Proceedings of the VLDB Endowment

Volume 15, No. 10 – June 2022

PVLDB – Proceedings of the VLDB Endowment

Volume 15, No. 10, June 2022.

All papers published in this issue will be presented at the 48th International Conference on Very Large Data Bases, Sydney, Australia, 2022.

Volume 15, Number 10, June 2022

Pages i – vii and 1978 - 2296

ISSN 2150-8097

Available at: http://www.pvldb.org and https://dl.acm.org/journal/pvldb

TABLE OF CONTENTS

Front Matter

Research Papers

LETTER FROM THE EDITORS IN CHIEF

We are pleased to present the tenth issue of PVLDB, Volume 15. This issue contains 24 papers in total including 21 regular research papers, 1 scalable data science (SDS) paper, and 2 experiments analysis & benchmark (EA&B) papers. A broad range of topics are covered in this issue including machine learning & applied AI for data management, graph data management, database performance, database engines, data privacy, and data quality.

For the first paper in this issue, Kakaraparthy et al. propose VIP hashing, a hash table method that uses lightweight mechanisms for learning the popularity distribution of keys and adapting to the skew in popularity on the fly. Next, Vincon et al. study near-data processing in database systems on native computational storage under HTAP Workloads. Echihabi et al. propose Hercules, a parallel tree-based technique for exact similarity search on massive disk-based data series collections. Siddiqui et al. present low-overhead data-driven techniques for filtering and costing indexes for scalable index tuning. Yang et al. optimize machine learning inference queries with correlative proxy models. Su et al. propose Banyan, a scoped dataflow engine for graph query service. Huang et al. study the problem of frequency estimation under both privacy and communication constraints, where the data is distributed among k parties. Ali et al. optimize inference serving on serverless platforms. Alkowaileet et al. propose techniques based on piggy-backing on Log-Structured Merge (LSM) tree events and tailored to document stores to store data in a columnar layout. Qiu et al. propose efficient solutions for shortest path counting on large road networks. Fu et al. introduce CELU-VFL, an efficient VFL training framework that exploits the local update technique to reduce the cross-party communication rounds. Zhu et al. propose DESIRE, an efficient dynamic cluster-based forest index for similarity search in multi-metric spaces. Kim et al. study the Attributed Bipartite Co-clustering (ABC) problem, which unifies the bipartite modularity optimization and attribute cohesiveness. Xiao et al. study the time series data encoding problem for efficient storage. Zhang et al. propose SA-LSM to use (S)urvival (A)nalysis for Log-Structure Merge Tree (LSM-tree) key-value (KV) stores. Ferragina et al. improve matrix-vector multiplication via lossless grammar-compressed matrices. Wu et al. propose NFL, a robust learned index via distribution transformation. Zare et al. design LEGOStore, an erasure coding (EC) based linearizable data store over geo-distributed public cloud data centers (DCs). Simpson et al. propose an information propagation model that captures important temporal aspects that have been well observed in the dynamics of fake news diffusion, in contrast with the diffusion of truth. Zhou et al. propose synergistic storage optimization techniques for duplication detection, page packing, and caching, to enhance database systems for model serving. Huang et al. present density-optimized intersection-free mapping and matrix multiplication for join-project operations. Jahangiri et al. design trade-offs for a robust dynamic hybrid hash join. Foufoulas et al. present YeSQL, an SQL extension with rich UDF support along with a pluggable architecture to easily integrate it with either server-based or embedded database engines. Ahmetaj et al. present magic shapes for SHACL validation.

All the papers in this issue will be presented at the 48th International Conference on Very Large Data Bases, 2022, in Sydney. We sincerely thank all the authors for submitting their work and all the reviewers for their outstanding service in reviewing the submissions. We hope that the reader will find this volume enjoyable.

Fatma Özcan, Juliana Freire and Xuemin Lin
Editors-in-Chief of PVLDB Volume 15
Program Chairs for VLDB 2022