Table of Contents

Modeling and ETL

An Analytics-Aware Conceptual Model for Evolving Graphs ............... 1
Amine Ghrib, Sabri Skhiri, Salim Jouili, and Esteban Zimányi

Declarative Expression and Optimization of Data-Intensive Flows ...... 13
Georgia Kougka and Anastasios Gounaris

Tabular Web Data: Schema Discovery and Integration ......................... 26
Prudhvi Janga and Karen C. Davis

Query Optimization and Parallelism

Uncoupled MapReduce: A Balanced and Efficient Data Transfer Model .................................................. 34
Jie Zhang, Maosen Sun, Jian Lin, and Li Zha

Efficient Evaluation of Ad-Hoc Range Aggregates .......................... 46
Christian Ammendola, Michael H. Böhlen, and Johann Gamper

SPIN: Concurrent Workload Scaling over Data Warehouses ................. 60
João Pedro Costa and Pedro Furtado

Spatial Data Warehouses and Applications

Lily: A Geo-Enhanced Library for Location Intelligence ..................... 72
Matteo Golfarelli, Marco Mantovani, Federico Ravaldi, and Stefano Rizzi

Discovering Co-location Patterns in Datasets with Extended Spatial Objects .................................................. 84
Aibek Adilmagambetov, Osmar R. Zaiane, and Alvaro Osornio-Vargas

SOLAP4epidemiologist: A Spatial Data Warehousing Application in Epidemiology Domain ........................................... 97
Assantina Cembalo, Michele Ferrucci, Francesca Maria Pisano, and Gianpaolo Pigliasco
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text Mining and OLAP</td>
<td></td>
</tr>
<tr>
<td>Document Difficulty Framework for Semi-automatic Text</td>
<td>110</td>
</tr>
<tr>
<td><strong>Miguel Martinez-Alvarez, Alejandro Bellogin, and Thomas Roeleke</strong></td>
<td></td>
</tr>
<tr>
<td>Knowledge-Free Table Summarization</td>
<td>122</td>
</tr>
<tr>
<td><strong>Dino Ienco, Yoann Pitarch, Pascal Poncelet, and Maguelonne Teisseire</strong></td>
<td></td>
</tr>
<tr>
<td>Recommendation and Prediction</td>
<td></td>
</tr>
<tr>
<td>Predicting Your Next OLAP Query Based on Recent Analytical Sessions</td>
<td>134</td>
</tr>
<tr>
<td>**Marie-Aude Aufaure, Nicolas Kuchmann-Beauper, Patrick Marcel,</td>
<td></td>
</tr>
<tr>
<td><strong>Stefano Rizzi, and Yves Vanrompay</strong></td>
<td></td>
</tr>
<tr>
<td>Where Will You Go? Mobile Data Mining for Next Place Prediction</td>
<td>146</td>
</tr>
<tr>
<td><strong>Joao Birtolo Gomes, Clifton Phua, and Shonali Krishnaswamy</strong></td>
<td></td>
</tr>
<tr>
<td>An Extended Local Hierarchical Classifier for Prediction of Protein</td>
<td>159</td>
</tr>
<tr>
<td>and Gene Functions</td>
<td></td>
</tr>
<tr>
<td><strong>Luiz Henrique de Campos Merschmann and Alex Alves Freitas</strong></td>
<td></td>
</tr>
<tr>
<td>Data Mining Optimization and Machine Learning Techniques</td>
<td></td>
</tr>
<tr>
<td>Supervised Dimensionality Reduction via Nonlinear Target Estimation</td>
<td>172</td>
</tr>
<tr>
<td><strong>Josif Grabocka, Lucas Drumond, and Lars Schmidt-Thieme</strong></td>
<td></td>
</tr>
<tr>
<td>Concurrent Execution of Data Mining Queries for Spatial Collocation</td>
<td>184</td>
</tr>
<tr>
<td>Pattern Discovery</td>
<td></td>
</tr>
<tr>
<td><strong>Pawel Boinski and Maciej Zakrzewicz</strong></td>
<td></td>
</tr>
<tr>
<td>Depth-First Traversal over a Mirrored Space for Non-redundant</td>
<td>196</td>
</tr>
<tr>
<td>Discriminative Itemsets</td>
<td></td>
</tr>
<tr>
<td><strong>Yoshitaka Kameya and Hiroki Asaoka</strong></td>
<td></td>
</tr>
<tr>
<td>Mining and Processing Data Stream</td>
<td></td>
</tr>
<tr>
<td>Stream Mining of Frequent Patterns from Delayed Batches of Uncertain</td>
<td>209</td>
</tr>
<tr>
<td>Data</td>
<td></td>
</tr>
<tr>
<td><strong>Fan Jiang and Carson Kai-Sang Leung</strong></td>
<td></td>
</tr>
<tr>
<td>Fast Causal Network Inference over Event Streams</td>
<td>222</td>
</tr>
<tr>
<td><strong>Saurav Acharya and Byung Suk Lee</strong></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>SSCJ: A Semi-Stream Cache Join Using a Front-Stage Cache Module</td>
<td>236</td>
</tr>
<tr>
<td>M. Asif Naeem, Gerald Weber, Gillian Dobbie, and Christof Lutteroth</td>
<td></td>
</tr>
<tr>
<td><strong>Clustering and Data Mining Applications</strong></td>
<td></td>
</tr>
<tr>
<td>Metrics to Support the Evaluation of Association Rule Clustering</td>
<td>248</td>
</tr>
<tr>
<td>Veronica Oliveira de Carvalho, Fabiano Fernandes dos Santos, and</td>
<td></td>
</tr>
<tr>
<td>Solange Oliveira Rezende</td>
<td></td>
</tr>
<tr>
<td>An Automatic Email Management Approach Using Data Mining Techniques</td>
<td>260</td>
</tr>
<tr>
<td>Gunjan Soni and C.I. Ezeife</td>
<td></td>
</tr>
<tr>
<td>A Data Mining-Based Wind Power Forecasting Method: Results for Wind</td>
<td>268</td>
</tr>
<tr>
<td>Power Plants in Turkey</td>
<td></td>
</tr>
<tr>
<td>Mehmet Barış Özkan, Dilek Küçük, Erman Terciyanlı,</td>
<td></td>
</tr>
<tr>
<td>Serkan Buhan, Turan Demirci, and Pinar Karagoz</td>
<td></td>
</tr>
<tr>
<td>Disease Occurrence Prediction Based on Spatio-temporal Characterization: A Mesoscale Study for Knowledge and Pattern Discovery</td>
<td>277</td>
</tr>
<tr>
<td>Vipul Raheja and K.S. Rajan</td>
<td></td>
</tr>
<tr>
<td>Population Estimation Mining Using Satellite Imagery</td>
<td>285</td>
</tr>
<tr>
<td>Kwankamon Dittakan, Frans Coenen, Rob Christley, and Mayra Wardeh</td>
<td></td>
</tr>
<tr>
<td><strong>Social Network and Graph Mining</strong></td>
<td></td>
</tr>
<tr>
<td>GMAC: A Seed-Insensitive Approach to Local Community Detection</td>
<td>297</td>
</tr>
<tr>
<td>Lianhang Ma, Hao Huang, Qiming He, Kevin Chiew, Jianan Wu, and Yanzhe Che</td>
<td></td>
</tr>
<tr>
<td>Finding Maximal Overlapping Communities</td>
<td>309</td>
</tr>
<tr>
<td>Eileen H.-C. Wei, Yun Sing Koh, and Gillian Dobbie</td>
<td></td>
</tr>
<tr>
<td>Minimal Vertex Unique Labelled Subgraph Mining</td>
<td>317</td>
</tr>
<tr>
<td>Wen Yu, Frans Coenen, Michele Zito, and Subhieh El Sallhi</td>
<td></td>
</tr>
<tr>
<td><strong>Event Sequence and Web Mining</strong></td>
<td></td>
</tr>
<tr>
<td>A Hybrid Graph-Based Method for Concept Rule Discovery</td>
<td>327</td>
</tr>
<tr>
<td>Alev Mutlu and Pinar Karagoz</td>
<td></td>
</tr>
<tr>
<td>Applications of Concurrent Access Patterns in Web Usage Mining</td>
<td>339</td>
</tr>
<tr>
<td>Jing Lu, Malcolm Keech, and Cuiqing Wang</td>
<td></td>
</tr>
</tbody>
</table>
Cell-at-a-Time Approach to Lazy Evaluation of Dimensional Aggregations .......................... 349
   Peter Thanisch, Jyrki Nummenmaa, Tapio Niemi, and Marko Niinimäki

MAF: A Method for Detecting Temporal Associations from Multiple Event Sequences .................. 359
   Han Liang

Author Index ......................................................... 373