Curriculum Vitae

Name	:	Konstantinos Tsichlas
Born	:	19th of May, 1976, Greece
Marital Status	:	Married with two children
Address	:	Ethnikis Antistasevs 16, Kalamaria, Thessaloniki Post Code 55133
Tel.	:	+30 2310991934
email	:	tsichlas@delab.csd.auth.gr
Webpage	:	http://delab.csd.auth.gr/˜tsichlas/

Brief Biography

I am a lecturer in the Informatics Department of Aristotle University of Thessaloniki since 2008. Since 2005 I am an adjunct Professor in Greek Open University as well. From 1/5/2011-31/10/2011 I was on sabbatical in the MADALGO (center for MAssive Data ALGOrithmics) institute of the Computer Science Department of the University of Aarhus in Denmark. From 4/1/2004 until 4/6/2005 I was a research assistant in the Computer Science Department of Kings College London in the Algorithm Design Group. I was awarded a Ph.D. diploma in 2004. I have completed my military service in the Greek army in 2003.

My research interests focus on the design and analysis of algorithms and data structures. In particular, I am working on fundamental data structures, like dictionaries and priority queues in various models of computations (PM, RAM, I/O model, distributed etc.), computational geometry problems, strings matching algorithms with applications to bioinformatics and music analysis, indexing problems of databases and combinatorial optimization problems in large-scale networks and with respect to query optimization. In addition, since 2011 I have initiated work on on lower bounds using various existing tools, on Natural Algorithms which I feel is an exciting new field and analysis of complex networks. Although these three areas seem different, my feel is there is a lot of common ground.

Education

April 2002 - April 2004	Ph.D. Diploma from the Computer Engineering and Informatics Department of the University of Patras (supervisor : A. Tsakalidis)
November 1999 - February 2002	M.Sc. Diploma in Computer Science and Technology by the Computer Engineering and Informatics Department of the University of Patras (9.8 / 10).
September 1994 - July 1999	5-year Diploma from the Computer Engineering and Informatics Department of the University of Patras (9.03 / 10).

Languages

- Greek (Native)
- English (Cambridge Certificate of Proficiency in English)

Scholarships - Honors

- Honor in **Thalis** contest organized by the Greek Mathematical Society in 1993.
- Award from the Technical Chamber of Greece for excellent performance during 1998-1999.
- One year scholarship in the M.Sc. programme *Computer Science and Technology* in the Computer Engineering and Informatics Department of the University of Patras for excellent performance during undergraduate studies.

Current Position

24/04/2008-...	Lecturer on *Computational Complexity* in the Informatics Department of Aristotle University of Thessaloniki, Greece.
01/12/2011-...	Adjunct Professor: *Foundations of Computer Science*, Greek Open University

Publications

Journals

1. **Optimal Solutions for the Temporal Precedence Problem.** G. S. Brodal, C. Makris, S. Sioutas, A. Tsakalidis and K. Tsichlas. *Algorithmica*, 33(4):494–510, 2002.

2. **Approximate String Matching with Gaps.** M. Crochemore, C. Iliopoulos, C. Makris, W. Rytter, A. Tsakalidis and K. Tsichlas. *Nordic Journal of Computing*, 9:54–65, 2002.

3. **Reflected Min-Max Heaps.** C. Makris, A. Tsakalidis and K. Tsichlas. *Information Processing Letters (IPL)*, 86(4):209–214, 2003.

4. **Optimal Finger Search Trees in the Pointer Machine.** G. S. Brodal, C. Makris, G. Lagogiannis, A. Tsakalidis and K. Tsichlas. *Journal of Computer and System Sciences, Special Issue on STOC 2002*, 67(2):381–418, 2003.

5. **Rectangle Enclosure Reporting in Linear Space Revisited.** G. Lagogiannis, C. Makris, Y. Panagis, S. Sioutas and K. Tsichlas. *Journal of Automata, Languages and Combinatorics*, 8(4):633–645, 2003.

6. **New Dynamic Balanced Search Trees with Worst-Case Constant Update Time.** G. Lagogiannis, C. Makris, Y. Panagis, S. Sioutas and K. Tsichlas. *Journal of Automata, Languages and Combinatorics*, 8(4):607–632, 2003.

7. **Geometric Retrieval of Grid Points in the RAM model.** C. Makris, S. Sioutas, A. Tsakalidis, J. Tsaknakis, K. Tsichlas and B. Vassiliadis. *Journal of Universal Computer Science*, 10(9):1325–1353, 2004.

8. **Computation of Repetitions and Regularities on Biological Weighted Sequences.** M. Christodoulakis, C. Iliopoulos, L. Mouchard, K. Perdikuri, A. Tsakalidis and K. Tsichlas. *Journal of Computational Biology*, 13(6):1214–1231, 2006.

9. **2-D Monotone Spatial Indexing Scheme with Optimal Update Time.** L. Drossos, S. Sioutas, K.Tsichlas and K.Ioannou. *Transactions on Systems*, ISSN: 1109-2777, 5(1):142–147, 2006.

10. **Locating Maximal Multirepeats in Multiple Strings Under Various Constraints.** A. Bakalis, C. Iliopoulos, C. Makris, S. Sioutas, E. Theodoridis, A.Tsakalidis and K.Tsichlas. *Computer Journal*, 50(2):178–185, 2007.

11. **Algorithms for Extracting Motifs from Biological Weighted Sequences.** C. Illiopoulos, K. Perdikuri, E. Theodoridis, A. Tsakalidis and K. Tsichlas. *Journal of Discrete Algorithms, Special Issue on SPIRE 2004*, 5(2):229–242, 2007.

12. **Efficient Access Methods for Temporal Interval Queries of Video Metadata.** S. Sioutas, K. Tsichlas, B. Vassiliadis and D.K. Tsolis. *Journal of Universal Computer Science*, 13(10): 1411–1433, 2007.

13. **Scheduling Algorithms for Procrastinators.** M. Bender, R. Clifford and K. Tsichlas. *Journal of Scheduling*, 11(2):95–104,2008.

14. **A New Approach on Indexing Mobile Objects on the Plane.** S. Sioutas, K. Tsakalidis, K. Tsichlas, C. Makris, Y. Manolopoulos. *Data Knowledgment Engineering,* 67(3): 362–380, 2008.

15. **Canonical Polygon Queries on the Plane: A New Approach.** S. Sioutas, D. Sofotassios, K. Tsichlas, D. Sotiropoulos and P. Vlamos. *Journal of Computers*, 4(9):913–919, 2009.

16. **An Experimental Performance Comparison for Indexing Mobile Objects on the Plane.** S. Sioutas, G. Papaloukopoulos, K. Tsichlas and Y. Manolopoulos. *Special issue of ACM-SIGAPP MEDES '09 on Collectively Intelligent Information and Knowledge Management, Journal on Organizational and Collective Intelligence (IJOCI)*, 1(4):78–96, 2010.

17. **ISB-Tree: A New Indexing Scheme with Efficient Expected Behaviour.** Ch. Makris, S. Sioutas, Tsakalidis, K. Tsichlas, Y. Ch. Zaroliagis. *Journal of Discrete Algorithms*, 8(4):373–387, 2010.

18. **Improved Bounds for Finger Search on a RAM.** A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *Algorithmica*, 66(2):249–286, 2013.

19. **On the Discovery of Group-Consistent Graph Substructure Patterns from Brain Networks.** N.D. Iakovidou, S.I. Dimitriadis, N.A. Laskaris, K. Tsichlas, Y. Manolopoulos.*Journal of Neuroscience*, 213(2):204–213, 2013.

20. **ART: Sub-Logarithmic Decentralized Range Query Processing with Probabilistic Guarantees.** S. Sioutas, P. Triantafillou, G. Papaloukopoulos, E. Sakkopoulos, K. Tsichlas, Y. Manolopoulos. *Distributed and Parallel Databases*, 31(1):71–109. 2013.

Chapters in Books

21. **New Upper Bounds on Various String Manipulation Problems.** C. Makris, Y. Panagis, K. Perdikuri, S. Sioutas, E. Theodoridis, A. Tsakalidis and K. Tsichlas. In *Text in Algorithms vol. 2: String Algorithmics*, eds. C. Iliopoulos and T. Lecroq, King's College Publications, ISBN 1-904987-0-2-8, pp. 171-193, 2004.

22. **Access Methods.** A.N. Papadopoulos, K. Tsichlas, A. Gounaris, Y. Manolopoulos. In Information Systems and Information Technology, Volume 2 (Computing Handbook Set, Third Edition,) edited by Heikki Topi and Allen Tucker. Boca Raton: Taylor and Francis, 2014 (forthcoming).

Conferences

23. **Approximate String Matching with Gaps.** M. Crochemore, C. Iliopoulos, C. Makris, W. Rytter, A. Tsakalidis and K. Tsichlas. *In Proc. of the World Multiconference on Systemics, Cybernetics and Informatics (SCI)*, vol. X, pp. 45-50, July 22–25, 2001.

24. **Time and Space Efficient Content Queries for Video Databases.** C. Makris, K. Perdikuri, S. Sioutas, A. Tsakalidis and K. Tsichlas. *In Proc. of the 1st International Workshop on Multimedia Data and Document Engineering (MDDE)*, pp. 1–8, 2001.

25. **Optimal Finger Search Trees in the Pointer Machine.** G.S. Brodal, C. Makris, G. Lagogiannis, A. Tsakalidis and K. Tsichlas. *In Proc. of the 34th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 583–591, 2002.

26. **Identifying Occurrences of Maximal Pairs in Multiple Strings.** C. Iliopoulos, C. Makris, S. Sioutas, A. Tsakalidis and K. Tsichlas. *In Proc. of the 13th Annual Symposium on Combinatorial Pattern Matching (CPM)*, pp. 133–143, 2002.

27. **Rectangle Enclosure Reporting in Linear Space Revisited.** G. Lagogiannis, Y. Panagis, S. Sioutas and K. Tsichlas. *In Proc. of the 13th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2002.

28. **New Dynamic Balanced Search Trees with Worst-Case Constant Update Time.** G. Lagogiannis, C. Makris, Y. Panagis and K. Tsichlas. *In Proc. of the 13th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2002.

29. **Data Structuring Applications for String Problems in Biological Sequences.** Y. Panagis, E. Theodoridis, K. Tsichlas. *In Proc. of the International Conference of Computational Methods in Science and Engineering (ICCMSE)*, pp. 479–483, 2003.

30. **Temporal Selection Queries in Video Databases.** S. Sioutas, C. Makris, G. Lagogiannis, E. Sakkopoulos, K. Tsichlas, V. Delis and A. Tsakalidis. *In Proc. of the 3rd International Workshop on Multimedia Data and Document Engineering (MDDE), collocated with VLDB*, 2003.

31. **Improved Bounds for Finger Search on a RAM.** A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 11th Annual European Symposium on Algorithms (ESA).*, LNCS 2832, pp. 325–336, 2003.

32. **The Pattern Matching Problem in Biological Weighted Sequences.** C. Iliopoulos, K. Perdikuri, A. Tsakalidis and K. Tsichlas. *In Proc. of FUN with Algorithms*, edited by Paolo Ferragina & Roberto Grossi, 106–117, 2004.

33. **On the Canonical k-vertex Polygon Spatial Retrieval Problem.** V. Bistiolas, S. Sioutas, D. Sofotassios and K. Tsichlas. *In Proc. of the 15th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2004.

34. **Motif Extraction from Weighted Sequences.** C. Illiopoulos, K. Perdikuri, E. Theodoridis, A. Tsakalidis and K. Tsichlas. *In Proc. of the 11th Symposium on String Processing and Information Retrieval (SPIRE)*, pp. 286–297, 2004.

35. **Searching for Regularities in Weighted Sequences.** M. Christodoulakis, C. Iliopoulos, K. Tsichlas and K. Perdikuri. *In Proc. of the International Conference of Computational Methods in Science and Engineering (ICCMSE)*, Lecture Series on Computer and Computational Sciences, pp. 701–704, 2004.

36. **Pattern Matching on Weighted Sequences.** M. Christodoulakis, C. Illiopoulos, L. Mouchard and K. Tsichlas. *In Proc. of Algorithms and Computational Methods for Biochemical and Evolutionary Networks (CompBionets)*, pp. 17–30, 2004.

37. **Algorithms for Extracting Structured Motifs from Biological Weighted Sequences.** C. Iliopoulos, K. Perdikuri, A. Tsakalidis and K. Tsichlas. In Proc. of the *the 16th Australasian Workshop on Combinatorial Algorithms (AWOCA)*, 2005.

38. **Finding Multirepeats in a Set of Strings.** A. Bakalis, C. Makris, S. Sioutas, E. Theodoridis and K. Tsichlas. In Proc. of the *International Conference of Computational Methods in Sciences and Engineering (ICCMSE)*, 2005.

39. **ISB-Tree: A New Indexing Scheme with Efficient Expected Behaviour.** A. Kaporis, C. Makris, G. Mayritsakis, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 16th Annual International Symposium on Algorithms and Computation (ISAAC)*, pp. 318–327, 2005.

40. **Dynamic Interpolation Search Revisited.** A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 382–394, 2006.

41. **Algorithms for Bitmasking Strings.** A. Bakalis, C. Iliopoulos, S. Sioutas and K. Tsichlas. Accepted for presentation in *International Conference of Computational Methods in Sciences and Engineering (ICCMSE)*, pp. 1061–1063, 2006.

42. **Purely Functional Worst Case Constant Time Catenable Sorted Lists.** G.S. Brodal, C. Makris and K. Tsichlas. *In Proc. of the 13th Annual European Symposium on Algorithms (ESA),* pp. 172–183, 2006.

43. **Indexing of mobile objects on the plane revisited.** S. Sioutas, K. Tsakalidis, K. Tsichlas, C. Makris and Y. Manolopoulos. *In Proc. of the 11th East-European Conference on Advances in Databases and Information Systems (ADBIS)*, pp. 189–204, 2007.

44. **New Expected Bounds for Three-sided Range Queries on the Plane.** G.S. Brodal, A. Kaporis, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 20th Int. Symposium on Algorithms and Computation (ISAAC)*, pp. 193–202, 2009.

45. **An Experimental Performance Comparison for Indexing Mobile Objects on the Plane.** S. Sioutas, G. Papaloukopoulos, K. Tsichlas and Y. Manolopoulos. *In Proc. of ACM-SIGAPP MEDES 2009*, pp. 210–217, 2009.

46. **A novel Distributed P2P Simulator Architecture: D-P2P-Sim.** S. Sioutas, G. Papaloukopoulos, E. Sakkopoulos, K. Tsichlas and Y. Manolopoulos. *In Proc. of ACM-CIKM*, pp. 2069–2070, 2009.

47. **Efficient Processing of 3-sided Range Queries with Probabilistic Guarantees.** A. Kaporis, A. Papadopoulos, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 13th International Conference on Database Theory (ICDT)*, pp. 34–43, 2010.

48. **ART–sub-logarithmic decentralized range query processing with probabilistic guarantees.** A. Kaporis, A. Papadopoulos, S. Sioutas, K. Tsakalidis and K. Tsichlas. *Brief announcement in Proc. of the 29th Annual Symposium on Principles of Distributed Computing (PODC)*, pp. 118–119, 2010.

49. **The D^2-tree: a new P2P deterministic data structure.** G.S. Brodal, S. Sioutas, K.Tsichlas and C. Zaroliagis. *In Proc. of the 21st Int. Symposium on Algorithms and Computation (ISAAC)*, pp. 1–12, 2010.

50. **Continuous Monitoring of Distance-based Outliers Over Streams.** M. Kontaki, A. Gounaris, Y. Manolopoulos, A. Papadopoulos and K. Tsichlas. *In Proc. of the 26th IEEE Int. Conference on Data Engineering (ICDE)*, pp. 135–146, 2011.

51. **NEFOS: Rapid Cache-Aware Range Query Processing with Probabilistic Guarantees.** S. Sioutas, K. Tsichlas, I. Karydis, Y. Manolopoulos and Y. Theodoridis. *In Proc. of the 22nd Int. Conference on Database and Expert Systems Applications (DEXA)*, pp. 62–77, 2011.

52. **Fully Persistent B-trees.** G.S. Brodal, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 23rd Symposium on Discerete Algorithms (SODA)*, pp. 602–614, 2012.

53. **DISCO: a New Algorithm for Detecting 3D Protein Structure Similarity.** N.D. Iakovidou, E. Tiakas, K. Tsichlas. *In Proc. of the 1st Workshop on Algorithms for Data and Text Mining in Bioinformatics (WADTMB) (Artificial Intelligence Applications and Innovations)*, 622–631, 2012.

54. **I/O-Efficient Orthogonal Planar Range Skyline Reporting and Catenable Priority Queues with Attrition.** C. Kejlberg-Rasmussen, Y. Tao, K. Tsakalidis, K. Tsichlas, J. Yoon. Accepted for presentation in *(PODS)*, 2013.

55. **Multi-objective optimization of data flows in a multi-cloud environment.** E. Tsamoura, A. Gounaris, K. Tsichlas. Accepted for presentation in *Workshop on Data analytics in the Cloud (DanaC)*, 2013.

56. **Continuous Outlier Detection in Data Streams: An Extensible Framework and State-Of-The-Art Algorithms.** D. Georgiadis, M. Kontaki, A. Gounaris, A. Papadopoulos, K. Tsichlas, Y. Manolopoulos. *SIGMOD Demonstration*, 2013.

Articles without Proceedings

57. **Continuous Monitoring of Distance-based Outliers Over Streams.** M. Kontaki, A. Gounaris, Y. Manolopoulos, A. Papadopoulos and K. Tsichlas. *Presented at the 10th Hellenic Data Management Symposium (HDMS)*, 2011.

58. **Fully Persistent B-trees.** G.S. Brodal, S. Sioutas, K. Tsakalidis and K. Tsichlas. *4th Workshop on Massive Data Algorithmics (MASSIVE)*, 2012.

59. **I/O Efficient Dynamic Planar Range Skyline Queries.** C. Kejlberg-Rasmussen, K. Tsakalidis, K. Tsichlas. *4th Workshop on Massive Data Algorithmics (MASSIVE)*, 2012.

Submitted Journal Articles

60. **A Space-Optimal Hidden Surface Removal Algorithm for Iso-Oriented Rectangles.** A. Tsakalidis and K. Tsichlas. Submitted to *Information Processing Letters.*

61. **Dynamic Interpolation Search Revisited.** A. Kaporis, Ch. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas, Ch. Zaroliagis. Submitted to *Transactions on Algorithms*.

62. **DISCO: A New Algorithm for Detecting 3D Protein Structure Similarity.** N. Iakovidou, E. Tiakas, K. Tsichlas and Y. Manolopoulos. Submitted to *ITDB*.

63. **The D^2-tree: a new P2P deterministic data structure.** G.S. Brodal, S. Sioutas, K.Tsichlas and C. Zaroliagis. Submitted to *Algorithmica*.

64. **Dynamic 3-sided Planar Range Queries with Expected Doubly Logarithmic Time.** G.S. Brodal, A.C. Kaporis, A.N. Papadopoulos, S. Sioutas, K. Tsakalidis and K. Tsichlas. Submitted to *Theoretical Computer Science (TCS)*.

Technical Reports

65. **A Brief Survey on Data Structuring Problems.** C. Makris, A. Tsakalidis and K. Tsichlas. TR No. 2000/07/02, Computer Technology Institute, University of Patras, 2000.

Lecture Notes

66. **Business Systems for Office Automation.** K. Tsichlas. Lecture notes for the course: Business Systems for Office Automation and Procedure Redesign. Department of Business Planning and Information Systems, Technological Educational Institute of Patras, 83 pages, 2006.

67. **Analysis of Algorithms.** I. Manolopoulos, S. Sioutas, A. Tsakalidis and K. Tsichlas. Lecture notes for the course: Analysis of Algorithms. Informatics Department, A.U.TH., 459 pages, 2011.

Thesis

68. **Data Structures for Hidden Surface Removal.** Kostas Tsichlas. Diploma Thesis (In Greek), Advisor Athanasios Tsakalidis, University of Patras, 1999.

69. **Redundant Number Systems and their Applications in Data Structures.** Kostas Tsichlas. M.Sc. Thesis (In Greek), Advisor Athanasios Tsakalidis, University of Patras, 2002.

70. **Efficient Data Structures and Applications.** Kostas Tsichlas. Ph.D. Thesis (In Greek), Advisor Athanasios Tsakalidis, University of Patras, 2004.

Participation in Program Committees

1. *London Stringology Day 2005*, Kings College London, U.K.
2. *Seventeenth ACM Conference on Hypertext and Hypermedia (ACM HyperText 2006)*, http://www.ht06.org, International Workshop on Architectures, Models and Infrastructures to Generate Semantics in Peer to Peer and Hypermedia Systems, Odense, Denmark, 23-25 August 2006.

Participation in Conferences/Workshops

September 2012 — *4th Workshop on Massive Data Algorithmics (MASSIVE)*, Ljubljana, Slovenia, 13/9/2012. (Presentation)

August 2011 — *Summer School on High-Dimensional Geometric Computing*, Aarhus University, Co-organized by MADALGO and CTIC, 8-11/8/2011.

June 2011 — *Synergies in Lower Bounds*, Aarhus University, Co-organized by MADALGO and CTIC, 28/6-1/7/2011.

December 2010 — *21st International Symposium on Algorithms and Computation (ISAAC 2010)*, Jeju Island, December 15-17, 2010. (Presentation)

April 2007 — *1st ARRIVAL Annual Workshop on Robust planning and Rescheduling in Railways*, Utrecht, April 18-20, 2007.

September 2006 — *ARRIVAL/Matheon Fall School 2006 on Timetabling and Line Planning*, Dabendorf near Berlin, September 26-29, 2006.

February 2005 — *London Stringology Day 2005*, Kings College London, U.K. (Presentation)

June 2002 — *EEF Summer School on Massive Data Sets*, BRICS, University of Aarhus, Denmark.

May 2002 — *34th Annual ACM Symposium on Theory of Computing (STOC)*, Montreal, Quebec, Canada. (Presentation)

July 2001 — *World Multiconference on Systemics, Cybernetics and Informatics (SCI 2001)*, Orlando, Florida, USA. (Presentation)

June 2001 — *33rd Annual ACM Symposium on Theory of Computing (STOC)*, Hersonissos, Crete, Greece.

July 2001 — *28th International Colloquium on Automata, Languages and Programming*, Hersonissos, Crete, Greece.

Reviewer in Conferences

2013 — 16th International Conference on Extending Database Technology (EDBT)
21th Conference on Algorithms and Complexity (CIAC)
12th International Symposium on Experimental Algorithms (SEA)

2012 — 15th Meeting on Algorithms Engineering and Experiments (ALENEX)
3rd Fourth Workshop on Tools for Automatic Program Analysis (TAPAS)

28th IEEE Intern. Conference on Data Engineering (ICDE)

2011 International Workshop on Combinatorial Algorithms (IWOCA)
19th European Symposium on Algorithm (ESA)
27th ACM Symposium On Applied Computing (SAC)

2010 18th Annual European Symposium on Algorithms (ESA)
28th Interational Symposium on Theoretical Aspects of Computer Science (STACS)
13th Workshop on Algorithm Engineering and Experiment (ALENEX)

2007 32nd International Symposium on Mathematical Foundations of Computer Science (MFCS)

2005 16th Annual Symposium on Combinatorial Pattern Matching (CPM)
31st International Colloquium on Automata, Languages and Programming(ICALP)
Euro-Par

2004 31st International Colloquium on Automata, Languages and Programming(ICALP)
11th Int. Conf. on String Processing and Information Retrieval (SPIRE)

2003 15th Annual Symposium On Discrete Algorithms(SODA)
2nd Workshop on Efficient and Experimental Algorithms (WEA)

2002 10th Annual European Symposium on Algorithms(ESA)
29th International Colloquium on Automata, Languages and Programming(ICALP)

2001 28th International Colloquium on Automata, Languages and Programming(ICALP)

Reviewer in Journals

2012 Journal of Experimental Algorithmics (JEA)

2011 Engineering Applications of Artificial Intelligence (EAAI)
Information Processing Letters (IPL)
Engineering Applications of Artificial Intelligence (EAAI)
Computational Geometry: Theory and Applications (CGTA)

2007 International Journal of Computer Mathematics (IJCM)
Journal of Discrete Algorithms (JDA)

2005 Information Processing Letters (IPL)

Talks

December 2011 *Massive Data: Many Problems, Some Solutions and More Questions.* Canterbury University, New Zealand.

May 2011 *Some Complex problems without Complexities.* Theory Seminar of MADALGO, University of Aarhus, Denmark.

May 2009 *Indexing Schemes for P2P Environments.* Talk after invitation at BAD 2009, Bristol, UK.

December 2008 *Indexing Schemes for P2P Environments.* Theory Seminar of MADALGO, University of Aarhus, Denmark.

February 2007 — *Turning Amortized to Worst-Case Data Structures: Techniques and Results.* Talk after invitation at BAD 2007, Bristol, UK.

December 2006 — *Design of Data Structures with Worst-Case Guarantees.* Aristotle University of Thessaloniki, Informatics Department.

May 2006 — *Finger Search Trees.* University of Ioannina, Informatics Department.

May 2005 — *Interpolation Search Trees Revisited.* In ALCOM seminars, University of Aarhus, Denmark.

February 2005 — *The Spoonerism Problem.* In the London Stringology Day organized by Kings College London, London, UK.

November 2004 — *Persistence and Localized Search.* In the Journal Club organized by Kings College London, London, UK.

June 2002 — *Optimal Finger Search Trees in the Pointer Machine.* In Annual seminar organized by Computer Technology Intitute (CTI), Patras, Greece.

Visits to Universities/Research Centres

2-8/12/2012 — Visit to the SUNY at Stony Brook, to the Prof. M. Bender.

1/5-31/10 2011 — On sabbatical to the MADALGO Institute, University of Aarhus, Denmark.

15-20/12/2008 — Visit to the University of Aarhus, after invitation by G.S. Brodal.

10-20/5/2005 — Visit to the University of Aarhus, after invitation by G.S. Brodal.

Teaching Experience

01/10/2011-.... — *Linear Algebra*, Aristotle University of Thessaloniki (AUTH) - Winter Semester, Obligatory Course

01/10/2008-.... — *Advanced Indexing Techniques*, Aristotle University of Thessaloniki (AUTH) - Spring Semester, M.Sc. Course

01/10/2008-.... — *Theory of Computation*, AUTH - Spring Semester, Elective Course

01/10/2008-.... — *Discrete Mathematics*, AUTH - Winter Semester, Obligatory Course

01/10/2008-.... — *Analysis of Algorithms*, AUTH - Winter Semester, Obligatory Course

01/12/2011-.... — Adjunct Professor: *Foundations of Computer Science*, Greek Open University

01/10/2005-31/7/2011 — Adjunct Professor: *Introduction to Informatics*, Greek Open University

01/10/2008-31/09/2011 — *Numerical Analysis*, AUTH - Winter Semester, Obligatory Course

01/10/2008-30/06/2009 — *Graph Theory*, AUTH - Spring Semester, Elective Course

01/09/2005-22/04/2008 — Teacher of Informatics in Secondary Education

01/10/2006-30/06/2007 Adjunct Lecturer: *Object Oriented Programming, Data Structures Laboratory*, Technological Education Institute of Mesologgi

06/10/2005-30/06/2006 Adjunct Lecturer: *Programming in C Laboratory, DataBases II*, Computer Engineering and Informatics Department, University of Patras

01/10/2005-30/06/2006 Adjunct Lecturer: *Data Structures Laboratory*,Technological Education Institute of Mesologgi

01/11/2005-30/06/2006 Adjunct Lecturer: *Business Systems in Automation of Office*, Technological Education Institute of Patras

01/10/2002-31/01/2003 Adjunct Lecturer: *Information Retrieval*, Technological Education Institute of Mesologgi

01/10/2002-31/01/2003 Adjunct Lecturer: *Data Structures in C Laboratory*, Technological Education Institute of Mesologgi

23/09/2002-14/02/2003 Did'askwn: *Introduction to Informatics of Business*, Technological Education Institute of Patras

Autumn 2001 Three Lectures in *Computer Graphics and Image Synthesis*, M.Sc. Course, Athens School of Fine Arts, Athens University

1/9/1999-31/01/2003 Teaching Assistant: *Advanced Data Structures and Computer Graphics*, Computer Engineering and Informatics Department, University of Patras

1/9/1999-31/01/2003 Teaching Assistant: *Data Structures*, Computer Engineering and Informatics Department, University of Patras

14/12/2000-21/02/2003 Adjunct Teacher in Secondary Education

Administrative Duties

2008 - ... Responsible for the creation of the course and examination timetables for the Informatics department of A.U.TH.

2010 Participation in a committee to change the curricula of the Informatics Department of A.U.TH.

Research Programs

01/01/2012-... Researcher in the Greek Research Program CLOUD9 (Thalis)

01/01/2012-31/12/2012 Single researcher in a program provided by A.U.TH. about reenforcement for research activity in basic research with title "Algorithmic Analysis"

01/09/2006-31/5/2009 Researcher in the European Research Program FET-IST: ARRIVAL (Algorithms for Robust and online Railway optimization: Improving the Validity and reliAbility of Large scale systems)

01/09/2005-31/12/2008 Researcher in the greek research program *Integrated Optimization Approaches in Large-Scale Networks* (PYTHAGORAS - EPEAEK II)

05/01/2004-04/06/2005 Research Assistant in Kings College London.

01/09/2002-31/01/2004 Participation in the research team of Research Unit 5 of CTI (Computer Technology Institute) in the research program STATOBJECT - Object Oriented Approach to Official Statistics, funded by GGET.

15/10/1999-15/04/2001 Participation in the research team in Research Unit 5 of CTI (Computer Technology Institute) in the research program *Clever Storage Systems* (PENED 1999).

Research Interests

- Algorithms and Data Structures for Main and External Memory
- Algorithms and Data Structures for P2P networks
- Computational Geometry
- String Algorithms - Bioinformatics
- Graph Algorithms
- Complexity (Lower bounds, interaction of Physics and Computer Science)
- Natural algorithms with emphasis on their analysis based on algorithmic tools.
- Complex Network Analysis

Some Statistics

	Before being a Lecturer	During my term as a Lecturer
Conferences	21	+13
Conferences without Proc.	0	+3
Journals	13	+7
Book Chapters	1	+1
Lecture Notes	1	+1
Number of Citations	25	+80

Self-citations are not included. The numbers are based on the results by Publish and Perish on 13/5/2013.

Publication Analysis

Journals

Conferences publications that have journal versions as well, are listed along with these journals.

1. **Optimal Solutions for the Temporal Precedence Problem.** G. S. Brodal, C. Makris, S. Sioutas, A. Tsakalidis and K. Tsichlas. *Algorithmica*, 33(4):494–510, 2002.

   In this paper the *Temporal Precedence Problem* on *Pure Pointer Machines* is considered. This problem asks for the design of a data structure, maintaining a set of stored elements and supporting the following two operations: *insert* and *precedes*. The operation $insert(a)$ introduces a new element a in the structure, while the operation $precedes(a, b)$ returns true iff element a was inserted before element b temporally. In Ranjan et al. a solution was provided to the problem with worst-case time complexity $O(\log\log n)$ per operation and $O(n \log\log n)$ space, where n is the number of elements inserted. It was also demonstrated that the *precedes* operation has a lower bound of $\Omega(\log\log n)$ for the *Pure Pointer Machine* model of computation. In this paper we present two simple solutions with linear space and worst-case constant insertion time. In addition, we describe two algorithms that can handle the $precedes(a, b)$ operation in $O(\log\log d)$ time, where d is the temporal distance between the elements a and b.

2. **Approximate String Matching with Gaps.** M. Crochemore, C. Iliopoulos, C. Makris, W. Rytter, A. Tsakalidis and K. Tsichlas. *Nordic Journal of Computing*, 9:54–65, 2002.

   Approximate String Matching with Gaps. M. Crochemore, C. Iliopoulos, C. Makris, W. Rytter, A. Tsakalidis and K. Tsichlas. *In Proc. of the World Multiconference on Systemics, Cybernetics and Informatics (SCI)*, vol. X, pp. 45-50, July 22–25, 2001.

   In this paper several new versions of approximate string matching with gaps are considered. The main characteristic of these new versions is the existence of gaps in the matching of a given pattern in a text. Algorithms are devised for each version and their time and space complexities are stated. These specific versions of approximate string matching have various applications in computerized music analysis.

3. **Reflected Min-Max Heaps.** C. Makris, A. Tsakalidis and K. Tsichlas. *Information Processing Letters (IPL)*, 86(4):209–214, 2003.

   In this paper a simple and efficient implementation of a min-max priority queue, reflected min-max priority queues, is presented. The main merits of our construction are threefold. First, the space utilization of the reflected min-max heaps is much better than the naive solution of putting two heaps back-to-back. Second, the methods applied in this structure can be easily used to transform ordinary priority queues into min-max priority queues. Third, when considering only the setting of min-max priority queues, we support merging in constant worst-case time which is a clear improvement over the best worst-case bounds achieved by Høyer.

4. **Optimal Finger Search Trees in the Pointer Machine.** G. S. Brodal, C. Makris, G. Lagogiannis, A. Tsakalidis and K. Tsichlas. *Journal of Computer and System Sciences, Special Issue on STOC 2002*, 67(2):381–418, 2003.

   Optimal Finger Search Trees in the Pointer Machine. G.S. Brodal, C. Makris, G. Lagogiannis, A. Tsakalidis and K. Tsichlas. *In Proc. of the 34th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 583–591, 2002.

   A new finger search tree with worst case constant update time in the Pointer Machine (PM) model of computation is proposed. This was a major problem in the field of Data Structures and was tantalizingly open for over twenty years, while many attempts by researchers were made to solve it. The result comes as a consequence of the innovative mechanism that guides the rebalancing operations, combined with incremental multiple splitting and fusion techniques over nodes.

5. **Rectangle Enclosure Reporting in Linear Space Revisited.** G. Lagogiannis, C. Makris, Y. Panagis, S. Sioutas and K. Tsichlas. *Journal of Automata, Languages and Combinatorics*, 8(4):633–645, 2003.

   Rectangle Enclosure Reporting in Linear Space Revisited. G. Lagogiannis, Y. Panagis, S. Sioutas and K. Tsichlas. *In Proc. of the 13th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2002.

   We present a new algorithm for reporting all the enclosures in a set of plane rectangles in $O(n \log n \log\log n + k \log\log n)$ time and linear space (k denotes the output size). The result is already known (it has already been achieved by two previous papers), however the proposed algorithm follows a different approach.

6. **New Dynamic Balanced Search Trees with Worst-Case Constant Update Time.** G. Lagogiannis, C. Makris, Y. Panagis, S. Sioutas and K. Tsichlas. *Journal of Automata, Languages and Combinatorics*, 8(4):607–632, 2003.

   New Dynamic Balanced Search Trees with Worst-Case Constant Update Time. G. Lagogiannis, C. Makris, Y. Panagis and K. Tsichlas. *In Proc. of the 13th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2002.

   We present new search trees with worst-case $O(1)$ update time and $O(\log n)$ search time, storing n elements in linear space, in the Pointer Machine (PM) model of computation. In addition, these trees can easily support finger searches in time $O(\log d)$ and update operations in worst-case $O(\log^* n)$ time. The parameter d represents the number of elements (distance) between the search element and an element pointed to by a pointer termed *finger*. Our data structure is based on a previous result by Fleischer that exhibits the same asymptotic time and space complexities for simple search trees. We improve on this result by handling deletions in an explicit way without using the standard trick of global rebuilding. This is the first search tree that combines worst-case update times with a local rebalancing scheme without using global rebuilding to tackle deletions. In addition, intuition is acquired from the construction of these trees as to why deletions are considered more difficult than insertions in the (a,b)-trees setting. Finally, we hope that these techniques may lead to a simpler version of the constant update finger search tree presented recently by Brodal et al.

7. **Geometric Retrieval of Grid Points in the RAM model.** C. Makris, S. Sioutas, A.

Tsakalidis, J. Tsaknakis, K. Tsichlas and B. Vassiliadis. *Journal of Universal Computer Science*, 10(9):1325–1353, 2004.

We consider the problem of d-dimensional searching ($d \geq 3$) for four query types: *range*, *partial range*, *exact match* and *partial match searching*. Let N be the number of points, s be the number of keys specified in a partial match and partial range query and t be the number of points retrieved. We present a data structure with worst case time complxities $O(t + \log^{d-2} N)$, $O(t + (d - s) + \log^s N)$, $O(d + \sqrt{\log N})$ and $O(t + (d - s) + s\sqrt{\log N})$ for each of the aforementioned query types respectively.

We also present a second, more concrete solution for exact and partial match queries, which achieves the same query time but has different space requirements. The proposed data structures are considered in the RAM model of computation.

8. **Computation of Repetitions and Regularities on Biological Weighted Sequences.** M. Christodoulakis, C. Iliopoulos, L. Mouchard, K. Perdikuri, A. Tsakalidis and K. Tsichlas. *Journal of Computational Biology*, 13(6):1214–1231, 2006.

   The Pattern Matching Problem in Biological Weighted Sequences. C. Iliopoulos, K. Perdikuri, A. Tsakalidis and K. Tsichlas. *In Proc. of FUN with Algorithms*, edited by Paolo Ferragina & Roberto Grossi, 106–117, 2004.

   Searching for Regularities in Weighted Sequences. M. Christodoulakis, C. Iliopoulos, K. Tsichlas and K. Perdikuri. *In Proc. of the International Conference of Computational Methods in Science and Engineering (ICCMSE)*, Lecture Series on Computer and Computational Sciences, pp. 701–704, 2004.

   Pattern Matching on Weighted Sequences. M. Christodoulakis, C. Illiopoulos, L. Mouchard and K. Tsichlas. *In Proc. of Algorithms and Computational Methods for Biochemical and Evolutionary Networks (CompBionets)*, pp. 17–30, 2004.

   Biological Weighted Sequences are used extensively in Molecular Biology as profiles for protein families, in the representation of binding sites and often for the representation of sequences produced by a shotgun sequencing strategy. In this paper, we address three fundamental problems in the area of Biologically Weighted Sequences: i) Computation of Repetitions, ii) Pattern Matching and iii) Computation of Regularities. To the best of our knowledge, this is the first time these problems are tackled in the relevant literature. Our algorithms can be used as basic building blocks for more sophisticated algorithms applied on weighted sequences.

9. **2-D Monotone Spatial Indexing Scheme with Optimal Update Time.** L. Drossos, S. Sioutas, K.Tsichlas and K.Ioannou. *Transactions on Systems*, ISSN: 1109-2777, 5(1):142–147, 2006.

   For monotone generated points on the plane we present the Dynamic Monotone Priority Search Tree (DMoPST) in main / external memory with $O(1)$ update time / $O(1)$ block transfers in worst-case. The external version of the structure above promises efficient applications in transaction time Databases systems.

10. **Locating Maximal Multirepeats in Multiple Strings Under Various Constraints.** A. Bakalis, C. Iliopoulos, C. Makris, S. Sioutas, E. Theodoridis, A.Tsakalidis and K.Tsichlas. *Computer Journal*, 50(2):178–185, 2007.

Identifying Occurrences of Maximal Pairs in Multiple Strings. C. Iliopoulos, C. Makris, S. Sioutas, A. Tsakalidis and K. Tsichlas. *In Proc. of the 13th Annual Symposium on Combinatorial Pattern Matching (CPM)*, pp. 133–143, 2002.

A multirepeat in a string is a substring (factor) that appears a predefined number of times. A multirepeat is maximal if it cannot be extended either to the right or to the left and produce a multirepeat. In this paper, we present algorithms for two different versions of the problem of finding maximal multirepeats in a set of strings. In the case of arbitrary gaps, we propose an algorithm with $O(\sigma N^2 n + \alpha)$ time complexity. When the gap is bounded in a small range c we propose an algorithm with $O((c^2 + \sigma^2) m N^2 n \log(Nn) + \alpha)$ time complexity. Here N is the number of strings, n is the mean length of each string, m is the multiplicity of the multirepeat and α is the number of reported occurrences. Our results extend previous work by considering sets of strings as well as by generalizing pairs to multirepeats.

11. **Algorithms for Extracting Motifs from Biological Weighted Sequences.** C. Illiopoulos, K. Perdikuri, E. Theodoridis, A. Tsakalidis and K. Tsichlas. *Journal of Discrete Algorithms, Special Issue on SPIRE 2004*, 5(2):229–242, 2007.

   Motif Extraction from Weighted Sequences. C. Illiopoulos, K. Perdikuri, E. Theodoridis, A. Tsakalidis and K. Tsichlas. *In Proc. of the 11th Symposium on String Processing and Information Retrieval (SPIRE)*, pp. 286–297, 2004.

   In this paper we present three algorithms for the *Motif Identification Problem* in Biological Weighted Sequences. The first algorithm extracts repeated motifs from a biological weighted sequence. The motifs correspond to repetitive words which are approximately equal, under a Hamming distance, with probability of occurrence $\geq 1/k$, where k is a small constant. The second algorithm extracts common motifs from a set of $N \geq 2$ weighted sequences. In this case, the motifs consists of words that must occur with probability $\geq 1/k$, in $1 \leq q \leq N$ distinct sequences of the set. The third algorithm extracts maximal pairs from a biological weighted sequence. A pair in a sequence is the occurrence of the same word twice. In addition, the algorithms presented in this paper improve previous work on these problems.

12. **Efficient Access Methods for Temporal Interval Queries of Video Metadata.** S. Sioutas, K. Tsichlas, B. Vassiliadis and D.K. Tsolis. *Journal of Universal Computer Science*, 13(10): 1411–1433, 2007.

   Time and Space Efficient Content Queries for Video Databases. C. Makris, K. Perdikuri, S. Sioutas, A. Tsakalidis and K. Tsichlas. *In Proc. of the 1st International Workshop on Multimedia Data and Document Engineering (MDDE)*, pp. 1–8, 2001.

   Temporal Selection Queries in Video Databases. S. Sioutas, C. Makris, G. Lagogiannis, E. Sakkopoulos, K. Tsichlas, V. Delis and A. Tsakalidis. *In Proc. of the 3rd International Workshop on Multimedia Data and Document Engineering (MDDE), collocated with VLDB*, 2003.

   Indexing video content is one of the most important problems in video databases. In this paper linear time and space algorithms for handling video metadata that represent objects or events present in various frames of the video sequence are presented. To accomplish this, we make a straightforward reduction of this problem to the intersection problem in Computational Geometry. Our first result is an improvement over the one of V. S. Subrahmanian [Subramanian, 1998] by a logarithmic factor in storage. This is achieved by using different basic data structures. Then, we present two other interesting time-efficient approaches. Finally a reduction to a special geometric problem is considered according to which we can achieve two optimal in time and space solutions in main and external memory model of computation respectively. We also present an extended experimental evaluation.

13. **Scheduling Algorithms for Procrastinators.** M. Bender, R. Clifford and K. Tsichlas. *Journal of Scheduling*, 11(2):95–104,2008.

    This paper presents scheduling algorithms for procrastinators, where the speed that a procrastinator executes a job increases as the due date approaches. We give optimal off-line scheduling policies for linearly increasing speed functions. We then explain the computational/numerical issues involved in implementing this policy. We next explore the online setting, showing that there exist adversaries that force any online scheduling policy to miss due dates. This impossibility result motivates the problem of minimizing the maximum interval stretch of any job; the interval stretch of a job is the job"s flow time divided by the job"s due date minus release time. We show that several common scheduling strategies, including the "hit-the-highest-nail" strategy beloved by procrastinators, have arbitrarily large maximum interval stretch. Then we give the "thrashing" scheduling policy and show that it is a $\Theta(1)$ approximation algorithm for the maximum interval stretch.

14. **A New Approach on Indexing Mobile Objects on the Plane.** S. Sioutas, K. Tsakalidis, K. Tsichlas, C. Makris, Y. Manolopoulos. *Data Knowledgment Engineering,* 67(3): 362–380, 2008.

    We present a set of time-efficient approaches to index objects moving on the plane to efficiently answer range queries about their future positions. Our algorithms are based on previously described solutions as well as on the employment of efficient access methods. Finally, an experimental evaluation is included that shows the performance, scalability and efficiency of our methods.

15. **Canonical Polygon Queries on the Plane: A New Approach.** S. Sioutas, D. Sofotassios, K. Tsichlas, D. Sotiropoulos and P. Vlamos. *Journal of Computers*, 4(9):913–919, 2009.

    On the Canonical k-vertex Polygon Spatial Retrieval Problem. V. Bistiolas, S. Sioutas, D. Sofotassios and K. Tsichlas. *In Proc. of the 15th Australian Workshop on Combinatorial Algorithms (AWOCA)*, 2004.

    The polygon retrieval problem on points is the problem of preprocessing a set of n points on the plane, so that given a polygon query, the subset of points lying inside it can be reported efficiently. It is of great interest in areas such as Computer Graphics, CAD applications, Spatial Databases and GIS developing tasks. In this paper we study the problem of canonical k-vertex polygon queries on the plane. A canonical k-vertex polygon query always meets the following specific property: a point retrieval query can be transformed into a linear number (with respect to the number of vertices) of point retrievals for orthogonal objects such as rectangles and triangles (throughout this work we call a triangle orthogonal iff two of its edges are axisparallel). We present two new algorithms for this problem. The first one requires $O(n \log^2 n)$ space and $O(k\frac{\log n}{\log\log n} + A$ query time. A simple modification scheme on the first algorithm lead us to a second solution, which consumes $O(n^2)$ space and $O(k\frac{\log n}{\log\log n} + A)$ query time, where A denotes the size of the answer and k is the number of vertices. The best previous solution for the general polygon retrieval problem uses $O(n^2)$ space and answers a query in $O(k \log n + A)$ time, where k is the number of vertices. It is also very complicated and difficult to be implemented in a standard imperative programming language such as C or C++.

16. **An Experimental Performance Comparison for Indexing Mobile Objects on the Plane.** S. Sioutas, G. Papaloukopoulos, K. Tsichlas and Y. Manolopoulos. *Special issue of ACM-SIGAPP MEDES 2009 on Collectively Intelligent Information and Knowledge Management, Journal on Organizational and Collective Intelligence (IJOCI)*, 1(4):78–96, 2010.

    An Experimental Performance Comparison for Indexing Mobile Objects on the Plane. S. Sioutas, G. Papaloukopoulos, K. Tsichlas and Y. Manolopoulos. *In Proc. of ACM-SIGAPP MEDES 2009*, pp. 210–217, 2009.

    In this paper, the authors present a time-efficient approach to index objects moving on the plane in order to answer range queries about their future positions. Each object is moving with non small velocity u, meaning that the velocity value distribution is skewed (Zipf) towards u_{min} in some range $[u_{min}, u_{max}]$, where u_{min} is a positive lower threshold. This algorithm enhances a previously described solution (Sioutas, Tsakalidis, Tsichlas, Makris and Manolopoulos, 2007) by accommodating the ISB-tree access method as presented in Kaporis et al. (2005). Experimental evaluation shows the improved performance, scalability, and efficiency of the new algorithm.

17. **ISB-Tree: A New Indexing Scheme with Efficient Expected Behaviour.** Ch. Makris, S. Sioutas, Tsakalidis, K. Tsichlas, Y. Ch. Zaroliagis. *Journal of Discrete Algorithms*, 8(4):373–387, 2010.

    ISB-Tree: A New Indexing Scheme with Efficient Expected Behaviour. A. Kaporis, C. Makris, G. Mayritsakis, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 16th Annual International Symposium on Algorithms and Computation (ISAAC)*, pp. 318–327, 2005.

    We present ISB-tree, a new indexing scheme that supports update operations (insertions and deletions) in $O(1)$ worst-case block transfers and search operations in $O(\log_B \log n)$ expected block transfers, where B represents the disk block size and n denotes the number of stored elements. The expected search bound holds with high probability for a large class of input distributions. The worst-case search bound of our indexing scheme is $O(\log_B n)$ block transfers. Our update and expected search bounds constitute a considerable improvement over the $O(\log_B n)$ worst-case block transfer bounds for search and update operations achieved by the B-tree and its numerous variants. Our indexing scheme is based on a novel combination of a new variant of the classical B-tree that we introduce here and call it *Lazy* B-tree, and on an externalization of a main memory data structure based on interpolation search.

18. **Improved Bounds for Finger Search on a RAM.** A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *Algorithmica*, 66(2):249–286, 2013.

    Improved Bounds for Finger Search on a RAM. A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 11th Annual European Symposium on Algorithms (ESA).*, LNCS 2832, pp. 325–336, 2003.

We present a new finger search tree with $O(\log\log d)$ expected search time in the Random Access Machine (RAM) model of computation for a large class of input distributions. The parameter d represents the number of elements (distance) between the search element and an element pointed to by a finger, in a finger search tree that stores n elements. Our data structure improves upon a previous result by Andersson and Mattsson that exhibits expected $O(\log\log n)$ search time by incorporating the distance d into the search time complexity, and thus removing the dependence on n. We are also able to show that the search time is $O(\log\log d + \phi(n))$ with high probability, where $\phi(n)$ is *any* slowly growing function of n. For the need of the analysis we model the updates by a "balls and bins" combinatorial game that is interesting in its own right as it involves insertions and deletions of balls according to an unknown distribution.

19. **On the Discovery of Group-Consistent Graph Substructure Patterns from Brain Networks.** N.D. Iakovidou, S.I. Dimitriadis, N.A. Laskaris, K. Tsichlas, Y. Manolopoulos.*Journal of Neuroscience*, 213(2):204–213, 2013.

    Complex networks constitute a recurring issue in the analysis of neuroimaging data. Recently, network motifs have been identified as patterns of interconnections since they appear in a significantly higher number than in randomized networks, in a given ensemble of anatomical or functional connectivity graphs. The current approach for detecting and enumerating motifs in brain networks requires a predetermined motif repertoire and can operate only with motifs of small size (consisting of few nodes). There is a growing interest in methodologies for frequent graph-based pattern mining in large graph datasets that can facilitate adaptive design of motifs. The results presented in this paper are based on the graph-based Substructure pattern mining (gSpan) algorithm and introduce a manifold of ways to exploit it for data-driven motif extraction in connectomics research.

    Functional connectivity graphs from electroencephalographic (EEG) recordings during resting state and mental calculations are used to demonstrate our approach. Relying on either time-invariant or time-evolving graphs, characteristic motifs associated with various frequency bands were derived and compared. With a suitable manipulation, the gSpan discovers motifs which are specific to performing mental arithmetics. Finally, the subject-dependent temporal signatures of motifs" appearance revealed the transient nature of the evolving functional connectivity (math-related motifs "come and go").

20. **ART: Sub-Logarithmic Decentralized Range Query Processing with Probabilistic Guarantees.** S. Sioutas, P. Triantafillou, G. Papaloukopoulos, E. Sakkopoulos, K. Tsichlas, Y. Manolopoulos. *Distributed and Parallel Databases*, 31(1):71–109. 2013.

    ART–sub-logarithmic decentralized range query processing with probabilistic guarantees. A. Kaporis, A. Papadopoulos, S. Sioutas, K. Tsakalidis and K. Tsichlas. *Brief announcement in Proc. of the 29th Annual Symposium on Principles of Distributed Computing (PODC)*, pp. 118–119, 2010.

    We focus on range query processing on large-scale, typically distributed infrastructures, such as clouds of thousands of nodes of shared-datacenters, of p2p distributed overlays, etc. In such distributed environments, efficient range query processing is the key for managing the distributed data sets per se, and for monitoring the infrastructure"s resources. We wish to develop an architecture that can support range queries in such large-scale decentralized environments and can scale in terms of the number of nodes as well as in terms of the data items stored. Of course, in the last few years there have been a number of solutions (mostly from researchers in the p2p domain) for designing such large-scale systems. However, these are inadequate for our purposes, since at the envisaged scales the classic logarithmic complexity (for point queries) is still too expensive while for range queries it is even more disappointing. In this paper we go one step further and achieve a sub-logarithmic complexity. We contribute the ART (Autonomous Range Tree) structure, which outperforms the most popular decentralized structures, including Chord (and some of its successors), BATON (and its successor) and Skip-Graphs. We contribute theoretical analysis, backed up by detailed experimental results, showing that the communication cost of query and update operations is $O(\log_b^2 \log N)$ hops, where the base b is a double-exponentially power of two and N is the total number of nodes. Moreover, ART is a fully dynamic and fault-tolerant structure, which supports the join/leave node operations in $O(\log \log N)$ expected w.h.p. number of hops. Our experimental performance studies include a detailed performance comparison which showcases the improved performance, scalability, and robustness of ART.

Conferences

1. **Data Structuring Applications for String Problems in Biological Sequences.** Y. Panagis, E. Theodoridis, K. Tsichlas. *In Proc. of the International Conference of Computational Methods in Science and Engineering (ICCMSE)*, pp. 479–483, 2003.

   In this work we present applications of data structuring techniques for string problems in biological sequences. We firstly consider the problem of approximate string matching with gaps and secondly the problem of identifying occurrences of maximal pairs in multiple strings. The proposed implementations can be used in many problems that arise in the field of Computational Molecular Biology.

2. **Algorithms for Extracting Structured Motifs from Biological Weighted Sequences.** C. Iliopoulos, K. Perdikuri, A. Tsakalidis and K. Tsichlas. In Proc. of the *the 16th Australasian Workshop on Combinatorial Algorithms (AWOCA)*, 2005.

   Biological Weighted Sequences are used extensively in Molecular Biology as profiles for protein families, in the representation of binding sites and often for the representation of sequences produced by a shotgun sequencing strategy. In this paper we present two algorithms. The first extracts structured motifs with $p = 2$ models from a set of $N \geq 2$ weighted sequences. The second algorithm extracts structured motifs with $p > 2$ models from a set of $N \geq 2$ weighted sequences. To the best of our knowledge, this is the first time these problems are tackled in the weighted sequences setting.

3. **Finding Multirepeats in a Set of Strings.** A. Bakalis, C. Makris, S. Sioutas, E. Theodoridis and K. Tsichlas. In Proc. of the *International Conference of Computational Methods in Sciences and Engineering (ICCMSE)*, 2005.

   A multirepeat in a string is a substring that appears many times. A multirepeat is maximal if it cannot be extended either to the right or to the left and produce a multirepeat. In this paper, we present algorithms for two different versions of the problem of finding maximal multirepeats in a set of strings. In the case of arbitrary gaps, we propose an algorithm with $O(N^2 n + \alpha)$ time complexity. When the gap is bounded in a small range and the multirepeat is composed of consecutive maximal repeats we propose an algorithm with

$O(mN^2 n \log(Nn) + \alpha)$ time complexity. Here N is the number of strings, n is the mean length of each string, m is multiplicity of the multirepeat and α is the number of reported occurrences. Throughout the paper the assumption is made that the size of the alphabet is constant.

4. **Dynamic Interpolation Search Revisited.** A. Kaporis, C. Makris, S. Sioutas, A. Tsakalidis, K. Tsichlas and C. Zaroliagis. *In Proc. of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 382–394, 2006.

   A new dynamic Interpolation Search (IS) data structure is presented that achieves $O(\log\log n)$ search time with high probability on unknown continuous or even discrete input distributions with measurable probability of key collisions, including power law and Binomial distributions. No such previous result holds for IS when the probability of key collisions is measurable. Moreover, our data structure exhibits $O(1)$ expected search time with high probability for a wide class of input distributions that contains all those for which $o(\log\log n)$ expected search time was previously known.

5. **Algorithms for Bitmasking Strings.** A. Bakalis, C. Iliopoulos, S. Sioutas and K. Tsichlas. Accepted for presentation in *International Conference of Computational Methods in Sciences and Engineering (ICCMSE)*, pp. 1061–1063, 2006.

   In this paper we present algorithms for efficiently bitmasking strings. In particular, we assume a pattern (the bitmask) consisting of zeroes and ones as well as a text over an alphabet Σ and the goal is to extract information efficiently when we apply the pattern on the text. Three algorithms are presented, each one based on different techniques, exhibiting various time complexities.

6. **Purely Functional Worst Case Constant Time Catenable Sorted Lists.** G.S. Brodal, C. Makris and K. Tsichlas. *In Proc. of the 13th Annual European Symposium on Algorithms (ESA),* pp. 172–183, 2006.

   We present a purely functional implementation of search trees that requires $O(\log n)$ time for search and update operations and supports the join of two trees in worst case constant time. Hence, we solve an open problem posed by Kaplan and Tarjan as to whether it is possible to envisage a data structure supporting simultaneously the join operation in $O(1)$ time and the search and update operations in $O(\log n)$ time.

7. **Indexing of mobile objects on the plane revisited.** S. Sioutas, K. Tsakalidis, K. Tsichlas, C. Makris and Y. Manolopoulos. *In Proc. of the 11th East-European Conference on Advances in Databases and Information Systems (ADBIS)*, pp. 189–204, 2007.

   We present a set of time-efficient approaches to index objects moving on the plane to efficiently answer range queries about their future positions. Our algorithms are based on previously described solutions as well as on the employment of efficient data structures. Finally, an experimental evaluation is included that shows the performance, scalability and efficiency of our methods.

8. **New Expected Bounds for Three-sided Range Queries on the Plane.** G.S. Brodal, A. Kaporis, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 20th Int. Symposium on Algorithms and Computation (ISAAC)*, pp. 193–202, 2009.

We consider the problem of maintaining dynamically a set of points on the plane and supporting range queries of the type $[a, b] \times (-\infty, c]$. We assume that the inserted points have their x-coordinates drawn from a *smooth* distribution, whereas the y-coordinates are arbitrarily distributed. The points to be deleted are selected uniformly at random among the inserted points. We present two solutions, one for the RAM model and one for the I/O model.

In the RAM model, we present a linear space data structure that supports queries in $O(\log\log n + t)$ expected with high probability time and updates in $O(\log\log n)$ expected amortized time, where n is the number of stored points and t is the size of the output of the query. The data structure is deterministic and the expectation is with respect to the input distribution. We juxtapose our results with the linear space data structure of Willard that attains $O\left(\frac{\log n}{\log\log n} + t\right)$ worst case range query time and $O\left(\frac{\log n}{\log\log n}\right)$ worst case or $O(\sqrt{\log n})$ randomized update time, without assumptions on the input distribution.

In the I/O model we achieve $O(\log\log_B n + t/B)$ expected with high probability query I/Os and $O(\log_B \log n)$ expected amortized update I/Os in linear space. This can be compared with the result of Arge et al. who achieve range queries in $O(\log_B n + t/B)$ I/Os worst case and updates in $O(\log_B n)$ I/Os amortized in linear space, without assumptions on the input distribution.

9. **A novel Distributed P2P Simulator Architecture: D-P2P-Sim.** S. Sioutas, G. Papaloukopoulos, E. Sakkopoulos, K. Tsichlas and Y. Manolopoulos. *In Proc. of ACM-CIKM*, pp. 2069–2070, 2009.

   In this paper we introduce a novel distributed simulation environment with GUI for P2P simulations (D-P2P-Sim). The key aim is to provide the appropriate integrated set of tools in a single software solution to evaluate the performance of various protocols. The basic architecture of the distributed P2P simulator is based on a multi-threading, asynchronous, message passing and distributed environment with graphical user interface to facilitate ease of use by both researchers and programmers. More information online: http://students.ceid.upatras.gr/ papalukg

10. **Efficient Processing of 3-sided Range Queries with Probabilistic Guarantees.** A. Kaporis, A. Papadopoulos, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 13th International Conference on Database Theory (ICDT)*, pp. 34–43, 2010.

    This work studies the problem of 2-dimensional searching for the 3-sided range query of the form $[a, b] \times (-\infty, c]$ in both main and external memory, by considering a variety of input distributions. A dynamic linear main memory solution is proposed, which answers 3-sided queries in $O(\log n + t)$ worst case time and scales with $O(\log\log n)$ expected with high probability update time, under continuous μ-random distributions of the x and y coordinates, where n is the current number of stored points and t is the size of the query output. Our expected update bound constitutes a considerable improvement over the $O(\log n)$ update time bound achieved by the classic Priority Search Tree of McCreight, as well as over the Fusion Priority Search Tree of Willard, which requires $O(\frac{\log n}{\log\log n})$ time for all operations. Moreover, we externalize this solution, gaining $O(\log_B n + t/B)$ worst case and $O(log_B logn)$ amortized expected with high probability I/Os for query and update operations respectively, where B is the disk block size. Then, combining the Modified Priority Search Tree with the Priority Search Tree, we achieve a query time of $O(\log\log n + t)$ expected with high probability and an update time of $O(\log\log n)$ expected with high probability, under the assumption that the x-coordinates are continuously drawn from a smooth distribution and the y-coordinates are continuously drawn from a more restricted class of distributions. The total space is linear. Finally, we externalize this solution, obtaining a dynamic data structure that answers 3-sided queries in $O(\log_B \log n + t/B)$ I/Os expected with high probability, and it can be updated in $O(\log_B \log n)$ I/Os amortized expected with high probability and consumes $O(n/B)$ space, under the same assumptions.

11. **The D^2-tree: a new P2P deterministic data structure.** G.S. Brodal, S. Sioutas, K.Tsichlas and C. Zaroliagis. *In Proc. of the 21st Int. Symposium on Algorithms and Computation (ISAAC)*, pp. 1–12, 2010.

    We present a new overlay, called the Deterministic Decentralized tree (D^2-tree). The D^2-tree compares favourably to other overlays for the following reasons: (a) it provides matching and better complexities, which are deterministic for the supported operations; (b) the management of nodes (peers) and elements are completely decoupled from each other; and (c) an efficient deterministic load-balancing mechanism is presented for the uniform distribution of elements into nodes, while at the same time probabilistic optimal bounds are provided for the congestion of operations at the nodes.

12. **Continuous Monitoring of Distance-based Outliers Over Streams.** M. Kontaki, A. Gounaris, Y. Manolopoulos, A. Papadopoulos and K. Tsichlas. *In Proc. of the 26th IEEE Int. Conference on Data Engineering (ICDE)*, pp. 135–146, 2011.

    Anomaly detection is considered an important data mining task, aiming at the discovery of elements (also known as *outliers*) that show significant diversion from the expected case. More specifically, given a set of objects the problem is to return the suspicious objects that deviate significantly from the typical behavior. As in the case of clustering, the application of different criteria lead to different definitions for an outlier. In this work, we focus on distance-based outliers: an object x is an outlier if there are less than k objects lying at distance at most R from x. The problem offers significant challenges when a stream-based environment is considered, where data arrive continuously and outliers must be detected on-the-fly. There are a few research works studying the problem of continuous outlier detection. However, non of these proposals meets the requirements of modern stream-based applications for the following reasons: (i) they demand a significant storage overhead, (ii) their efficiency is limited and (iii) they lack flexibility. In this work, we propose new algorithms for continuous outlier monitoring in data streams, based on sliding windows. Our techniques are able to reduce the required storage overhead, run faster than previously proposed techniques and offer significant flexibility. Experiments performed on real-life as well as synthetic data sets verify our theoretical study.

13. **NEFOS: Rapid Cache-Aware Range Query Processing with Probabilistic Guarantees.** S. Sioutas, K. Tsichlas, I. Karydis, Y. Manolopoulos and Y. Theodoridis. *In Proc. of the 22nd Int. Conference on Database and Expert Systems Applications (DEXA)*, pp. 62–77, 2011.

    We present NEFOS (NEsted FOrest of balanced treeS), a new cache-aware indexing scheme that supports insertions and deletions in $O(1)$ worst-case block transfers for rebalancing operations (given and update position) and searching in $O(\log_B \log n)$ expected block transfers,

(B= disk block size and n= number of stored elements). The expected search bound holds with high probability for any (unknown) *realistic* input distribution. Our expected search bound constitutes an improvement over the $O(\log_B \log n)$ expected bound for search achieved by the ISB-tree (Interpolation Search B-tree), since the latter holds with high probability for the class of *smooth* only input distributions. We define any unknown distribution as realistic if the smoothness doesn't appear in the whole data set, still it may appear locally in small spatial neighborhoods. This holds for a variety of real-life non-smooth distributions like skew, zipfian, powlaw, beta e.t.c.. The latter is also verified by an accompanying experimental study. Moreover, NEFOS is a B-parametrized concrete structure, which works for both I/O and RAM model, without any kind of transformation or adaptation. Also, it is the first time an expected sub-logarithmic bound for search operation was achieved for a broad family of non-smooth input distributions.

14. **Fully Persistent B-trees.** G.S. Brodal, S. Sioutas, K. Tsakalidis and K. Tsichlas. *In Proc. of the 23rd Symposium on Discerete Algorithms (SODA)*, pp. 602–614, 2012.

    We present an I/O-efficient fully persistent B-Tree that supports range searches in $O(\log_B n + t/B)$ I/Os and updates in $O(\log_B n)$ amortized I/Os, using space $O(m/B)$ disk blocks, where n denotes the number of elements in the version of the data structure accessed, m the total number of updates, t the size of the output of a range query and B the disk block size. To achieve the result, we first present a new B-Tree implementation, which supports searches and updates in $O(\log_B n)$ I/Os and where every update makes in the worst case $O(1)$ changes to the data structure. We make this B-Tree fully persistent using an adaptation of the *node-splitting* method, introduced by Driscoll et al, to the I/O model. The new method can be applied to any external memory pointer based data structure with indegree bounded by $O(1)$ and outdegree bounded by $O(B)$, such that the overhead is $O(1)$ amortized I/Os and $O(1/B)$ blocks amortized space per update step. The access to a block of the ephemeral data structure can be done with $O(1)$ I/Os in the worst case. The result improves the previous fully persistent B-Tree of Lanka and Mays by a factor of $O(\log_B n)$ for the update and $O(\log_B m)$ for the range query complexity and answers affirmatively a question posed by Vitter, whether a B-Tree can be made fully persistent with constant overhead.

15. **DISCO: a New Algorithm for Detecting 3D Protein Structure Similarity.** N.D. Iakovidou, E. Tiakas, K. Tsichlas. *In Proc. of the 1st Workshop on Algorithms for Data and Text Mining in Bioinformatics (WADTMB) (Artificial Intelligence Applications and Innovations)*, 622–631, 2012.

    Protein structure similarity is one of the most important aims pursued by bioinformatics and structural biology, nowadays. Although quite a few similarity methods have been proposed lately, yet fresh algorithms that fulfill new preconditions are needed to serve this purpose. In this paper, we provide a new similarity measure for $3D$ protein structures that detects not only similar structures but also similar substructures to a query protein, supporting both multiple and pairwise comparison procedures and combining many comparison characteristics. In order to handle similarity queries we utilize efficient and effective indexing techniques such as M-trees and we provide interesting results using real, previously tested protein data sets.

16. **I/O-Efficient Orthogonal Planar Range Skyline Reporting and Catenable Priority Queues with Attrition.** C. Kejlberg-Rasmussen, Y. Tao, K. Tsakalidis, K. Tsichlas, J. Yoon. Accepted for presentation in *(PODS)*, 2013.

We study the static and dynamic *planar range skyline reporting problem* in the external memory model with block size B, under a linear space budget. The problem asks for an $O(n/B)$ space data structure that stores n points in the plane, and supports reporting the k maximal input points (a.k.a. *skyline*) among the points that lie within a given query rectangle $Q = [\alpha_1, \alpha_2] \times [\beta_1, \beta_2]$. When Q is *3-sided*, i.e. one of its edges is grounded, two variants arise: *top-open* for $\beta_2 = \infty$ and *left-open* for $\alpha_1 = -\infty$ (symmetrically *bottom-open* and *right-open*) queries.

We present optimal static data structures for *top-open* queries, for the cases where the universe is $\mathbb{R}^2$, a $U \times U$ grid, and rank space $[O(n)]^2$. We also show that *left-open* queries are harder, as they require $\Omega((n/B)^\epsilon + k/B)$ I/Os for $\epsilon > 0$, when only linear space is allowed. We show that the lower bound is tight, by a structure that supports 4-sided queries in matching complexities. Interestingly, these lower and upper bounds coincide with those of the *planar orthogonal range reporting problem*, i.e., the skyline requirement does not alter the problem difficulty at all!

Finally, we present the first dynamic linear space data structure that supports top-open queries in $O(\log_{2B^\epsilon} n + k/B^{1-\epsilon})$ and updates in $O(\log_{2B^\epsilon} n)$ worst case I/Os, for $\epsilon \in [0, 1]$. This also yields a linear space data structure for 4-sided queries with optimal query I/Os and $O(\log(n/B))$ amortized update I/Os. We consider of independent interest the main component of our dynamic structures, a new real-time I/O-efficient and catenable variant of the fundamental structure *priority queue with attrition* by Sundar.

17. **Multi-objective optimization of data flows in a multi-cloud environment.** E. Tsamoura, A. Gounaris, K. Tsichlas. Accepted for presentation in *Workshop on Data analytics in the Cloud (DanaC)*, 2013.

    As cloud-based solutions have become one of the main choices for intensive data analysis both for business decision making and scientific purposes, users face the problem of choosing among different cloud providers. In this work, we deal with data analysis flows that can be split in stages, and each stage can run on multiple cloud infrastructures. For each stage, a cloud provider may make a bid in the form of a continuous function in the time delay-monetary cost domain. The goal is to compute the optimal combination of bids according to how much a user is prepared to pay for the total time delay to execute the analysis task. The contributions of this work are (i) to provide a solution that can be computed in pseudo-polynomial time and with bounded relative error for the generic case; (ii) to provide exact polynomial solutions for specific cases; and (iii) to experimentally evaluate our proposal against other techniques. Our extensive results show that we can increase user satisfaction up to an order of magnitude compared to existing heuristics.

18. **Continuous Outlier Detection in Data Streams: An Extensible Framework and State-Of-The-Art Algorithms.** D. Georgiadis, M. Kontaki, A. Gounaris, A. Papadopoulos, K. Tsichlas, Y. Manolopoulos. *SIGMOD Demonstration*, 2013.

    Anomaly detection is an important data mining task, aiming at the discovery of elements that show significant diversion from the expected behavior; such elements are termed as outliers. One of the most widely employed criteria for determining whether an element is an outlier is based on the number of neighboring elements within a fixed distance (R), against a fixed threshold (k). Such outliers are referred to as distance-based outliers and are the focus of this work. In this demo, we show both an extendible framework for outlier detection algorithms and specific outlier detection algorithms for the demanding case where outlier detection is continuously performed over a data stream. More specifically: i) first we demonstrate a novel flavor of an open-source publicly available tool for Massive Online Analysis (MOA) that is endowed with capabilities to encapsulate algorithms that continuously detect outliers and ii) second, we present four online outlier detection algorithms. Two of these algorithms have been designed by the authors of this demo, with a view to improving on key aspects related to outlier mining, such as running time, flexibility and space requirements.