

# 28th Annual Symposium on Combinatorial Pattern Matching

CPM 2017, July 4–6, 2017, Warsaw, Poland

Edited by

Juha Kärkkäinen

Jakub Radoszewski

Wojciech Rytter



*Editors*

Juha Kärkkäinen  
Department of Computer Science  
University of Helsinki, Finland  
juha.karkkainen@cs.helsinki.fi

Jakub Radoszewski  
Faculty of Mathematics,  
Informatics and Mechanics  
University of Warsaw, Poland  
jrad@mimuw.edu.pl

Wojciech Rytter  
Faculty of Mathematics,  
Informatics and Mechanics  
University of Warsaw, Poland  
rytter@mimuw.edu.pl

*ACM Classification 1998*

E.1 Data Structures, E.2 Data Storage Representations, E.4 Coding and Information Theory, F. Theory of Computation, G.2 Discrete Mathematics, H. Information Systems, I.7 Document and Text Processing

**ISBN 978-3-95977-039-2**

*Published online and open access by*

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <http://www.dagstuhl.de/dagpub/978-3-95977-039-2>.

*Publication date*

July, 2017

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

*License*

This work is licensed under a Creative Commons Attribution 3.0 Unported license (CC-BY 3.0): <http://creativecommons.org/licenses/by/3.0/legalcode>.



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Digital Object Identifier: 10.4230/LIPIcs.CPM.2017.0

ISBN 978-3-95977-039-2

ISSN 1868-8969

<http://www.dagstuhl.de/lipics>

## ■ Contents

|                                                                      |      |
|----------------------------------------------------------------------|------|
| Preface                                                              |      |
| <i>Juha Kärkkäinen, Jakub Radoszewski, and Wojciech Rytter</i> ..... | 0:ix |

### Invited Talks

|                                                              |         |
|--------------------------------------------------------------|---------|
| Wheeler Graphs: Variations on a Theme by Burrows and Wheeler |         |
| <i>Giovanni Manzini</i> .....                                | 1:1–1:1 |
| Recompression of SLPs                                        |         |
| <i>Artur Jeż</i> .....                                       | 2:1–2:1 |
| Shortest Superstring                                         |         |
| <i>Marcin Mucha</i> .....                                    | 3:1–3:1 |

### Regular Papers

|                                                                                                                                                             |            |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Document Listing on Repetitive Collections with Guaranteed Performance                                                                                      |            |
| <i>Gonzalo Navarro</i> .....                                                                                                                                | 4:1–4:13   |
| Path Queries on Functions                                                                                                                                   |            |
| <i>Travis Gagie, Meng He, and Gonzalo Navarro</i> .....                                                                                                     | 5:1–5:15   |
| Deterministic Indexing for Packed Strings                                                                                                                   |            |
| <i>Philip Bille, Inge Li Gørtz, and Frederik Rye Skjoldjensen</i> .....                                                                                     | 6:1–6:11   |
| Representing the Suffix Tree with the CDAWG                                                                                                                 |            |
| <i>Djamal Belazzougui and Fabio Cunial</i> .....                                                                                                            | 7:1–7:13   |
| Position Heaps for Parameterized Strings                                                                                                                    |            |
| <i>Diptarama, Takashi Katsura, Yuhei Otomo, Kazuyuki Narisawa, and Ayumi Shinohara</i> .....                                                                | 8:1–8:13   |
| On-Line Pattern Matching on Similar Texts                                                                                                                   |            |
| <i>Roberto Grossi, Costas S. Iliopoulos, Chang Liu, Nadia Pisanti, Solon P. Pissis, Ahmad Retha, Giovanna Rosone, Fatima Vayani, and Luca Versari</i> ..... | 9:1–9:14   |
| A Family of Approximation Algorithms for the Maximum Duo-Preservation String Mapping Problem                                                                |            |
| <i>Bartłomiej Dudek, Paweł Gawrychowski, and Piotr Ostropolski-Nalewaja</i> .....                                                                           | 10:1–10:14 |
| Revisiting the Parameterized Complexity of Maximum-Duo Preservation String Mapping                                                                          |            |
| <i>Christian Komusiewicz, Mateus de Oliveira Oliveira, and Meirav Zehavi</i> .....                                                                          | 11:1–11:17 |
| Clique-Based Lower Bounds for Parsing Tree-Adjoining Grammars                                                                                               |            |
| <i>Karl Bringmann and Philip Wellnitz</i> .....                                                                                                             | 12:1–12:14 |
| Communication and Streaming Complexity of Approximate Pattern Matching                                                                                      |            |
| <i>Tatiana Starikovskaya</i> .....                                                                                                                          | 13:1–13:11 |

28th Annual Symposium on Combinatorial Pattern Matching (CPM 2017).

Editors: Juha Kärkkäinen, Jakub Radoszewski, and Wojciech Rytter

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

|                                                                                                                                                                             |            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| The Longest Filled Common Subsequence Problem<br><i>Mauro Castelli, Riccardo Dondi, Giancarlo Mauri, and Italo Zoppis</i> .....                                             | 14:1–14:13 |
| Lempel-Ziv Compression in a Sliding Window<br><i>Philip Bille, Patrick Hagge Cording, Johannes Fischer, and Inge Li Gørtz</i> .....                                         | 15:1–15:11 |
| Time-Space Trade-Offs for Lempel-Ziv Compressed Indexing<br><i>Philip Bille, Mikko Berggren Ettiienne, Inge Li Gørtz, and Hjalte Wedel Vildhøj</i> ..                       | 16:1–16:17 |
| From LZ77 to the Run-Length Encoded Burrows-Wheeler Transform, and Back<br><i>Alberto Policriti and Nicola Prezza</i> .....                                                 | 17:1–17:10 |
| Longest Common Extensions with Recompression<br><i>Tomohiro I</i> .....                                                                                                     | 18:1–18:15 |
| Fast and Simple Jumbled Indexing for Binary Run-Length Encoded Strings<br><i>Luis Cunha, Simone Dantas, Travis Gagie, Roland Wittler, Luis Kowada, and Jens Stoye</i> ..... | 19:1–19:9  |
| Faster STR-IC-LCS Computation via RLE<br><i>Keita Kuboi, Yuta Fujishige, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda</i> .....                                      | 20:1–20:12 |
| Gapped Pattern Statistics<br><i>Philippe Duchon, Cyril Nicaud, and Carine Pivoteau</i> .....                                                                                | 21:1–21:12 |
| Computing All Distinct Squares in Linear Time for Integer Alphabets<br><i>Hideo Bannai, Shunsuke Inenaga, and Dominik Köppl</i> .....                                       | 22:1–22:18 |
| Palindromic Length in Linear Time<br><i>Kirill Borozdin, Dmitry Kosolobov, Mikhail Rubinchik, and Arseny M. Shur</i> .....                                                  | 23:1–23:12 |
| Tight Bounds on the Maximum Number of Shortest Unique Substrings<br><i>Takuya Mieno, Shunsuke Inenaga, Hideo Bannai, and Masayuki Takeda</i> .....                          | 24:1–24:11 |
| Can We Recover the Cover?<br><i>Amihood Amir, Avivit Levy, Moshe Lewenstein, Ronit Lubin, and Benny Porat</i> ...                                                           | 25:1–25:15 |
| Approximate Cover of Strings<br><i>Amihood Amir, Avivit Levy, Ronit Lubin, and Ely Porat</i> .....                                                                          | 26:1–26:14 |
| Beyond Adjacency Maximization: Scaffold Filling for New String Distances<br><i>Laurent Bulteau, Guillaume Fertin, and Christian Komusiewicz</i> .....                       | 27:1–27:17 |
| On the Weighted Quartet Consensus Problem<br><i>Manuel Lafond and Céline Scornavacca</i> .....                                                                              | 28:1–28:18 |
| Optimal Omnitig Listing for Safe and Complete Contig Assembly<br><i>Massimo Cairo, Paul Medvedev, Nidia Obscura Acosta, Romeo Rizzi, and Alexandru I. Tomescu</i> .....     | 29:1–29:12 |
| Dynamic Elias-Fano Representation<br><i>Giulio Ermanno Pibiri and Rossano Venturini</i> .....                                                                               | 30:1–30:14 |
| Synergistic Solutions on MultiSets<br><i>Jérémy Barbay, Carlos Ochoa, and Srinivasa Rao Satti</i> .....                                                                     | 31:1–31:14 |