

Table of Contents

Invited Papers

- The DBLP Computer Science Bibliography:
Evolution, Research Issues, Perspectives1
Michael Ley
- From Searching Text to Querying XML Streams 11
Dan Suciu
- String Matching Problems from Bioinformatics
Which Still Need Better Solutions (Extended Abstract) 27
Gaston H. Gonnet

String Matching

- Optimal Exact String Matching Based on Suffix Arrays 31
Mohamed Ibrahim Abouelhoda, Enno Ohlebusch, and Stefan Kurtz
- Faster String Matching with Super-Alphabets 44
Kimmo Fredriksson
- On the Size of DASG for Multiple Texts 58
Maxime Crochemore and Zdeněk Troníček

String Processing

- Sorting by Prefix Transpositions 65
Zanoni Dias and João Meidanis
- Efficient Computation of Long Similar Subsequences 77
Abdullah N. Arslan and Ömer Eğecioğlu
- Stemming Galician Texts 91
*Nieves R. Brisaboa, Carlos Callón, Juan-Ramón López,
Ángeles S. Places, and Goretti Sanmartín*
- Firing Policies for an Arabic Rule-Based Stemmer 98
Imad A. Al Sughaiyer and Ibrahim A. Al Kharashi

Web Ranking and Link Analysis

- Enhancing the Set-Based Model Using Proximity Information 104
Bruno Póssas, Nivio Ziviani, and Wagner Meira, Jr.

Web Structure, Dynamics and Page Quality 117
Ricardo Baeza-Yates, Felipe Saint-Jean, and Carlos Castillo

A Theoretical Analysis of Google's PageRank 131
Luca Pretto

Machine Learning Approach for Homepage Finding Task 145
Wensi Xi, Edward A. Fox, Roy P. Tan, and Jiang Shu

Pattern Matching

Tree Pattern Matching for Linear Static Terms 160
Cedric Chauve

Processing Text Files as Is: Pattern Matching over Compressed Texts,
Multi-byte Character Texts, and Semi-structured Texts 170
*Masayuki Takeda, Satoru Miyamoto, Takuya Kida, Ayumi Shinohara,
Shuichi Fukamachi, Takeshi Shinohara, and Setsuo Arikawa*

Pattern Matching over Multi-attribute Data Streams 187
Lilian Harada

Digital Libraries and Applications

Java MARIAN: From an OPAC to a Modern Digital Library System 194
*Marcos André Gonçalves, Paul Mather, Jun Wang, Ye Zhou, Ming Luo,
Ryan Richardson, Rao Shen, Liang Xu, and Edward A. Fox*

A Framework for Generating Attribute Extractors for Web Data Sources .. 210
*Davi de Castro Reis, Robson Braga Araújo, Altigran S. da Silva,
and Berthier A. Ribeiro-Neto*

Multiple Example Queries in Content-Based Image Retrieval 227
Seyed M.M. Tahaghoghi, James A. Thom, and Hugh E. Williams

Focussed Structured Document Retrieval 241
Gabrialla Kazai, Mounia Lalmas, and Thomas Roelleke

Towards a More Comprehensive Comparison of Collaborative Filtering
Algorithms 248
*Cristina N. González-Caro, Maritza L. Calderón-Benavides,
José de J. Pérez-Alcázar, Juan C. García-Díaz, and Joaquín Delgado*

Approximate Searching

Fully Dynamic Spatial Approximation Trees 254
Gonzalo Navarro and Nora Reyes

String Matching with Metric Trees Using an Approximate Distance	271
<i>Ilaria Bartolini, Paolo Ciaccia, and Marco Patella</i>	
Probabilistic Proximity Searching Algorithms Based on Compact Partitions	284
<i>Benjamin Bustos and Gonzalo Navarro</i>	
 Indexing Techniques	
t -Spanners as a Data Structure for Metric Space Searching	298
<i>Gonzalo Navarro, Rodrigo Paredes, and Edgar Chávez</i>	
Compact Directed Acyclic Word Graphs for a Sliding Window	310
<i>Shunsuke Inenaga, Ayumi Shinohara, Masayuki Takeda, and Setsuo Arikawa</i>	
Indexing Text Using the Ziv-Lempel Trie	325
<i>Gonzalo Navarro</i>	
 Author Index	 337