# **2024 IEEE 40th International Conference on Data Engineering (ICDE 2024)**

Utrecht, Netherlands 13-17 May 2024

Pages 1-718



IEEE Catalog Number: ISBN:

CFP24026-POD 979-8-3503-1716-9

## Copyright © 2024 by the Institute of Electrical and Electronics Engineers, Inc. All Rights Reserved

*Copyright and Reprint Permissions*: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

#### \*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

IEEE Catalog Number:	CFP24026-POD
ISBN (Print-On-Demand):	979-8-3503-1716-9
ISBN (Online):	979-8-3503-1715-2
ISSN:	1063-6382

#### Additional Copies of This Publication Are Available From:

Curran Associates, Inc 57 Morehouse Lane Red Hook, NY 12571 USA Phone: (845) 758-0400 Fax: (845) 758-2633 E-mail: curran@proceedings.com Web: www.proceedings.com



## 2024 IEEE 40th International Conference on Data Engineering (ICDE) ICDE 2024

### **Table of Contents**

Welcome Message from the ICDE 2024 Chairs	lxix
Organizing Committee	. lxxi
Program Committees	xxiii

### **Research Papers**

DeepMapping: Learned Data Mapping for Lossless Compression and Efficient Lookup
<ul> <li>PURPLE: Making a Large Language Model a Better SQL Writer</li></ul>
LeaderKV: Improving Read Performance of KV Stores via Learned Index and Decoupled KV Table. 29 Yi Wang (Shenzhen University), Jianan Yuan (Shenzhen University), Shangyu Wu (City University of Hong Kong), Huan Liu (Shenzhen University), Jiaxian Chen (Shenzhen University), Chenlin Ma (Shenzhen University), and Jianbin Qin (Shenzhen University)
<ul> <li>TRAP: Tailored Robustness Assessment for Index Advisors via Adversarial Perturbation</li></ul>
<ul> <li>Duet: Efficient and Scalable Hybrid neUral rElation Understanding</li></ul>

<ul> <li>Knowledge Graph Enhanced Multimodal Transformer for Image-Text Retrieval</li></ul>
<ul> <li>Functionality-Aware Database Tuning via Multi-task Learning</li></ul>
COSTREAM: Learned Cost Models for Operator Placement in Edge-Cloud Environments
SiloFuse: Cross-Silo Synthetic Data Generation with Latent Tabular Diffusion Models
<ul> <li>Explainable Database Management System Configuration Tuning Through Counterfactuals</li></ul>
In Situ Neural Relational Schema Matcher
Towards Exploratory Query Optimization for Template-Based SQL Workloads
SPES: Towards Optimizing Performance-Resource Trade-Off for Serverless Functions

<ul> <li>KGLiDS: A Platform for Semantic Abstraction, Linking, and Automation of Data Science</li></ul>	179
Efficiently Estimating Mutual Information Between Attributes Across Tables Aécio Santos (New York University), Flip Korn (Google Research), and Juliana Freire (New York University)	193
Effective Entry-Wise Flow for Molecule Generation	207
<ul> <li>HYPPO: Using Equivalences to Optimize Pipelines in Exploratory Machine Learning</li></ul>	221
HITSNDIFFS: From Truth Discovery to Ability Discovery by Recovering Matrices with the Consecutive Ones Property	235
<ul> <li>Cross-Domain-Aware Worker Selection with Training for Crowdsourced Annotation</li></ul>	249
<ul> <li>Graph Contrastive Learning for Truth Inference</li></ul>	263

<ul> <li>Task Recommendation in Spatial Crowdsourcing: A Trade-off Between Diversity and Coverage 276 Liwei Deng (University of Electronic Science and Technology of China, China), Yan Zhao (Aalborg University, Denmark), Yue Cui (The Hong Kong University of Science and Technology, China), Yuyang Xia (University of Electronic Science and Technology of China, China), Jin Chen (University of Electronic Science and Technology of China, China), and Kai Zheng (University of Electronic Science and Technology of China, China)</li> </ul>
<ul> <li>MACRO: Incentivizing Multi-Leader Game-Based Pareto-Efficient Crowdsourcing for Video</li> <li>Analytics</li></ul>
Cooperative Global Path Planning for Multiple Platforms
Cross Online Assignment of Hybrid Task in Spatial Crowdsourcing
RA^3: A Human-in-the-Loop Framework for Interpreting and Improving Image Captioning with Relation-Aware Attribution Analysis
Efficient Example-Guided Interactive Graph Search
<ul> <li>Wait to be Faster: A Smart Pooling Framework for Dynamic Ridesharing</li></ul>
Adaptive Recursive Query Optimization368Anna Herlihy (EPFL, Switzerland), Guillaume Martres (EPFL, Switzerland), Anastasia Ailamaki (EPFL, Switzerland), and Martin Odersky (EPFL, Switzerland)

Ontology-Mediated Query Answering Using Graph Patterns with Conditions	<u>)</u>
<ul> <li>An Efficient Algorithm for Continuous Complex Event Matching Using Bit-Parallelism</li></ul>	;
<ul> <li>Personalized PageRanks over Dynamic Graphs – The Case for Optimizing Quality of Service 409</li> <li>Zulun Zhu (Nanyang Technological University, Singapore), Siqiang Luo</li> <li>(Nanyang Technological University, Singapore), Wenqing Lin (Tencent,</li> <li>China), Sibo Wang (The Chinese University of Hong Kong, China),</li> <li>Dingheng Mo (Nanyang Technological University, Singapore), and Chunbo</li> <li>Li (Nanyang Technological University, Singapore)</li> </ul>	)
PyTond: Efficient Python Data Science on the Shoulders of Databases	}
Efficient Fault Tolerance for Pipelined Query Engines via Write-Ahead Lineage	;
Independent Range Sampling on Interval Data	)
Incremental Fusion: Unifying Compiled and Vectorized Query Execution	2
The Indistinguishability Query       475         Ashwin Lall (Denison University, USA)	5
Range Cache: An Efficient Cache Component for Accelerating Range Queries on LSM-Based         Key-Value Stores       488         Xiaoliang Wang (University of Science and Technology of China, China),         Peiquan Jin (University of Science and Technology of China, China),         Yongping Luo (University of Science and Technology of China, China),         and Zhaole Chu (University of Science and Technology of China, China)	3
Optimizing Context-Enhanced Relational Joins	L
IndeXY: A Framework for Constructing Indexes Larger than Memory	;

Are ID Embeddings Necessary? Whitening Pre-Trained Text Embeddings for Effective Sequential Recommendation	0
Lingzi Zhang (Nanyang Technological University, Singapore), Xin Zhou (Nanyang Technological University, Singapore), Zhiwei Zeng (Nanyang Technological University, Singapore), and Zhiqi Shen (Nanyang Technological University, Singapore)	U
<ul> <li>Structure- and Logic-Aware Heterogeneous Graph Learning for Recommendation</li></ul>	4
Graph Augmentation for Recommendation	7
<ul> <li>From Chaos to Clarity: Time Series Anomaly Detection in Astronomical Observations</li></ul>	0
<ul> <li>Enhancing Topic Interpretability for Neural Topic Modeling Through Topic-Wise Contrastive</li> <li>Learning</li></ul>	4
<ul> <li>Improve ROI with Causal Learning and Conformal Prediction</li></ul>	8
<ul> <li>Hide Your Model: A Parameter Transmission-free Federated Recommender System</li></ul>	1

<ul> <li>TimeDRL: Disentangled Representation Learning for Multivariate Time-Series</li></ul>
Boosting Meaningful Dependency Mining with Clustering and Covariance Analysis
Uncovering the Propensity Identification Problem in Debiased Recommendations
<ul> <li>Scaling up Multivariate Time Series Pre-Training with Decoupled Spatial-Temporal</li> <li>Representations</li></ul>
<ul> <li>Meta-Optimized Structural and Semantic Contrastive Learning for Graph Collaborative</li> <li>Filtering</li></ul>
Efficient Set-Based Order Dependency Discovery with a Level-Wise Hybrid Strategy
<ul> <li>Meta-Optimized Joint Generative and Contrastive Learning for Sequential Recommendation705 Yongjing Hao (Soochow University, China), Pengpeng Zhao (Soochow University, China), Junhua Fang (Soochow University, China), Jianfeng Qu (Soochow University, China), Guanfeng Liu (Macquarie University, Australia), Fuzhen Zhuang (Beihang University, China), Victor S. Sheng (Texas Tech University, USA), and Xiaofang Zhou (The Hong Kong University of Science and Technology, China)</li> </ul>

Multi-modal Siamese Network for Few-Shot Knowledge Graph Completion
Local-Global History-Aware Contrastive Learning for Temporal Knowledge Graph Reasoning 733 Wei Chen (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), Huaiyu Wan (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), Yuting Wu (Beijing Jiaotong University, China), Shuyuan Zhao (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China), Jiayaqi Cheng (Beijing Jiaotong University, China), Yuxin Li (Beijing Jiaotong University, China), and Youfang Lin (Beijing Jiaotong University, China; Beijing Key Laboratory of Traffic Data Analysis and Mining, China)
Learning Multi-pattern Normalities in the Frequency Domain for Efficient Time Series Anomaly Detection
<ul> <li>Modeling User Attention in Music Recommendation</li></ul>
A Robust Prioritized Anomaly Detection when Not All Anomalies are of Primary Interest
<ul> <li>Enhancing Quantitative Reasoning Skills of Large Language Models Through Dimension</li> <li>Perception</li></ul>
SSDRec: Self-Augmented Sequence Denoising for Sequential Recommendation

<ul> <li>BSL: Understanding and Improving Softmax Loss for Recommendation</li></ul>	316
<ul> <li>Online Detection of Outstanding Quantiles with QuantileFilter</li></ul>	331
<ul> <li>When Multi-behavior Meets Multi-interest: Multi-behavior Sequential Recommendation with</li> <li>Multi-interest Self-Supervised Learning</li> <li>Binquan Wu (South China University of Technology, China), Yu Cheng</li> <li>(South China University of Technology, China), Haitao Yuan (South</li> <li>China University of Technology, China), and Qianli Ma (South China</li> <li>University of Technology, China)</li> </ul>	345
E2GCL: Efficient and Expressive Contrastive Learning on Graph Neural Networks	359
Ambiguous Entity Oriented Targeted Document Detection	374
TS3Net: Triple Decomposition with Spectrum Gradient for Long-Term Time Series Analysis 8 Xiangkai Ma (Nanjing University, China), Xiaobin Hong (Nanjing University, China), Sanglu Lu (Nanjing University, China), and Wenzhong Li (Nanjing University, China)	387
An Efficient Fuzzy Stream Clustering Method Based on Granular-Ball Structure	901

<ul> <li>W-GBC: An Adaptive Weighted Clustering Method Based on Granular-Ball Structure</li></ul>
RobFL: Robust Federated Learning via Feature Center Separation and Malicious Center         Detection       926         Ting Zhou (Shandong University, China), Ning Liu (Shandong University,         China), Bo Song (Shandong University, China), Hongtao Lv (Shandong         University, China), Deke Guo (National University of Defense         Technology, China), and Lei Liu (Shandong University, China; Shandong         Research Institute of Industrial Technology, China)
Towards Task-Conflicts Momentum-Calibrated Approach for Multi-Task Learning
Hybrid Evaluation for Occlusion-Based Explanations on CNN Inference Queries
Enhancing the Performance of Bandit-Based Hyperparameter Optimization
Unraveling the 'Anomaly' in Time Series Anomaly Detection: A Self-Supervised Tri-Domain Solution
A Robust Low-Rank Tensor Decomposition and Quantization Based Compression Method

A Coarse-to-Fine Framework for Entity-Relation Joint Extraction
KGLink: A Column Type Annotation Method that Combines Knowledge Graph and pre-Trained         Language Model       1023         Yubo Wang (HKUST, China), Hao Xin (HKUST, China), and Lei Chen       1023         (HKUST(GZ) & HKUST, China)       1023
Learning k-Determinantal Point Processes for Personalized Ranking
A Unified Replay-Based Continuous Learning Framework for Spatio-Temporal Prediction on Streaming Data
Representation Learning of Tangled Key-Value Sequence Data for Early Classification
A Two Phase Recall-and-Select Framework for Fast Model Selection
BTS: Load-Balanced Distributed Union-Find for Finding Connected Components with Balanced Tree Structures
Interpretable Knowledge Tracing via Response Influence-Based Counterfactual Reasoning 1103 Jiajun Cui (East China Normal University, China), Minghe Yu (Northeastern University, China), Bo Jiang (East China Normal University, China), Aimin Zhou (East China Normal University, China), Jianyong Wang (Tsinghua University, China), and Wei Zhang (East China Normal University, China)

<ul> <li>Stable Heterogeneous Treatment Effect Estimation Across Out-of-Distribution Populations 1117</li> <li>Yuling Zhang (Tsinghua University), Anpeng Wu (Zhejiang University;</li> <li>Mohamed bin Zayed University of Artificial Intelligence), Kun Kuang</li> <li>(Zhejiang University), Liang Du (Interactive Entertainment Group,</li> <li>Tencent), Zixun Sun (Interactive Entertainment Group, Tencent), and</li> <li>Zhi Wang (Tsinghua University)</li> </ul>
Towards Cross-Domain Continual Learning1131Marcus de Carvalho (Nanyang Technological University, Singapore),1131Mahardhika Pratama (University of South Australia, Australia), Zhang116Jie (Nanyang Technological University, Singapore), Haoyan Chua117(Nanyang Technological University, Singapore), Haoyan Chua117(Singapore Institute of Manufacturing Technology (SIMTech) Agency for117Science, Technology and Research (A*STAR), Singapore)117
DROPP: Structure-Aware PCA for Ordered Data: A General Method and Its Applications in Climate Research and Molecular Dynamics
<ul> <li>Scalable Overspeed Item Detection in Streams</li></ul>
<ul> <li>GradGCL: Gradient Graph Contrastive Learning</li></ul>
<ul> <li>ST-ABC: Spatio-Temporal Attention-Based Convolutional Network for Multi-scale Lane-Level</li> <li>Traffic Prediction</li></ul>
CPDG: A Contrastive Pre-Training Method for Dynamic Graph Neural Networks

Graph Anomaly Detection at Group Level: A Topology Pattern Enhanced Unsupervised Approach .... 1213

Xing Ai (The Hong Kong Polytechnic University, HKSAR), Jialong Zhou (The Hong Kong Polytechnic University, HKSAR), Yulin Zhu (The Hong Kong Polytechnic University, HKSAR), Gaolei Li (Shanghai Jiao Tong University, China), Tomasz P. Michalak (University of Warsaw, Poland), Xiapu Luo (The Hong Kong Polytechnic University, HKSAR), and Kai Zhou (The Hong Kong Polytechnic University, HKSAR)	
<ul> <li>Temporal-Frequency Masked Autoencoders for Time Series Anomaly Detection</li></ul>	8
REGER: Reordering Time Series Data for Regression Encoding	2
<ul> <li>SAGDFN: A Scalable Adaptive Graph Diffusion Forecasting Network for Multivariate Time</li> <li>Series Forecasting</li></ul>	5
Knowledge-Enhanced Recommendation with User-Centric Subgraph Network	9
<ul> <li>MUSE-Net: Disentangling Multi-Periodicity for Traffic Flow Forecasting</li></ul>	2
Model Selection with Model Zoo via Graph Learning	6

Logical Relation Modeling and Mining in Hyperbolic Space for Recommendation
<ul> <li>HeteFedRec: Federated Recommender Systems with Model Heterogeneity</li></ul>
A Compact and Accurate Sketch for Estimating a Large Range of Set Difference Cardinalities 1338 Peng Jia (Xi'an Jiaotong University, China), Pinghui Wang (Xi'an Jiaotong University, China), Rundong Li (Xi'an Jiaotong University, China), Junzhou Zhao (Xi'an Jiaotong University, China), Junlan Feng (China Mobile Group Design Institute, China), Xidian Wang (China Mobile Research, China), and Xiaohong Guan (Xi'an Jiaotong University, China; Tsinghua University, China)
A Unified Model for Spatio-Temporal Prediction Queries with Arbitrary Modifiable Areal Units
Across Images and Graphs for Question Answering
LightLT: A Lightweight Representation Quantization Framework for Long-Tail Data
TSec: An Efficient and Effective Framework for Time Series Classification

<ul> <li>McCatch: Scalable Microcluster Detection in Dimensional and Nondimensional Datasets</li></ul>
Contrastive Learning for Fraud Detection from Noisy Labels
Adapting Large Language Models by Integrating Collaborative Semantics for Recommendation . 1435 Bowen Zheng (Renmin University of China, China), Yupeng Hou (University of California San Diego, United States), Hongyu Lu (Tencent, China), Yu Chen (Tencent, China), Wayne Xin Zhao (Renmin University of China, China), Ming Chen (Tencent, China), and Ji-Rong Wen (Renmin University of China, China)
Effective Data Selection and Replay for Unsupervised Continual Learning
<ul> <li>Target-Agnostic Source-free Domain Adaptation for Regression Tasks</li></ul>
<ul> <li>Fast Parallel Recovery for Transactional Stream Processing on Multicores</li></ul>
<ul> <li>PP-Stream: Toward High-Performance Privacy-Preserving Neural Network Inference via</li> <li>Distributed Stream Processing</li></ul>
AdaEdge: A Dynamic Compression Selection Framework for Resource Constrained Devices 1506 Chunwei Liu (Massachusetts Institute of Technology), John Paparrizos

(The Ohio State University), and Aaron Elmore (University of Chicago)

A Predictive Profiling and Performance Modeling Approach for Distributed Stream Processing in Edge
Hasan Geren (RMIT University, Australia), Nasrin Sohrabi (Deakin University, Australia), Zahir Tari (RMIT University, Australia), and Nour Moustafa (UNSW/ADFA, Australia)
Joint Mobile Edge Caching and Pricing: A Mean-Field Game Approach
<ul> <li>Online Container Caching with Late-Warm for IoT Data Processing</li></ul>
COUPLE: Orchestrating Video Analytics on Heterogeneous Mobile Processors
Multiple Continuous Top-K Queries Over Data Stream
CodingSketch: A Hierarchical Sketch with Efficient Encoding and Recursive Decoding
Everything Everyway All at Once - Time Traveling Debugging for Stream Processing         Applications       1606         Timo Räth (TU Ilmenau, Germany), Marius Schlegel (TU Ilmenau,       1606         Germany), and Kai-Uwe Sattler (TU Ilmenau, Germany)       1606
LDPRecover: Recovering Frequencies from Poisoning Attacks Against Local Differential Privacy
Differentially Private Graph Neural Networks for Link Prediction

Secure and Practical Functional Dependency Discovery in Outsourced Databases
SecMdp: Towards Privacy-Preserving Multimodal Deep Learning in End-Edge-Cloud
CARGO: Crypto-Assisted Differentially Private Triangle Counting Without Trusted Servers 1671 Shang Liu (Kyoto University), Yang Cao (Hokkaido University), Takao Murakami (Institute of Statistical Mathematics), Jinfei Liu (Zhejiang University), and Masatoshi Yoshikawa (Osaka Seikei University)
<ul> <li>Real-Time Trajectory Synthesis with Local Differential Privacy</li></ul>
<ul> <li>Privacy-Preserving Traffic Flow Release with Consistency Constraints</li></ul>
Unraveling Privacy Risks of Individual Fairness in Graph Neural Networks
Sketches-Based Join Size Estimation Under Local Differential Privacy
PrivShape: Extracting Shapes in Time Series Under User-Level Local Differential Privacy 1739 Yulian Mao (Southern University of Science and Technology; The Hong Kong Polytechnic University), Qingqing Ye (The Hong Kong Polytechnic University), Haibo Hu (The Hong Kong Polytechnic University), Qi Wang (Southern University of Science and Technology), and Kai Huang (Macau University of Science and Technology)
<ul> <li>SparDL: Distributed Deep Learning Training with Efficient Sparse Communication</li></ul>

MetaSQL: A Generate-Then-Rank Framework for Natural Language to SQL Translation	1765
<ul> <li>Feed: Towards Personalization-Effective Federated Learning</li></ul>	1779
T-Rex (Tree-Rectangles): Reformulating Decision Tree Traversal as Hyperrectangle Enclosure1 Meghana Madhyastha (Johns Hopkins University, USA), Tamas Budavari (Johns Hopkins University, USA), Vladmir Braverman (Rice University, USA), Joshua Vogelstein (Johns Hopkins University, USA), and Randal Burns (Johns Hopkins University, USA)	1792
<ul> <li>FeatAug: Automatic Feature Augmentation From One-to-Many Relationship Tables</li></ul>	1805
AutoMC: Automated Model Compression Based on Domain Knowledge and Progressive Search 1 Chunnan Wang (Harbin Institute of Technology, China), Hongzhi Wang (Harbin Institute of Technology, China), and Xiangyu Shi (Harbin Institute of Technology, China)	1819
<ul> <li>Task-Oriented GNNs Training on Large Knowledge Graphs for Accurate and Efficient Modeling</li> <li>1833         <ul> <li>Hussein Abdallah (Concordia University), Waleed Afandi (Concordia University), Panos Kalnis (KAUST), and Essam Mansour (Concordia University)</li> </ul> </li> </ul>	
Clients Help Clients: Alternating Collaboration for Semi-Supervised Federated Learning	1847
<ul> <li>AutoFeat: Transitive Feature Discovery over Join Paths</li></ul>	1861
<ul> <li>Triple-d: Denoising Distant Supervision for High-Quality Data Creation</li></ul>	1874

Efficient Partial Order Based Transaction Processing for Permissioned Blockchains Shuai Zhao (Beijing Institute of Technology, China), Zhiwei Zhang (Beijing Institute of Technology, China), Junkai Wang (Beijing Institute of Technology, China), Ye Yuan (Beijing Institute of Technology, China), Meihui Zhang (Beijing Institute of Technology, China), Meihui Zhang (Beijing Institute of Technology, China), Guoren Wang (Beijing Institute of Technology, China), and Jiang Xiao (Huazhong University of Science and Technology, China)	1888
TELL: Efficient Transaction Execution Protocol Towards Leaderless Consensus Xing Tong (East China Normal University, China), Zheming Ye (East China Normal University, China), Zhao Zhang (East China Normal University, China), Cheqing Jin (East China Normal University, China), and Aoying Zhou (East China Normal University, China)	1902
SpotLess: Concurrent Rotational Consensus Made Practical Through Rapid View Synchronization Dakai Kang (University of California, Davis), Sajjad Rahnama (University of California, Davis), Jelle Hellings (McMaster University), and Mohammad Sadoghi (University of California, Davis)	1916
PrestigeBFT: Revolutionizing View Changes in BFT Consensus Algorithms with Reputation Mechanisms Gengrui Zhang (University of Toronto), Fei Pan (University of Toronto), Sofia Tijanic (University of Toronto), and Hans-Arno Jacobsen (University of Toronto)	1930
<ul> <li>Porygon: Scaling Blockchain via 3D Parallelism</li></ul>	1944
Authenticated Keyword Search on Large-Scale Graphs in Hybrid-Storage Blockchains Siyu Li (Beijing Institute of Technology, China), Zhiwei Zhang (Beijing Institute of Technology, China), Jiang Xiao (Huazhong University of Science and Technology, China), Meihui Zhang (Beijing Institute of Science and Technology, China), Meihui Zhang (Beijing Institute of Technology, China), Ye Yuan (Beijing Institute of Technology, China), and Guoren Wang (Beijing Institute of Technology, China)	1958
<ul> <li>MuFuzz: Sequence-Aware Mutation and Seed Mask Guidance for Blockchain Smart Contract</li> <li>Fuzzing</li> <li>Peng Qian (Zhejiang University, China; Goplus Security, China), Hanjie</li> <li>Wu (Zhejiang Gongshang University, China), Zeren Du (Zhejiang</li> <li>Gongshang University, China), Turan Vural (University of California,</li> <li>USA), Dazhong Rong (Zhejiang University, China), Zheng Cao (Zhejiang</li> <li>University, China; Goplus Security, China), Lun Zhang (Goplus</li> <li>Security, China), Yanbin Wang (Zhejiang University, China), Jianhai</li> <li>Chen (Zhejiang University, China), and Qinming He (Zhejiang</li> <li>University, China)</li> </ul>	1972

Authenticated Subgraph Matching in Hybrid-Storage Blockchains
<ul> <li>V^2FS: A Verifiable Virtual Filesystem for Multi-chain Query Authentication</li></ul>
Lion: Minimizing Distributed Transactions through Adaptive Replica Provision
<ul> <li>FC: Adaptive Atomic Commit via Failure Detection</li></ul>
ZeroTune: Learned Zero-Shot Cost Models for Parallelism Tuning in Stream Processing
MergeSFL: Split Federated Learning with Feature Merging and Batch Size Regulation

<ul> <li>SharDAG: Scaling DAG-Based Blockchains via Adaptive Sharding</li></ul>
<ul> <li>Boosting Write Performance of KV Stores: An NVM-Enabled Storage Collaboration Approach 2082</li> <li>Yi Wang (Shenzhen University, China), Jiajian He (Shenzhen University, China), Kaoyi Sun (Shenzhen University, China), Yunhao Dong (Shenzhen University, China), Jiaxian Chen (Shenzhen University, China), Chenlin Ma (Shenzhen University, China), Amelie Chi Zhou (Hong Kong Baptist University, China), and Rui Mao (Shenzhen University, China)</li> </ul>
Log Replaying for Real-Time HTAP: An Adaptive Epoch-Based Two-Stage Framework
FSD-Inference: Fully Serverless Distributed Inference with Scalable Cloud Communication
Graph Computation with Adaptive Granularity

<ul> <li>FedCross: Towards Accurate Federated Learning via Multi-Model Cross-Aggregation</li></ul>
Mitigating Subgroup Unfairness in Machine Learning Classifiers: A Data-Driven Approach 2151 Yin Lin (University of Michigan), Samika Gupta (University of Michigan), and H. V. Jagadish (University of Michigan)
Non-Invasive Fairness in Learning Through the Lens of Data Drift
Preventing the Popular Item Embedding Based Attack in Federated Recommendations
Explainable Disparity Compensation for Efficient Fair Ranking
Generating Explanations to Understand and Repair Embedding-Based Entity Alignment
Enhancing the Rationale-Input Alignment for Self-Explaining Rationalization

Model Trip: Enhancing Privacy and Fairness in Model Fusion Across Multi-Federations for Trustworthy Global Healthcare
Qian Chen (Institute of Computing Technology, Chinese Academy of
Sciences; University of Chinese Academy of Sciences; Beijing Key
Laboratory of Mobile Computing and Pervasive Device), Yiqiang Chen
(Institute of Computing Technology, Chinese Academy of Sciences;
University of Chinese Academy of Sciences; Beijing Key Laboratory of
Mobile Computing and Pervasive Device; Peng Cheng Laboratory), Bingjie
Yan (Institute of Computing Technology, Chinese Academy of Sciences;
University of Chinese Academy of Sciences; Beijing Key Laboratory of
Mobile Computing and Pervasive Device), Xinlong Jiang (Institute of
Computing Technology, Chinese Academy of Sciences; University of
Chinese Academy of Sciences; Beijing Key Laboratory of Mobile
Computing and Pervasive Device), Xiaojin Zhang (Huazhong University of
Science and Technology), Yan Kang (WeBank), Teng Zhang (Institute of
Computing Technology, Chinese Academy of Sciences; University of
Chinese Academy of Sciences; Beijing Key Laboratory of Mobile
Computing and Pervasive Device), Wuliang Huang (Institute of Computing
Technology, Chinese Academy of Sciences; University of Chinese Academy
of Sciences; Beijing Key Laboratory of Mobile Computing and Pervasive
Device), Chenlong Gao (Institute of Computing Technology, Chinese
Academy of Sciences; University of Chinese Academy of Sciences;
Beijing Key Laboratory of Mobile Computing and Pervasive Device), Livin Fan (MaBank) and Oigna Yang (MaBank: Hong Kong University of
Lixin Fan (WeBank), and Qiang Yang (WeBank; Hong Kong University of Science and Technology)
Why-Not Explainable Graph Recommender
Hervé-Madelein Attolou (ETIS, CY Cergy Paris University, France), Kataving Transmuchi (ETIS, CY Cargo Paris University, France), Kastas
Katerina Tzompanaki (ETIS, CY Cergy Paris University, France), Kostas
Stefanidis (Tampere University, Finland), and Dimitris Kotzinos (ETIS, CY Cergy Paris University, France)
GAGE: Genetic Algorithm-Based Graph Explainer for Malware Analysis
Mohd Saqib (McGill University, Canada), Benjamin C. M. Fung (McGill
University, Canada), Philippe Charland (Defence R&D, Canada), and
Andrew Walenstein (BlackBerry Limited, Canada)
Accurate Explanation Model for Image Classifiers Using Class Association Embedding
Ruitao Xie (Shenzhen Institute of Advanced Technology, Chinese Academy
of Sciences, China; University of Chinese Academy of Sciences, China),
Jingbang Chen (Shenzhen Institute of Advanced Technology, Chinese
Academy of Sciences, China), Limai Jiang (Shenzhen Institute of
Advanced Technology, Chinese Academy of Sciences, China; University of
Chinese Academy of Sciences, China), Rui Xiao (Shenzhen Institute of
Advanced Technology, Chinese Academy of Sciences, China), Yi Pan
(Shenzhen Institute of Advanced Technology, Chinese Academy of
Sciences, China; Shenzhen Key Laboratory of Intelligent
Bioinformatics, China), and Yunpeng Cai (Shenzhen Institute of

Advanced Technology, Chinese Academy of Sciences, China)

<ul> <li>FairGen: Towards Fair Graph Generation</li></ul>	
Fast, Robust and Interpretable Participant Contribution Estimation for Federated Learning 229 Yong Wang (Tsinghua University), Kaiyu Li (Tsinghua University), Yuyu Luo (HKUST (GZ)), Guoliang Li (Tsinghua University), Yunyan Guo (Tsinghua University), and Zhuo Wang (Tsinghua University)	:98
<ul> <li>Exploring Optimal Parameters for Expected Results on Radius-Bounded k-Core Queries</li></ul>	12
<ul> <li>Explaining Entity Matching with Clusters of Words</li></ul>	25
<ul> <li>Fair Top-k Query on Alpha-Fairness</li></ul>	38
<ul> <li>Butterfly Counting over Bipartite Graphs with Local Differential Privacy</li></ul>	51
<ul> <li>Temporal Graph Generation Featuring Time-Bound Communities</li></ul>	65
Breaking the Entanglement of Homophily and Heterophily in Semi-Supervised Node Classification	379
Efficient Core Decomposition over Large Heterogeneous Information Networks	93

Accelerating SpMV for Scale-Free Graphs with Optimized Bins YuAng Chen (The Chinese University of Hong Kong) and Jeffrey Xu Yu (The Chinese University of Hong Kong)	2407
PlatoD2GL: An Efficient Dynamic Deep Graph Learning System for Graph Neural Network Training on Billion-Scale Graphs Xing Huang (Tencent Inc.), Dandan Lin (Shenzhen Institute of Computing Sciences), Weiyi Huang (Tencent Inc.), Shijie Sun (Tencent Inc.), Jie Wen (Tencent Inc.), and Chuan Chen (Tencent Inc.)	2421
Quantum Algorithms for the Maximum K-Plex Problem Xiaofan Li (Nanyang Technological University, Singapore), Gao Cong (Nanyang Technological University, Singapore), and Rui Zhou (Swinburne University of Technology, Australia)	2435
<ul> <li>Fast Iterative Graph Computing with Updated Neighbor States</li></ul>	2449
Querying Numeric-Constrained Shortest Distances on Road Networks Mingyu Yang (The Hong Kong University of Science and Technology (Guangzhou)), Wentao Li (The Hong Kong University of Science and Technology (Guangzhou)), Wei Wang (The Hong Kong University of Science and Technology (Guangzhou)), Dong Wen (The University of New South Wales), and Lu Qin (AAII, FEIT, University of Technology Sydney)	2463
<ul> <li>Mining Quasi-Periodic Communities in Temporal Network</li></ul>	2476
<ul> <li>GraphRARE: Reinforcement Learning Enhanced Graph Neural Network with Relative Entropy Tianhao Peng (Beihang University, China), Wenjun Wu (Beihang University, China), Haitao Yuan (Nanyang Technological University, Singapore), Zhifeng Bao (RMIT University, Australia), Zhao Pengrui (Beihang University, China), Xin Yu (Beihang University, China), Xuetao Lin (Beihang University, China), Yu Liang (Beijing University of Technology, China), and Yanjun Pu (Beijing University of Technology, China)</li> </ul>	2489
Querying Historical Cohesive Subgraphs over Temporal Bipartite Graphs Shunyang Li (University of New South Wales), Kai Wang (Antai College of Economics and Management, Shanghai Jiao Tong University), Xuemin Lin (Antai College of Economics and Management, Shanghai Jiao Tong University), Wenjie Zhang (University of New South Wales), Yizhang He (University of New South Wales), and Long Yuan (Nanjing University of Science and Technology)	2503

AdaFGL: A New Paradigm for Federated Node Classification with Topology Heterogeneity 2517 Xunkai Li (Beijing Institute of Technology, China), Zhengyu Wu (Beijing Institute of Technology, China), Wentao Zhang (Peking University; National Engineering Labratory for Big Data Analytics and Applications, China), Henan Sun (Beijing Institute of Technology, China), Rong-Hua Li (Beijing Institute of Technology, China), and Guoren Wang (Beijing Institute of Technology, China)
<ul> <li>Positive Communities on Signed Graphs That Are Not Echo Chambers: A Clique-Based Approach</li> <li>2531 <ul> <li>Alexander Zhou (Hong Kong University of Science and Technology), Yue</li> <li>Wang (Shenzhen Institute of Computing Sciences), Lei Chen (Hong Kong</li> <li>University of Science and Technology (Guangzhou)), and M. Tamer Özsu</li> <li>(University of Waterloo)</li> </ul> </li> </ul>
<ul> <li>Maximal Biclique Enumeration: A Prefix Tree Based Approach</li></ul>
Batch Hop-Constrained s-t Simple Path Query Processing in Large Graphs
<ul> <li>On Searching Maximum Directed (k, ℓ)-Plex</li></ul>
Masked Graph Modeling with Multi-view Contrast
Multi-view Teacher with Curriculum Data Fusion for Robust Unsupervised Domain Adaptation .2598 Yuhao Tang (Peking University, China), Junyu Luo (Peking University, China), Ling Yang (Peking University, China), Xiao Luo (University of California Los Angeles), Wentao Zhang (Peking University, China), and Bin Cui (Peking University, China)
GSHOP: Towards Flexible Pricing for Graph Statistics

LearnSC: An Efficient and Unified Learning-Based Framework for Subgraph Counting Problem . 2625 Wenzhe Hou (National University of Defense Technology, China), Xiang Zhao (National University of Defense Technology, China), and Bo Tang (Southern University of Science and Technology, China)
AFTER: Adaptive Friend Discovery for Temporal-Spatial and Social-Aware XR
Reducing Resource Usage for Continuous Model Updating and Predictive Query Answering in       2653         Graph Streams       2653         Qu Liu (University of Massachusetts, Lowell, USA), Adam King       2653         (University of Massachusetts, Lowell, USA), and Tingjian Ge       2000         (University of Massachusetts, Lowell, USA)       2000
Graph Anomaly Detection with Domain-Agnostic Pre-Training and Few-Shot Adaptation
NC-ALG: Graph-Based Active Learning Under Noisy Crowd
Fast Multilayer Core Decomposition and Indexing2695Dandan Liu (Harbin Institute of Technology, China), Run-An Wang(Harbin Institute of Technology, China), Zhaonian Zou (HarbinInstitute of Technology, China), and Xin Huang (Hong Kong BaptistUniversity, China)
CINA: Curvature-Based Integrated Network Alignment with Hypergraph
Open-World Semi-Supervised Learning for Node Classification
<ul> <li>Scalable Community Search with Accuracy Guarantee on Attributed Graphs</li></ul>

From Motif to Path: Connectivity and Homophily Qihao Wang (University of Illinois Urbana-Champaign, USA), Hongtai Cao (University of Illinois Urbana-Champaign, USA), Xiaodong Li (The University of Hong Kong, China), Kevin Chen-Chuan Chang (University of Illinois Urbana-Champaign, USA), and Reynold Cheng (The University of Hong Kong, China)	2751
<ul> <li>Self-Training GNN-Based Community Search in Large Attributed Heterogeneous Information Networks</li> <li>Yuan Li (North China University of Technology, China), Xiuxu Chen (North China University of Technology, China), Yuhai Zhao (Northeastern University, China), Wen Shan (Singapore University of Social Sciences, Singapore), Zhengkui Wang (Singapore Institute of Technology, Singapore), Guoli Yang (Advanced Institute of Big Data, China), and Guoren Wang (Beijing Institute of Technology, China)</li> </ul>	.2765
HGAMLP: Heterogeneous Graph Attention MLP with De-Redundancy Mechanism Yuxuan Liang (Peking University), Wentao Zhang (Peking University), Zeang Sheng (Peking University), Ling Yang (Peking University), Jiawei Jiang (Wuhan University), Yunhai Tong (Peking University), and Bin Cui (Peking University)	2779
FocusCore Decomposition of Multilayer Graphs Run-An Wang (Harbin Institute of Technology, China), Dandan Liu (Harbin Institute of Technology, China), and Zhaonian Zou (Harbin Institute of Technology, China)	2792
<ul> <li>Search to Fine-Tune Pre-Trained Graph Neural Networks for Graph-Level Tasks</li></ul>	2805
BOURNE: Bootstrapped Self-Supervised Learning Framework for Unified Graph Anomaly Detection Jie Liu (Northwestern Polytechnical University, China), Mengting He (Northwestern Polytechnical University, China), Xuequn Shang (Northwestern Polytechnical University, China), Jieming Shi (The Hong Kong Polytechnic University, China), Bin Cui (Peking University, China), and Hongzhi Yin (The University of Queensland, Australia)	2820
Discovering Personalized Characteristic Communities in Attributed Graphs Yudong Niu (Singapore Management University, Singapore), Yuchen Li (Singapore Management University, Singapore), Panagiotis Karras (University of Copenhagen, Denmark), Yanhao Wang (East China Normal University, China), and Zhao Li (Zhejiang Lab, China)	2834
TP-GNN: Continuous Dynamic Graph Neural Network for Graph Classification Jie Liu (Southwest University, China), Jiamou Liu (University of Auckland, New Zealand), Kaiqi Zhao (University of Auckland, New Zealand), Yanni Tang (University of Auckland, New Zealand), and Wu Chen (Southwest University, China)	2848

<ul> <li>GraphHI: Boosting Graph Neural Networks for Large-Scale Graphs</li></ul>
DiscoGNN: A Sample-Efficient Framework for Self-Supervised Graph Representation Learning 2876 Jun Xia (Zhejiang University, China), Shaorong Chen (Westlake University), Yue Liu (Westlake University), Zhangyang Gao (Westlake University), Jiangbin Zheng (Westlake University), Xihong Yang (Westlake University), and Stan Z. Li (Westlake University)
Incorporating Dynamic Temperature Estimation into Contrastive Learning on Graphs
Newton Sketches: Estimating Node Intimacy in Dynamic Graphs Using Newton's Law of Cooling 2904
Qizhi Chen (Peking University, China), Ke Wang (Yale University, USA), Aoran Li (Peking University, China), Tong Yang (Peking University, China), and Bin Cui (Peking University, China)
Counting Butterflies in Fully Dynamic Bipartite Graph Streams
BIM: Improving Graph Neural Networks with Balanced Influence Maximization
SES: Bridging the Gap Between Explainability and Prediction of Graph Neural Networks
Efficient Cross-Layer Community Search in Large Multilayer Graphs

Large Subgraph Matching: A Comprehensive and Efficient Approach for Heterogeneous Graphs ...... 2972

Hongtai Cao (University of Illinois at Urbana-Champaign, USA), Qihao Wang (University of Illinois at Urbana-Champaign, USA), Xiaodong Li (University of Hong Kong, Hong Kong SAR), Matin Najafi (University of Hong Kong, Hong Kong SAR), Kevin Chen-Chuan Chang (University of Illinois at Urbana-Champaign, USA), and Reynold Cheng (University of Hong Kong, Hong Kong SAR)
<ul> <li>Adaptive Hypergraph Network for Trust Prediction</li></ul>
Bottom-up k-Vertex Connected Component Enumeration by Multiple Expansion
<ul> <li>Wings: Efficient Online Multiple Graph Pattern Matching</li></ul>
SGCL: Semantic-Aware Graph Contrastive Learning with Lipschitz Graph Augmentation
Accelerating Scalable Graph Neural Network Inference with Node-Adaptive Propagation
<ul> <li>Graph Condensation for Inductive Node Representation Learning</li></ul>

Graphix: "One User's JSON is Another User's Graph"	070
CSM-TopK: Continuous Subgraph Matching with TopK Density Constraints	084
<ul> <li>Efficient Maximal Temporal Plex Enumeration</li></ul>	8098
Denoising High-Order Graph Clustering	3111
A Revisit to Graph Neighborhood Cardinality Estimation	125
Attributed Network Embedding in Streaming Style	3138
<ul> <li>Faster Depth-First Subgraph Matching on GPUs</li></ul>	,151
G^2-AIMD: A Memory-Efficient Subgraph-Centric Framework for Efficient Subgraph Finding on GPUs	3164
GPUs	104
Fine-Grained Anomaly Detection on Dynamic Graphs via Attention Alignment	178

Accelerating Biclique Counting on GPU
<ul> <li>GPU-Accelerated Batch-Dynamic Subgraph Matching</li></ul>
I/O Efficient Max-Truss Computation in Large Static and Dynamic Graphs
Efficient Multi-Query Oriented Continuous Subgraph Matching
Label Constrained Reachability Queries on Time Dependent Graphs
Time-Constrained Continuous Subgraph Matching Using Temporal Information for Filtering and Backtracking
Adaptive Truss Maximization on Large Graphs: A Minimum Cut Approach

<ul> <li>SACH: Significant-Attributed Community Search in Heterogeneous Information Networks 3283 Yanghao Liu (CAS Key Laboratory of AI Safety &amp; Security, Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Fangda Guo (CAS Key Laboratory of AI Safety &amp; Security, Institute of Computing Technology, Chinese Academy of Sciences, China), Bingbing Xu (CAS Key Laboratory of AI Safety &amp; Security, Institute of Computing Technology, Chinese Academy of Sciences, China), Peng Bao (Beijing Jiaotong University, China), Huawei Shen (CAS Key Laboratory of AI Safety &amp; Security, Institute of Computing Technology, Chinese Academy of Sciences, China), and Xueqi Cheng (CAS Key Laboratory of AI Safety &amp; Security, Institute of Computing Technology, Chinese Academy of Sciences, China)</li> </ul>
TimeSGN: Scalable and Effective Temporal Graph Neural Network3297Yuanyuan Xu (University of New South Wales), Wenjie Zhang (University of New South Wales), Ying Zhang (Zhejiang Gongshang University), Maria Orlowska (Polish-Japanese Institute of Information Technology), and Xuemin Lin (Shanghai Jiao Tong University)
Variable-Length Path Query Evaluation Based on Worst-Case Optimal Joins
NewSP: A New Search Process for Continuous Subgraph Matching over Dynamic Graphs
Querying Cohesive Subgraph Regarding Span-Constrained Triangles on Temporal Graphs 3338 Chuhan Hu (Wuhan University, China), Ming Zhong (Wuhan University, China; †Zhejiang Lab, China), Yuanyuan Zhu (Wuhan University, China), Tieyun Qian (Wuhan University, China), Ting Yu (Zhejiang Lab, China), Hongyang Chen (Zhejiang Lab, China), Mengchi Liu (South China Normal University, China), and Jeffrey X. Yu (The Chinese University of Hong Kong, China)
Generating Robust Counterfactual Witnesses for Graph Neural Networks
Generative and Contrastive Paradigms Are Complementary for Graph Self-Supervised Learning 3364 Yuxiang Wang (Wuhan University, China), Xiao Yan (Centre for Perceptual and Interactive Intelligence (CPII)), Chuang Hu (Wuhan University, China), Quanqing Xu (OceanBase), Chuanhui Yang (OceanBase), Fangcheng Fu (Peking University, China), Wentao Zhang (Peking University, China), Hao Wang (Wuhan University, China), Bo Du (Wuhan University, China), and Jiawei Jiang (Wuhan University)

<ul> <li>FedMix: Boosting with Data Mixture for Vertical Federated Learning</li> <li>Yihang Cheng (University of Science and Technology of China, China),</li> <li>Lan Zhang (University of Science and Technology of China, China; Hefei</li> <li>Comprehensive National Science Center, China), Junyang Wang</li> <li>(University of Science and Technology of China, China), Xiaokai Chu</li> <li>(Tencent, China), Dongbo Huang (Tencent, China), and Lan Xu (Tencent,</li> <li>China)</li> </ul>	3379
DMRNet: Effective Network for Accurate Discharge Medication Recommendation Jiyun Shi (Beijing Institute of Technology), Yuqiao Wang (Beijing Institute of Technology), Chi Zhang (Beijing Institute of Technology), Zhaojing Luo (Beijing Institute of Technology), Chengliang Chai (Beijing Institute of Technology), and Meihui Zhang (Beijing Institute of Technology)	3393
BClean: A Bayesian Data Cleaning System Jianbin Qin (Shenzhen University), Sifan Huang (Shenzhen University), Yaoshu Wang (Shenzhen Institute of Computing Science), Jing Zhu (Shenzhen University), Yifan Zhang (Shenzhen University), Yukai Miao (Zhongguancun Laboratory), Rui Mao (Shenzhen University), Makoto Onizuka (Osaka University), and Chuan Xiao (Osaka University, Nagoya University)	.3407
MultiEM: Efficient and Effective Unsupervised Multi-table Entity Matching Xiaocan Zeng (Zhejiang University), Pengfei Wang (Zhejiang University), Yuren Mao (Zhejiang University), Lu Chen (Zhejiang University), Xiaoze Liu (Zhejiang University), and Yunjun Gao (Zhejiang University)	3421
A Critical Re-evaluation of Record Linkage Benchmarks for Learning-Based Matching Algorithms George Papadakis (National & Kapodistrian University of Athens, Greece), Nishadi Kirielle (The Australian National University, Australia), Peter Christen (The Australian National University, Australia), and Themis Palpanas (Universite Paris Cite & French University Institute, France)	3435
Online Query-Based Data Pricing with Time-Discounting Valuations Yicheng Fu (Zhejiang University, China), Xiaoye Miao (Zhejiang University, China), Huanhuan Peng (Zhejiang University, China), Chongning Na (Zhejiang Lab, China), Shuiguang Deng (Zhejiang University, China), and Jianwei Yin (Zhejiang University, China)	3449
Representation Learning for Entity Alignment in Knowledge Graph: A Design Space Exploration Peng Huang (Beijing Institute of Technology, China), Meihui Zhang (Beijing Institute of Technology, China), Ziyue Zhong (Beijing Institute of Technology, China), Chengliang Chai (Beijing Institute of Technology, China), and Ju Fan (Renmin University of China, China)	3462
Fairness-Aware Data Preparation for Entity Matching Nima Shahbazi (University of Illinois Chicago), Jin Wang (Megagon labs), Zhengjie Miao (Simon Fraser University), and Nikita Bhutani (Megagon Labs)	3476

Mitigating Data Sparsity in Integrated Data Through Text Conceptualization	490
<ul> <li>Measuring Approximate Functional Dependencies: A Comparative Study</li></ul>	505
Efficient Relaxed Functional Dependency Discovery with Minimal Set Cover	519
Gen-T: Table Reclamation in Data Lakes	532
Discovering Denial Constraints in Dynamic Datasets	546
Towards Semantic Consistency: Dirichlet Energy Driven Robust Multi-modal Entity Alignment 3 Yuanyi Wang (Beijing University of Posts and Telecommunications, China), Haifeng Sun (Beijing University of Posts and Telecommunications, China), Jiabo Wang (Beijing University of Posts and Telecommunications, China), Jingyu Wang (Beijing University of Posts and Telecommunications, China), Wei Tang (Beijing University of Posts and Telecommunications, China), Qi Qi (Beijing University of Posts and Telecommunications, China), Shaoling Sun (China Mobile (Suzhou) Software Technology Co., Ltd., China), and Jianxin Liao (Beijing University of Posts and Telecommunications, China)	559
Share: Stackelberg-Nash Based Data Markets	573
Interactive Trimming Against Evasive Online Data Manipulation Attacks: A Game-Theoretic         Approach       3         Yue Fu (The Hong Kong Polytechnic University), Qingqing Ye (The Hong         Kong Polytechnic University), Rong Du (The Hong Kong Polytechnic         University), and Haibo Hu (The Hong Kong Polytechnic University)	587

<ul> <li>Label Noise Correction for Federated Learning: A Secure, Efficient and Reliable</li> <li>Realization</li></ul>
Mitigating Data Scarcity in Supervised Machine Learning Through Reinforcement Learning Guided Data Generation
Dual-Teacher De-Biasing Distillation Framework for Multi-domain Fake News Detection
Are There Fundamental Limitations in Supporting Vector Data Management in Relational Databases? A Case Study of PostgreSQL
Compression and In-Situ Query Processing for Fine-Grained Array Lineage
TSDDISCOVER: Discovering Data Dependency for Time Series Data
Time Series Data Cleaning Under Expressive Constraints on Both Rows and Columns
<ul> <li>Cost-Effective In-Context Learning for Entity Resolution: A Design Space Exploration</li></ul>
A Multi-task Learning Framework for Reading Comprehension of Scientific Tabular Data

Enabling Efficient NVM-Based Text Analytics Without Decompression
F-TADOC: FPGA-Based Text Analytics Directly on Compression with HLS
Robust External Hash Aggregation in the Solid State Age
Neos: A NVMe-GPUs Direct Vector Service Buffer in User Space
TEngine: A Native Distributed Table Storage Engine
DmRPC: Disaggregated Memory-Aware Datacenter RPC for Data-Intensive Applications
RapidGKC: GPU-Accelerated K-mer Counting
Sylvie: 3D-Adaptive and Universal System for Large-Scale Graph Neural Network Training 3823 Meng Zhang (Nanyang Technological University; NTU; Shanghai AI Laboratory), Qinghao Hu (Nanyang Technological University; NTU; Shanghai AI Laboratory), Cheng Wan (Georgia Institute of Technology), Haozhao Wang (NTU), Peng Sun (Shanghai AI Laboratory; SenseTime), Yonggang Wen (NTU), and Tianwei Zhang (NTU)
UltraPrecise: A GPU-Based Framework for Arbitrary-Precision Arithmetic in Database Systems 3837 Xin Li (Shandong University, China), Mengbai Xiao (Shandong University, China), Dongxiao Yu (Shandong University, China), Rubao Lee (Freelance, USA), and Xiaodong Zhang (The Ohio State University, USA)

<ul> <li>Exploiting Persistent CPU Cache for Scalable Persistent Hash Index</li></ul>
LTPG: Large-Batch Transaction Processing on GPUs with Deterministic Concurrency Control 3865 Jianpeng Wei (Northeastern University, China), Yu Gu (Northeastern University, China), Tianyi Li (Aalborg University, Denmark), Jianzhong Qi (The University of Melbourne, Australia), Chuanwen Li (Northeastern University, China), Yanfeng Zhang (Northeastern University, China), Christian S Jensen (Aalborg University, Denmark), and Ge Yu (Northeastern University, China)
Why Files If You Have a DBMS?3878Lam-Duy Nguyen (Technische Universität München) and Viktor Leis(Technische Universität München)
STEM: Streaming-Based FPGA Acceleration for Large-Scale Compactions in LSM KV
Improving the Relationship Between B+-Tree and Memory Allocator for Persistent Memory 3906 Wei Yan (Xi'an Jiaotong University, China) and Xingjun Zhang (Xi'an Jiaotong University, China)
Accelerating Aggregation using a Real Processing-in-Memory System
CLIMBER: Pivot-Based Approximate Similarity Search over Big Data Series
<ul> <li>Hill-Cache: Adaptive Integration of Recency and Frequency in Caching with Hill-Climbing 3947</li> <li>Yunfan Li (East China Normal University, China), Huiqi Hu (East China</li> <li>Normal University, China), Chaojing Lei (Zhejiang University, China),</li> <li>Xuan Zhou (East China Normal University, China), and Weining Qian</li> <li>(East China Normal University, China)</li> </ul>
Efficient Approximate Maximum Inner Product Search over Sparse Vectors
Riveter: Adaptive Query Suspension and Resumption Framework for Cloud Native Databases 3975 Rui Liu (University of Chicago), Aaron Elmore (University of Chicago), Michael Franklin (University of Chicago), and Sanjay Krishnan (University of Chicago)

<ul> <li>Mirage: Generating Enormous Databases for Complex Workloads</li></ul>	1989
Joint Directory, File and IO Trace Feature Extraction and Feature-Based Trace Regeneration for Enterprise Storage Systems	1002
Robust Auto-Scaling with Probabilistic Workload Forecasting for Cloud Databases	016
CheckMate: Evaluating Checkpointing Protocols for Streaming Dataflows	030
<ul> <li>BenchTemp: A General Benchmark for Evaluating Temporal Graph Neural Networks</li></ul>	044
<ul> <li>Fast Query Answering by Labeling Index on Uncertain Graphs</li></ul>	1058
<ul> <li>Scavenger: Better Space-Time Trade-Offs for Key-Value Separated LSM-Trees</li></ul>	1072

Yang Guo (The Chinese University of Hong Kong), Zhiqi Wang (The Chinese University of Hong Kong), Jin Xue (The Chinese University of	)86
<ul> <li>Hong Kong), and Zili Shao (The Chinese University of Hong Kong)</li> <li>Reverse Regret Query</li></ul>	.00
Resistance Eccentricity in Graphs: Distribution, Computation and Optimization	.13
BushStore: Efficient B+Tree Group Indexing for LSM-Tree in Non-Volatile Memory	.27
Cross Online Ride-Sharing for Multiple-Platform Cooperations in Spatial Crowdsourcing	40
Cooperative Air-Ground Instant Delivery by UAVs and Crowdsourced Taxis	.53
<ul> <li>Urban Sensing for Multi-destination Workers via Deep Reinforcement Learning</li></ul>	67
<ul> <li>Semi-Asynchronous Online Federated Crowdsourcing</li></ul>	.80

<ul> <li>FUDJ: Flexible User-Defined Distributed Joins</li></ul>
IVE: Accelerating Enumeration-Based Subgraph Matching via Exploring Isolated Vertices
<ul> <li>Approximate Skyline Index for Constrained Shortest Pathfinding with Theoretical Guarantee 4222</li> <li><i>Ziyi Liu (The Hong Kong University of Science and Technology, China),</i></li> <li><i>Lei Li (The Hong Kong University of Science and Technology, China),</i></li> <li><i>Mengxuan Zhang (The Hong Kong University of Science and Technology, China),</i></li> <li><i>Mengxuan Zhang (The Hong Kong University of Science and Technology, China),</i></li> <li><i>Mengxuan Zhang (The Hong Kong University of Science and Technology, China),</i></li> <li><i>Xiaofang Zhou (The Hong Kong University of Science and Technology, China),</i></li> <li><i>China)</i></li> </ul>
<ul> <li>CAGRA: Highly Parallel Graph Construction and Approximate Nearest Neighbor Search for GPUs</li> <li>4236</li> <li>Hiroyuki Ootomo (NVIDIA, Japan), Akira Naruse (NVIDIA, Japan), Corey</li> <li>Nolet (NVIDIA, USA), Ray Wang (NVIDIA, China), Tamas Feher (NVIDIA,</li> <li>Germany), and Yong Wang (NVIDIA, China)</li> </ul>
<ul> <li>VisionEmbedder: Bit-Level-Compact Key-Value Storage with Constant Lookup, Rapid Updates, and Rare Failure</li></ul>
Efficient Reverse k Approximate Nearest Neighbor Search over High-Dimensional Vectors
HJG: An Effective Hierarchical Joint Graph for ANNS in Multi-metric Spaces
Dynamic Data Layout Optimization with Worst-Case Guarantees

<ul> <li>QCFE: An Efficient Feature Engineering for Query Cost Estimation</li></ul>
Chameleon: Towards Update Efficient Learned Indexing for Locally Skewed Data
<ul> <li>FOSS: A Self-Learned Doctor for Query Optimizer</li></ul>
MFIX: An Efficient and Reliable Index Advisor via Multi-Fidelity Bayesian Optimization
VDTuner: Automated Performance Tuning for Vector Data Management Systems
<ul> <li>TrendSharing: a Framework to Discover and Follow the Trends for Shared Mobility Services 4370</li> <li>Jiexi Zhan (East China Normal University, China), Han Wu (East China</li> <li>Normal University, China), Peng Cheng (East China Normal University,</li> <li>China), Libin Zheng (Sun Yat-sen University, China), Lei Chen</li> <li>(HKUST(GZ) and HKUST, China), Chen Jason Zhang (The Hong Kong</li> <li>Polytechnic University, China), Xuemin Lin (Shanghai Jiaotong</li> <li>University, China), and Wenjie Zhang (The University of New South</li> <li>Wales, Australia)</li> </ul>
Collectively Simplifying Trajectories in a Database: A Query Accuracy Driven Approach

Efficient Learning-based Top-k Representative Similar Subtrajectory Query
Urban Region Representation Learning with Attentive Fusion
LightTR: A Lightweight Framework for Federated Trajectory Recovery
<ul> <li>Learning Time-Aware Graph Structures for Spatially Correlated Time Series Forecasting</li></ul>
Deep Dirichlet Process Mixture Model for Non-Parametric Trajectory Clustering
Parameterized Decision-Making with Multi-modality Perception for Autonomous Driving 4463 Yuyang Xia (University of Electronic Science and Technology of China, China), Shuncheng Liu (University of Electronic Science and Technology of China, China), Quanlin Yu (University of Electronic Science and Technology of China, China), Liwei Deng (University of Electronic Science and Technology of China, China), You Zhang (University of Michigan, USA), Han Su (University of Electronic Science and Technology of China, China), and Kai Zheng (University of Electronic Science and Technology of China, China)

CausalTAD: Causal Implicit Generative Model for Debiased Online Trajectory Anomaly
Detection
Wenbin Li (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Di
Yao (Institute of Computing Technology, Chinese Academy of Sciences,
China), Chang Gong (Institute of Computing Technology, Chinese Academy
of Sciences, China; University of Chinese Academy of Sciences, China),
Xiaokai Chu (Institute of Computing Technology, Chinese Academy of
Sciences, China), Quanliang Jing (Institute of Computing Technology,
Chinese Academy of Sciences, China), Xiaolei Zhou (DIDI Global Inc.),
Yuxuan Zhang (DIDI Global Inc.), Yunxia Fan (DIDI Global Inc.), and Jiwaning Bi (Institute of Commuting Technology, Chinese Academy, of
Jingping Bi (Institute of Computing Technology, Chinese Academy of Sciences, China)
Learning to Hash for Trajectory Similarity Computation and Search
Liwei Deng (University of Electronic Science and Technology of China,
China), Yan Zhao (Aalborg University, Denmark), Jin Chen (University
of Electronic Science and Technology of China, China), Shuncheng Liu
(University of Electronic Science and Technology of China, China),
Yuyang Xia (University of Electronic Science and Technology of China,
China), and Kai Zheng (University of Electronic Science and Technology
of China, China)
Ocean: Online Clustering and Evolution Analysis for Dynamic Streaming Data
Chunhui Feng (Soochow University, China), Junhua Fang (Soochow University, China), Yue Xia (Soochow University, China), Pingfu Chao
(Soochow University, China), Pengpeng Zhao (Soochow University,
China), Jiajie Xu (Soochow University, China), and Xiaofang Zhou (The
Hong Kong University of Science and Technology, China)
SWISP: Distributed Convoy Mining via Sliding Window-Based Indexing and Sub-Track
Partitioning
Chenxu Wang (Xi'an Jiaotong University, China), Xin Yang (Xi'an
Jiaotong University, China), Tianyi Li (Aalborg University, Denmark),
Jiaxing Wei (Xi'an Jiaotong University, China), Pinghui Wang (Xi'an Jiaotong University, China), Hongzhen Xiang (Xi'an Jiaotong
University, China), and Christain S. Jensen (Aalborg University,
Denmark)
Querying Shortest Path on Large Time-Dependent Road Networks with Shortcuts
Zengyang Gong (The Hong Kong University of Science and Technology,
China), Yuxiang Zeng (Beihang University, China), and Lei Chen (The
Hong Kong University of Science and Technology (Guangzhou), Guangzhou, Ching: The Hong Kong University of Science and Technology, Ching)
China; The Hong Kong University of Science and Technology, China)

<ul> <li>FRESH: Towards Efficient Graph Queries in an Outsourced Graph</li></ul>
<ul> <li>Managing the Future: Route Planning Influence Evaluation in Transportation Systems</li></ul>
<ul> <li>Scalable Distance Labeling Maintenance and Construction for Dynamic Small-World Networks 4573</li> <li><i>Xinjie Zhou (The Hong Kong University of Science and Technology,</i></li> <li><i>China), Mengxuan Zhang (The Hong Kong University of Science and</i></li> <li><i>Technology, China), Lei Li (The Hong Kong University of Science and</i></li> <li><i>Technology, China), and Xiaofang Zhou (The Hong Kong University of Science and Technology, China),</i></li> </ul>
Congestion-Mitigating Spatiotemporal Routing in Road Networks
A Just-In-Time Framework for Continuous Routing
QSRP: Efficient Reverse k-Ranks Query Processing on High-Dimensional Embeddings
FedCTQ: A Federated-based Framework for Accurate and Efficient Contact Tracing Query 4628 Zhihao Zeng (Zhejiang University), Ziquan Fang (Zhejiang University), Lu Chen (Zhejiang University), Yunjun Gao (Zhejiang University), Kai Zheng (University of Electronic Science and Technology of China), and Gang Chen (Zhejiang University)

Alleviating the Inconsistency of Multimodal Data in Cross-Modal Retrieval
<ul> <li>Firzen: Firing Strict Cold-Start Items with Frozen Heterogeneous and Homogeneous Graphs</li> <li>for Recommendation</li></ul>
Reconsidering Tree Based Methods for k-Maximum Inner-Product Search: The LRUS-CoverTree .4671 Hengzhao Ma (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China), Jianzhong Li (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China), and Yong Zhang (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, China)
Cross-Insight Trader: A Trading Approach Integrating Policies with Diverse Investment Horizons for Portfolio Management
Unsupervised Multimodal Graph Contrastive Semantic Anchor Space Dynamic Knowledge Distillation Network for Cross-Media Hash Retrieval
HIT: Solving Partial Index Tracking via Hierarchical Reinforcement Learning
<ul> <li>FieldSwap: Data Augmentation for Effective Form-Like Document Extraction</li></ul>

LT^2R: Learning to Online Learning to Rank for Web Search	
MUST: An Effective and Scalable Framework for Multimodal Search of Target Modality	
Online Anomaly Detection over Live Social Video Streaming	
Computing All Restricted Skyline Probabilities on Uncertain Datasets	
<ul> <li>M4: A Framework for Per-Flow Quantile Estimation</li></ul>	
DISCO: A Dynamically Configurable Sketch Framework in Skewed Data Streams	
BitMatcher: Bit-Level Counter Adjustment for Sketches	
Space-Efficient Indexes for Uncertain Strings	

<ul> <li>GLO: Towards Generalized Learned Query Optimization</li></ul>	843
A Fully On-Disk Updatable Learned Index	856
Routing-Guided Learned Product Quantization for Graph-Based Approximate Nearest Neighbor Search	870
Qiang Yue (Hangzhou Dianzi University, China), Xiaoliang Xu (Hangzhou Dianzi University, China), Yuxiang Wang (Hangzhou Dianzi University, China), Yikun Tao (Hangzhou Dianzi University, China), and Xuliyuan Luo (Hangzhou Dianzi University, China)	
Guided SQL-Based Data Exploration with User Feedback	884
<ul> <li>ShrinkHPO: Towards Explainable Parallel Hyperparameter Optimization</li></ul>	897
LBSC: A Cost-Aware Caching Framework for Cloud Databases	911
<ul> <li>DACE: A Database-Agnostic Cost Estimator</li></ul>	925
Enhancing LSM-Tree Key-Value Stores for Read-Modify-Writes via Key-Delta Separation	938

<ul> <li>TMan: A High-Performance Trajectory Data Management System Based on Key-Value Stores 4951 Huajun He (Southwest Jiaotong University, China; JD iCity, JD Technology, China; JD Intelligent Cities Research, China), Zihang Xu (Southwest Jiaotong University, China), Ruiyuan Li (Chongqing University, China), Jie Bao (JD iCity, China; JD Technology, China), Tianrui Li (Southwest Jiaotong University, China), and Zheng Yu (Southwest Jiaotong University, China; JD iCity, JD Technology, China; JD Intelligent Cities Research, China)</li> </ul>
<ul> <li>Kondo: Efficient Provenance-Driven Data Debloating</li></ul>
<ul> <li>Preserving Topological Feature with Sign-of-Determinant Predicates in Lossy Compression: A</li> <li>Case Study of Vector Field Critical Points</li></ul>
<ul> <li>FreqyWM: Frequency Watermarking for the New Data Economy</li></ul>
Multi-modality is All You Need for Transferable Recommender Systems
Multi-view Attentive Variational Learning for Group Recommendation
<ul> <li>Corruption Robust Dynamic Pricing in Liner Shipping Under Capacity Constraint</li></ul>

AdapTraj: A Multi-source Domain Generalization Framework for Multi-agent Trajectory Prediction
Tangwen Qian (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yile Chen (Nanyang Technological University, Singapore), Gao Cong (Nanyang Technological University, Singapore), Yongjun Xu (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), and Fei Wang (Institute of Computing Technology, Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China)
Towards Effective Next POI Prediction: Spatial and Semantic Augmentation with Remote         Sensing Data       5061         Nan Jiang (Beijing University of Posts and Telecommunications, China),       5061         Haitao Yuan (Nanyang Technological University, Singapore), Jianing Si       5061         (Beijing University of Posts and Telecommunications, China),       5061         Chen (Beijing University of Posts and Telecommunications, China),       5061         Shangguang Wang (Beijing University of Posts and Telecommunications, China),       5061         China)       5061
KartGPS: Knowledge Base Update with Temporal Graph Pattern-Based Semantic Rules 5075 Hao Xin (HKUST, China) and Lei Chen (HKUST(GZ) & HKUST, China)
Optimizing Probabilistic Box Embeddings with Distance Measures
A Multi-view Clustering Algorithm for Short Text

# Industry and Application Papers

GaussDB-Global: A Geographically Distributed Database System Puya Memarzia (Huawei Technologies Ltd., Canada), Huaxin Zhang (Huawei Technologies Ltd., Canada), Kelvin Ho (Huawei Technologies Ltd., Canada), Ronen Grosman (Huawei Technologies Ltd., Canada), and Jiang Wang (Huawei Technologies Ltd., China)	5111
Towards a Shared-Storage-Based Serverless Database Achieving Seamless Scale-up and Read Scale-out	. 5119
Yingqiang Zhang (Alibaba Group), Xinjun Yang (Alibaba Group), Hao Chen (Alibaba Group), Feifei Li (Alibaba Group), Jiawei Xu (Alibaba Group), Jie Zhou (Alibaba Group), Xudong Wu (Alibaba Group), and Qiang Zhang (Alibaba Group)	

Optimized Locking in SQL Azure
<ul> <li>Separation Is for Better Reunion: Data Lake Storage at Huawei</li></ul>
Deep Learning with Spatiotemporal Data: A Deep Dive into GeotorchAI
DATALORE: Can a Large Language Model Find All Lost Scrolls in a Data Repository?
ETUDE - Evaluating the Inference Latency of Session-Based Recommendation Models at Scale 5177 Barrie Kersbergen (bol.com; University of Amsterdam), Olivier Sprangers (University of Amsterdam), Frank Kootte (bol.com), Shubha Guha (University of Amsterdam), Maarten de Rijke (University of Amsterdam), and Sebastian Schelter (University of Amsterdam; Ahold Delhaize)
CoachLM: Automatic Instruction Revisions Improve the Data Quality in LLM Instruction Tuning
GaussML: An End-to-End In-Database Machine Learning System

<ul> <li>Xorbits: Automating Operator Tiling for Distributed Data Science</li></ul>	11
Couler: Unified Machine Learning Workflow Optimization in Cloud	24
<ul> <li>AntDT: A Self-Adaptive Distributed Training Framework for Leader and Straggler Nodes</li></ul>	38
<ul> <li>Addressing the Nested Data Processing Gap: JSONiq Queries on Snowflake Through Snowpark . 525 Dan Graur (ETH Zürich, Switzerland), Remo Röthlisberger (ETH Zürich, Switzerland), Adrian Jenny (ETH Zürich, Switzerland), Ghislain Fourny (ETH Zürich, Switzerland), Filip Drozdowski (Snowflake, USA), Choden Konigsmark (Snowflake, USA), Ingo Müller (ETH Zürich, Switzerland), and Gustavo Alonso (ETH Zürich, Switzerland)</li> </ul>	52
<ul> <li>Bw<sup>^</sup>e-Tree: an Evolution of Bw-Tree on Fast Storage</li></ul>	66
Resource Allocation with Service Affinity in Large-Scale Cloud Environments	80
Online Index Recommendation for Slow Queries	94
On Tuning Raft for IoT Workload in Apache IoTDB	07

Enabling Roll-up and Drill-Down Operations in News Exploration with Knowledge Graphs for Due Diligence and Risk Management	)
Multifaceted Reformulations for Null & Low Queries and its Parallelism with Counterfactuals	7
An Effective, Efficient, and Stable Framework for Query Clustering	1
A Framework for Continuous kNN Ranking of EV Chargers with Estimated Components	1

## Tutorials

Large Language Models: Principles and Practice	5354
Bipartite Graph Analytics: Current Techniques and Future Trends	358
Privacy-Aware Analysis Based on Data Series	365
Robust Query Optimization in the Era of Machine Learning: State-of-the-Art and Future         Directions       5         Amin Kamali (University of Ottawa, Canada), Verena Kantere (University         of Ottawa, Canada), and Calisto Zuzarte (IBM Canada Ltd., Canada)	371
Quantum Data Management: From Theory to Opportunities	376
An Interactive Dive into Time-Series Anomaly Detection	382

A Comprehensive Tutorial on over 100 Years of Diagrammatic Representations of Logical	
Statements and Relational Queries	5387
Wolfgang Gatterbauer (Northeastern University, USA)	

### Demonstrations

Entity/Relationship Profiling
GA-Tag: Data Enrichment with an Automatic Tagging System Utilizing Large Language Models 5397 Genki Kusano (NEC Corporation, Japan)
Comparing Personalized Relevance Algorithms for Directed Graphs
<ul> <li>FSM-Explorer: An Interactive Tool for Frequent Subgraph Pattern Mining from a Big Graph 5405 Jalal Khalil (St. Cloud State University), Da Yan (Indiana University Bloomington), Lyuheng Yuan (Indiana University Bloomington), Jiao Han (Indiana University Bloomington), Saugat Adhikari (Indiana University Bloomington), Cheng Long (Nanyang Technological University), and Yang Zhou (Auburn University)</li> </ul>
<ul> <li>TASKS: A Real-Time Query System for Instant Error-Tolerant Spatial Keyword Queries on Road</li> <li>Networks</li></ul>
VASIM: Vertical Autoscaling Simulator Toolkit
<ul> <li>Demonstration of FeVisQA: Free-Form Question Answering over Data Visualization</li></ul>
CleanEr: Interactive, Query-Guided Error Mitigation for Data Cleaning Systems
<ul> <li>Wearables for Health (W4H) Toolkit for Acquisition, Storage, Analysis and Visualization of</li> <li>Data from Various Wearable Devices</li></ul>

Chat2Query: A Zero-Shot Automatic Exploratory Data Analysis System with Large Language Models
China Normal University), Boyan Niu (PingCAP), Zheming Ni (PingCAP), Kai Xu (PingCAP), Jiajun Huang (PingCAP), Jianwei Wan (PingCAP), Shengbo Ma (PingCAP), Bing Wang (PingCAP), Donghui Zhang (PingCAP), Liu Tang (PingCAP), and Qi Liu (PingCAP)
EADS: An Early Anomaly Detection System for Sensor-Based Multivariate Time Series
d_symb Playground: An Interactive Tool to Explore Large Multivariate Time Series Datasets 5437 Sylvain W. Combettes (Université Paris-Saclay, Université Paris Cité, Centre Borelli), Paul Boniol (Université Paris-Saclay, Université Paris Cité, Centre Borelli), Charles Truong (Université Paris-Saclay, Université Paris Cité, Centre Borelli), and Laurent Oudre (Université Paris-Saclay, Université Paris Cité, Centre Borelli)
ADecimo: Model Selection for Time Series Anomaly Detection
ChatGraph: Chat with Your Graphs
A Fast Plan Enumerator for Recursive Queries
KGSEC: A Modular Framework for Knowledge Graph Schema Extraction and Comparison 5453 Petros Skoufis ("Athena" Research Center, Greece) and Dimitrios Skoutas ("Athena" Research Center, Greece)
<ul> <li>QFusor: A UDF Optimizer Plugin for SQL Databases</li></ul>
ARTS: A System for Aggregate Related Table Search
Explaining Expert Search Systems with ExES

RAGE Against the Machine: Retrieval-Augmented LLM Explanations
FairCR - An Evaluation and Recommendation System for fair Classification Algorithms
GraphLingo: Domain Knowledge Exploration by Synchronizing Knowledge Graphs and Large Language Models
MixedSearch: An Interactive System of Searching for the Best Tuple with Mixed Attributes 5481 Weicheng Wang (Hong Kong University of Science and Technology), Min Xie (Shenzhen Institute of Computing Sciences), and Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)
MorphStream: Scalable Processing of Transactions over Streams
FONT: A Flexible Polystore Evaluation Platform
CAMO: Explaining Consensus Across MOdels
<ul> <li>Pyneapple-R: Scalable and Expressive Spatial Regionalization</li></ul>

# **Data Engineering Future Technologies Papers**

SQL++: We Can Finally Relax!	. 5501
SQL++: We Can Finally Relax! Michael Carey (University of California, USA), Don Chamberlin (IBM	
Research (Retired), USA), Almann Goo (Amazon Web Services, USA), Kian	
Win Ong (Meta, USA), Yannis Papakonstantinou (Google Cloud, USA),	
Chris Suver (Amazon.com, USA), Sitaram Vemulapalli (Couchbase, Inc.,	
USA), and Till Westmann (Couchbase, Inc., USA)	

Data Flow Architectures for Data Processing on Modern Hardware
Personal Manifold: Management of Personal Data in the Age of Large Language Models
Applications and Challenges for Large Language Models: From Data Management Perspective 5530 Meihui Zhang (Beijing Institute of Technology), Zhaoxuan Ji (Beijing Institute of Technology), Zhaojing Luo (Beijing Institute of Technology), Yuncheng Wu (National University of Singapore), and Chengliang Chai (Beijing Institute of Technology)
Routing with Massive Trajectory Data
When Data Pricing Meets Non-Cooperative Game Theory5548Yuran Bi (Zhejiang University), Yihang Wu (Zhejiang University),5548Jinfei Liu (Zhejiang University), and Li Xiong (Emory University)
Secure Normal Form: Mediation Among Cross Cryptographic Leakages in Encrypted Databases . 5560 Shufan Zhang (University of Waterloo), Xi He (University of Waterloo), Ashish Kundu (Cisco Research), Sharad Mehrotra (University of California, Irvine), and Shantanu Sharma (New Jersey Institute of Technology)
Reactive Knowledge Management
LakeHarbor: Making Structures First-Class Citizens in Data Lakes
A CXL-Powered Database System: Opportunities and Challenges
Bifrost: A Future Graph Database Runtime
V2V: Efficiently Synthesizing Video Results for Video Queries
Higher-Order SQL Lambda Functions       5622         Maximilian Schüle (University of Bamberg) and Jakob Hornung       (University of Bamberg)

## PhD Symposium Papers

PR-GNN: Enhancing PoC Report Recommendation with Graph Neural Network
Cascade: Optimal Transaction Scheduling for High-Contention Workloads
Construction and Enhancement of an RNA-Based Knowledge Graph for Discovering new RNA Drugs 5639 <i>Emanuele Cavalleri (AnacletoLab, Italy)</i>
Enhancing Data Systems Performance by Exploiting SSD Concurrency & Asymmetry 5644 Tarikul Islam Papon (Boston University, USA)
Differential Analysis for System Provenance

## Lightning Talks

Evaluating Text-to-SQL Model Failures on Real-World Data
Synergies between Graph Data Management and Machine Learning in Graph Data Pipeline 5655 Arijit Khan (Aalborg University, Denmark)
Large Language Models as Storage for SQL Querying
Accelerating Deletion Interventions on OLAP Workload
User Learning In Interactive Data Exploration
Multivariate Similarity Search – A Call for a new Breed of Similarity Search Algorithms 5662 Odysseas Papapetrou (Eindhoven University of Technology, The Netherlands) and Jens E. d'Hondt (Eindhoven University of Technology, The Netherlands)
Towards Streaming Consistency Management       5663         Samuele Langhi (Lyon 1 University, France), Angela Bonifati (Lyon 1       5663         University, France), and Riccardo Tommasini (INSA Lyon, France)
Unveiling Dis-Integration

Cross-Source ML Model Training
<ul> <li>Why Model-Based Lossy Compression is Great for Wind Turbine Analytics</li></ul>
Towards Explainability in Retrieval-Augmented LLMs
Benchmarking Data Management Systems for Microservices
Exploring the Space of Model Comparisons
On Native Location-Optimized Data Systems
Observations and Opportunities in Solving Large-Scale Graph Data Processing Challenges at ByteDance by Using Heterogeneous Hardware

#### **TKDE Posters**

Data Lakes: A Survey of Functions and Systems (Extended Abstract)	679
OOD-GNN: Out-of-Distribution Generalized Graph Neural Network (Extended Abstract)	681
<ul> <li>Hierarchical Adaptive Pooling by Capturing High-Order Dependency for Graph Representation</li> <li>Learning (Extended Abstract)</li></ul>	683
PLAME: Piecewise-Linear Approximate Measure for Additive Kernel SVM (Extended Abstract) . 50 Tsz Nam Chan (Shenzhen University), Zhe Li (Alibaba Cloud), Leong Hou U (University of Macau), and Reynold Cheng (The University of Hong Kong)	685

Short-Text Author Linking Through Multi-Facet Temporal-Textual Embedding (Extended Abstract)
Saeed Najafipour (Iran University of Science and Technology, Iran), Saeid Hosseini (Sohar University, Oman), Wen Hua (The Hong Kong Polytechnic University), Mohammad Reza Kangavari (Iran University of Science and Technology, Iran), and Xiaofang Zhou (Hong Kong University of Science and Technology)
DKWS: A Distributed System for Keyword Search on Massive Graphs (Extended Abstract)
Multi-Grained Semantics-Aware Graph Neural Networks (Extended abstract)
Distilled Neural Networks for Efficient Learning to Rank
<ul> <li>Higher-Order Truss Decomposition in Graphs (Extended Abstract)</li></ul>
Finding the Maximum k-Balanced Biclique on Weighted Bipartite Graphs (Extended Abstract) 5697 Yiwei Zhao (Nanjing University of Aeronautics and Astronautics, China), Zi Chen (Nanjing University of Aeronautics and Astronautics, China), Chen Chen (University of Wollongong, Australia), Xiaoyang Wang (The University of New South Wales, Australia), Xuemin Lin (Shanghai Jiao Tong University, China), and Wenjie Zhang (The University of New South Wales, Australia)
Enabling Efficient, Verifiable, and Secure Conjunctive Keyword Search in Hybrid-Storage Blockchains
Hybrid Regret Minimization: A Submodular Approach (Extended Abstract)

A Neural Database for Answering Aggregate Queries on Incomplete Relational Data (Extended Abstract)
Mutual Information-Guided GA for Bayesian Network Structure Learning (Extended Abstract) 5705 Kefei Yan (Jiangnan University, China), Wei Fang (Jiangnan University, China), Hengyang Lu (Jiangnan University, China), Xin Zhang (Jiangnan University, China), Jun Sun (Jiangnan University, China), and Xiaojun Wu (Jiangnan University, China)
Efficient Discovery of Functional Dependencies on Massive Data (Extended Abstract) 5707 Xiaolong Wan (Harbin Institute of Technology, China), Xixian Han (Harbin Institute of Technology, China), Jinbao Wang (Harbin Institute of Technology, China), and Jianzhong Li (Harbin Institute of Technology, China)
Neural Similarity Search on Supergraph Containment (Extended Abstract)
Contact Tracing over Uncertain Indoor Positioning Data (Extended Abstract)
Efficient Semi-External SCC Computation (Extended Abstract)
<ul> <li>Value-Wise ConvNet for Transformer Models: An Infinite Time-Aware Recommender System</li> <li>(Extended Abstract)</li></ul>
Contrastive Graph Representations for Logical Formulas Embedding (Extended Abstract) 5717 Qika Lin (Xi'an Jiaotong University), Jun Liu (Shaanxi Province Key Laboratory of Satellite and Terrestrial Network Tech. R&D, Xi'an Jiaotong University), Lingling Zhang (Shaanxi Province Key Laboratory of Satellite and Terrestrial Network Tech. R&D, Xi'an Jiaotong University), Yudai Pan (Xi'an Jiaotong University), Xin Hu (Xi'an Jiaotong University), Fangzhi Xu (Xi'an Jiaotong University), and Hongwei Zeng (Xi'an Jiaotong University)

CUBE: Causal Intervention-Based Counterfactual Explanation for Prediction Models (Extended Abstract)
Xinyue Shao (Harbin Institute of Technology, China), Hongzhi Wang (Harbin Institute of Technology, China), Xiang Chen (King Abdullah University of Science and Technology, Saudi Arabia), Xiao Zhu (Harbin Institute of Technology, China), and Yan Zhang (Harbin Institute of Technology, China)
Data Level Privacy Preserving: A Stochastic Perturbation Approach Based on Differential Privacy (Extended Abstract)
of Science and Technology, China), Li Han (East China Normal University, China), Ming Ding (Data61, CSIRO, China), Raghav Bhaskar (Data61, CSIRO, China), and Jun Li (Nanjing University of Science and Technology, China)
Incremental Graph Computation: Anchored Vertex Tracking in Dynamic Social Networks (Extended Abstract)
Pushing ML Predictions into DBMSs (Extended Abstract)
Searching Personalized k-wing in Bipartite Graphs (Extented Abstract)
Complex Event Summarization Using Multi-social Attribute Correlation (Extended Abstract) 5729 Xi Chen (Soochow University, China), Xiangmin Zhou (RMIT University, Australia), Jeffrey Chan (RMIT University, Australia), Lei Chen (Hong Kong University of Science and Technology, China), Timos Sellis ("Athena" Research and Innovation Center, Greece), and Yanchun Zhang (Victoria University, Australia)
Efficient Community Search in Edge-Attributed Graphs (Extended Abstract)
<ul> <li>GPU-Based Efficient Parallel Heuristic Algorithm for High-Utility Itemset Mining in Large</li> <li>Transaction Datasets (Extended Abstract)</li></ul>

An Investigation of SMOTE Based Methods for Imbalanced Datasets With Data Complexity Analysis (Extended Abstract)
An Experimental Survey of Missing Data Imputation Algorithms (Extended Abstract)
Differentiable and Scalable Generative Adversarial Models for Data Imputation (Extended Abstract)
Yangyang Wu (Zhejiang University, China), Jun Wang (Hong Kong University of Science and Technology, China), Xiaoye Miao (Zhejiang University, China), Wenjia Wang (Hong Kong University of Science and Technology, China), and Jianwei Yin (Zhejiang University, China)
Matching Knowledge Graphs in Entity Embedding Spaces: An Experimental Study [Extended Abstract]
Weixin Zeng (Laboratory for Big Data and Decision, National University of Defense Technology, China), Xiang Zhao (Laboratory for Big Data and Decision, National University of Defense Technology, China), Zhen Tan (Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, China), Jiuyang Tang (Laboratory for Big Data and Decision, National University of Defense Technology, China), and Xueqi Cheng (Institute of Computing Technology, CAS, China)
Topic Modeling on Document Networks with Dirichlet Optimal Transport Barycenter (Extended         Abstract)
<ul> <li>Trajectory-Aware Task Coalition Assignment in Spatial Crowdsourcing (Extended Abstract) 5745</li> <li>Yuan Xie (Hunan University, China; National University of Singapore),</li> <li>Fan Wu (Hunan University, China), Xu Zhou (Hunan University, China),</li> <li>Wensheng Luo (Chinese University of Hong Kong), Yifang Yin (Institute</li> <li>for Infocomm Research (I^2R), A*STAR, Singapore), Roger Zimmermann</li> <li>(National University of Singapore), Keqin Li (State University of New</li> <li>York, USA), and Kenli Li (Hunan University, China)</li> </ul>
Recommending Learning Objects Through Attentive Heterogeneous Graph Convolution and Operation-Aware Neural Network (Extended Abstract)

Simultaneous Deep Clustering and Feature Selection via K-Concrete Autoencoder	5749
SoulMate: Short-Text Author Linking Through Multi-aspect Temporal-Textual Embedding (Extended Abstract)	5751
Saeed Najafipour (Iran University of Science and Technology, Iran), Saeid Hosseini (Sohar University, Oman), Wen Hua (Hong Kong Polytechnic University, Hong Kong), Mohammad Reza Kangavari (Iran University of Science and Technology, Iran), and Xiaofang Zhou (Hong Kong University of Science and Technology, Hong Kong)	0701

#### Author Index