weirui Tu (4Paradigm Inc.)</i>	
FastSGG: Efficient Social Graph Generation Using a Degree Distribution Generation Model	564
<i>Chaokun Wang (Tsinghua University, China), Binbin Wang (Tsinghua University, China), Bingyang Huang (Tsinghua University, China), Shaoxu Song (Tsinghua University, China), and Zai Li (Kwai Inc., China)</i>	
Noah: Neural-Optimized A* Search Algorithm for Graph Edit Distance Computation	576
<i>Lei Yang (Peking University, China) and Lei Zou (Peking University, China; National Engineering Laboratory for Big Data Analysis Technology and Application (PKU), China)</i>	

Indexing

TS-Benchmark: A Benchmark for Time Series Databases	588
<i>Yuanzhe Hao (Renmin University of China, China), Xiongpai Qin (Renmin University of China, China), Yueguo Chen (Renmin University of China, China), Yaru Li (Renmin University of China, China), Xiaoguang Sun (Renmin University of China, China), Yu Tao (Renmin University of China, China), Xiao Zhang (Renmin University of China, China), and Xiaoyong Du (Renmin University of China, China)</i>	
DBA Bandits: Self-Driving Index Tuning Under ad-hoc, Analytical Workloads with Safety Guarantees	600
<i>R. Malinga Perera (University of Melbourne), Bastian Oetomo (University of Melbourne), Benjamin I. P. Rubinstein (University of Melbourne), and Renata Borovica-Gajic (University of Melbourne)</i>	
Less is More: De-Amplifying I/Os for Key-Value Stores with a Log-Assisted LSM-Tree	612
<i>Kecheng Huang (Shandong University), Zhiping Jia (Shandong University), Zhaoyan Shen (Shandong University), Zili Shao (The Chinese University of Hong Kong), and Feng Chen (Louisiana State University)</i>	
Multidimensional Adaptive & Progressive Indexes	624
<i>Matheus Agio Nerone (CWI), Pedro Holanda (CWI), Eduardo Cunha de Almeida (UFPR), and Stefan Manegold (CWI)</i>	
Hash Adaptive Bloom Filter	636
<i>Rongbiao Xie (Nanjing University, China), Meng Li (Nanjing University, China), Zheyu Miao (Zhejiang University, China; Alibaba Group, China), Rong Gu (Nanjing University, China), He Huang (Sochow University, China), Haipeng Dai (Nanjing University, China), and Guihai Chen (Nanjing University, China)</i>	

HST+: An Efficient Index for Embedding Arbitrary Metric Spaces	648
<i>Yuxiang Zeng (The Hong Kong University of Science and Technology, China), Yongxin Tong (Beihang University, China), and Lei Chen (The Hong Kong University of Science and Technology, China)</i>	

Spatial and Temporal Data Management 2

Flow Computation in Temporal Interaction Networks	660
<i>Chrysanthi Kosyfaki (University of Ioannina, Greece), Nikos Mamoulis (University of Ioannina, Greece), Evaggelia Pitoura (University of Ioannina, Greece), and Panayiotis Tsaparas (University of Ioannina, Greece)</i>	
Leveraging Temporal and Topological Selectivities in Temporal-Clique Subgraph Query Processing	672
<i>Kaijie Zhu (Eindhoven University of Technology, The Netherlands; NDSC, China), George Fletcher (Eindhoven University of Technology, The Netherlands), and Nikolay Yakovets (Eindhoven University of Technology, The Netherlands)</i>	
Trajectory Simplification with Reinforcement Learning	684
<i>Zheng Wang (Nanyang Technological University, Singapore), Cheng Long (Nanyang Technological University, Singapore), and Gao Cong (Nanyang Technological University, Singapore)</i>	
E ² DTC: An End to End Deep Trajectory Clustering Framework via Self-Training	696
<i>Ziquan Fang (Zhejiang University, China), Yuntao Du (Zhejiang University, China), Lu Chen (Zhejiang University, China), Yujia Hu (Zhejiang University, China), Yunjun Gao (Zhejiang University, China; Alibaba–Zhejiang University Joint Institute of Frontier Technologies, China), and Gang Chen (Zhejiang University, China)</i>	
REPOSE: Distributed Top-k Trajectory Similarity Search with Local Reference Point Tries	708
<i>Bolong Zheng (Huazhong University of Science and Technology, China), Lianggui Weng (Huazhong University of Science and Technology, China), Xi Zhao (Huazhong University of Science and Technology, China), Kai Zeng (Alibaba Group, China), Xiaofang Zhou (Hong Kong University of Science and Technology, China), and Christian S. Jensen (Aalborg University, Denmark)</i>	
Durable Top-K Instant-Stamped Temporal Records with User-Specified Scoring Functions	720
<i>Junyang Gao (Google Inc.), Stavros Sintos (University of Chicago), Pankaj K. Agarwal (Duke University), and Jun Yang (Duke University)</i>	

Data Management on New Hardware

The Case for In-Memory OLAP on "Wimpy" Nodes	732
<i>Andrew Crotty (Brown University), Alex Galakatos (Brown University), Connor Luckett (Brown University), and Ugur Cetintemel (Brown University)</i>	

DyCuckoo: Dynamic Hash Tables on GPUs	744
<i>Yuchen Li (Singapore Management University; Zhejiang University), Qiwei Zhu (Zhejiang University), Zheng Lyu (Alibaba Group), Zhongdong Huang (Zhejiang University), and Jianling Sun (Zhejiang University)</i>	
Programming an SSD Controller to Support Batched Writes for Variable-Size Pages	756
<i>Jaeyoung Do (Microsoft Research, USA), Chen Luo (University of California, Irvine, USA), and David Lomet (Microsoft Research, USA)</i>	
Predict and Write: Using K-Means Clustering to Extend the Lifetime of NVM Storage	768
<i>Saeed Kargar (University of California, Santa Cruz), Heiner Litz (University of California, Santa Cruz), and Faisal Nawab (University of California, Santa Cruz)</i>	
Discriminative Admission Control for Shared-Everything Database Under Mixed OLTP Workloads....	780
<i>Donghui Wang (East China Normal University, China), Peng Cai (East China Normal University, China), Weining Qian (East China Normal University, China), and Aoying Zhou (East China Normal University, China)</i>	
Efficiently Reclaiming Space in a Log Structured Store	792
<i>David Lomet (Independent Researcher) and Chen Luo (Snowflake Inc.)</i>	

Stream Data Management 1

LogLog Filter: Filtering Cold Items within a Large Range over High Speed Data Streams	804
<i>Peng Jia (MOEKLINNS Lab, Xi'an Jiaotong University, Xi'an, China), Pinghui Wang (Shenzhen Research Institute of Xi'an Jiaotong University, Shenzhen, China; MOEKLINNS Lab, Xi'an Jiaotong University, Xi'an, China), Junzhou Zhao (MOEKLINNS Lab, Xi'an Jiaotong University, Xi'an, China), Ye Yuan (Beijing Institute of Technology, China), Jing Tao (MOEKLINNS Lab, Xi'an Jiaotong University, Xi'an, China), and Xiaohong Guan (MOEKLINNS Lab, Xi'an Jiaotong University, Xi'an, Shaanxi, China; Shenzhen Research Institute of Xi'an Jiaotong University, China; Tsinghua University, China)</i>	
SliceNStitch: Continuous CP Decomposition of Sparse Tensor Streams	816
<i>Taehyung Kwon (KAIST, South Korea), Inkyu Park (KAIST, South Korea), Dongjin Lee (KAIST, South Korea), and Kijung Shin (KAIST, South Korea)</i>	
DISC: Density-Based Incremental Clustering by Striding over Streaming Data	828
<i>Bogyong Kim (Seoul National University, Korea), Kyoseung Koo (Seoul National University, Korea), Juhun Kim (Seoul National University, Korea), and Bongki Moon (Seoul National University, Korea)</i>	
Robust Factorization of Real-World Tensor Streams with Patterns, Missing Values, and Outliers	840
<i>Dongjin Lee (Korea Advanced Institute of Science and Technology, South Korea) and Kijung Shin (Korea Advanced Institute of Science and Technology, South Korea)</i>	
Single Point Incremental Fourier Transform on 2D Data Streams	852
<i>Muhammad Saad (University of Zurich, Switzerland), Abraham Bernstein (University of Zurich, Switzerland), Michael Böhlen (University of Zurich, Switzerland), and Daniele Dell'Aglio (Aalborg University, Denmark; University of Zurich, Switzerland)</i>	

SALSA: Self-Adjusting Lean Streaming Analytics	864
<i>Ran Ben Basat (University College London), Gil Einziger (Ben Gurion University), Michael Mitzenmacher (Harvard University), and Shay Vargafik (VMware Research)</i>	

Knowledge Discovery

NewsLink: Empowering Intuitive News Search with Knowledge Graphs	876
<i>Yueji Yang (National University of Singapore), Yuchen Li (Singapore Management University), and Anthony K. H. Tung (National University of Singapore)</i>	
On Disambiguating Authors: Collaboration Network Reconstruction in a Bottom-up Manner	888
<i>Na Li (East China Normal University, China), Renyu Zhu (East China Normal University, China), Xiaoxu Zhou (East China Normal University, China), Xiangnan He (University of Science and Technology of China, China), Wenyuan Cai (Shanghai Hypers Data Technology Inc., China), Ming Gao (East China Normal University, China), and Aoying Zhou (East China Normal University, China)</i>	
A Bootstrapping Approach to Optimize Random Walk Based Statistical Estimation over Graphs ...	900
<i>Pei Yi (Chongqing University, China), Hong Xie (Chongqing University, China), Yongkun Li (University of Science & Technology of China, China), and John C.S. Lui (The Chinese University of Hong Kong)</i>	
Leveraging Meta-Path Contexts for Classification in Heterogeneous Information Networks	912
<i>Xiang Li (East China Normal University, China), Danhao Ding (The University of Hong Kong, China), Ben Kao (The University of Hong Kong, China), Yizhou Sun (University of California, USA), and Nikos Mamoulis (University of Ioannina, Greece)</i>	
Property Graph Schema Optimization for Domain-Specific Knowledge Graphs	924
<i>Rana Alotaibi (University of California San Diego, USA), Chuan Lei (IBM Research - Almaden, USA), Abdul Quamar (IBM Research - Almaden, USA), Vasilis Efthymiou (IBM Research - Almaden, USA), and Fatma Özcan (IBM Research - Almaden, USA)</i>	
Fast Core-Based Top-k Frequent Pattern Discovery in Knowledge Graphs	936
<i>Jian Zeng (Harbin Institute of Technology; Southern University of Science and Technology), Leong Hou U (University of Macau), Xiao Yan (Southern University of Science and Technology; Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology), Mingji Han (Southern University of Science and Technology), and Bo Tang (Southern University of Science and Technology; Research Institute of Trustworthy Autonomous Systems, Southern University of Science and Technology; University of Macau)</i>	

Query Processing and Optimization 1

The Logarithmic Dynamic Cuckoo Filter	948
<i>Fan Zhang (National Engineering Research Center for Big Data Technology and System, Cluster and Grid Computing Lab, Huazhong University of Science and Technology, China), Hanhua Chen (National Engineering Research Center for Big Data Technology and System, Cluster and Grid Computing Lab, Huazhong University of Science and Technology, China), Hai Jin (National Engineering Research Center for Big Data Technology and System, Cluster and Grid Computing Lab, Huazhong University of Science and Technology, China), and Pedro Reviriego (Universidad Carlos III de Madrid, Avenida de la Universidad, Spain)</i>	
Continuously Bulk Loading over Range Partitioned Tables for Large Scale Historical Data	960
<i>Peng Cai (East China Normal University, China), Xuan Zhou (East China Normal University, China), and Aoying Zhou (East China Normal University, China)</i>	
Eclipse: Generalizing kNN and Skyline	972
<i>Jinfei Liu (Zhejiang University), Li Xiong (Emory University), Qiuchen Zhang (Emory University), Jian Pei (Simon Fraser University), and Jun Luo (Machine Intelligence Center, Lenovo)</i>	
Memory-Efficient Key/Foreign-Key Join Size Estimation via Multiplicity and Intersection Size	984
<i>Magnus Müller (University of Mannheim, Germany), Daniel Flachs (University of Mannheim, Germany), and Guido Moerkotte (University of Mannheim, Germany)</i>	
Authenticated Keyword Search in Scalable Hybrid-Storage Blockchains	996
<i>Ce Zhang (Hong Kong Baptist University, Hong Kong), Cheng Xu (Hong Kong Baptist University, Hong Kong), Haixin Wang (Hong Kong Baptist University, Hong Kong), Jianliang Xu (Hong Kong Baptist University, Hong Kong), and Byron Choi (Hong Kong Baptist University, Hong Kong)</i>	
NestGPU: Nested Query Processing on GPU	1008
<i>Sofoklis Floratos (The Ohio State University, USA), Mengbai Xiao (The Ohio State University, USA), Hao Wang (The Ohio State University, USA), Chengxin Guo (Renmin University of China, China), Yuan Yuan (Google Inc., USA), Rubao Lee (RateUp Inc., USA), and Xiaodong Zhang (The Ohio State University, USA)</i>	

Data Management on New Hardware

Aria: Tolerating Skewed Workloads in Secure In-Memory Key-Value Stores	1020
<i>Fan Yang (Tsinghua University), Youmin Chen (Tsinghua University), Youyou Lu (Tsinghua University), Qing Wang (Tsinghua University), and Jiwu Shu (Tsinghua University)</i>	
CruiseDB: An LSM-Tree Key-Value Store with Both Better Tail Throughput and Tail Latency	1032
<i>Junkai Liang (Key Laboratory of DEKE, MOE, China; Renmin University of China, China) and Yunpeng Chai (Key Laboratory of DEKE, MOE, China; Renmin University of China, China)</i>	

FPGA for Aggregate Processing: The Good, The Bad, and The Ugly	1044
<i>Zubeyr Furkan Eryilmaz (University of Wisconsin-Madison), Aarati Kakaraparthi (University of Wisconsin-Madison), Jignesh M. Patel (University of Wisconsin-Madison), Rathijit Sen (Microsoft Gray Systems Lab), and Kwanghyun Park (Microsoft Gray Systems Lab)</i>	

Stream Data Management 2

Fingerprinting Concepts in Data Streams with Supervised and Unsupervised Meta-Information ..	1056
<i>Ben Halstead (The University of Auckland, New Zealand), Yun Sing Koh (The University of Auckland, New Zealand), Patricia Riddle (The University of Auckland, New Zealand), Mykola Pechenizkiy (Eindhoven University of Technology, The Netherlands), Albert Bifet (University of Waikato, New Zealand and LTCL, Telecom Paris, IP-Paris, France), and Russel Pears (Auckland University of Technology, New Zealand)</i>	
Concept Drift Detection from Multi-class Imbalanced Data Streams	1068
<i>Lukasz Korycki (Virginia Commonwealth University, USA) and Bartosz Krawczyk (Virginia Commonwealth University, USA)</i>	
DisMASTD: An Efficient Distributed Multi-aspect Streaming Tensor Decomposition	1080
<i>Keyu Yang (Zhejiang University, China), Yunjun Gao (Zhejiang University, China; Alibaba-Zhejiang University Joint Institute of Frontier Technologies, China), Yifeng Shen (Zhejiang University, China), Baihua Zheng (Singapore Management University, Singapore), and Lu Chen (Aalborg University, Denmark)</i>	

Stream Data Management 3

EDGE: Entity-Diffusion Gaussian Ensemble for Interpretable Tweet Geolocation Prediction	1092
<i>Bo Hui (Auburn University), Haiquan Chen (California State University, Sacramento), Da Yan (University of Alabama at Birmingham), and Wei-Shinn Ku (Auburn University)</i>	
Efficient Relation-Aware Scoring Function Search for Knowledge Graph Embedding	1104
<i>Shimin Di (The Hong Kong University of Science and Technology, Hong Kong SAR, China), Quanming Yao (4Paradigm Inc. Hong Kong SAR, China; Tsinghua University, China), Yongqi Zhang (4Paradigm Inc. Hong Kong SAR, China), and Lei Chen (The Hong Kong University of Science and Technology, Hong Kong SAR, China)</i>	
InfoShield: Generalizable Information-Theoretic Human-Trafficking Detection	1116
<i>Meng-Chieh Lee (National Chiao Tung University), Catalina Vajiac (Carnegie Mellon University), Aayushi Kulshrestha (McGill University and Mila), Sacha Levy (McGill University and Mila), Namyong Park (Carnegie Mellon University), Cara Jones (Marinus Analytics), Reihaneh Rabbany (McGill University and Mila), and Christos Faloutsos (Carnegie Mellon University)</i>	
An Efficient Approach for Cross-Silo Federated Learning to Rank	1128
<i>Yansheng Wang (Beihang University, China), Yongxin Tong (Beihang University, China), Dingyuan Shi (Beihang University, China), and Ke Xu (Beihang University, China)</i>	

Efficient Construction of Nonlinear Models over Normalized Data	1140
<i>Zhaoyue Cheng (University of Toronto), Nick Koudas (University of Toronto), Zhe Zhang (York University), and Xiaohui Yu (York University)</i>	
Workload-Aware Materialization for Efficient Variable Elimination on Bayesian Networks	1152
<i>Cigdem Aslay (Aarhus University, Denmark), Martino Ciaperoni (Aalto University, Finland), Aristides Gionis (KTH Royal Institute of Technology, Sweden), and Michael Mathioudakis (University of Helsinki, Finland)</i>	

Spatial and Temporal Data

A Distance-Based Scheme for Reducing Bandwidth in Distributed Geometric Monitoring	1164
<i>Yuval Alfassi (University of Haifa, Israel), Moshe Gabel (University of Toronto, Canada), Gal Yehuda (Technion - Israel Institute of Technology, Israel), and Daniel Keren (University of Haifa, Israel)</i>	
SAKE: Spatial Question Answering over Knowledge Graph Based on Embedding Techniques	1176
<i>Huan Li (Aalborg University, Denmark), Hua Lu (Roskilde University, Denmark), Lidan Shou (Zhejiang University, China), Ke Chen (Zhejiang University, China), and Gang Chen (Zhejiang University, China)</i>	
LHist: Towards Learning Multi-Dimensional Histogram for Massive Spatial Data	1188
<i>Qiyu Liu (HKUST), Yanyan Shen (Shanghai Jiao Tong University, China), and Lei Chen (HKUST)</i>	
Data-Driven Fairness-Aware Vehicle Displacement for Large-Scale Electric Taxi Fleets	1200
<i>Guang Wang (Rutgers University), Shuxin Zhong (Rutgers University), Shuai Wang (Southeast University), Fei Miao (University of Connecticut), Zheng Dong (Wayne State University), and Desheng Zhang (Rutgers University)</i>	
On Efficient and Scalable Time-Continuous Spatial Crowdsourcing	1212
<i>Ting Wang (University of Science and Technology of China, China), Xike Xie (University of Science and Technology of China, China), Xin Cao (University of New South Wales, Australia), Torben Pedersen (Aalborg University, Denmark), Yang Wang (University of Science and Technology of China, China), and Mingjun Xiao (University of Science and Technology of China, China)</i>	
Spatial-Temporal Similarity for Trajectories with Location Noise and Sporadic Sampling	1224
<i>Guanyao Li (The Hong Kong University of Science and Technology), Chih-Chieh Hung (National Chung Hsing university), Mengyun Liu (The Hong Kong University of Science and Technology), Linfei Pan (The Hong Kong University of Science and Technology), Wen-Chih Peng (National Yang Ming Chiao Tung University), and S.-H. Gary Chan (The Hong Kong University of Science and Technology)</i>	

Data Integration and Cleaning 2

Learning to Characterize Matching Experts	1236
<i>Roei Shraga (Technion, Israel), Ofra Amir (Technion, Israel), and Avigdor Gal (Technion, Israel)</i>	

End-to-end Task Based Parallelization for Entity Resolution on Dynamic Data	1248
<i>Leonardo Gazzarri (University of Stuttgart, Germany) and Melanie Herschel (National University of Singapore, Singapore, University of Stuttgart, Germany)</i>	
KDDLog: Performance and Scalability in Knowledge Discovery by Declarative Queries with Aggregates	1260
<i>Youfu Li (University of California, Los Angeles), Jin Wang (University of California, Los Angeles), Mingda Li (University of California, Los Angeles), Ariyam Das (University of California, Los Angeles), Jiaqi Gu (University of California, Los Angeles), and Carlo Zaniolo (University of California, Los Angeles)</i>	
Cost-Effective Variational Active Entity Resolution	1272
<i>Alex Bogatu (University of Manchester, UK; Peak AI Ltd.), Norman W. Paton (University of Manchester, UK), Mark Douthwaite (Peak AI Ltd.), Stuart Davie (Peak AI Ltd.), and André Freitas (University of Manchester, UK; Idiap Research Institute, Switzerland)</i>	
Structured Object Matching Across Web Page Revisions	1284
<i>Tobias Bleifuß (Hasso Plattner Institute, University of Potsdam), Leon Bornemann (Hasso Plattner Institute, University of Potsdam), Dmitri V. Kalashnikov (AT&T Labs – Research), Felix Naumann (Hasso Plattner Institute, University of Potsdam), and Divesh Srivastava (AT&T Chief Data Office)</i>	
Automating Entity Matching Model Development	1296
<i>Pei Wang (Simon Fraser University), Weiling Zheng (Simon Fraser University), Jiannan Wang (Simon Fraser University), and Jian Pei (Simon Fraser University)</i>	

Graph Data Management 3

A Framework to Quantify Approximate Simulation on Graph Data	1308
<i>Xiaoshuang Chen (University of New South Wales, Australia), Longbin Lai (Alibaba Group, China), Lu Qin (University of Technology Sydney, Australia), Xuemin Lin (University of New South Wales, Australia), and Boge Liu (University of New South Wales, Australia)</i>	
PEFP: Efficient k-hop Constrained s-t Simple Path Enumeration on FPGA	1320
<i>Zhengmin Lai (East China Normal University, China), You Peng (The University of New South Wales, Australia), Shiyu Yang (East China Normal University, China), Xuemin Lin (The University of New South Wales, Australia), and Wenjie Zhang (The University of New South Wales, Australia)</i>	
DPTL+: Efficient Parallel Triangle Listing on Batch-Dynamic Graphs	1332
<i>Michael Yu (University of New South Wales, Australia), Lu Qin (AAIL, University of Technology Sydney, Australia), Ying Zhang (AAIL, University of Technology Sydney, Australia), Wenjie Zhang (University of New South Wales, Australia), and Xuemin Lin (University of New South Wales, Australia)</i>	

Finding a Summary for All Maximal Cliques	1344
<i>Xiaofan Li (Swinburne University of Technology, Australia), Rui Zhou (Swinburne University of Technology, Australia), Lu Chen (Swinburne University of Technology, Australia), Yong Zhang (Tsinghua University, China), Chengfei Liu (Swinburne University of Technology, Australia), Qiang He (Swinburne University of Technology, Australia), and Yun Yang (Swinburne University of Technology, Australia)</i>	
An Efficient Algorithm for the Anchored k-Core Budget Minimization Problem	1356
<i>Kaixin Liu (BNRist, DCST, RIIT, Institute of Internet Industry, Tsinghua University, China), Sibowang (The Chinese University of Hong Kong, China), Yong Zhang (BNRist, DCST, RIIT, Institute of Internet Industry, Tsinghua University, China), and Chunxiao Xing (BNRist, DCST, RIIT, Institute of Internet Industry, Tsinghua University, China)</i>	
Scalable Graph Isomorphism: Combining Pairwise Color Refinement and Backtracking via Compressed Candidate Space	1368
<i>Geonmo Gu (Seoul National University, South Korea), Yehyun Nam (Seoul National University, South Korea), Kunsoo Park (Seoul National University, South Korea), Zvi Galil (Georgia Institute of Technology, USA), Giuseppe F. Italiano (LUISS University, Italy), and Wook-Shin Han (Pohang University of Science and Technology (POSTECH), Korea)</i>	

Distributed Data Management 2

Scalable Model-Based Management of Correlated Dimensional Time Series in ModelarDB+	1380
<i>Søren Kejser Jensen (Aalborg University, Denmark), Torben Bach Pedersen (Aalborg University, Denmark), and Christian Thomsen (Aalborg University, Denmark)</i>	
RCC: Resilient Concurrent Consensus for High-Throughput Secure Transaction Processing	1392
<i>Suyash Gupta (University of California, Davis), Jelle Hellings (University of California, Davis), and Mohammad Sadoghi (University of California, Davis)</i>	
WipDB: A Write-in-Place Key-Value Store that Mimics Bucket Sort	1404
<i>Xingsheng Zhao (University of Texas at Arlington, USA), Song Jiang (University of Texas at Arlington, USA), and Xingbo Wu (University of Illinois at Chicago, USA)</i>	
Lock Violation for Fault-Tolerant Distributed Database System	1416
<i>Hua Guo (Renmin University of China, China), Xuan Zhou (Shanghai Engineering Research Center of Big Data Management, East China Normal University, China), and Le Cai (Alibaba Group, US)</i>	
Efficient Control Flow in Dataflow Systems: When Ease-of-Use Meets High Performance	1428
<i>Gábor E. Gévay (Technische Universität Berlin (TU Berlin)), Tilmann Rabl (Hasso Plattner Institute, Uni Potsdam), Sebastian Breß (Snowflake Inc.), Loránd Madai-Tahy (Technische Universität Berlin (TU Berlin)), Jorge-Arnulfo Quiané-Ruiz (Technische Universität Berlin (TU Berlin), DFKI, Berlin), and Volker Markl (Technische Universität Berlin (TU Berlin), DFKI, Berlin)</i>	

Samya: A Geo-Distributed Data System for High Contention Aggregate Data	1440
<i>Sujaya Maiyya (University of California Santa Barbara), Ishtiyaque Ahmad (University of California Santa Barbara), Divyakant Agrawal (University of California Santa Barbara), and Amr El Abbadi (University of California Santa Barbara)</i>	

Distributed Data Management 3

FAST: FPGA-Based Subgraph Matching on Massive Graphs	1452
<i>Xin Jin (East China Normal University, China), Zhengyi Yang (University Of New South Wales, Australia), Xuemin Lin (University Of New South Wales, Australia), Shiyu Yang (Guangzhou University, China), Lu Qin (University of Technology Sydney, Australia), and You Peng (University Of New South Wales, Australia)</i>	
A+ Indexes: Tunable and Space-Efficient Adjacency Lists in Graph Database Management Systems	1464
<i>Amine Mhedhbi (University of Waterloo), Pranjal Gupta (University of Waterloo), Shahid Khaliq (University of Waterloo), and Semih Salihoglu (University of Waterloo)</i>	
Explaining Missing Data in Graphs: A Constraint-Based Approach	1476
<i>Qi Song (Amazon.com), Peng Lin (Washington State University), Hanchao Ma (Case Western Reserve University), and Yinghui Wu (Case Western Reserve University, Pacific Northwest National Laboratory)</i>	
Influence Maximization Based on Dynamic Personal Perception in Knowledge Graph	1488
<i>Ya-Wen Teng (Research Center for Information Technology Innovation, Academia Sinica, Taiwan), Yishuo Shi (Wenzhou University, China), Chih-Hua Tai (National Taipei University, Taiwan), De-Nian Yang (Research Center for Information Technology Innovation, Academia Sinica, Taiwan; Institute of Information Science, Academia Sinica, Taiwan), Wang-Chien Lee (The Pennsylvania State University, USA), and Ming-Syan Chen (National Taiwan University, Taiwan)</i>	
Privacy Preserving Strong Simulation Queries on Large Graphs	1500
<i>Lyu Xu (Hong Kong Baptist University, Hong Kong), Jiabin Jiang (Hong Kong Baptist University, Hong Kong), Byron Choi (Hong Kong Baptist University, Hong Kong), Jianliang Xu (Hong Kong Baptist University, Hong Kong), and Sourav S Bhowmick (Nanyang Technological University, Singapore)</i>	
Trillion-Scale Graph Processing Simulation Based on Top-Down Graph Upscaling	1512
<i>Himchan Park (Korea Advanced Institute of Science and Technology, Republic of Korea), Jinjun Xiong (IBM Thomas J. Watson Research Center, USA), and Min-Soo Kim (Korea Advanced Institute of Science and Technology, Republic of Korea)</i>	

Recommender Systems

Multi-facet Recommender Networks with Spherical Optimization	1524
<i>Yanchao Tan (Zhejiang University, China), Carl Yang (Emory University, United States), Xiangyu Wei (Zhejiang University, China), Yun Ma (Zhejiang University, China), and Xiaolin Zheng (Zhejiang University, China)</i>	
Group-Buying Recommendation for Social E-Commerce	1536
<i>Jun Zhang (Tsinghua University, China), Chen Gao (Tsinghua University, China), Depeng Jin (Tsinghua University, China), and Yong Li (Tsinghua University, China)</i>	
Reliable Recommendation with Review-Level Explanations	1548
<i>Yanzhang Lyu (Xi'an Jiaotong University), Hongzhi Yin (The University of Queensland), Jun Liu (Xi'an Jiaotong University), Mengyue Liu (Xi'an Jiaotong University), Huan Liu (Xi'an Jiaotong University), and Shizhuo Deng (Northeastern University)</i>	
Variational Self-Attention Network for Sequential Recommendation	1559
<i>Jing Zhao (Soochow University, China), Pengpeng Zhao (Soochow University, China), Lei Zhao (Soochow University, China), Yanchi Liu (Rutgers University, USA), Victor S. Sheng (Texas Tech University, USA), and Xiaofang Zhou (The Hong Kong University of Science and Technology, China)</i>	
Knowledge-Aware Group Representation Learning for Group Recommendation	1571
<i>Zhiyi Deng (University of Electronic Science and Technology of China, China), Changyu Li (University of Electronic Science and Technology of China, China), Shujin Liu (University of Electronic Science and Technology of China, China), Waqar Ali (University of Electronic Science and Technology of China, China), and Jie Shao (University of Electronic Science and Technology of China, China; Sichuan Artificial Intelligence Research Institute, China)</i>	
Attacking Black-box Recommendations via Copying Cross-Domain User Profiles	1583
<i>Wenqi Fan (The Hong Kong Polytechnic University), Tyler Derr (Vanderbilt University), Xiangyu Zhao (Michigan State University), Yao Ma (Michigan State University), Hui Liu (Michigan State University), Jianping Wang (City University of Hong Kong), Jiliang Tang (Michigan State University), and Qing Li (The Hong Kong Polytechnic University)</i>	

Query Processing and Optimization 2

Approximating Multidimensional Range Counts with Maximum Error Guarantees	1595
<i>Michael Shekelyan (University of Warwick, United Kingdom), Anton Dignoes (Free University of Bozen-Bolzano, Italy), Johann Gamper (Free University of Bozen-Bolzano, Italy), and Minos Garofalakis (ATHENA Research Center & Technical University of Crete, Greece)</i>	
LATEST: Learning-Assisted Selectivity Estimation Over Spatio-Textual Streams	1607
<i>Mayur Patil (University of California, Riverside) and Amr Magdy (University of California, Riverside)</i>	

ProMIPS: Efficient High-Dimensional c-Approximate Maximum Inner Product Search with a Lightweight Index	1619
<i>Yang Song (Northeastern University, China), Yu Gu (Northeastern University, China), Rui Zhang (www.ruizhang.info), and Ge Yu (Northeastern University, China)</i>	
A Fully Dynamic Algorithm for k-Regret Minimizing Sets	1631
<i>Yanhao Wang (University of Helsinki), Yuchen Li (Singapore Management University), Raymond Chi-Wing Wong (The Hong Kong University of Science and Technology), and Kian-Lee Tan (National University of Singapore)</i>	
Optimizing Error-Bounded Lossy Compression for Scientific Data by Dynamic Spline Interpolation	1643
<i>Kai Zhao (University of California, Riverside), Sheng Di (Argonne National Laboratory, USA), Maxim Dmitriev (Saudi Aramco, Saudi Arabia), Thierry Tonellot (Saudi Aramco, Saudi Arabia), Zizhong Chen (University of California, Riverside, USA), and Franck Cappello (Argonne National Laboratory, USA)</i>	
MLCask: Efficient Management of Component Evolution in Collaborative Data Analytics Pipelines	1655
<i>Zhaojing Luo (National University of Singapore), Sai Ho Yeung (National University of Singapore), Meihui Zhang (Beijing Institute of Technology), Kaiping Zheng (National University of Singapore), Lei Zhu (National University of Singapore), Gang Chen (Zhejiang University), Feiyi Fan (ICTCAS), Qian Lin (National University of Singapore), Kee Yuan Ngiam (National University Health System, Singapore), and Beng Chin Ooi (National University of Singapore)</i>	

Search and Retrieval

Improving Constrained Search Results By Data Melioration	1667
<i>Ido Guy (eBay Research), Tova Milo (Tel Aviv University), Slava Novgorodov (eBay Research), and Brit Youngmann (Tel Aviv University)</i>	
G-TADOC: Enabling Efficient GPU-Based Text Analytics without Decompression	1679
<i>Feng Zhang (Renmin University of China), Zaifeng Pan (Shanghai Jiao Tong University), Yanliang Zhou (Renmin University of China), Jidong Zhai (Tsinghua University), Xipeng Shen (North Carolina State University), Onur Mutlu (ETH Zürich), and Xiaoyong Du (Renmin University of China)</i>	
Fast Similarity Computation for t-SNE	1691
<i>Yasuhiro Fujiwara (NTT, Japan), Yasutoshi Ida (NTT, Japan), Sekitoshi Kanai (NTT, Japan), Atsutoshi Kumagai (NTT, Japan), and Naonori Ueda (NTT, Japan)</i>	
Rapid Approximate Aggregation with Distribution-Sensitive Interval Guarantees	1703
<i>Stephen Macke (University of Illinois (UIUC)), Maryam Aliakbarpour (MIT), Ilias Diakonikolas (University of Wisconsin, Madison), Aditya Parameswaran (UC Berkeley), and Ronitt Rubinfeld (MIT)</i>	

Optimally Summarizing Data by Small Fact Sets for Concise Answers to Voice Queries	1715
<i>Immanuel Trummer (Cornell University, USA) and Connor Anderson (Cornell University, USA)</i>	
Automatic Webpage Briefing	1727
<i>Yimeng Dai (The University of Melbourne, Australia), Rui Zhang (www.ruizhang.info), and Jianzhong Qi (The University of Melbourne, Australia)</i>	

Spatial and Temporal Data Management 3

EnhanceNet: Plugin Neural Networks for Enhancing Correlated Time Series Forecasting	1739
<i>Razvan-Gabriel Cirstea (Aalborg University, Denmark), Tung Kieu (Aalborg University, Denmark), Chenjuan Guo (Aalborg university, Denmark), Bin Yang (Aalborg University, Denmark), and Sinno Jialin Pan (Nanyang Technological University, Singapore)</i>	
Forecasting Ambulance Demand with Profiled Human Mobility via Heterogeneous Multi-graph Neural Networks	1751
<i>Zhaonan Wang (The University of Tokyo; National Institute of Advanced Industrial Science and Technology), Tianqi Xia (The University of Tokyo; National Institute of Advanced Industrial Science and Technology), Renhe Jiang (The University of Tokyo; Southern University of Science and Technology), Xin Liu (National Institute of Advanced Industrial Science and Technology; AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory), Kyoung-Sook Kim (National Institute of Advanced Industrial Science and Technology; AIST-Tokyo Tech Real World Big-Data Computation Open Innovation Laboratory), Xuan Song (The University of Tokyo; Southern University of Science and Technology), and Ryosuke Shibasaki (The University of Tokyo)</i>	
Efficient Constrained Shortest Path Query Answering with Forest Hop Labeling	1763
<i>Ziyi Liu (The University of Queensland, Australia), Lei Li (The University of Queensland, Australia), Mengxuan Zhang (The University of Queensland, Australia), Wen Hua (The University of Queensland, Australia), Pingfu Chao (The University of Queensland, Australia), and Xiaofang Zhou (The Hong Kong University of Science and Technology, Hong Kong)</i>	
TASM: A Tile-Based Storage Manager for Video Analytics	1775
<i>Maureen Daum (University of Washington), Brandon Haynes (Gray Systems Lab, Microsoft), Dong He (University of Washington), Amrita Mazumdar (University of Washington), and Magdalena Balazinska (University of Washington)</i>	
A Two-Layer Partitioning for Non-Point Spatial Data	1787
<i>Dimitrios Tsitsigkos (University of Ioannina), Konstantinos Lampropoulos (University of Ioannina), Panagiotis Bouros (Johannes Gutenberg University Mainz), Nikos Mamoulis (University of Ioannina), and Manolis Terrovitis (Athena Research Center)</i>	
Spangle: A Distributed In-Memory Processing System for Large-Scale Arrays	1799
<i>Sangchul Kim (Seoul National University, Korea), Bogyong Kim (Seoul National University, Korea), and Bongki Moon (Seoul National University, Korea)</i>	

Short Papers

Short Papers 1

Memory-Efficient Database Fragment Allocation for Robust Load Balancing when Nodes Fail	1811
<i>Halfpap Stefan (Hasso Plattner Institute, Germany) and Rainer Schlosser (Hasso Plattner Institute, Germany)</i>	
An Empirical Experiment on Deep Learning Models for Predicting Traffic Data	1817
<i>Hyunwook Lee (Ulsan National Institute of Science and Technology), Cheonbok Park (NAVER), Seungmin Jin (Ulsan National Institute of Science and Technology), Hyeslin Chu (Ulsan National Institute of Science and Technology), Jaegul Choo (Korea Advanced Institute of Science and Technology), and Sungahn Ko (Ulsan National Institute of Science and Technology)</i>	
Evaluating List Intersection on SSDs for Parallel I/O Skipping	1823
<i>Jianguo Wang (Purdue University), Chunbin Lin (Amazon), Yannis Papakonstantinou (University of California San Diego), and Steven Swanson (University of California San Diego)</i>	
Performance Characterization of HTAP Workloads	1829
<i>Utku Sirin (EPFL), Sandhya Dwarkadas (University of Rochester), and Anastasia Ailamaki (EPFL)</i>	
Accelerating the Yinyang K-Means Algorithm Using the GPU	1835
<i>Colin Taylor (Northern Arizona University, USA) and Michael Gowanlock (Northern Arizona University, USA)</i>	
SlimStore: A Cloud-Based Deduplication System for Multi-version Backups	1841
<i>Zihao Zhang (East China Normal University, China), Huiqi Hu (East China Normal University, China), Zhihui Xue (Alibaba Group, China), Changcheng Chen (Alibaba Group, China), Yang Yu (East China Normal University, China), Cuiyun Fu (Alibaba Group, China), Xuan Zhou (East China Normal University, China), and Feifei Li (Alibaba Group, China)</i>	
Meepo: Sharded Consortium Blockchain	1847
<i>Peilin Zheng (Sun Yat-sen University), Quanqing Xu (Ant Group), Zibin Zheng (Sun Yat-sen University), Zhiyuan Zhou (Ant Group), Ying Yan (Ant Group), and Hui Zhang (Ant Group)</i>	
SciChain: Blockchain-Enabled Lightweight and Efficient Data Provenance for Reproducible Scientific Computing	1853
<i>Abdullah Al-Mamun (University of Nevada), Feng Yan (University of Nevada), and Dongfang Zhao (University of Nevada)</i>	
Accelerating Similarity-Based Mining Tasks on High-Dimensional Data by Processing-in-Memory	1859
<i>Fang Wang (Hong Kong Polytechnic University), Man Lung Yiu (Hong Kong Polytechnic University), and Zili Shao (the Chinese University of Hong Kong)</i>	

DS ² : Handling Data Skew Using Data Stealings over High-Speed Networks	1865
<i>Zeyu He (East China Normal University), Zhifang Li (East China Normal University), Xiaoshuang Peng (East China Normal University), and Chuliang Weng (East China Normal University)</i>	
Efficient Matrix Factorization on Heterogeneous CPU-GPU Systems	1871
<i>Yuanhang Yu (Zhejiang Gongshang University, China; University of Technology Sydney, Australia), Dong Wen (University of Technology Sydney, Australia), Ying Zhang (University of Technology Sydney, Australia), Xiaoyang Wang (Zhejiang Gongshang University, China), Wenjie Zhang (The University of New South Wales, Australia), and Xuemin Lin (The University of New South Wales, Australia)</i>	
Rethink the Linearizability Constraints of Raft for Distributed Key-Value Stores	1877
<i>Yangyang Wang (Key Laboratory of Data Engineering and Knowledge Engineering, MOE, China; Renmin University of China, China), Zikai Wang (Key Laboratory of Data Engineering and Knowledge Engineering, MOE, China; Renmin University of China, China), Yunpeng Chai (Key Laboratory of Data Engineering and Knowledge Engineering, MOE, China; Renmin University of China), and Xin Wang (Tianjin University, China)</i>	
SING: Sequence Indexing Using GPUs	1883
<i>Botao Peng (Université de Paris, Institute of Computing Technology, Chinese Academy of Sciences & LIPADE), Panagiota Fatourou (University of Crete), and Themis Palpanas (LIPADE, Université de Paris & French University Institute (IUF))</i>	
TLBtree: A Read/Write-Optimized Tree Index for Non-Volatile Memory	1889
<i>Yongping Luo (University of Science and Technology of China, China), Peiquan Jin (University of Science and Technology of China, China), Qinglin Zhang (Tencent, China), and Bin Cheng (Tencent, China)</i>	
Utilizing Delta Trees for Efficient, Iterative Exploration and Transformation of Semi-Structured Contents	1895
<i>Nico Schäfer (TU Kaiserslautern, Germany) and Sebastian Michel (TU Kaiserslautern, Germany)</i>	
Joint Index, Sorting, and Compression Optimization for Memory-Efficient Spatio-Temporal Data Management	1901
<i>Keven Richly (Hasso Plattner Institute, Germany), Rainer Schlosser (Hasso Plattner Institute, Germany), and Martin Boissier (Hasso Plattner Institute, Germany)</i>	
High-Performance Smart Contracts Concurrent Execution for Permissioned Blockchain Using SGX	1907
<i>Min Fang (East China Normal University, China), Zhao Zhang (East China Normal University, China; Guilin University of Electronic Technology, China), Cheqing Jin (East China Normal University, China), and Aoying Zhou (East China Normal University, China)</i>	

Short Papers 2

Estimating the Extent of the Effects of Data Quality through Observations	1913
<i>Daniele Foroni (Huawei ERC), Matteo Lissandrini (Aalborg University), and Yannis Velegarakis (University of Trento and Utrecht University)</i>	

Decoupled Instance-Label Extreme Multi-label Classification with Skew Coordinate Feature Space	1919
<i>Jihyeon Song (Seoul National University, Korea) and Bongki Moon (Seoul National University, Korea)</i>	
Hierarchical Tree-Based Sequential Event Prediction with Application in the Aviation Accident Report	1925
<i>Xinyu Zhao (Arizona State University, USA), Hao Yan (Arizona State University, USA), and Yongming Liu (Arizona State University, USA)</i>	
Multi-behavior Enhanced Recommendation with Cross-Interaction Collaborative Relation Modeling	1931
<i>Lianghao Xia (South China University of Technology), Chao Huang (JD Finance America Corporation), Yong Xu (South China University of Technology), Peng Dai (JD Finance America Corporation), Mengyin Lu (JD Finance America Corporation), and Liefeng Bo (JD Finance America Corporation)</i>	
Ranking Data Slices for ML Model Validation: A Shapley Value Approach	1937
<i>Eitan Frachi (IBM Research, Israel), Ramasuri Narayanam (IBM Research, India), and Lokesh Nagalapatti (IBM Research, India)</i>	
From Minimum Change to Maximum Density: On S-Repair Under Integrity Constraints	1943
<i>Yu Sun (Tsinghua University, China) and Shaoxu Song (Tsinghua University, China)</i>	
Managing Consent for Data Access in Shared Databases	1949
<i>Osnat Drien (Bar Ilan University), Antoine Amarilli (LTCI, Télécom Paris, Institut Polytechnique de Paris), and Yael Amsterdamer (Bar-Ilan University)</i>	
Summarizing Provenance of Aggregate Query Results in Relational Databases	1955
<i>Omar Alomeir (University of British Columbia, Canada), Eugenie Yujing Lai (University of British Columbia, Canada), Mostafa Milani (University of British Columbia, Canada), and Rachel Pottinger (University of British Columbia, Canada)</i>	
Patterns Count-Based Labels for Datasets	1961
<i>Yuval Moskovitch (University of Michigan) and H. V. Jagadish (University of Michigan)</i>	
PROTEUS: Predictive Explanation of Anomalies	1967
<i>Nikolaos Myrtakis (University of Crete, Greece), Ioannis Tsamardinos (University of Crete, Greece), and Vassilis Christophides (ETIS Lab ENSEA, France)</i>	
Ranking Desired Tuples by Database Exploration	1973
<i>Xuedi Qin (Tsinghua University), Chengliang Chai (Tsinghua University), Yuyu Luo (Tsinghua University), Tianyu Zhao (Tsinghua University), Nan Tang (Qatar Computing Research Institute), Guoliang Li (Tsinghua University), Jianhua Feng (Tsinghua University), Xiang Yu (Tsinghua University), and Mourad Ouzzani (Qatar Computing Research Institute)</i>	
CIAO: An Optimization Framework for Client-Assisted Data Loading	1979
<i>Cong Ding (Peking University), Dixin Tang (The University of Chicago), Xi Liang (The University of Chicago), Aaron Elmore (The University of Chicago), and Sanjay Krishnan (The University of Chicago)</i>	

Optimizing Multiple Multi-way Stream Joins	1985
<i>Manuel Dossinger (TU Kaiserslautern, Germany) and Sebastian Michel (TU Kaiserslautern, Germany)</i>	
Updatable Materialization of Approximate Constraints	1991
<i>Steffen Kläbe (TU Ilmenau, Germany), Kai-Uwe Sattler (TU Ilmenau, Germany), and Stephan Baumann (Actian Germany GmbH, Germany)</i>	
Ranking Papers by Their Short-Term Scientific Impact	1997
<i>Ilias Kanellos (Athena R.C., Greece), Thanasis Vergoulis (Athena R.C., Greece), Dimitris Sacharidis (ULB, Belgium), Theodore Dalamagas (Athena R.C., Greece), and Yannis Vassiliou (NTU, Greece)</i>	
Substring Similarity Search with Synonyms	2003
<i>Gwangho Song (Seoul National University, South Korea), Kyuseok Shim (Seoul National University, South Korea), and Hongrae Lee (Google, USA)</i>	

Short Papers 3

CaSIE: Canonicalize and Informative Selection of the OpenIE System	2009
<i>Hao Xin (The Hong Kong University of Science and Technology, China), Xueling Lin (The Hong Kong University of Science and Technology, China), and Lei Chen (The Hong Kong University of Science and Technology, China)</i>	
Node2LV: Squared Lorentzian Representations for Node Proximity	2015
<i>Shanshan Feng (Inception Institute of Artificial Intelligence, UAE), Lisi Chen (KAUST, Saudi Arabia), Kaiqi Zhao (University of Auckland, New Zealand), Wei Wei (Huazhong University of Science and Technology, China), Fan Li (Fraunhofer Singapore, Singapore), and Shuo Shang (KAUST, Saudi Arabia)</i>	
Privacy-Preserving Sequential Publishing of Knowledge Graphs	2021
<i>Anh-Tu Hoang (University of Insubria, Italy), Barbara Carminati (University of Insubria, Italy), and Elena Ferrari (University of Insubria, Italy)</i>	
Cluster-and-Conquer: When Randomness Meets Graph Locality	2027
<i>George Giakkoupis (Univ Rennes, Inria, CNRS, IRISA, France), Anne-Marie Kermarrec (EPFL, Switzerland), Olivier Ruas (Inria, Univ. Lille, France), and François Taïani (Univ Rennes, Inria, CNRS, IRISA, France)</i>	
DDHH: A Decentralized Deep Learning Framework for Large-Scale Heterogeneous Networks	2033
<i>Mubashir Imran (The University of Queensland, Australia), Hongzhi Yin (The University of Queensland, Australia), Tong Chen (The University of Queensland, Australia), Zi Huang (The University of Queensland, Australia), Xiangliang Zhang (King Abdullah University of Science and Technology, Saudi Arabia), and Kai Zheng (University of Electronic Science and Technology of China, China)</i>	

EnsemFDet: An Ensemble Approach to Fraud Detection Based on Bipartite Graph	2039
<i>Yuxiang Ren (Florida State University, USA), Hao Zhu (Australian National University, USA), Jiawei Zhang (Florida State University, USA), Peng Dai (JD Finance America Corporation, USA), and Liefeng Bo (JD Finance America Corporation, USA)</i>	
HuGE: An Entropy-Driven Approach to Efficient and Scalable Graph Embeddings	2045
<i>Peng Fang (Huazhong University of Science and Technology, China), Fang Wang (Huazhong University of Science and Technology, China), Zhan Shi (Huazhong University of Science and Technology, China), Hong Jiang (Huazhong University of Science and Technology, China), Dan Feng (Huazhong University of Science and Technology, China), and Lei Yang (Huazhong University of Science and Technology, China)</i>	
Hypercore Maintenance in Dynamic Hypergraphs	2051
<i>Qi Luo (Shandong University, P.R. China), Dongxiao Yu (Shandong University, P.R. China), Zhipeng Cai (Georgia State University, USA), Xuemin Lin (University of New South Wales, Australia), and Xiuzhen Cheng (Shandong University, P.R. China)</i>	
Selective Edge Shedding in Large Graphs Under Resource Constraints	2057
<i>Yiling Zeng (Nankai University, China), Chunyao Song (Nankai University, China; Nanjing University of Posts and Telecommunications, China), and Tingjian Ge (University of Massachusetts Lowell, USA)</i>	
Social Visibility Optimization in OSNs with Anonymity Guarantees: Modeling, Algorithms and Applications	2063
<i>Shiyuan Zheng (The Chinese University of Hong Kong), Hong Xie (Chongqing University), and John C.S. Lui (The Chinese University of Hong Kong)</i>	
Stealthy Targeted Data Poisoning Attack on Knowledge Graphs	2069
<i>Prithu Banerjee (University of British Columbia), Lingyang Chu (Huawei Canada Technologies Co. Ltd.), Yong Zhang (Huawei Canada Technologies Co. Ltd.), Laks V.S. Lakshmanan (University of British Columbia), and Lanjun Wang (Huawei Canada Technologies Co. Ltd.)</i>	
Structure-Aware Parameter-Free Group Query via Heterogeneous Information Network Transformer	2075
<i>Hsi-Wen Chen (National Taiwan University, Taiwan), Hong-Han Shuai (National Chiao Tung University, Taiwan), De-Nian Yang (Academia Sinica, Taiwan), Wang-Chien Lee (Pennsylvania State University, United States), Chuan Shi (Beijing University of Posts and Telecommunications, China), Philip S. Yu (University of Illinois at Chicago, United States), and Ming-Syan Chen (National Taiwan University, Taiwan)</i>	

Taking Heuristic Based Graph Edge Partitioning One Step Ahead via OffStream Partitioning Approach	2081
<i>Tewodros Ayall (School of Computer Science and Engineering; University of Electronic Science and Technology of China, China), Hancong Duan (School of Computer Science and Engineering; University of Electronic Science and Technology of China, China), Changhong Liu (School of Computer Science and Engineering; University of Electronic Science and Technology of China, China), Fantahun Gereme (School of Computer Science and Engineering; University of Electronic Science and Technology of China, China), Mohammed Abegaz (School of Computer Science and Engineering; University of Electronic Science and Technology of China, China), and Mesay Deleli (University of Electronic Science and Technology of China, China)</i>	
Fast Distributed Complex Join Processing	2087
<i>Hao Zhang (The Chinese University of Hong Kong), Miao Qiao (The University of Auckland), Jeffrey Xu Yu (The Chinese University of Hong Kong), and Hong Cheng (The Chinese University of Hong Kong)</i>	
Top-k Community Similarity Search Over Large Road-Network Graphs	2093
<i>Niranjan Rai (Kent State University) and Xiang Lian (Kent State University)</i>	
Batching and Matching for Food Delivery in Dynamic Road Networks	2099
<i>Manas Joshi (IIT, Delhi, India), Arshdeep Singh (IIT, Delhi, India), Sayan Ranu (IIT, Delhi, India), Amitabha Bagchi (IIT, Delhi, India), Priyank Karia (Swiggy, India), and Puneet Kala (Swiggy, India)</i>	
Sequential Recommendation on Dynamic Heterogeneous Information Network	2105
<i>Tao Xie (Sun Yat-sen University, China), Yangjun Xu (Sun Yat-sen University, China), Liang Chen (Sun Yat-sen University, China), Yang Liu (Sun Yat-sen University, China), and Zibin Zheng (Sun Yat-sen University, China)</i>	

Short Papers 4

Towards the Smart use of Embedding and Instance Features for Property Matching	2111
<i>Daniel Ayala (Universidad de Sevilla, Spain), Inma Hernández (Universidad de Sevilla, Spain), David Ruiz (Universidad de Sevilla, Spain), and Erhard Rahm (Leipzig University, Germany)</i>	
AutoOD: Neural Architecture Search for Outlier Detection	2117
<i>Yuening Li (Texas A&M University, USA), Zhengzhang Chen (NEC Laboratories America, USA), Daochen Zha (Texas A&M University, USA), Kaixiong Zhou (Texas A&M University, USA), Haifeng Jin (Texas A&M University, USA), Haifeng Chen (NEC Laboratories America, USA), and Xia Hu (Texas A&M University, USA)</i>	
Catching them red-Handed: Real-Time Aggression Detection on Social Media	2123
<i>Herodotos Herodotou (Cyprus University of Technology, Cyprus), Despoina Chatzakou (Centre for Research and Technology Hellas, Greece), and Nicolas Kourtellis (Telefonica Research, Spain)</i>	

Gallat: A Spatiotemporal Graph Attention Network for Passenger Demand Prediction	2129
<i>Yuandong Wang (Beihang University, China), Hongzhi Yin (The University of Queensland, Australia), Tong Chen (The University of Queensland), Chunyang Liu (Didichuxing, China), Ben Wang (Didichuxing, China), Tianyu Wo (Beihang University, China), and Jie Xu (The University of Leeds)</i>	
Heterogeneous Information Assisted Bandit Learning: Theory and Application	2135
<i>Xiaoying Zhang (The Chinese University of Hong Kong), Hong Xie (Chongqing University, China), and John C.S. Lui (The Chinese University of Hong Kong)</i>	
Package Pick-up Route Prediction via Modeling Couriers' Spatial-Temporal Behaviors	2141
<i>Haomin Wen (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), Youfang Lin (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China; CAAC Key Laboratory of Intelligent Passenger Service of Civil Aviation, China), Fan Wu (Cainiao Supply Chain Management Co., Ltd., China), Huaiyu Wan (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China; CAAC Key Laboratory of Intelligent Passenger Service of Civil Aviation, China), Shengnan Guo (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), Lixia Wu (Cainiao Supply Chain Management Co., Ltd., China), Chao Song (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), and Yinghui Xu (Cainiao Supply Chain Management Co., Ltd., China)</i>	
Palette: Towards Multi-source Model Selection and Ensemble for Reuse	2147
<i>Yiming Li (Hong Kong University of Science and Technology, China), Yanyan Shen (Shanghai Jiao Tong University, China), and Lei Chen (Hong Kong University of Science and Technology, China)</i>	
Querying for Interactions	2153
<i>Yannis Xarchakos (University of Toronto, Canada) and Nick Koudas (University of Toronto, Canada)</i>	
An Autonomous Materialized View Management System with Deep Reinforcement Learning	2159
<i>Yue Han (Tsinghua University, China), Guoliang Li (Tsinghua University, China), Haitao Yuan (Tsinghua University, China), and Ji Sun (Tsinghua University, China)</i>	
Revisiting Data Prefetching for Database Systems with Machine Learning Techniques	2165
<i>Yu Chen (Tsinghua University, China), Yong Zhang (Tsinghua University, China), Jiacheng Wu (Tsinghua University, China), Jin Wang (University of California), and Chunxiao Xing (Tsinghua University, China)</i>	
Self-Supervised Deep Metric Learning for Pointsets	2171
<i>Pattaramanee Arsomngern (Vidyasirimedhi Institute of Science and Technology, Thailand), Cheng Long (Nanyang Technological University, Singapore), Supasorn Suwajanakorn (Vidyasirimedhi Institute of Science and Technology, Thailand), and Sarana Nutanong (Vidyasirimedhi Institute of Science and Technology, Thailand)</i>	

ValueNet: A Natural Language-to-SQL System that Learns from Database Information	2177
<i>Ursin Brunner (ZHAW Zurich University of Applied Sciences, Switzerland) and Kurt Stockinger (ZHAW Zurich University of Applied Sciences, Switzerland)</i>	
T3S: Effective Representation Learning for Trajectory Similarity Computation	2183
<i>Peilun Yang (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Hanchen Wang (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Ying Zhang (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Lu Qin (Centre for Artificial Intelligence, University of Technology Sydney, Australia), Wenjie Zhang (The University of New South Wales, Australia), and Xuemin Lin (The University of New South Wales, Australia)</i>	
TrajForesee: How Limited Detailed Trajectories Enhance Large-Scale Sparse Information to Predict Vehicle Trajectories?	2189
<i>Kangjia Shao (University of Science and Technology of China, China), Yang Wang (University of Science and Technology of China, China), Zhengyang Zhou (University of Science and Technology of China, China), Xike Xie (University of Science and Technology of China, China), and Guang Wang (Rutgers University, US)</i>	
Concurrency Control Based on Transaction Clustering	2195
<i>Xuebin Su (Harbin Institute of Technology, China), Hongzhi Wang (Harbin Institute of Technology, China), and Yan Zhang (Harbin Institute of Technology, China)</i>	
A Learning to Tune Framework for LSH	2201
<i>Xiu Tang (Zhejiang University), Sai Wu (Zhejiang University), Gang Chen (Zhejiang University), Jingyang Gao (Alibaba Group), Wei Cao (Alibaba Group), and Zhifei Pang (Zhejiang University)</i>	
TIRA in Baidu Image Advertising	2207
<i>Tan Yu (Baidu Search), Xuemeng Yang (Baidu Search), Yan Jiang (Baidu Search), Hongfang Zhang (Baidu Search), and Ping Li (Baidu Search)</i>	

Short Papers 5

Description Generation for Points of Interest	2213
<i>Meng Zhou (Peking University, China), Jingbo Zhou (Baidu Research, China), Yanjie Fu (University of Central Florida, United States), Zhaochun Ren (Shandong University, China), Xiaoli Wang (Xiamen University, China), and Hui Xiong (Rutgers University, United States)</i>	
CrowdAtlas: Estimating Crowd Distribution within the Urban Rail Transit System	2219
<i>Jinlong E (Nanyang Technological University, Singapore; NTU-Alibaba Singapore Joint Research Institute), Mo Li (Nanyang Technological University, Singapore; NTU-Alibaba Singapore Joint Research Institute), and Jianqiang Huang (Alibaba Group, China; NTU-Alibaba Singapore Joint Research Institute)</i>	

DAEMON: Unsupervised Anomaly Detection and Interpretation for Multivariate Time Series	2225
<i>Xuanhao Chen (University of Electronic Science and Technology of China, China), Liwei Deng (University of Electronic Science and Technology of China, China), Feiteng Huang (Cloud Database Innovation Lab of Cloud BU, Huawei Technologies Co., Ltd), Chengwei Zhang (Cloud Database Innovation Lab of Cloud BU, Huawei Technologies Co., Ltd), Zongquan Zhang (Cloud Database Innovation Lab of Cloud BU, Huawei Technologies Co., Ltd), Yan Zhao (Aalborg University), and Kai Zheng (University of Electronic Science and Technology of China, China)</i>	
Knowledge-Based Dynamic Systems Modeling: A Case Study on Modeling River Water Quality .	2231
<i>Namyong Park (Carnegie Mellon University), MinHyeok Kim (LG Electronics), Xuan Hoai Nguyen (AI Academy Vietnam), R.I. (Bob) McKay (Australian National University), and Dong-Kyun Kim (K-water Research Institute)</i>	
Collecting Geospatial Data with Local Differential Privacy for Personalized Services	2237
<i>Daeyoung Hong (Seoul National University, South Korea), Woohwan Jung (Seoul National University, South Korea), and Kyuseok Shim (Seoul National University, South Korea)</i>	
Experimental Study of Big Raster and Vector Database Systems	2243
<i>Samriddhi Singla (University of California), Ahmed Eldawy (University of California), Tina Diao (Stanford University), Ayan Mukhopadhyay (Stanford University), and Elia Scudiero (University of California)</i>	
Predicting the Impact of Disruptions to Urban Rail Transit Systems	2249
<i>Xiaoyun Mo (Nanyang Technological University, Singapore), Chu Cao (Nanyang Technological University, Singapore), Mo Li (Nanyang Technological University, Singapore), and David Z.W. Wang (Nanyang Technological University, Singapore)</i>	
An Actor-Critic Ensemble Aggregation Model for Time-Series Forecasting	2255
<i>Amal Saadallah (Artificial intelligence Group, TU Dortmund, Germany), Maryam Tavakol (Artificial intelligence Group, TU Dortmund, Germany), and Katharina Morik (Artificial intelligence Group, TU Dortmund, Germany)</i>	
Crowdrebate: An Effective Platform to Get More Rebate for Customers	2261
<i>Wangze Ni (The Hong Kong University of Science and Technology, China), Nian Chen (Hong Kong University of Science and Technology, China), Peng Cheng (East China Normal University, China), Lei Chen (The Hong Kong University of Science and Technology, China), and Xuemin Lin (The University of New South Wales, Australia; East China Normal University, China)</i>	
GRAB: Finding Time Series Natural Structures via A Novel Graph-Based Scheme	2267
<i>Yi Lu (Fudan University), Peng Wang (Fudan University), Bo Tang (Southern University of Science and Technology), Shen Liang (Fudan University), Chen Wang (Tsinghua University), Wei Wang (Fudan University), and Jianmin Wang (Tsinghua University)</i>	
Near-Optimal Fixed-Route Scheduling for Crowdsourced Transit System	2273
<i>Hanlin Li (University of Macau, China), Xiaowei Wu (University of Macau, China), Leong Hou U (University of Macau, China), and Kun Pang Kou (University of Macau, China)</i>	

SPEAR: Dynamic Spatio-Temporal Query Processing over High Velocity Data Streams	2279
<i>Furqan Baig (Stony Brook University), Dejun Teng (Stony Brook University), Jun Kong (Georgia State University), and Fusheng Wang (Stony Brook University)</i>	
The LSM RUM-Tree: A Log Structured Merge R-Tree for Update-Intensive Spatial Workloads	2285
<i>Jaewoo Shin (Purdue University), Jianguo Wang (Purdue University), and Walid G. Aref (Purdue University)</i>	
Top-K Publish/Subscribe for Ride Hitching	2291
<i>Yafei Li (Zhengzhou University, China), Hongyan Gu (Zhengzhou University, China), Rui Chen (Harbin Engineering University, China), Jianliang Xu (Hong Kong Baptist University, China), and Mingliang Xu (Zhengzhou University, China)</i>	
Towards Efficient MaxBRNN Computation for Streaming Updates	2297
<i>Wentao Ning (Southern University of Science and Technology, China), Xiao Yan (Southern University of Science and Technology, China), and Bo Tang (Southern University of Science and Technology, China)</i>	
User Profiling Based on Nonlinguistic Audio Data	2303
<i>Jiaxing Shen (The Hong Kong Polytechnic University), Oren Lederman (Massachusetts Institute of Technology), Jiannong Cao (The Hong Kong Polytechnic University), Shaojie Tang (The University of Texas at Dallas), and Alex Pentland (Massachusetts Institute of Technology)</i>	
Semantic Search Pipeline: From Query Expansion to Concept Forging	2309
<i>Elizabeth Soper (Vail Systems, Inc., USA; State University of New York at Buffalo, USA), Jordan Hosier (Vail Systems, Inc., USA; Northwestern University, USA), Dustin Bales (Vail Systems, Inc., USA), and Vijay K. Gurbani (Vail Systems, Inc., USA; Illinois Institute of Technology, USA)</i>	

TKDE Posters

Leveraging Currency for Repairing Inconsistent and Incomplete Data (Extended Abstract)	2315
<i>Xiaoou Ding (Harbin Institute of Technology, China), Hongzhi Wang (Harbin Institute of Technology, China), Jiakuan Su (Harbin Institute of Technology, China), Muxian Wang (Harbin Institute of Technology, China), Jianzhong Li (Harbin Institute of Technology, China), and Hong Gao (Harbin Institute of Technology, China)</i>	
A Collective Approach to Scholar Name Disambiguation (Extended Abstract)	2317
<i>Dongsheng Luo (Pennsylvania State University), Shuai Ma (Beihang University, China), Yaowei Yan (Pennsylvania State University), Chunmin Hu (Beihang University, China), Xiang Zhang (Pennsylvania State University), and Jinpeng Huai (Beihang University, China)</i>	
Analyzing In-Memory NoSQL Landscape (Extended Abstract)	2319
<i>Masoud Hemmatpour (Politecnico di Torino), Bartolomeo Montrucchio (Politecnico di Torino), Maurizio Rebaudengo (Politecnico di Torino), and Mohammad Sadoghi (University of California Davis)</i>	

LogStore: A Workload-Aware, Adaptable Key-Value Store on Hybrid Storage Systems (Extended Abstract)	2321
<i>Prashanth Menon (Carnegie Mellon University), Thamir M. Qadah (Umm Al-Qura University), Tilmann Rabl (TU Berlin), Mohammad Sadoghi (University of California Davis), and Hans-Arno Jacobsen (University of Toronto)</i>	
Entity Alignment for Knowledge Graphs with Multi-order Convolutional Networks (Extended Abstract)	2323
<i>Thanh Tam Nguyen (Ho Chi Minh City University of Technology (HUTECH)), Thanh Trung Huynh (Griffith University), Hongzhi Yin (The University of Queensland), Van Vinh Tong (Hanoi University of Science and Technology), Darnbi Sakong (Griffith University), Bolong Zheng (Huazhong University of Science and Technology), and Quoc Viet Hung Nguyen (Griffith University)</i>	
Compressed Indexes for Fast Search of Semantic Data (Extended Abstract)	2325
<i>Giulio Ermanno Pibiri (ISTI-CNR, Italy), Raffaele Perego (ISTI-CNR, Italy), and Rossano Venturini (University of Pisa, Italy)</i>	
FastDTW is Approximate and Generally Slower than the Algorithm it Approximates (Extended Abstract)	2327
<i>Renjie Wu (University of California, Riverside, USA) and Eamonn Keogh (University of California, Riverside, USA)</i>	
ESA-Stream: Efficient Self-Adaptive Online Data Stream Clustering (Extended Abstract)	2329
<i>Yanni Li (Xidian University, China), Hui Li (Xidian University, China), Zhi Wang (Xidian University, China), Bing Liu (University of Illinois, USA), Jiangtao Cui (Xidian University, China), and Hang Fei (Xidian University, China)</i>	
CuWide: Towards Efficient Flow-Based Training for Sparse Wide Models on GPUs	2330
<i>Xupeng Miao (Peking University; Tencent Inc.), Lingxiao Ma (Peking University), Zhi Yang (Peking University), Yingxia Shao (BUPT), Bin Cui (Peking University), Lele Yu (Tencent Inc.), and Jiawei Jiang (ETH Zurich)</i>	
Reliability Maximization in Uncertain Graphs (Extended Abstract)	2332
<i>Xiangyu Ke (Nanyang Technological University, Singapore), Arijit Khan (Nanyang Technological University, Singapore), Mohammad Al Hasan (Indiana University-Purdue University), and Rojin Rezoansangsari (Nanyang Technological University, Singapore)</i>	
MaxiZone: Maximizing Influence Zone over Geo-Textual Data (Extended Abstract)	2334
<i>Qing Liu (Hong Kong Baptist University, Hong Kong), Ziyuan Zhu (Zhejiang University, China), Jianliang Xu (Hong Kong Baptist University, Hong Kong), and Yunjun Gao (Zhejiang University, China)</i>	
Efficient Shapelet Discovery for Time Series Classification (Extended Abstract)	2336
<i>Guozhong Li (Hong Kong Baptist University), Byron Choi (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), Sourav S Bhowmick (Nanyang Technological University), Kwok-Pan Chun (Hong Kong Baptist University), and Grace Lai-Hung Wong (The Chinese University of Hong Kong)</i>	

A Generic Ontology Framework for Indexing Keyword Search on Massive Graphs (Extended Abstract)	2338
<i>Jiaxin Jiang (Hong Kong Baptist University), Byron Choi (Hong Kong Baptist University), Jianliang Xu (Hong Kong Baptist University), and Sourav S Bhowmick (Nanyang Technological University, Singapore)</i>	
LShape Partitioning: Parallel Skyline Query Processing using MapReduce (Extended Abstract)....	2340
<i>Heri Wijayanto (Asia University), Wenlu Wang (Texas A&M University-Corpus Christi), Wei-Shinn Ku (Auburn University), and Arbee L.P. Chen (Asia University)</i>	
Index-Based Solutions for Efficient Density Peak Clustering (Extended Abstract)	2342
<i>Zafaryab Rasool (Swinburne University of Technology, Australia), Rui Zhou (Swinburne University of Technology, Australia), Lu Chen (Swinburne University of Technology, Australia), Chengfei Liu (Swinburne University of Technology, Australia), and Jiajie Xu (Soochow University, China)</i>	
A Hybrid Data Cleaning Framework Using Markov Logic Networks (Extended Abstract)	2344
<i>Congcong Ge (Zhejiang University, China), Yunjun Gao (Zhejiang University, China), Xiaoye Miao (Zhejiang University, China), Bin Yao (Shanghai Jiao Tong University, China), and Haobo Wang (Zhejiang University, China)</i>	
Truss-Based Structural Diversity Search in Large Graphs (Extended Abstract)	2346
<i>Jinbin Huang (Hong Kong Baptist University, China), Xin Huang (Hong Kong Baptist University, China), and Jianliang Xu (Hong Kong Baptist University, China)</i>	
Towards Query Pricing on Incomplete Data (Extended Abstract)	2348
<i>Xiaoye Miao (Zhejiang University, China), Yunjun Gao (Zhejiang University, China), Lu Chen (Zhejiang University, China), Huanhuan Peng (Zhejiang University), Jianwei Yin (Zhejiang University, China), and Qing Li (Hong Kong Polytechnic University, China)</i>	
Effective Keyword Search in Weighted Graphs (Extended Abstract)	2350
<i>Mehdi Kargar (Ryerson University, Canada), Lukasz Golab (University of Waterloo, Canada), Divesh Srivastava (AT&T Chief Data Office, USA), Jaroslav Szlichta (Ontario Tech University, Canada), and Morteza Zihayat (Ryerson University, Canada)</i>	
Distributed Density Peaks Clustering Revisited (Extended Abstract)	2352
<i>Jing Lu (Northeastern University, China), Yuhai Zhao (Northeastern University, China), Kian-Lee Tan (National University of Singapore, Singapore), and Zhengkui Wang (Singapore Institute of Technology, Singapore)</i>	
Discovering Relaxed Functional Dependencies Based on Multi-attribute Dominance [Extended Abstract]	2354
<i>Loredana Caruccio (University of Salerno, Italy), Vincenzo Deufemia (University of Salerno, Italy), Felix Naumann (Hasso Plattner Institute, Germany), and Giuseppe Polese (University of Salerno, Italy)</i>	

Constrained Truth Discovery (Extended Abstract)	2356
<i>Chen Ye (Hangzhou Dianzi University, China), Hongzhi Wang (Harbin Institute of Technology & Peng Cheng Laboratory, China), Kangjie Zheng (Harbin Institute of Technology, China), YouKang Kong (Harbin Institute of Technology, China), Rong Zhu (Damo Academy, Alibaba Group, China), Jing Gao (SUNY Buffalo, USA), and Jianzhong Li (Harbin Institute of Technology, China)</i>	

Tutorials

Fairness in Rankings and Recommenders: Models, Methods and Research Directions	2358
<i>Evaggelia Pitoura (University of Ioannina, Greece), Kostas Stefanidis (Tampere University, Finland), and Georgia Koutrika (Athena Research Center, Greece)</i>	
Countering Bias in Personalized Rankings: From Data Engineering to Algorithm Development ...	2362
<i>Ludovico Boratto (EURECAT - Centre Tecnològic de Catalunya, Spain) and Mirko Marras (École Polytechnique Fédérale de Lausanne (EPFL), Switzerland)</i>	
Workload-Aware Performance Tuning for Autonomous DBMSs	2365
<i>Zhengdong Yan (University of Helsinki, Finland), Jiaheng Lu (University of Helsinki, Finland), Naresh Chainani (Amazon AWS, USA), and Chunbin Lin (Amazon AWS, USA)</i>	
High-Dimensional Similarity Search for Scalable Data Science	2369
<i>Karima Echiabi (Mohammed VI Polytechnic University), Kostas Zoumpatianos (Harvard University), and Themis Palpanas (LIPADE, University of Paris)</i>	
Evaluation of Duplicate Detection Algorithms: From Quality Measures to Test Data Generation	2373
<i>Fabian Panse (Universität Hamburg, Germany) and Felix Naumann (Hasso Plattner Institute, University of Potsdam, Germany)</i>	
Nullius in Verba: Reproducibility for Database Systems Research, Revisited	2377
<i>Wolfgang Mauerer (Technical University of Applied Sciences Regensburg, Siemens AG, Corporate Research, Germany) and Stefanie Scherzinger (University of Passau, Germany)</i>	

Industrial Papers

Query Optimization and Processing

Exploratory Data Analysis in SAP IQ Using Query-Time Sampling	2381
<i>Xiao Meng (SAP Labs, Canada) and Güneş Aluç (SAP Labs, Canada)</i>	
Swift: Reliable and Low-Latency Data Processing at Cloud Scale	2387
<i>Bo Wang (Alibaba Group), Zhenyu Hou (Alibaba Group), Yangyu Tao (Alibaba Group), Yifeng Lu (Alibaba Group), Chao Li (Alibaba Group), Tao Guan (Alibaba Group), Xiaowei Jiang (Alibaba Group), and Jinlei Jiang (Tsinghua University, China)</i>	

DBSpinner: Making a Case for Iterative Processing in Databases	2399
<i>Sofoklis Floratos (The Ohio State University, USA), Ahmad Ghazal (n/a), Jason Sun (ByteDance US Lab), Jianjun Chen (ByteDance US Lab), and Xiaodong Zhang (The Ohio State University, USA)</i>	
Prefix-Graph: A Versatile Log Parsing Approach Merging Prefix Tree with Probabilistic Graph	2411
<i>Guojun Chu (Beijing University of Posts and Telecommunications, China), Jingyu Wang (Beijing University of Posts and Telecommunications, China), Qi Qi (Beijing University of Posts and Telecommunications, China), Haifeng Sun (Beijing University of Posts and Telecommunications, China), Shimin Tao (Huawei, China), and Jianxin Liao (Beijing University of Posts and Telecommunications, China)</i>	
Microlearner: A Fine-Grained Learning Optimizer for Big Data Workloads at Microsoft	2423
<i>Alekh Jindal (Gray Systems Lab, Microsoft), Shi Qiao (Azure Data, Microsoft), Rathijit Sen (Gray Systems Lab, Microsoft), and Hireen Patel (Azure Data, Microsoft)</i>	
Query Rewriting via Cycle-Consistent Translation for E-Commerce Search	2435
<i>Yiming Qiu (JD.com, China), Kang Zhang (JD.com, China), Han Zhang (JD.com, China), Songlin Wang (JD.com, China), Sulong Xu (JD.com, China), Yun Xiao (JD.com Silicon Valley Research Center, United States), Bo Long (JD.com, China), and Wen-Yun Yang (Silicon Valley Research Center, United States)</i>	

E-Commerce and Recommendations

Learnings from a Retail Recommendation System on Billions of Interactions at bol.com	2447
<i>Barrie Kersbergen (Ahold Delhaize Research & AIRLab, University of Amsterdam) and Sebastian Scheller (Ahold Delhaize Research & AIRLab, University of Amsterdam)</i>	
Adversarial Mixture of Experts with Category Hierarchy Soft Constraint	2453
<i>Zhuojian Xiao (JD.com), Yunjiang Jiang (JD.com), Guoyu Tang (JD.com), Lin Liu (JD.com), Sulong Xu (JD.com), Yun Xiao (JD.com), and Weipeng Yan (JD.com)</i>	
Explore User Neighborhood for Real-Time E-Commerce Recommendation	2464
<i>Xu Xie (Peking University), Fei Sun (Alibaba Group), Xiaoyong Yang (Alibaba Group), Zhao Yang (Alibaba Group), Jinyang Gao (Alibaba Group), Wenwu Ou (Alibaba Group), and Bin Cui (Peking University; Center for Data Science, Peking University & National Engineering Laboratory for Big Data Analysis and Applications; Peiking University(Qingdao), China)</i>	
Billion-Scale Pre-Trained E-Commerce Product Knowledge Graph Model	2476
<i>Wen Zhang (Zhejiang University, China), Chi-Man Wong (Alibaba Group, University of Macao), Ganqiang Ye (Zhejiang University, China), Bo Wen (Zhejiang University, China), Wei Zhang (Alibaba Group, China), and Huajun Chen (Zhejiang University, China)</i>	

Purchase Intent Forecasting with Convolutional Hierarchical Transformer Networks	2488
<i>Chao Huang (JD Finance America Corporation), Jiashu Zhao (Wilfrid Laurier University), and Dawei Yin (Baidu Inc)</i>	
ATNN: Adversarial Two-Tower Neural Network for New Item's Popularity Prediction in E-Commerce	2499
<i>Xin Shen (Nanyang Technological University, Singapore), Zhao Li (Alibaba Group, China), Pengcheng Zou (Alibaba Group, China), Cheng Long (Nanyang Technological University, Singapore), Jie Zhang (Nanyang Technological University, Singapore), Jiajun Bu (Zhejiang University, China), and Jingren Zhou (Alibaba Group, China)</i>	

Spatial/Temporal/Cloud/Edge

Learning to Optimize Industry-Scale Dynamic Pickup and Delivery Problems	2511
<i>Xijun Li (MIRA Lab, USTC; Noah's Ark Lab), Weilin Luo (Noah's Ark Lab; Beihang University), Mingxuan Yuan (Noah's Ark Lab), Jun Wang (Noah's Ark Lab; University College London), Jiatwen Lu (Noah's Ark Lab), Jie Wang (MIRA Lab, USTC), Jinhu Lü (Beihang University), and Jia Zeng (Noah's Ark Lab)</i>	
The IoT Meta-Control Firewall	2523
<i>Soteris Constantinou (University of Cyprus, Cyprus), Andreas Konstantinidis (Frederick University, Cyprus; University of Cyprus, Cyprus), Demetrios Zeinalipour-Yazti (University of Cyprus, Cyprus), and Panos K. Chrysanthis (University of Pittsburgh, USA; University of Cyprus, Cyprus)</i>	
GeoDart: A System for Discovering Maps Discrepancies	2535
<i>Ayush Bandil (University of Washington Tacoma, USA), Vaishali Girdhar (University of Washington Tacoma, USA), Hieu Chau (University of Washington Tacoma, USA), Mohamed Ali (University of Washington Tacoma, USA), Abdeltawab Hendawi (University of Rhode Island Kingston, USA), Harsh Govind (Microsoft Corporation, USA), Peiwei Cao (Microsoft Corporation, USA), and Ashley Song (Microsoft Corporation, USA)</i>	
Implementing Rigid Temporal Geometries in Moving Object Databases	2547
<i>Maxime Schoemans (Université Libre de Bruxelles, Belgium), Mahmoud Sakr (Université Libre de Bruxelles, Belgium; Ain Shams University, Egypt), and Esteban Zimányi (Université Libre de Bruxelles, Belgium)</i>	
IntelliTag: An Intelligent Cloud Customer Service System Based on Tag Recommendation	2559
<i>Minghui Yang (Ant Group, China), Shaosheng Cao (Ant Group, China), Binbin Hu (Ant Group, China), Xianling Chen (Ant Group, China), Hengbin Cui (Ant Group, China), Zhiqiang Zhang (Ant Group, China), Jun Zhou (Ant Group, China), and Xiaolong Li (Ant Group, China)</i>	
IPS: Unified Profile Management for Ubiquitous Online Recommendations	2571
<i>Rui Shi (ByteDance, Inc.), Yang Liu (ByteDance, Inc.), Jianjun Chen (ByteDance, Inc.), Xuan Zou (ByteDance, Inc.), Yanbin Chen (ByteDance, Inc.), Minghua Fan (ByteDance, Inc.), Zhihao Cai (ByteDance, Inc.), Guanghui Zhang (ByteDance, Inc.), Zhiwen Li (ByteDance, Inc.), and Yuming Liang (ByteDance, Inc.)</i>	

Analysis and ML over Graphs

Turbo: Fraud Detection in Deposit-free Leasing Service via Real-Time Behavior Network Mining	2583
<i>Sihao Hu (Zhejiang University, China; Alibaba Group, China), Xuhong Zhang (Zhejiang University, China), Junfeng Zhou (Zhejiang University, China), Shouling Ji (Zhejiang University, China), Jiaqi Yuan (Zhejiang University, China), Zhao Li (Alibaba Group, China), Zhipeng Wang (Jimi Store, China), Chen Qi (Zhejiang University, China), Qinming He (Zhejiang University, China), and Liming Fang (Nanjing University of Aeronautics and Astronautics, and Astronautics, China)</i>	
Large-Scale Fake Click Detection for E-Commerce Recommendation Systems	2595
<i>Jingdong Li (East China Normal University, China), Zhao Li (Alibaba Group, China), Jiaming Huang (Alibaba Group, China), Ji Zhang (Zhejiang Lab, China), Xiaoling Wang (East China Normal University, China), Xingjian Lu (East China Normal University, China), and Jingren Zhou (Alibaba Group, China)</i>	
Improving Conversational Recommender System by Pretraining Billion-Scale Knowledge Graph	2607
<i>Chi-Man Wong (Alibaba Group; University of Macau), Fan Feng (Alibaba Group), Wen Zhang (Zhejiang University, China), Chi-Man Vong (University of Macau), Hui Chen (Alibaba Group), Yichi Zhang (Alibaba Group), Peng He (Alibaba Group), Huan Chen (Alibaba Group), Kun Zhao (Alibaba Group), and Huaqun Chen (Zhejiang University, China)</i>	
Efficient and Scalable Structure Learning for Bayesian Networks: Algorithms and Applications	2613
<i>Rong Zhu (Alibaba Group, China), Andreas Pfadler (Alibaba Group, China), Ziniu Wu (Alibaba Group, China; Oxford University, UK), Yuxing Han (Alibaba Group, China), Xiaoke Yang (Alibaba Group, China), Feng Ye (Alibaba Group, China), Zhenping Qian (Alibaba Group, China), Jingren Zhou (Alibaba Group, China), and Bin Cui (Peking University, China)</i>	
ReLink: Complete-Link Industrial Record Linkage Over Hybrid Feature Spaces	2625
<i>Salil Rajeev Joshi (American Express AI Labs), Arpan Somani (American Express AI Labs), and Shourya Roy (Flipkart)</i>	
Distributed Company Control in Company Shareholding Graphs	2637
<i>Andrea Gulino (Politecnico di Milano), Stefano Ceri (Politecnico di Milano), Emanuel Sallinger (TU Wien and Univ. of Oxford), Georg Gottlob (TU Wien and Univ. of Oxford), and Luigi Bellomarini (Banca d'Italia)</i>	

DEMOS

Demos 1

SPARQLit: Interactive SPARQL Query Refinement	2649
<i>Yael Amsterdamer (Bar-Ilan University, Israel) and Yehuda Callen (Bar-Ilan University, Israel)</i>	

SubDEX: Exploring Ratings in Subjective Databases	2653
<i>Sihem Amer-Yahia (CNRS, Univ. Grenoble Alpes), Tova Milo (Tel Aviv University), and Brit Youngmann (Tel Aviv University)</i>	
SOUP: A Fleet Management System for Passenger Demand Prediction and Competitive Taxi Supply	2657
<i>Qi Hu (Huazhong University of Science and Technology, China), Lingfeng Ming (Huazhong University of Science and Technology, China), Ruijie Xi (Huazhong University of Science and Technology, China), Lu Chen (Zhejiang University, China), Christian Jensen (Aalborg University, Denmark), and Bolong Zheng (Huazhong University of Science and Technology, China)</i>	
VADETIS: An Explainable Evaluator for Anomaly Detection Techniques	2661
<i>Abdelouahab Khelifati (University of Fribourg, Switzerland), Mourad Khayati (University of Fribourg, Switzerland), Philippe Cudré-Mauroux (University of Fribourg, Switzerland), Adrian Hännli (University of Bern, Switzerland), Qian Liuz (TU Berlin, Germany), and Manfred Hauswirthz (TU Berlin, Germany)</i>	
CoWiz: Interactive Covid-19 Visualization Based on Multilayer Network Analysis	2665
<i>Kunal Samant (The University of Texas at Arlington), Endrit Memeti (The University of Texas at Arlington), Abhishek Santra (The University of Texas at Arlington), Enamul Karim (The University of Texas at Arlington), and Sharma Chakravarthy (The University of Texas at Arlington)</i>	
SpeakNav: A Voice-Based Navigation System via Route Description Language Understanding ...	2669
<i>Lei Bi (Huazhong University of Science and Technology, China), Juan Cao (Huazhong University of Science and Technology, China), Guohui Li (Huazhong University of Science and Technology, China), Quoc Viet Hung Nguyen (Griffith University, Australia), Christian S. Jensen (Aalborg University, Denmark), and Bolong Zheng (Huazhong University of Science and Technology, China)</i>	
QeNoBi: A System for QuErying and miNing Behavioral Patterns	2673
<i>Abdelouahab Chibah (CNRS, Univ. Grenoble Alpes Saint Martin D'Hères, France), Sihem Amer-Yahia (CNRS, Univ. Grenoble Alpes Saint Martin D'Hères, France), and Laure Berti-Equille (IRD ESPACE-DEV, France)</i>	

Demos 2

CREATE: Clinical Report Extraction and Annotation Technology	2677
<i>Yichao Zhou (University of California, Los Angeles), Wei-Ting Chen (University of California, Los Angeles), Bowen Zhang (University of California, Los Angeles), David Lee (University of California, Los Angeles), J. Harry Caufield (University of California, Los Angeles), Kai-Wei Chang (University of California, Los Angeles), Yizhou Sun (University of California, Los Angeles), Peipei Ping (University of California, Los Angeles), and Wei Wang (University of California, Los Angeles)</i>	

UniKG: A Unified Interoperable Knowledge Graph Database System	2681
<i>Baozhu Liu (Tianjin University, China), Xin Wang (Tianjin University, China), Pengkai Liu (Tianjin University, China), Sizhuo Li (Tianjin University, China), Qiang Fu (Tianjin University, China), and Yunpeng Chai (Renmin University of China, China)</i>	
A Cockpit for the Development and Evaluation of Autonomous Database Systems	2685
<i>Jan Kossmann (Hasso Plattner Institute, University of Potsdam, Germany), Martin Boissier (Hasso Plattner Institute, University of Potsdam, Germany), Alexander Dubrawski (Hasso Plattner Institute, University of Potsdam, Germany), Fabian Heseding (Hasso Plattner Institute, University of Potsdam, Germany), Caterina Mandel (Hasso Plattner Institute, University of Potsdam, Germany), Udo Pigorsch (Hasso Plattner Institute, University of Potsdam, Germany), Max Schneider (Hasso Plattner Institute, University of Potsdam, Germany), Til Schniese (Hasso Plattner Institute, University of Potsdam, Germany), Mona Sobhani (Hasso Plattner Institute, University of Potsdam, Germany), Petr Tsayun (Hasso Plattner Institute, University of Potsdam, Germany), Katharina Wille (Hasso Plattner Institute, University of Potsdam, Germany), Michael Perscheid (Hasso Plattner Institute, University of Potsdam, Germany), Matthias Uflacker (Hasso Plattner Institute, University of Potsdam, Germany), and Hasso Plattner (Hasso Plattner Institute, University of Potsdam, Germany)</i>	
Automated Data Science for Relational Data	2689
<i>Thanh Lam Hoang (IBM Research, Ireland), Beat Buesser (IBM Research, Ireland), Hong Min (IBM Research, USA), Ngoc Minh Tran (IBM Research, Ireland), Martin Wistuba (IBM Research, Ireland), Udayan Khurana (IBM Research, USA), Gregory Bramble (IBM Research, USA), Theodoros Salonidis (IBM Research, USA), Dakuo Wang (IBM Research, USA), and Horst Samulowitz (IBM Research, USA)</i>	
Josch: Managing Schemas for NoSQL Document Stores	2693
<i>Michael Fruth (University of Passau, Germany), Kai Dauberschmidt (University of Passau, Germany), and Stefanie Scherzinger (University of Passau, Germany)</i>	
DeBinelle: Semantic Patches for Coupled Database-Application Evolution	2697
<i>Stefanie Scherzinger (University Passau, Germany), Wolfgang Mauerer (Technical University of Applied Sciences Regensburg; Siemens AG, Corporate Research, Germany), and Haridimos Kondylakis (FORTH_ICS, Greece)</i>	
ConCaT: Construction of Category Trees from Search Queries in E-Commerce	2701
<i>Uri Avron (Tel Aviv University), Shay Gershtein (Tel Aviv University), Ido Guy (eBay Research), Tova Milo (Tel Aviv University), and Slava Novgorodov (eBay Research)</i>	

Demos 3

A System for Efficiently Hunting for Cyber Threats in Computer Systems Using Threat Intelligence	2705
<i>Peng Gao (University of California, Berkeley), Fei Shao (Case Western Reserve University), Xiaoyuan Liu (University of California, Berkeley), Xusheng Xiao (Case Western Reserve University), Haoyuan Liu (University of California, Berkeley), Zheng Qin (Nanjing University), Fengyuan Xu (Nanjing University), Prateek Mittal (Princeton University), Sanjeev R. Kulkarni (Princeton University), and Dawn Song (University of California, Berkeley)</i>	
Odlaw: A Tool for Retroactive GDPR Compliance	2709
<i>Connor Luckett (Brown University), Andrew Crotty (Brown University), Alex Galakatos (Brown University), and Ugur Cetintemel (Brown University)</i>	
PITA: Privacy Through Provenance Abstraction	2713
<i>Daniel Deutch (Tel Aviv University), Ariel Frankenthal (Tel Aviv University), Amir Gilad (Duke University), and Yuval Moskovitch (University of Michigan)</i>	
The F4U System for Understanding the Effects of Data Quality	2717
<i>Daniele Foroni (Huawei ERC), Matteo Lissandrini (Aalborg University), and Yannis Velegarakis (University of Trento and Utrecht University)</i>	
FloraVision: A Spatial Crowd-Based Learning System for California Native Plants	2721
<i>George Constantinou (Integrated Media Systems Center, University of Southern California, USA), Onur Orhan (Integrated Media Systems Center, University of Southern California, USA), Roopal Kondepudi (Integrated Media Systems Center, University of Southern California, USA), Hyunjae Cho (Integrated Media Systems Center, University of Southern California, USA), Seon Ho Kim (Integrated Media Systems Center, University of Southern California, USA), Abdullah Alfarrarjeh (German Jordanian University, Jordan), and Cyrus Shahabi (Integrated Media Systems Center, University of Southern California, USA)</i>	
Clouseau: Blockchain-Based Data Integrity for HDFS Clusters	2725
<i>Alyzia Konsta (National Technical University of Athens), Ioannis Mytilinis (École Polytechnique Fédérale de Lausanne (EPFL)), Katerina Doka (National Technical University of Athens), Sotiris Niarchos (National Technical University of Athens), and Nectarios Koziris (National Technical University of Athens)</i>	
REACT: Real-Time Contact Tracing and Risk Monitoring via Privacy-Enhanced Mobile Tracking	2729
<i>Yanan Da (Emory University, USA), Ritesh Ahuja (University of Southern California, USA), Li Xiong (Emory University, USA), and Cyrus Shahabi (University of Southern California, USA)</i>	

Ph.D Symposium

Edge Sparsification for Graphs via Meta-Learning	2733
<i>Guihong Wan (The University of Texas at Dallas)</i>	

MoniLog: An Automated Log-Based Anomaly Detection System for Cloud Computing Infrastructures	2739
<i>Arthur Vervaeet (ISEP - Institut Supérieur d'Electronique de Paris / 3DS OUTSCALE, France)</i>	
Graph Based Approach to Real-Time Metro Passenger Flow Anomaly Detection	2744
<i>Weiqi Zhang (Hong Kong University of Science and Technology, Hong Kong)</i>	
Combining Anatomical Constraints and Deep Learning for 3-D CBCT Dental Image Multi-label Segmentation	2750
<i>Jiayu Huang (Arizona State University, USA), Hao Yan (Arizona State University, USA), Jing Li (Georgia Institute of Technology, USA), and Frank Setzer (University of Pennsylvania, USA)</i>	
Tensor Topic Models with Graphs and Applications on Individualized Travel Patterns	2756
<i>Ziyue Li (The Hong Kong University of Science and Technology, Hong Kong), Fugee Tsung (The Hong Kong University of Science and Technology, Hong Kong), Hao Yan (Arizona State University, U.S.A), and Chen Zhang (Tsinghua University, China)</i>	
BERT-Based Dynamic Clustering of Subway Stations Using Flow Information	2762
<i>Man Li (Hong Kong University of Science and Technology, China)</i>	
Author Index	2767