

Table of Contents

“Search Is a Solved Problem” and Other Annoying Fallacies (Invited Talk)	1
<i>David Hawking</i>	
Approximate Runs – Revisited (Invited Talk)	2
<i>Gad M. Landau</i>	
Engineering Radix Sort for Strings	3
<i>Juha Kärkkäinen and Tommi Rantala</i>	
Faster Text Fingerprinting	15
<i>Roman Kolpakov and Mathieu Raffinot</i>	
Context-Sensitive Grammar Transform: Compression and Pattern Matching	27
<i>Shirou Maruyama, Yohei Tanaka, Hiroshi Sakamoto, and Masayuki Takeda</i>	
Improved Variable-to-Fixed Length Codes	39
<i>Shmuel T. Klein and Dana Shapira</i>	
Term Impacts as Normalized Term Frequencies for BM25 Similarity Scoring	51
<i>Vo Ngoc Anh, Raymond Wan, and Alistair Moffat</i>	
The Effect of Weighted Term Frequencies on Probabilistic Latent Semantic Term Relationships	63
<i>Laurence A.F. Park and Kotagiri Ramamohanarao</i>	
Comparison of s-gram Proximity Measures in Out-of-Vocabulary Word Translation	75
<i>Anni Järvelin and Antti Järvelin</i>	
Speeding Up Pattern Matching by Text Sampling	87
<i>Francisco Claude, Gonzalo Navarro, Hannu Peltola, Leena Salmela, and Jorma Tarhio</i>	
Mismatch Sampling	99
<i>Raphaël Clifford, Klim Efremenko, Benny Porat, Ely Porat, and Amir Rothschild</i>	
Sliding CDAWG Perfection	109
<i>Martin Senft and Tomáš Dvořák</i>	
Self-indexing Natural Language	121
<i>Nieves R. Brisaboa, Antonio Fariña, Gonzalo Navarro, Angeles S. Places, and Eduardo Rodríguez</i>	

New Perspectives on the Prefix Array	133
<i>W.F. Smyth and Shu Wang</i>	
Indexed Hierarchical Approximate String Matching.....	144
<i>Luís M.S. Russo, Gonzalo Navarro, and Arlindo L. Oliveira</i>	
An Efficient Linear Space Algorithm for Consecutive Suffix Alignment under Edit Distance (<i>Short Preliminary Paper</i>)	155
<i>Heikki Hyvrö</i>	
Run-Length Compressed Indexes Are Superior for Highly Repetitive Sequence Collections	164
<i>Jouni Sirén, Niko Välimäki, Veli Mäkinen, and Gonzalo Navarro</i>	
Practical Rank/Select Queries over Arbitrary Sequences	176
<i>Francisco Claude and Gonzalo Navarro</i>	
Clique Analysis of Query Log Graphs.....	188
<i>Alexandre P. Francisco, Ricardo Baeza-Yates, and Arlindo L. Oliveira</i>	
Out of the Box Phrase Indexing	200
<i>Frederik Transier and Peter Sanders</i>	
Approximated Pattern Matching with the L_1, L_2 and L_∞ Metrics	212
<i>Ohad Lipsky and Ely Porat</i>	
Interchange Rearrangement: The Element-Cost Model	224
<i>Oren Kapah, Gad M. Landau, Avivit Levy, and Nitsan Oz</i>	
$\delta\gamma$ -Parameterized Matching	236
<i>Inbok Lee, Juan Mendivelso, and Yoan J. Pinzón</i>	
Pattern Matching with Pair Correlation Distance	249
<i>Benny Porat, Ely Porat, and Asaf Zur</i>	
Some Approximations for Shortest Common Nonsubsequences and Supersequences	257
<i>Vadim G. Timkovsky</i>	
On the Structure of Small Motif Recognition Instances	269
<i>Christina Boucher, Daniel G. Brown, and Stephane Durocher</i>	
Exact Distribution of a Spaced Seed Statistic for DNA Homology Detection	282
<i>Gary Benson and Denise Y.F. Mak</i>	
Author Index	295