

# ACM SIGMOD/PODS POSTER SESSION 1

Tuesday 3:30 – 5:00

Poster A

Poster B

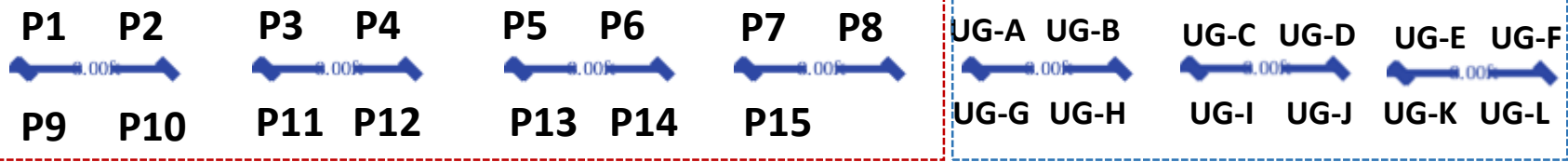
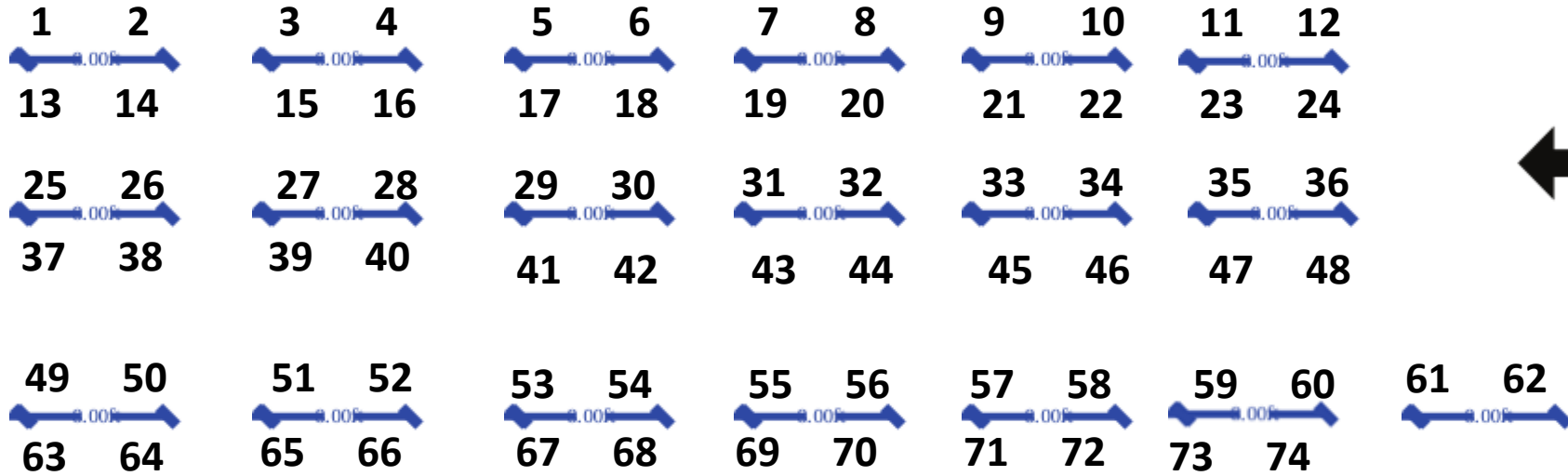
4'

8'

Grand C

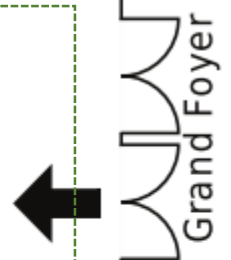
Grand B

SIGMOD

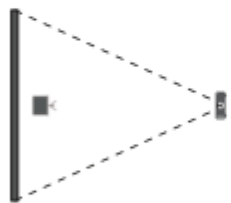


PODS

Undergrad competition



Grand Foyer



Grand A

# ACM SIGMOD/PODS POSTER SESSION 2

Wednesday 3:30 – 5:00

Poster A

Poster B

4'

8'

Grand C

Grand B

## SIGMOD

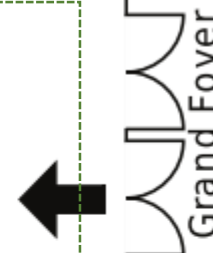
75	76	77	78	79	80	81	82	83	84	85	86		
87	88	89	90	91	92	93	94	95	96	97	98		
99	100	101	102	103	104	105	106	107	108	109	110		
111	112	113	114	115	116	117	118	119	120	121	122		
123	124	125	126	127	128	129	130	131	132	133	134		
135	136	137	138	139	140	141	142	143	144	145	146	147	148

TO-A	TO-B	PG-A	PG-B		
TO-C	TO-D	TO-E	PG-C	PG-D	PG-E

TODS – Programming Contest

P16	P17	P18	P19	P20	P21	P22	P23
P24	P25	P26	P27	P28	P29	P30	P31

PODS



Grand Foyer

Grand A

## Poster Number | Title

- 1 | Learning Linear Regression Models over Factorized Joins
- 2 | To Join or Not to Join? Thinking Twice about Joins before Feature Selection
- 3 | Real-time Video Recommendation Exploration
- 4 | Towards Globally Optimal Crowdsourcing Quality Management
- 5 | Building the Enterprise Fabric for Big Data with Vertica and Spark Integration
- 6 | Probabilistic Truss Decomposition: Concepts and Algorithms
- 7 | Efficient and Progressive Group Steiner Tree Search
- 8 | Publishing Attributed Social Graphs with Formal Privacy Guarantees
- 9 | Publishing Graph Degree Distribution with Node Differential Privacy
- 10 | Principled evaluation of differentially private algorithms using DPBench
- 11 | PrivTree: A Differentially Private Algorithm for Hierarchical Decompositions
- 12 | Adaptive Indexing over Encrypted Data
- 13 | Practical Private Range Search Revisited
- 14 | Privacy Preserving Subgraph Matching on Large Graphs in Cloud
- 15 | The Snowflake Elastic Data Warehouse
- 16 | Closing the Schema and Performance Gap between SQL and NoSQL
- 17 | Have Your Data and Query It Too: From Key-Value Caching to Big Data Management
- 18 | Ambry: LinkedIn's Scalable Geo-Distributed Object Store
- 19 | SQL Schema Design
- 20 | TableShare: Results from a Multi-Year SQL-as-a-Service Experiment
- 21 | Automatic Generation of Normalized Relational Schemas from Key-Value Data
- 22 | Data Blocks: Hybrid OLTP and OLAP on Compressed Storage using both Vectorization and Compilation
- 23 | GeckoFTL: Scalable Flash Translation Techniques For Very Large Flash Devices
- 24 | SHARE Interface in Flash Storage for Databases
- 25 | Accelerating Relational Databases by Leveraging Remote Memory and RDMA
- 26 | FPTree: A Hybrid SCM-DRAM Persistent and Concurrent B-Tree for Storage Class Memory
- 27 | Micro-architectural Analysis of In-memory OLTP
- 28 | iBFS: Concurrent Breadth-First Graph Traversal on GPUs
- 29 | Tornado: A System for Real-Time Iterative Analysis over Evolving Data
- 30 | System-X: A Relational Engine for Graph Processing
- 31 | GTS: A Fast and Scalable Graph Processing Method based on Streaming Topology to GPUs
- 32 | Graph Analytics Through Fine-Grained Parallelism
- 33 | Hybrid Pulling/Pushing for I/O-Efficient Distributed and Iterative Graph Computing
- 34 | Scalable Pattern Sharing over Event Streams
- 35 | How to win a hot dog eating contest
- 36 | Sharing-Aware Outlier Analytics over High-Volume Data Streams
- 37 | THEMIS: Fairness in Federated Stream Processing under Overload
- 38 | SABER: Window-Based Hybrid Stream Processing for Heterogeneous Architectures
- 39 | Range Thresholding on Streams
- 40 | Bridging the Archipelago between Row-Stores and Column-Stores for Hybrid Workloads
- 41 | An Effective Syntax for Bounded Relational Queries
- 42 | Wander Join: Online Aggregation via Random Walks
- 43 | Lazily Approximating Complex AdHoc Queries in Big Data Clusters
- 44 | A Study of Sorting Algorithms on Approximate Memory
- 45 | Distributed Wavelet Thresholding for Maximum Error Metrics
- 46 | Sample + Seek: Approximating Aggregates with Distribution Precision Guarantee
- 47 | Stop-and-Stare: Optimal Sampling Algorithms for Viral Marketing in Billion-scale Networks
- 48 | Spheres of Influence for More Effective Viral Marketing
- 49 | Continuous Influence Maximization: What Discounts Should We Offer to Social Network Users?
- 50 | Holistic Influence Maximization: Combining Scalability and Efficiency with Opinion-Aware Models
- 51 | Potential and Pitfalls of Domain-Specific Information Extraction at Web Scale
- 52 | Robust and Noise Resistant Wrapper Induction
- 53 | Goods: Organizing Google's Datasets
- 54 | Multi-Source Uncertain Entity Resolution at Yad Vashem: Transforming Holocaust Victim Reports into People
- 55 | A Hybrid Approach to Functional Dependency Discovery
- 56 | Ontological Pathfinding: Mining First-Order Knowledge from Large Knowledge Bases
- 57 | Extracting Databases from Dark Data with DeepDive
- 58 | Estimating the Impact of Unknown Unknowns on Aggregate Query Results
- 59 | Constraint-Variance Tolerant Data Repairing
- 60 | Interactive and Deterministic Data Cleaning
- 61 | Sequential Data Cleaning: A Statistical Approach
- 62 | Learning-Aided Cleansing for Indoor RFID Data
- 63 | PrivateClean: Data Cleaning and Differential Privacy
- 64 | RDFind: Scalable Conditional Inclusion Dependency Discovery in RDF Datasets
- 65 | Cost-Effective Crowdsourced Entity Resolution: A Partial-Order Approach
- 66 | Topic Exploration in Spatio-Temporal Document Collections
- 67 | ParTime: Parallel Temporal Aggregation
- 68 | Data Polygamy: The Many-Many Relationships among Urban Spatio-Temporal Data Sets
- 69 | Distributed Top-k Temporal Joins
- 70 | AT-GIS: Highly Parallel Spatial Query Processing with Associative Transducers
- 71 | Towards Best Region Search with Submodular Monotone Aggregate Function
- 72 | Simba: Efficient In-Memory Spatial Analytics
- 73 | Realtime Data Processing at Facebook
- 74 | SparkR: Scaling R Programs with Spark
- UG-A | CodeOntology: a Java semantic database
- UG-B | Constructing Join Histograms from Histograms with q-error Guarantees
- UG-C | Vectorizing an In Situ Query Engine
- UG-D | M3: Scaling Up Machine Learning via Memory Mapping
- UG-E | Minimizing Average Regret Ratio in Database
- UG-F | Graph Summarization for Geo-correlated Trends Detection in Social Networks
- UG-G | K-means Split Revisited: Well-grounded Approach and Experimental Evaluation
- UG-H | Research Contribution as a Measure of Influence
- UG-I | Adaptive Denormalization
- UG-J | Adaptive Data Skipping in Main-Memory Systems
- UG-K | Exploring Visualization of Data Transforms
- UG-L | Searching Web Data using MinHash LSH

Poster Number | Title

- P1 | Minimization of Tree Pattern Queries
- P2 | Incremental View Maintenance For Collection Programming
- P3 | Aggregations over Generalized Hypertree Decompositions
- P4 | Bounded Query Rewriting Using Views
- P5 | Red Spider Meets a Rainworm: Conjunctive Query Finite Determinacy Is Undecidable
- P6 | Towards a Worst-Case I/O-Optimal Algorithm for Acyclic Joins
- P7 | On the Complexity of Inner Product Similarity Join
- P8 | Verification of Hierarchical Artifact Systems
- P9 | Recency-Bounded Verification of Dynamic Database-Driven Systems
- P10 | Making SQL Queries Correct on Incomplete Databases: A Feasibility Study
- P11 | Designing a Query Language for RDF: Marrying Open and Closed Worlds
- P12 | Towards Tight Bounds for the Streaming Set Cover Problem
- P13 | An Optimal Algorithm for Heavy Hitters in Insertion Streams and Related Problems
- P14 | Better Algorithms for Counting Triangles in Data Streams
- P15 | Locating a Small Cluster Privately

Poster Number | Title

75 | VectorH: taking SQL-on-Hadoop to the next level

76 | Adaptive Logging: Optimizing Logging and Recovery Costs in Distributed In-memory Databases

77 | Big Data Analytics with Datalog Queries on Spark

78 | An Efficient MapReduce Cube Algorithm for Varied Data Distributions

79 | Diversified Subgraph Querying in a Large Graph

80 | Graph Indexing for Shortest-Path Finding over Dynamic Sub-Graphs

81 | Efficient Subgraph Matching by Postponing Cartesian Products

82 | Adding Counting Quantifiers to Graph Patterns

83 | DualSim: Parallel Subgraph Enumeration in a Massive Graph on a Single Machine

84 | Distributed Set Reachability

85 | Fast Multi-column Sorting in Main-Memory Column-Stores

86 | Elastic Pipelining In an In-Memory Database Cluster

87 | Page As You Go: Piecewise Columnar Access In SAP HANA

88 | Hybrid Garbage Collection for Multi-Version Concurrency Control in SAP HANA

89 | Bridging the Archipelago between Row-Stores and Column-Stores for Hybrid Workloads

90 | UpBit: Scalable In-Memory Updatable Bitmap Indexing

91 | FluxQuery: An Execution Framework for Highly Interactive Query Workloads

92 | iOLAP: Incremental OLAP for Interactive Analysis on Big Data

93 | Dynamic Prefetching of Data Tiles for Interactive Visualization

94 | Expressive Query Construction through Direct Manipulation of Nested Relational Results

95 | Shasta: Interactive Reporting At Scale

96 | Datometry Hyper-Q: Bridging the Gap Between Real Time and Historical Analytics

97 | Time Adaptive Sketches (Ada-Sketches) for Summarizing Data Streams

98 | Streaming Algorithms for Robust Distinct Elements

99 | Augmented Sketch: Faster and More Accurate Stream Processing

100 | Matrix Sketching Over Sliding Windows

101 | Graph Stream Summarization: From Big Bang to Big Crunch

102 | Scalable Approximate Query Tracking over Highly Distributed Data Streams

103 | A Hybrid B+-Tree as Solution for In-Memory Indexing on CPU-GPU Heterogeneous Computing Platforms

104 | Low-Overhead Asynchronous Checkpointing in Main-Memory Database Systems

105 | T-Part: Partitioning of Transactions for Onward-Pushing in Deterministic Database Systems

106 | Reducing the Storage Overhead of Main-Memory OLTP Databases with Hybrid Indexes

107 | Scaling Multi-core OLTP Under High Contention

108 | DBSherlock: A Performance Diagnostic Tool for Transactional Databases

109 | TARDiS: Making Sense of Weak Consistency

110 | TicToc: Time-Traveling Optimistic Concurrency Control

111 | Unleashing Parallelism in Multi-core Databases via Dependency Tracking

112 | Towards a Non-2PC Transaction Management in Distributed Database Systems

113 | ERMIA: Fast memory-optimized database system for heterogeneous workloads

114 | Transaction Healing: Scaling Optimistic Concurrency Control on Multicores

115 | Enabling Incremental Query Re-Optimization

116 | Sampling-Based Query Re-Optimization

117 | A Fast Randomized Algorithm for Multi-Objective Query Optimization

118 | Operator and Query Progress Estimation in Microsoft SQL Server Live Query Statistics

119 | Optimization of Nested Queries using the NF2 Algebra

120 | Extracting Equivalent SQL From Imperative Code in Database Applications

121 | Generating Preview Tables for Entity Graphs

122 | Speedup Graph Processing by Graph Ordering

123 | ROLL: Fast In-Memory Generation of Gigantic Scale-free Networks

124 | Functional Dependencies for Graphs

125 | SLING: A Near-Optimal Index Structure for SimRank

126 | Query Planning for Evaluating SPARQL Property Paths

127 | Robust Query Processing in Co-Processor-accelerated Databases

128 | How to Architect a Query Compiler

129 | Automated Demand-driven Resource Scaling in Relational Database-as-a-Service

130 | GPL: A GPU-based Pipelined Query Processing Engine

131 | Towards a Hybrid Design for Fast Query Processing in DB2 with BLU Acceleration...

132 | An Experimental Comparison of Thirteen Relational Equi-Joins in Main Memory

133 | Top-k Relevant Semantic Place Retrieval on Spatial RDF Data

134 | Local Similarity Search for Unstructured Text

135 | Similarity Join over Array Data

136 | LazyLSH: Approximate Nearest Neighbor Search for Multiple Distance Functions with a Single Index

137 | Set-based Similarity Search for Time Series

138 | Range-based Obstructed Nearest Neighbor Queries

139 | Data Cleaning: Overview and Emerging Challenges

140 | Big Graph Analytics Systems

141 | Microblogs Data Management Systems: Querying, Analysis, and Visualization

142 | Design Tradeoffs of Data Access Methods

143 | The Challenges of Global-scale Data Management

144 | Provenance: On and Behind the Screens

145 | Querying Geo-Textual Data: Spatial Keyword Queries and Beyond

146 | Introduction to Spark 2.0 for Database Researchers

147 | Automatic Entity Recognition and Typing in Massive Text Data

148 | Semistructured Models, Queries and Algebras in the Big Data Era

TODS-A | Workload-Driven Antijoin Cardinality Estimation

TODS-B | Reducing Layered Database Applications to their Essence through Vertical Integration

TODS-C | BEVA: An Efficient Query Processing Algorithm for Error Tolerant Autocompletion

TODS-D | Synchronization of Queries and Views Upon Schema Evolutions: A Survey

TODS-E | The Exact Complexity of the First-Order Logic Definability Problem

PROG-A | SIGMOD Programming Contest Finalist/Peking University

PROG-B | SIGMOD Programming Contest Finalist/University of Engineering and Technology – Vietnam National University

PROG-C | SIGMOD Programming Contest Finalist/The University of Tokyo

PROG-D | SIGMOD Programming Contest Finalist/National and Kapodistrian University of Athens

PROG-E | SIGMOD Programming Contest Finalist/Technische Universität München

## SIGMOD/PODS Posters Session 2

Poster Number | Title

P16 | FAQ: Questions Asked Frequently

P17 | Shortest Paths and Distances with Differential Privacy

P18 | Schema validation via streaming circuits

P19 | Variability in data streams

P20 | Streaming Space Complexity of Nearly All Functions of One Variable on Frequency Vectors

P21 | Efficient Top-k Indexing via General Reductions

P22 | Anti-Persistence on Persistent Storage: History-Independent Sparse Arrays and Dictionaries

P23 | Counting Answers to Existential Positive Queries: A Complexity Classification

P24 | Worst-case Optimal Algorithms for Conjunctive Queries with Functional Dependencies

P25 | Semantic Acyclicity Under Constraints

P26 | Tractable Lineages on Treelike Instances: Limits and Extensions

P27 | Parallel Algorithms for Constructing Range and Nearest-Neighbor Searching Data Structures

P28 | Space Lower Bounds for Itemset Frequency Sketches

P29 | Are Few Bins Enough: Testing Histogram Distributions

P30 | Range-Max Queries on Uncertain Data

P31 | Fast Algorithms for Parsing Sequences of Parentheses with Few Errors