



Keynotes

- 1  **Human Computation**
(Luis von Ahn)
 - 3  **The Database Architecture Jigsaw Puzzle**
(Martin L. Kersten)
 - 5  **PhotoSpread: A Spreadsheet for Managing Photos**
(Hector Garcia-Molina)
-

Panels

- 7  **Declarative, Domain-Specific Languages — Elegant Simplicity or a Hammer in Search of a Nail?**
(Samuel Madden, Johannes Gehrke)
- 8  **Cloud Computing — Was Thomas Watson Right After All?**
(Raghu Ramakrishnan)
- 9  **Scientific Data Management: An Orphan in the Database Community?**
(Randal Burns)



Research Session 1: Schema Mappings and Matching

- 10  **Muse: Mapping Understanding and deSign by Example**
(Bogdan Alexe, Laura Chiticariu, Renée J. Miller, Wang-Chiew Tan)
- 20  **Usage-Based Schema Matching**
(Hazem Elmeleegy, Mourad Ouzzani, Ahmed K. Elmagarmid)
- 30  **Clip: A Visual Language for Explicit Schema Mappings**
(Alessandro Raffio, Daniele Braga, Stefano Ceri, Paolo Papotti, Mauricio A. Hernández)
-

Research Session 2: Advanced Query Processing

- 40  **Transformation-Based Framework for Record Matching**
(Arvind Arasu, Surajit Chaudhuri, Raghav Kaushik)
- 50  **Querying Data Under Access Limitations**
(Andrea Cali, Davide Martinenghi)
- 60  **Constant-Time Query Processing**
(Vijayshankar Raman, Garret Swart, Lin Qiao, Frederick Reiss, Vijay Dialani, Donald Kossmann, Inderpal Narang, Richard Sidle)



Research Session 3: Spatial/Temporal Queries and Schema Matching

- 70  **A Hybrid Prediction Model for Moving Objects**
(*Hoyoung Jeung, Qing Liu, Heng Tao Shen, Xiaofang Zhou*)
- 80  **A Generic Framework for Continuous Motion Pattern Query Evaluation**
(*Petko Bakalov, Vassilis J. Tsotras*)
- 90  **Querying Complex Spatio-Temporal Sequences in Human Motion Databases**
(*Yueguo Chen, Shouxu Jiang, Beng Chin Ooi, Anthony K.H. Tung*)
- 100  **Common Influence Join: A Natural Join Operation for Spatial Pointsets**
(*Man Lung Yiu, Nikos Mamoulis, Panagiotis Karras*)
- 110  **Matching Schemas in Online Communities: A Web 2.0 Approach**
(*Robert McCann, Warren Shen, AnHai Doan*)
- 120  **Validating Multi-Column Schema Matchings by Type**
(*Bing Tian Dai, Nick Koudas, Divesh Srivastava, Anthony K.H. Tung, Suresh Venkatasubramanian*)
-

Research Session 4: Data Correlation and Clustering

- 130  **CARE: Finding Local Linear Correlations in High Dimensional Data**
(*Xiang Zhang, Feng Pan, Wei Wang*)
- 140  **Trajectory Outlier Detection: A Partition-and-Detect Framework**
(*Jae-Gil Lee, Jiawei Han, Xiaolei Li*)
- 150  **A Framework for Clustering Uncertain Data Streams**
(*Charu C. Aggarwal, Philip S. Yu*)



Research Session 5: Data Mining

- 160  **Mining Approximate Order Preserving Clusters in the Presence of Noise**
(Mengsheng Zhang, Wei Wang, Jinze Liu)
- 169  **Direct Discriminative Pattern Mining for Effective Classification**
(Hong Cheng, Xifeng Yan, Jiawei Han, Philip S. Yu)
- 179  **Verifying and Mining Frequent Patterns from Large Windows Over Data Streams**
(Barzan Mozafari, Hetal Thakkar, Carlo Zaniolo)
-

Research Session 6: Query Processing I

- 189  **On Signatures for Communication Graphs**
(Graham Cormode, Flip Korn, S. Muthukrishnan, Yihua Wu)
- 199  **Robust Stratified Sampling Plans for Low Selectivity Queries**
(Shantanu Joshi, Christopher Jermaine)
- 209  **Training Linear Discriminant Analysis in Linear Time**
(Deng Cai, Xiaofei He, Jiawei Han)
-

Research Session 7: Large-Scale Query Evaluation

- 218  **Parallel Evaluation of Composite Aggregate Queries**
(Lei Chen, Christopher Olston, Raghu Ramakrishnan)
- 228  **Efficient Computation of Diverse Query Results**
(Erik Vee, Utkarsh Srivastava, Jayavel Shanmugasundaram, Prashant Bhat, Sihem Amer Yahia)
- 237  **Toward Expressive and Scalable Sponsored Search Auctions**
(David J. Martin, Johannes Gehrke, Joseph Y. Halpern)



Research Session 8: Advanced Data Structures

- 247  **Lattice Histograms: A Resilient Synopsis Structure**
(*Panagiotis Karras, Nikos Mamoulis*)
- 257  **Efficient Merging and Filtering Algorithms for Approximate String Searches**
(*Chen Li, Jiaheng Lu, Yiming Lu*)
- 267  **Fast Indexes and Algorithms for Set Similarity Selection Queries**
(*Marios Hadjieleftheriou, Amit Chandel, Nick Koudas, Divesh Srivastava*)
-

Research Session 9: Security and Privacy I

- 277  **Privacy: Theory Meets Practice on the Map**
(*Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, Lars Vilhuber*)
- 287  **Auditing SQL Queries**
(*Rajeev Motwani, Shubha U. Nabar, Dilys Thomas*)
- 297  **Optimal Boolean Matrix Decomposition: Application to Role Engineering**
(*Haibing Lu, Jaideep Vaidya, Vijayalakshmi Atluri*)



Research Session 10: Similarity Search

- 307  **Efficient Similarity Search Using the Earth Mover's Distance for Large Multimedia Databases**
(Ira Assent, Marc Wichterich, Tobias Meisen, Thomas Seidl)
- 317  **Similarity Search in Arbitrary Subspaces Under L_p -Norm**
(Xiang Lian, Lei Chen)
- 327  **Nearest Neighbor Retrieval Using Distance-Based Hashing**
(Vassilis Athitsos, Michalis Potamias, Panagiotis Papapetrou, George Kollios)
- 337  **On Monitoring the Top- k Unsafe Places**
(Donghui Zhang, Yang Du, Ling Hu)
- 346  **Compact Similarity Joins**
(Brent Bryan, Frederick Eberhardt, Christos Faloutsos)
- 356  **Standing Out in a Crowd: Selecting Attributes for Maximum Visibility**
(Muhammed Miah, Gautam Das, Vagelis Hristidis, Heikki Mannila)
-

Research Session 11: Security and Privacy II

- 366  **SpaceTwist: Managing the Trade-Offs Among Location Privacy, Query Performance, and Query Accuracy in Mobile Services**
(Man Lung Yiu, Christian S. Jensen, Xuegang Huang, Hua Lu)
- 376  **Never Walk Alone: Uncertainty for Anonymity in Moving Objects Databases**
(Osman Abul, Francesco Bonchi, Mirco Nanni)
- 386  **On Unifying Privacy and Uncertain Data Models**
(Charu C. Aggarwal)



Research Session 12: Data Stream Processing

- 396  **Efficient Construction of Compact Shedding Filters for Data Stream Processing**
(Buğra Gedik, Kun-Lung Wu, Philip S. Yu)
- 406  **A Security Punctuation Framework for Enforcing Access Control on Streaming Data**
(Rimma V. Nehme, Elke A. Rundensteiner, Elisa Bertino)
- 416  **Randomized Synopses for Query Assurance on Data Streams**
(Ke Yi, Feifei Li, Marios Hadjieleftheriou, George Kollios, Divesh Srivastava)
-

Research Session 13: Information Extraction

- 426  **LOCUST: An Online Analytical Processing Framework for High Dimensional Classification of Data Streams**
(Charu C. Aggarwal, Philip S. Yu)
- 436  **Dynamic Materialization of Query Views for Data Warehouse Workloads**
(Thomas Phan, Wen-Syan Li)
- 446  **Injector: Mining Background Knowledge for Data Anonymization**
(Tiancheng Li, Ninghui Li)
- 456  **RiTE: Providing On-Demand Data for Right-Time Data Warehousing**
(Christian Thomsen, Torben Bach Pedersen, Wolfgang Lehner)
- 466  **Automatic Extraction of Useful Facet Hierarchies from Text Databases**
(Wisam Dakka, Panagiotis G. Ipeirotis)
- 476  **A Sampling-Based Approach to Information Recovery**
(Junyi Xie, Jun Yang, Yuguo Chen, Haixun Wang, Philip S. Yu)



Research Session 14: Data Privacy in Social Networks

- 486  **Mining (Social) Network Graphs to Detect Random Link Attacks**
(Nisheeth Shrivastava, Anirban Majumder, Rajeev Rastogi)
- 496  **A Hybrid Approach to Private Record Linkage**
(Ali Inan, Murat Kantarcioglu, Elisa Bertino, Monica Scannapieco)
- 506  **Preserving Privacy in Social Networks Against Neighborhood Attacks**
(Bin Zhou, Jian Pei)
-

Research Session 15: Query Processing II

- 516  **Increasing the Expressivity of Conditional Functional Dependencies Without Extra Complexity**
(Loreto Bravo, Wenfei Fan, Floris Geerts, Shuai Ma)
- 526  **Efficient Constraint Monitoring Using Adaptive Thresholds**
(Srinivas Kashyap, Jeyashankher Ramamirtham, Rajeev Rastogi, Pushpraj Shukla)
- 536  **Dominant Graph: An Efficient Indexing Structure to Answer Top-k Queries**
(Lei Zou, Lei Chen)
-

Research Session 16: Skyline and Top-K Queries

- 546  **Parallel Distributed Processing of Constrained Skyline Queries by Filtering**
(Bin Cui, Hua Lu, Quanqing Xu, Lijiang Chen, Yafei Dai, Yongluan Zhou)
- 556  **Skyline Query Processing for Incomplete Data**
(Mohamed E. Khalefa, Mohamed F. Mokbel, Justin J. Levandoski)
- 566  **Top-k Spatial Joins of Probabilistic Objects**
(Vebjorn Ljosa, Ambuj K. Singh)
-



Research Session 17: Distributed and Federated Databases

- 576  **The Cost of Serializability on Platforms That Use Snapshot Isolation**
(Mohammad Alomari, Michael Cahill, Alan Fekete, Uwe Röhm)
- 586  **Network-Aware Join Processing in Global-Scale Database Federations**
(Xiaodan Wang, Randal Burns, Andreas Terzis, Amol Deshpande)
- 596  **Robust Runtime Optimization of Data Transfer in Queries Over Web Services**
(Anastasios Gounaris, Christos Yfoulis, Rizos Sakellariou, Marios D. Dikaiakos)
-

Research Session 18: XML Data Processing

- 606  **XML Processing in DHT Networks**
(Serge Abiteboul, Ioana Manolescu, Neoklis Polyzotis, Nicoleta Preda, Chong Sun)
- 616  **Efficient Processing of XML Update Streams**
(Leonidas Fegaras)
- 626  **XML Prefiltering as a String Matching Problem**
(Christopher Koch, Stefanie Scherzinger, Michael Schmidt)
-

Research Session 19: Search in Databases

- 636  **Optimizing SQL Queries Over Text Databases**
(Alpa Jain, AnHai Doan, Luis Gravano)
- 646  **Building Community Wikipedias: A Machine-Human Partnership Approach**
(Pedro DeRose, Xiaoyong Chai, Byron J. Gao, Warren Shen, AnHai Doan, Philip Bohannon, Xiaojin Zhu)
- 656  **Keyword Search on Spatial Databases**
(Ian De Felipe, Vagelis Hristidis, Naphtali Rishe)



Research Session 20: Data Streams

- 666  **Simultaneous Equation Systems for Query Processing on Continuous-Time Data Streams**
(Yanif Ahmad, Olga Papaemmanouil, Uğur Çetintemel, Jennie Rogers)
- 676  **Runtime Semantic Query Optimization for Event Stream Processing**
(Luping Ding, Songting Chen, Elke A. Rundensteiner, Junichi Tatemura, Wang-Pin Hsiung, K. Selçuk Candan)
- 686  **Efficient Online Subsequence Searching in Data Streams Under Dynamic Time Warping Distance**
(Mi Zhou, Man Hon Wong)
-

Research Session 21: Data Privacy and Security

- 696  **Efficient Data Authentication in an Environment of Untrusted Third-Party Distributors**
(Mikhail J. Atallah, YounSun Cho, Ashish Kundu)
- 705  **OptRR: Optimizing Randomized Response Schemes for Privacy-Preserving Data Mining**
(Zhengli Huang, Wenliang Du)
- 715  **On the Anonymization of Sparse High-Dimensional Data**
(Gabriel Ghinita, Yufei Tao, Panos Kalnis)
- 725  **On Anti-Corruption Privacy Preserving Publication**
(Yufei Tao, Xiaokui Xiao, Jiexing Li, Donghui Zhang)
- 735  **Similar Document Detection with Limited Information Disclosure**
(Wei Jiang, Mummoorthy Murugesan, Chris Clifton, Luo Si)
- 744  **k-Anonymization Revisited**
(Aristides Gionis, Arnon Mazza, Tamir Tassa)



Research Session 22: Distributed and Parallel Query Processing

- 754  **Flow Algorithms for Parallel Query Optimization**
(*Amol Deshpande, Lisa Hellerstein*)
- 764  **Self-Join Size Estimation in Large-Scale Distributed Data Systems**
(*Theoni Pitoura, Peter Triantafillou*)
- 774  **Sideways Information Passing for Push-Style Query Processing**
(*Zachary G. Ives, Nicholas E. Taylor*)
-

Research Session 23: Sensor Networks and Data Streams

- 784  **COLR-Tree: Communication-Efficient Spatio-Temporal Indexing for a Sensor Data Web Portal**
(*Yanif Ahmad, Suman Nath*)
- 794  **Region Sampling: Continuous Adaptive Sampling on Sensor Networks**
(*Song Lin, Benjamin Arai, Dimitrios Gunopulos, Gautam Das*)
- 804  **Fast and Highly-Available Stream Processing Over Wide Area Networks**
(*Jeong-Hyon Hwang, Uğur Çetintemel, Stan Zdonik*)
-

Research Session 24: XML Query Processing

- 814  **Approximate Joins for Data-Centric XML**
(*Nikolaus Augsten, Michael Böhlen, Curtis Dyreson, Johann Gamper*)
- 824  **The Space Complexity of Processing XML Twig Queries Over Indexed Documents**
(*Mirit Shalem, Ziv Bar-Yossef*)
- 833  **XML Index Recommendation with Tight Optimizer Coupling**
(*Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, Fei Chiang, Andrey Balmin, Kevin Beyer, Calisto Zuzarte*)



Research Session 25: Continuous Data Processing

- 843  **A Motion-Aware Approach to Continuous Retrieval of 3D Objects**
(*Mohammed Eunus Ali, Rui Zhang, Egemen Tanin, Lars Kulik*)
- 853  **Continuous Content-Based Copy Detection Over Streaming Videos**
(*Ying Yan, Beng Chin Ooi, Aoying Zhou*)
- 863  **Continuous Intersection Joins Over Moving Objects**
(*Rui Zhang, Dan Lin, Kotagiri Ramamohanarao, Elisa Bertino*)
-

Research Session 26: Query Processing

- 873  **Multiple Materialized View Selection for XPath Query Rewriting**
(*Nan Tang, Jeffrey Xu Yu, M. Tamer Özsu, Byron Choi, Kam-Fai Wong*)
- 883  **Explaining and Reformulating Authority Flow Queries**
(*Ramakrishna Varadarajan, Vagelis Hristidis, Louiqa Raschid*)
- 893  **An Efficient Algorithm for Answering Graph Reachability Queries**
(*Yangjun Chen, Yibin Chen*)
- 903  **Summarizing Graph Patterns**
(*Yong Liu, Jianzhong Li, Hong Gao*)
- 913  **Fast Graph Pattern Matching**
(*Jiefeng Cheng, Jeffrey Xu Yu, Bolin Ding, Philip S. Yu, Haixun Wang*)
- 923  **Stop Chasing Trends: Discovering High Order Models in Evolving Data**
(*Shixi Chen, Haixun Wang, Shuigeng Zhou, Philip S. Yu*)



Research Session 27: Information Extraction and Semantic Search

- 933  **An Algebraic Approach to Rule-Based Information Extraction**
(*Frederick Reiss, Sriram Raghavan, Rajasekar Krishnamurthy, Huaiyu Zhu, Shivakumar Vaithyanathan*)
- 943  **Efficient Information Extraction Over Evolving Text Data**
(*Fei Chen, AnHai Doan, Jun Yang, Raghu Ramakrishnan*)
- 953  **NAGA: Searching and Ranking Knowledge**
(*Gjergji Kasneci, Fabian M. Suchanek, Georgiana Ifrim, Maya Ramanath, Gerhard Weikum*)
-

Research Session 28: Uncertainty and Approximated Data

- 963  **TALE: A Tool for Approximate Large Graph Matching**
(*Yuanyuan Tian, Jignesh M. Patel*)
- 973  **Probabilistic Verifiers: Evaluating Constrained Nearest-Neighbor Queries Over Uncertain Data**
(*Reynold Cheng, Jinchuan Chen, Mohamed F. Mokbel, Chi-Yin Chow*)
- 983  **Fast and Simple Relational Processing of Uncertain Data**
(*Lyublena Antova, Thomas Jansen, Christoph Koch, Dan Olteanu*)



Research Session 29: Database Systems

- 993  **SHAROE: A Data Sharing Platform for Outsourced Enterprise Storage Environments**
(*Aameek Singh, Ling Liu*)
- 1003  **Accelerating Lookups in P2P Systems Using Peer Caching**
(*Supratim Deb, Prakash Linga, Rajeev Rastogi, Anand Srinivasan*)
- 1013  **Diagnosing Estimation Errors in Page Counts Using Execution Feedback**
(*Surajit Chaudhuri, Vivek Narasayya, Ravishankar Ramamurthy*)
- 1023  **Exploiting Lineage for Confidence Computation in Uncertain and Probabilistic Databases**
(*Anish Das Sarma, Martin Theobald, Jennifer Widom*)
- 1033  **Greedy List Intersection**
(*Robert Krauthgamer, Aranyak Mehta, Vijayshankar Raman, Atri Rudra*)
- 1043  **Handling Non-Linear Polynomial Queries Over Dynamic Data**
(*Shetal Shah, Krithi Ramamritham*)
- 1053  **Database Support for Probabilistic Attributes and Tuples**
(*Sarvjeet Singh, Chris Mayfield, Rahul Shah, Sunil Prabhakar, Susanne Hambrusch, Jennifer Neville, Reynold Cheng*)
-

Research Session 30: Data Management in Science

- 1062  **Scalable Rule-Based Gene Expression Data Classification**
(*Mark A. Iwen, Willis Lang, Jignesh M. Patel*)
- 1072  **Querying and Managing Provenance Through User Views in Scientific Workflows**
(*Olivier Biton, Sarah Cohen-Boulakia, Susan B. Davidson, Carmem S. Hara*)
- 1082  **Spatial Outsourcing for Location-Based Services**
(*Yin Yang, Stavros Papadopoulos, Dimitris Papadias, George Kollios*)



Research Session 31: Preference Queries

- 1092  **P-Cube: Answering Preference Queries in Multi-Dimensional Space**
(Dong Xin, Jiawei Han)
- 1101  **Efficient Rewriting Algorithms for Preference Queries**
(Periklis Georgiadis, Ioannis Kapantaidakis, Vassilis Christophides, Elhadji Mamadou Nguer, Nicolas Spyrtas)
- 1111  **A Fast Similarity Join Algorithm Using Graphics Processing Units**
(Michael D. Lieberman, Jagan Sankaranarayanan, Hanan Samet)
-

Research Session 32: Data Streams

- 1121  **MobiQual: QoS-Aware Load Shedding in Mobile CQ Systems**
(Buğra Gedik, Kun-Lung Wu, Philip S. Yu, Ling Liu)
- 1131  **Approximate Clustering on Distributed Data Streams**
(Qi Zhang, Jinze Liu, Wei Wang)
- 1140  **Handling Uncertain Data in Array Database Systems**
(Tingjian Ge, Stan Zdonik)
- 1150  **Just-In-Time Processing of Continuous Queries**
(Yin Yang, Dimitris Papadias)
- 1160  **Online Filtering, Smoothing and Probabilistic Modeling of Streaming Data**
(Bhargav Kanagal, Amol Deshpande)
- 1170  **Butterfly: Protecting Output Privacy in Stream Mining**
(Ting Wang, Ling Liu)
- 1180  **XStream: A Signal-Oriented Data Stream Management System**
(Lewis Girod, Yuan Mei, Ryan Newton, Stanislav Rost, Arvind Thiagarajan, Hari Balakrishnan, Samuel Madden)



Industry Session 1: Advanced Query Processing Techniques

- 1190  **Optimizing Star Join Queries for Data Warehousing in Microsoft SQL Server**
(César A. Galindo-Legaria, Torsten Grabs, Sreenivas Gukal, Steve Herbert, Aleksandras Surna, Shirley Wang, Wei Yu, Peter Zabback, Shin Zhang)
- 1200  **Integration of Server, Storage and Database Stack: Moving Processing Towards Data**
(Lin Qiao, Vijayshankar Raman, Inderpal Narang, Prashant Pandey, David Chambliss, Gene Fuh, James Ruddy, Ying-Lin Chen, Kou-Horng Yang, Fen-Ling Lin)
- 1209  **Accurate and Efficient Inter-Transaction Dependency Tracking**
(Tzi-cker Chiueh, Shweta Bajpai)
-

Industry Session 2: Beyond Queries: Annotation, Scripting, and Inference

- 1219  **Native Web Browser Enabled SVG-Based Collaborative Multimedia Annotation for Medical Images**
(Fusheng Wang, Cornelius Rabsch, Peiya Liu)
- 1229  **XQSE: An XQuery Scripting Extension for the AquaLogic Data Services Platform**
(Vinayak Borkar, Michael Carey, Daniel Engovatov, Dmitry Lychagin, Till Westmann, Warren Wong)
- 1239  **Implementing an Inference Engine for RDFS/OWL Constructs and User-Defined Rules in Oracle**
(Zhe Wu, George Eadon, Souripriya Das, Eugene Inseok Chong, Vladimir Kolovski, Melliyal Annamalai, Jagannathan Srinivasan)



Industry Session 3: Dynamic, Flexible Systems

- 1249  **Towards Declarative Queries on Adaptive Data Structures**
(Nicolas Bruno, Pablo Castro)
- 1259  **Filtered Indices and Their Use in Flexible Schema Scenarios**
(Srini Acharya, César A. Galindo-Legaria, Milind Joshi, Babu Krishnaswamy, Stefano Stefani, Pawel Terlecki)
- 1267  **Maintaining Connectivity in Dynamic Multimodal Network Models**
(Petko Bakalov, Erik Hoel, Wee-Liang Heng, Vassilis J. Tsotras)
-

Industry Session 4: ETL

- 1277  **An Enhanced Extract-Transform-Load System for Migrating Data in Telecom Billing**
(Himanshu Agrawal, Girish Chafle, Sunil Goyal, Sumit Mittal, Sougata Mukherjea)
- 1287  **An Overview of SQL Support in Workflow Products**
(Marko Vrhovnik, Holger Schwarz, Sylvia Radeschütz, Bernhard Mitschang)
- 1297  **A Scalable Scheme for Bulk Loading Large RDF Graphs into Oracle**
(Souripriya Das, Eugene Inseok Chong, Zhe Wu, Melliya Annamalai, Jagannathan Srinivasan)
- 1307  **Orchid: Integrating Schema Mapping and ETL**
(Stefan Dessloch, Mauricio A. Hernández, Ryan Wisnesky, Ahmed Radwan, Jindan Zhou)



Poster Session 1

- 1317  **Merging Hierarchies Using Object Placement**
(Kai Zhao, Robert Ikeda, Hector Garcia-Molina)
- 1319  **Model Management Engine for Data Integration with Reverse-Engineering Support**
(Michael N. Gubanov, Philip A. Bernstein, Alexander Moshchuk)
- 1322  **Extracting Loosely Structured Data Records Through Mining Strict Patterns**
(Yipu Wu, Jing Chen, Qing Li)
- 1325  **Detection of Shape Anomalies: A Probabilistic Approach Using Hidden Markov Models**
(Zheng Liu, Jeffrey Xu Yu, Lei Chen, Di Wu)
- 1328  **Monitoring Network Evolution Using MDL**
(Jure Ferlež, Christos Faloutsos, Jure Leskovec, Dunja Mladenić, Marko Grobelnik)
- 1331  **Database Exploration Using Join Paths**
(Cecilia M. Procopiuc, Divesh Srivastava)
- 1334  **What-If OLAP Queries with Changing Dimensions**
(L.V.S. Lakshmanan, A. Russakovsky, V. Sashikanth)
- 1337  **A General Framework for Fast Co-Clustering on Large Datasets Using Matrix Decomposition**
(Feng Pan, Xiang Zhang, Wei Wang)
- 1340  **MUSE: Multi-Represented Similarity Estimation**
(Hans-Peter Kriegel, Peter Kunath, Alexey Pryakhin, Matthias Schubert)
- 1343  **FLAME: Shedding Light on Hidden Frequent Patterns in Sequence Datasets**
(Sandeep Tata, Jignesh M. Patel)
- 1346  **Mining Search-Phrase Definitions from Item Descriptions**
(Hung V. Nguyen, Hasan Davulcu)



- 1349  **Rights Protection of Trajectory Datasets**
(*Claudio Lucchese, Michail Vlachos, Deepak Rajan, Philip S. Yu*)
-

Poster Session 2

- 1352  **Privacy Preserving Joins**
(*Yaping Li, Minghua Chen*)
- 1355  **Link Privacy in Social Networks**
(*Aleksandra Korolova, Rajeev Motwani, Shubha U. Nabar, Ying Xu*)
- 1358  **Protecting Databases from Query Flood Attacks**
(*Anna Cinzia Squicciarini, Ivan Paloscia, Elisa Bertino*)
- 1361  **Secure Delta-Publishing of XML Content**
(*Mohamed Nabeel, Elisa Bertino*)
- 1364  **On the Optimal Selection of k in the k -Anonymity Problem**
(*Rinku Dewri, Indrajit Ray, Indrakshi Ray, Darrell Whitley*)
- 1367  **Anonymizing Streaming Data for Privacy Protection**
(*Jianzhong Li, Beng Chin Ooi, Weiping Wang*)
- 1370  **Privately Updating Suppression and Generalization Based k -Anonymous Databases**
(*Alberto Trombetta, Wei Jiang, Elisa Bertino, Lorenzo Bossi*)
- 1373  **InstantDB: Enforcing Timely Degradation of Sensitive Data**
(*Nicolas Anciaux, Luc Bouganim, Harold van Heerde, Philippe Pucheral, Peter M.G. Apers*)
- 1376  **CASTLE: A Delay-Constrained Scheme for k_s -Anonymizing Data Streams**
(*Jianneng Cao, Barbara Carminati, Elena Ferrari, Kian Lee Tan*)
- 1379  **Exponentially Decayed Aggregates on Data Streams**
(*Graham Cormode, Flip Korn, Srikanta Tirthapura*)
- 1382  **Efficient Aggregate Computation Over Data Streams**
(*Kanthi Nagaraj, K.V.M. Naidu, Rajeev Rastogi, Scott Satkin*)



- 1385  **Frequency Estimation Over Sliding Windows**
(Linfeng Zhang, Yong Guan)
- 1388  **Online Failure Forecast for Fault-Tolerant Data Stream Processing**
(Xiaohui Gu, Spiros Papadimitriou, Philip S. Yu, Shu-Ping Chang)
-

Poster Session 3

- 1391  **On Supporting Kleene Closure Over Event Streams**
(Daniel Gyllstrom, Jagrati Agrawal, Yanlei Diao, Neil Immerman)
- 1394  **MOOLAP: Towards Multi-Objective OLAP**
(Shyam Antony, Ping Wu, Divyakant Agrawal, Amr El Abbadi)
- 1397  **On Skylining with Flexible Dominance Relation**
(Tian Xia, Donghui Zhang, Yufei Tao)
- 1400  **Designing Random Sample Synopses with Outliers**
(Philipp Rösch, Rainer Gemulla, Wolfgang Lehner)
- 1403  **Efficiently Answering Probabilistic Threshold Top-k Queries on Uncertain Data**
(Ming Hua, Jian Pei, Wenjie Zhang, Xuemin Lin)
- 1406  **Efficient Processing of Top-k Queries in Uncertain Databases**
(Ke Yi, Feifei Li, George Kollios, Divesh Srivastava)
- 1409  **Adjoined Dimension Column Clustering to Improve Data Warehouse Query Performance**
(Xuedong Chen, Patrick O'Neil, Elizabeth O'Neil)
- 1412  **Adaptive Segmentation for Scientific Databases**
(Milena Ivanova, Martin L. Kersten, Niels Nes)
- 1415  **QShuffler: Getting the Query Mix Right**
(Mumtaz Ahmad, Ashraf Aboulmaga, Shivnath Babu, Kamesh Munagala)



-
- 1418  **Adaptive Bitmap Indexes for Space-Constrained Systems**
(*Rishi Rakesh Sinha, Marianne Winslett, Kesheng Wu, Kurt Stockinger, Arie Shoshani*)
- 1421  **Skyline-Based Peer-to-Peer Top-k Query Processing**
(*Akrivi Vlachou, Christos Doulkeridis, Kjetil Nørnvåg, Michalis Vazirgiannis*)
- 1424  **Adaptive Approximate Similarity Searching Through Metric Social Networks**
(*Jan Sedmidubský, Stanislav Bartoň, Vlastislav Dohnal, Pavel Zezula*)
- 1427  **Fixed-Precision Approximate Continuous Aggregate Queries in Peer-to-Peer Databases**
(*Farnoush Banaei-Kashani, Cyrus Shahabi*)
-

Poster Session 4

- 1430  **Faster Join Enumeration for Complex Queries**
(*Guido Moerkotte, Thomas Neumann*)
- 1433  **PermJoin: An Efficient Algorithm for Producing Early Results in Multi-Join Query Plans**
(*Justin J. Levandoski, Mohamed E. Khalefa, Mohamed F. Mokbel*)
- 1436  **Processing Event-Monitoring Queries in Sensor Networks**
(*Vassilis Stoumpos, Antonios Deligiannakis, Yannis Kotidis, Alex Delis*)
- 1439  **An Architecture for Query Optimization in Sensor Networks**
(*Ixent Galpin, Christian Y.A. Brenninkmeijer, Farhana Jabeen, Alvaro A.A. Fernandes, Norman W. Paton*)
- 1442  **Optimizing Complex Event Processing Over RFID Data Streams**
(*Qun Chen, Zhanhuai Li, Hailong Liu*)
- 1445  **Efficient Data Interpretation and Compression Over RFID Streams**
(*Richard Cocci, Thanh Tran, Yanlei Diao, Prashant Shenoy*)
- 1448  **Outlier-Aware Data Aggregation in Sensor Networks**
(*Antonios Deligiannakis, Vassilis Stoumpos, Yannis Kotidis, Vasilis Vassalos, Alex Delis*)



-
- 1451  **Efficient Query Processing in Large Traffic Networks**
(Hans-Peter Kriegel, Peer Kröger, Peter Kunath, Matthias Renz, Tim Schmidt)
- 1454  **Optimal-Nearest-Neighbor Queries**
(Yunjun Gao, Jing Zhang, Gencai Chen, Qing Li, Shen Liu, Chun Chen)
- 1457  **Convoy Queries in Spatio-Temporal Databases**
(Hoyoung Jeung, Heng Tao Shen, Xiaofang Zhou)
- 1460  **On High Dimensional Indexing of Uncertain Data**
(Charu C. Aggarwal, Philip S. Yu)
- 1462  **Pattern Matching Over Cloaked Time Series**
(Xiang Lian, Lei Chen, Jeffrey Xu Yu)
-

Poster Session 5

- 1465  **Satisfying Complex Data Needs using Pull-Based Online Monitoring of Volatile Data Sources**
(Haggai Roitman, Avigdor Gal, Louiqa Raschid)
- 1468  **Processing Diagnosis Queries: A Principled and Scalable Approach**
(Shivnath Babu, Songyun Duan, Kamesh Munagala)
- 1471  **Capturing Approximated Data Delivery Tradeoffs**
(Haggai Roitman, Avigdor Gal, Louiqa Raschid)
- 1474  **Skippy: Enabling Long-Lived Snapshots of the Long-Lived Past**
(Ross Shaull, Liuba Shriru, Hao Xu)
- 1477  **Adjourn State Concurrency Control Avoiding Time-Out Problems in Atomic Commit Protocols**
(Sebastian Obermeier, Stefan Böttcher, Martin Hett, Panos K. Chrysanthis, George Samaras)
- 1480  **Probabilistic Event Extraction from RFID Data**
(Nodira Khoussainova, Magdalena Balazinska, Dan Suciu)



- 1483  **Mobile Filter: Exploring Migration of Filters for Error-Bounded Data Collection in Sensor Networks**
(*Dan Wang, Jianliang Xu, Jiangchuan Liu, Feng Wang*)
- 1486  **PageChaser: A Tool for the Automatic Correction of Broken Web Links**
(*Atsuyuki Morishima, Akiyoshi Nakamizo, Toshinari Iida, Shigeo Sugimoto, Hiroyuki Kitagawa*)
- 1489  **Index Design for Dynamic Personalized PageRank**
(*Amit Pathak, Soumen Chakrabarti, Manish Gupta*)
- 1492  **Quality-Aware Retrieval of Data Objects from Autonomous Sources for Web-Based Repositories**
(*Houtan Shirani-Mehr, Chen Li, Gang Liang, Michal Shmueli-Scheuer*)
- 1495  **Efficient Discovery of Authoritative Resources**
(*Ravi Kumar, Kevin Lang, Cameron Marlow, Andrew Tomkins*)
- 1498  **Automatically Extracting Form Labels**
(*Hoa Nguyen, Eun Yong Kang, Juliana Freire*)
-

Poster Session 6

- 1501  **Answering Keyword Queries on XML Using Materialized Views**
(*Ziyang Liu, Yi Chen*)
- 1504  **An Inflationary Fixed Point Operator in XQuery**
(*Loredana Afanasiev, Torsten Grust, Maarten Marx, Jan Rittinger, Jens Teubner*)
- 1507  **Grouping and Optimization of XPath Expressions in System RX**
(*Andrey Balmin, Fatma Özcan, Ashutosh Singh, Edison Ting*)
- 1510  **Hierarchical Indexing Approach to Support XPath Queries**
(*Nan Tang, Jeffrey Xu Yu, M. Tamer Özsu, Kam-Fai Wong*)



- 1513  **A Query Processing Architecture for an XML Data Warehouse**
(*Nuwee Wiwatwattana, H.V. Jagadish*)
- 1516  **A Clustered Index Approach to Distributed XPath Processing**
(*Georgia Koloniari, Evaggelia Pitoura*)
- 1519  **AxPRE Summaries: Exploring the (Semi-)Structure of XML Web Collections**
(*Mariano P. Consens, Flavio Rizzolo, Alejandro A. Vaisman*)
- 1522  **Correlation-Based Attribute Outlier Detection in XML**
(*Judice L.Y. Koh, Mong Li Lee, Wynne Hsu, Wee Tiong Ang*)
- 1525  **Ontology-Aware Search on XML-Based Electronic Medical Records**
(*Fernando Farfán, Vagelis Hristidis, Anand Ranganathan, Redmond P. Burke*)
- 1528  **Query-Aware Partitioning for Monitoring Massive Network Data Streams**
(*Theodore Johnson, S. Muthukrishnan, Vladislav Shkapenyuk, Oliver Spatscheck*)
- 1531  **Optimizing Hierarchical Access in OLAP Environment**
(*Lipyeow Lim, Bishwaranjan Bhattacharjee*)
- 1534  **Automatic Result Verification for the Functional Testing of a Query Language**
(*Carsten Binnig, Donald Kossmann, Eric Lo, Angel Saenz-Badillos*)
- 1537  **Improving Information Access for a Community of Practice Using Business Process as Context**
(*Yu Deng, Murthy Devarakonda, Nithya Rajamani, Wlodek Zadrozny*)



Demo Session 1

- 1540  **DescribeX: Interacting with AxPRE Summaries**
(M.S. Ali, Mariano P. Consens, Shahan Khatchadourian, Flavio Rizzolo)
- 1544  **COSTA: Adaptive Indexing for Terms in a Large-Scale Distributed System**
(Aoying Zhou, Rong Zhang, Quang Hieu Vu, Weining Qian)
- 1548  **IMPrECISE: Good-is-Good-Enough Data Integration**
(Ander de Keijzer, Maurice van Keulen)
- 1552  **SPARK: A Keyword Search Engine on Relational Databases**
(Yi Luo, Wei Wang, Xuemin Lin)
- 1556  **Astoria: A Programming Model for Data on the Web**
(Pablo Castro, Anil Nori)
- 1560  **Environmental Tomography: Ubiquitous Sensing with Mobile Devices**
(Stacy Patterson, Bassam Bamieh, Amr El Abbadi)
- 1564  **OptimAX: Efficient Support for Data-Intensive Mash-Ups**
(Serge Abiteboul, Ioana Manolescu, S. Zoupanos)
- 1568  **Graphitti: An Annotation Management System for Heterogeneous Objects**
(Sandeep Gupta, Christopher Condit, Amarnath Gupta)



Demo Session 2

- 1572  **Distributed Monitoring of Peer-to-Peer Systems**
(*Serge Abiteboul, Bogdan Marinoiu, Pierre Bourhis*)
- 1576  **SMART: A System for Online Monitoring Large Volumes of Network Traffic**
(*Aoying Zhou, Ying Yan, Xueqing Gong, Jianlong Chang, Dai Dai*)
- 1580  **Toward Simulation-Based Optimization in Data Stream Management Systems**
(*Christoph Heinz, Jürgen Krämer, Tobias Riemenschneider, Bernhard Seeger*)
- 1584  **PGM/F: A Framework for the Optimization of Data Processing in Business Processes**
(*Marko Vrhovnik, Oliver Suhre, Stephan Ewen, Holger Schwarz*)
- 1588  **Fuzzy Multi-Dimensional Search in the Wayfinder File System**
(*Christopher Peery, Wei Wang, Amélie Marian, Thu D. Nguyen*)
- 1592  **MED: A Multimedia Event Database for 3D Crime Scene Representation and Analysis**
(*Marcin Kwietniewski, Stephanie Wilson, Anna Topol, Sunbir Gill, Jarek Gryz, Michael Jenkin, Piotr Jasiobedzki, Ho-Kong Ng*)
- 1596  **DIPBench Toolsuite: A Framework for Benchmarking Integration Systems**
(*Matthias Böhm, Dirk Habich, Wolfgang Lehner, Uwe Wloka*)



Demo Session 3

- 1600  **Managing Biological Data Using bdbms**
(*Mohamed Y. Eltabakh, Mourad Ouzzani, Walid G. Aref, Ahmed K. Elmagarmid, Yasin Laura-Silva, Muhammad U. Arshad, David Salt, Ivan Baxter*)
- 1604  **Mining Gene Expression Database for Primary Human Disease Tissues**
(*Andrew Campen, Yuni Xia, Dan Rigsby, Ying Guo, Xingdong Feng, Eric W. Su, Mathew Palakal, Shuyu Li*)
- 1608  **Mining Views: Database Views for Data Mining**
(*Hendrik Blockeel, Toon Calders, Elisa Fromont, Bart Goethals, Adriana Prado*)
- 1612  **PictureBook: A Text-and-Image Summary System for Web Search Result**
(*Mei Wang, Hongtao Xu, Guoyu Hao, Xiangdong Zhou, Wei Wang, Qi Zhang, Baile Shi*)
- 1616  **Enhanced Business Intelligence Using EROCS**
(*M. Bhide, V. Chakravarthy, A. Gupta, H. Gupta, M. Mohania, K. Puniyani, P. Roy, S. Roy, V. Sengar*)
- 1620  **T-Time: Threshold-Based Data Mining on Time Series**
(*Johannes Aßfalg, Hans-Peter Kriegel, Peer Kröger, Peter Kunath, Alexey Pryakhin, Matthias Renz*)
- 1624  **RAD: A Scalable Framework for Annotator Development**
(*Sanjeet Khaitan, Ganesh Ramakrishnan, Sachindra Joshi, Anup Chalamalla*)
- 1628  **SPOT: A System for Detecting Projected Outliers from High-Dimensional Data Streams**
(*Ji Zhang, Qigang Gao, Hai Wang*)