<table>
<thead>
<tr>
<th>Topic</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keynotes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Stream Mining and Resource Adaptive Computation</td>
<td>Philip S. Yu</td>
<td>1</td>
</tr>
<tr>
<td>Purpose Based Access Control for Privacy Protection in Database Systems</td>
<td>Elisa Bertino</td>
<td>2</td>
</tr>
<tr>
<td>Complex Networks and Network Data Mining</td>
<td>Deyi Li</td>
<td>3</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indexing DNA Sequences Using q-Grams</td>
<td>Xia Cao, Shuai Cheng Li, Anthony K.H. Tung</td>
<td>4</td>
</tr>
<tr>
<td>PADS: Protein Structure Alignment Using Directional Shape Signatures</td>
<td>S. Alireza Aghili, Divyakant Agrawal, Amr El Abbadi</td>
<td>17</td>
</tr>
<tr>
<td>LinkageTracker: A Discriminative Pattern Tracking Approach to</td>
<td>Li Lin, Limsoon Wong, Tzeyun Leong, Pohsan Lai</td>
<td>30</td>
</tr>
<tr>
<td>Linkage Disequilibrium Mapping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watermarking and Encryption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Query Optimization in Encrypted Database Systems</td>
<td>Hakan Hacigümüş, Bala Iyer, Sharad Mehrotra</td>
<td>43</td>
</tr>
<tr>
<td>Watermarking Spatial Trajectory Database</td>
<td>Xiaoming Jin, Zhihao Zhang, Jianmin Wang, Deyi Li</td>
<td>56</td>
</tr>
<tr>
<td>Effective Approaches for Watermarking XML Data</td>
<td>Wilfred Ng, Ho-Lam Lau</td>
<td>68</td>
</tr>
<tr>
<td>XML Query Processing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Unifying Framework for Merging and Evaluating XML Information</td>
<td>Ho-Lam Lau, Wilfred Ng</td>
<td>81</td>
</tr>
</tbody>
</table>
Efficient Evaluation of Partial Match Queries for XML Documents
Using Information Retrieval Techniques
Young-Ho Park, Kyu-Young Whang, Byung Suk Lee,
Wook-Shin Han .......................................................... 95

PathStack*: A Holistic Path Join Algorithm for Path Query with
Not-Predicates on XML Data
Enhua Jiao, Tok Wang Ling, Chee-Yong Chan .................. 113

XML Coding and Metadata Management

An Improved Prefix Labeling Scheme: A Binary String Approach for
Dynamic Ordered XML
Changqing Li, Tok Wang Ling .................................... 125

Efficiently Coding and Indexing XML Document
Zhongming Han, Congting Xi, Jiajin Le .......................... 138

XQuery-Based TV-Anytime Metadata Management
Jong-Hyun Park, Byung-Kyu Kim, Yong-Hee Lee, Min-Woo Lee,
Min-Ok Jung, Ji-Hoon Kang ................................. 151

Data Mining

Effective Database Transformation and Efficient Support Computation
for Mining Sequential Patterns
Chung-Wen Cho, Yi-Hung Wu, Arbee L.P. Chen ............... 163

Mining Succinct Systems of Minimal Generators of Formal Concepts
Guozhu Dong, Chunyu Jiang, Jian Pei, Jinyan Li,
Limsoon Wong ............................................................. 175

A General Approach to Mining Quality Pattern-Based Clusters from
Microarray Data
Daxin Jiang, Jian Pei, Aidong Zhang ......................... 188

Data Generation and Understanding

Real Datasets for File-Sharing Peer-to-Peer Systems
Shen Tat Goh, Panos Kalnis, Spiridon Bakiras,
Kian-Lee Tan .......................................................... 201

SemEQUAL: Multilingual Semantic Matching in Relational Systems
A. Kumaran, Jayant R. Haritsa .................................. 214
A Metropolis Sampling Method for Drawing Representative Samples from Large Databases

Hong Guo, Wen-Chi Hou, Feng Yan, Qiang Zhu ........................................ 226

Panel

Stay Current and Relevant in Data Mining Research

Haixun Wang, Wei Wang ................................................................. 239

Music Retrieval

An Efficient Approach to Extracting Approximate Repeating Patterns in Music Databases

Ning-Han Liu, Yi-Hung Wu, Arbee L.P. Chen ................................. 240

On Efficient Music Genre Classification

Jialie Shen, John Shepherd, Anne H.H Ngu ................................. 253

Effectiveness of Note Duration Information for Music Retrieval

Iman S.H. Suyoto, Alexandra L. Uitdenbogerd .......................... 265

Query Processing in Subscription Systems

A Self-Adaptive Model to Improve Average Response Time of Multiple-Event Filtering for Pub/Sub System

Botao Wang, Wang Zhang, Masaru Kitsuregawa .......................... 276

Filter Indexing: A Scalable Solution to Large Subscription Based Systems

Wanxia Xie, Shamkant B. Navathe, Sushil K. Prasad .......................... 288

Caching Strategies for Push-Based Broadcast Considering Consecutive Data Accesses with Think-Time

Wataru Uchida, Takahiro Hara, Shojiro Nishio .......................... 300

Extending XML

XDO2: A Deductive Object-Oriented Query Language for XML

Wei Zhang, Tok Wang Ling, Zhuo Chen, Gillian Dobbie ................. 311

Extending XML with Nonmonotonic Multiple Inheritance

Guoren Wang, Mengchi Liu .............................................................. 323
### Table of Contents

#### Database Design with Equality-Generating Dependencies
*Junhu Wang* .................................................. 335

#### Web Services

WDEE: Web Data Extraction by Example  
*Zhao Li, Wee Kong Ng* ........................................... 347

Concept-Based Retrieval of Alternate Web Services  
*Dunlu Peng, Sheng Huang, Xiaoling Wang, Aoying Zhou* .......... 359

WSQuery: XQuery for Web Services Integration  
*Zhimao Guo, Xiaoling Wang, Aoying Zhou* ....................... 372

#### High-Dimensional Indexing

A New Indexing Method for High Dimensional Dataset  
*Jiyuan An, Yi-Ping Phoebe Chen, Qinying Xu, Xiaofang Zhou* ...... 385

BM⁺-Tree: A Hyperplane-Based Index Method for High-Dimensional Metric Spaces  
*Xiangmin Zhou, Guoren Wang, Xiaofang Zhou, Ge Yu* ............. 398

Approaching the Efficient Frontier: Cooperative Database Retrieval Using High-Dimensional Skylines  
*Wolf-Tilo Balke, Jason Xin Zheng, Ulrich Güntzer* ............... 410

#### Sensor and Stream Data Processing

False-Negative Frequent Items Mining from Data Streams with Bursting  
*Zhihong Chong, Jeffrey Xu Yu, Hongjun Lu, Zhengjie Zhang, Aoying Zhou* ................................................. 422

Adaptively Detecting Aggregation Bursts in Data Streams  
*Aoying Zhou, Shouke Qin, Weining Qian* ........................ 435

Communication-Efficient Implementation of Join in Sensor Networks  
*Vishal Chowdhary, Himanshu Gupta* .............................. 447

#### Database Performance Issues

Zoned-RAID for Multimedia Database Servers  
*Ali E. Dashti, Seon Ho Kim, Roger Zimmermann* .................. 461
Randomized Data Allocation in Scalable Streaming Architectures
  Kun Fu, Roger Zimmermann ........................................ 474

Trace System of iSCSI Storage Access and Performance Improvement
  Saneyasu Yamaguchi, Masato Oguchi, Masaru Kitsuregawa .......... 487

CoCACHE: Query Processing Based on Collaborative Caching in P2P
  Systems
  Weining Qian, Linhao Xu, Shuigeng Zhou, Aoying Zhou .......... 498

Clustering, Classification and Data Warehouses

Multi-represented kNN-Classification for Large Class Sets
  Hans-Peter Kriegel, Alexey Pryakhin, Matthias Schubert .......... 511

Enhancing SNNB with Local Accuracy Estimation and Ensemble
  Techniques
  Zhipeng Xie, Qing Zhang, Wynne Hsu, Mong Li Lee ............... 523

MMPClust: A Skew Prevention Algorithm for Model-Based Document
  Clustering
  Xiaoguang Li, Ge Yu, Daling Wang .................................. 536

Designing and Using Views to Improve Performance of Aggregate
  Queries
  Foto Afrati, Rada Chirkova, Shalu Gupta, Charles Loftis .......... 548

Large Relations in Node-Partitioned Data Warehouses
  Pedro Furtado .......................................................... 555

Data Mining and Web Data Processing

Mining Frequent Tree-Like Patterns in Large Datasets
  Tzung-Shi Chen, Shih-Chun Hsu ....................................... 561

An Efficient Approach for Mining Fault-Tolerant Frequent Patterns
  Based on Bit Vector Representations
  Jia-Ling Koh, Pei-Wy Yo ............................................... 568

NNF: An Effective Approach in Medicine Paring Analysis of Traditional
  Chinese Medicine Prescriptions
  Chuan Li, Changjie Tang, Jing Peng, Jianjun Hu,
  Yongguang Jiang, Xiaojia Yong ...................................... 576
## Table of Contents

From XML to Semantic Web  
*Changqing Li, Tok Wang Ling* ........................................ 582

A Hybrid Approach for Refreshing Web Page Repositories  
*Mohammad Ghodsi, Oktie Hassanzadeh, Shahab Kamali, Morteza Monemizadeh* ........................................ 588

Schema Driven and Topic Specific Web Crawling  
*Qi Guo, Hang Guo, Zhiqiang Zhang, Jing Sun, Jianhua Feng* ............ 594

### Moving Object Databases

Towards Optimal Utilization of Main Memory for Moving Object Indexing  
*Bin Cui, Dan Lin, Kian-Lee Tan* ........................................ 600

Aqua: An Adaptive QUery-Aware Location Updating Scheme for Mobile Objects  
*Jing Zhou, Hong Va Leong, Qin Lu, Ken C.K. Lee* ...................... 612

A Spatial Index Using MBR Compression and Hashing Technique for Mobile Map Service  
*Jin-Deog Kim, Sang-Ho Moon, Jin-Oh Choi* ............................ 625

### Temporal Databases

Indexing and Querying Constantly Evolving Data Using Time Series Analysis  
*Yuni Xia, Sunil Prabhakar, Jianzhong Sun, Shan Lei* ................. 637

Mining Generalized Spatio-Temporal Patterns  
*Junmei Wang, Wynne Hsu, Mong Li Lee* .............................. 649

Exploiting Temporal Correlation in Temporal Data Warehouses  
*Ying Feng, Hua-Gang Li, Divyakant Agrawal, Amr El Abbadi* .......... 662

### Semantics

Semantic Characterization of Real World Events  
*Aparna Nagargadde, Sridhar Varadarajan, Krithi Ramamritham* .... 675

Learning Tree Augmented Naive Bayes for Ranking  
*Liangxiao Jiang, Harry Zhang, Zhihua Cai, Jiang Su* ................. 688
   Peixiang Zhao, Ming Zhang, Dongqing Yang, Shiwei Tang ............... 699

XML Update and Query Patterns

XANDY: Detecting Changes on Large Unordered XML Documents Using Relational Databases
   Erwin Leonardi, Sourav S. Bhowmick, Sanjay Madria ................. 711

FASST Mining: Discovering Frequently Changing Semantic Structure from Versions of Unordered XML Documents
   Qiankun Zhao, Sourav S. Bhowmick ..................................... 724

Mining Positive and Negative Association Rules from XML Query Patterns for Caching
   Ling Ghen, Sourav S. Bhowmick, Liang-Tien Chia ....................... 736

Join Processing and View Management

Distributed Intersection Join of Complex Interval Sequences
   Hans-Peter Kriegel, Peter Kunath, Martin Pfeifle, Matthias Renz .... 748

Using Prefix-Trees for Efficiently Computing Set Joins
   Ravindranath Jampani, Vikram Pudi .................................... 761

Maintaining Semantics in the Design of Valid and Reversible SemiStructured Views
   Ya Bing Chen, Tok Wang Ling, Mong Li Lee ............................ 773

Spatial Databases

DCbot: Finding Spatial Information on the Web
   Mihály Jakob, Matthias Grossmann, Daniela Nicklas, Bernhard Mitschang ............................................. 779

Improving Space-Efficiency in Temporal Text-Indexing
   Kjetil Nørvåg, Albert Overskeid Nybø .................................. 791

Nearest Neighbours Search Using the PM-Tree
   Tomáš Skopal, Jaroslav Pokorný, Vášclav Snášel ....................... 803
## Enhancing Database Services

Deputy Mechanism for Workflow Views  
*Zhe Shan, Qing Li, Yi Luo, Zhiyong Peng* .............................. 816

Automatic Data Extraction from Data-Rich Web Pages  
*Dongdong Hu, Xiaofeng Meng* ........................................... 828

Customer Information Visualization via Customer Map  
*Ji Young Woo, Sung Min Bae, Chong Un Pyon, Sang Chan Park* .... 840

Finding and Analyzing Database User Sessions  
*Qingsong Yao, Aijun An, Xiangji Huang* ............................. 851

## Recovery and Correctness

Time-Cognizant Recovery Processing for Embedded Real-Time Databases  
*Guoqiong Liao, Yunsheng Liu, Yingyuan Xiao* .......................... 863

An Efficient Phantom Protection Method for Multi-dimensional Index Structures  
*Seok Il Song, Seok Jae Lee, Tae Ho Kang, Jae Soo Yoo* ............. 875

CMC: Combining Multiple Schema-Matching Strategies Based on Credibility Prediction  
*KeWei Tu, Yong Yu* .................................................. 888

## XML Databases and Indexing

Translating XQuery to SQL Based on Query Forests  
*Ya-Hui Chang, Greg Liu, Sue-Shain Wu* ............................... 894

A New Indexing Structure to Speed Up Processing XPath Queries  
*Jeong Hee Hwang, Van Trang Nguyen, Keun Ho Ryu* .................. 900

Translate Graphical XML Query Language to SQLX  
*Wei Ni, Tok Wang Ling* ............................................... 907

GTree: An Efficient Grid-Based Index for Moving Objects  
*Xiaoyuan Wang, Qing Zhang, Weiwei Sun* ............................ 914