

Matteo Comin

Curriculum Vitae

Personal Data:

Name: Matteo Comin
Data of birth: 02/20/1977
Place of birth: Venice, Italy
Citizenship: Italian
Military status: Honorable discharge
E-mail: ciompin@dei.unipd.it

Professional Experience:

- 2007 October - Ricercatore (Tenured Assistant Professor), Department of Information Engineering, University of Padova.
- 2007 March - September. Research Fellow at the Department of Information Engineering, University of Padova.
- 2006 June - October. Research intern in the group “Genographic”, supervisor Dr. Ajay Royyuru, IBM T.J. Watson Research Center, Yorktown Heights, New York, U.S..
- 2006 April. Member of the Organizing Committee of the “Tenth Annual International Conference on Research in Computational Molecular Biology, RECOMB 2006”, April 2-5, 2006, Venice, Italy.
- 2004 September - December. Research intern in the group “Bioinformatics and Pattern Discovery”, supervisor Dr. Isidore Rigoutsos, IBM T.J. Watson Research Center, Yorktown Heights, New York, U.S..
- 2004 - 2006. Doctoral Student, Ph.D program in Computer Science, Department of Information Engineering, University of Padova.
- 2003 April - December. Research Collaborator, F.I.R.B. Project “Enabling platforms for high-performance computational grids”, Supervisor: prof. A. Apostolico, MI.U.R., 2003-2005.
- 2002 October - November. Research Fellowship: Research on Algorithms and Data Compression, Computer Sciences Department, Purdue University, Indiana, US.

Involved as Research Assistant in the following projects:

- Ateneo Project “Computational assessment of Protein-Protein interaction networks: target prediction and validation as guide for modern system biology”, University of Padova 2008-10.
- Italy-Israel F.I.R.B Project “Pattern Discovery Algorithms in Discrete Structures, with Applications to Bioinformatics”, MI.U.R. 2006-2009.
- P.R.I.N.: “Data mining methods for e-business applications”, MI.U.R. 2005-2006.
- F.I.R.B.: “Enabling platforms for high-performance computational grids”, Ministry of University and Research (MI.U.R.) 2003-2005.
- F.I.R.B.: “Bioinformatics for genomics and proteomics”, MI.U.R. 2003-2005.

Education:

- 2007 - Ph.D. in Computer Science, Department of Information Engineering, University of Padova.
- 2004 - 2006. Doctoral Student, Ph.D program in Computer Science, Department of Information Engineering, University of Padova (XIX cycle).
Thesis title: “Motif Discovery and Data Compression: Models, Methods and Applications”.
Supervisor: prof. Alberto Apostolico.
- 2003 - Qualifying for engineering profession.
- 2003 - Laurea degree in Computer Science, Department of Information Engineering University of Padova. Supervisor: Prof. Alberto Apostolico.
- 1996 - High School Diploma in Computer Science, C. Zuccante Institute, Mestre (VE).

Didactic and Seminarial Activities:

- a.a. 2008/2009 - Course “Computer Architectures”, Dept. of Information Engineering, University of Padova.
- a.a. 2007/2008 - Lecturer within the course “Three-Dimensional Data Elaboration”, Prof. C. Guerra, Dept. of Information Engineering, University of Padova.
- a.a. 2007/2008 - Lecturer within the course “Computer Networks”, Prof. C. Guerra, Dept. of Information Engineering, University of Padova.
- a.a. 2005/2006 - Lecturer within the course “Distributed Systems”, Prof. C. Ferrari, Dept. of Information Engineering, University of Padova.
- a.a. 2005/2006 - Lecturer within the course “Three-Dimensional Data Elaboration”, Prof. C. Guerra, Dept. of Information Engineering, University of Padova.

- a.a. 2005/2006 - Lecturer within the course “Computer Networks”, Prof. C. Guerra, Dept. of Information Engineering, University of Padova.
- a.a. 2004/2005 - Lecturer within the course “Bioinformatics and Computational Biology”, Prof. A. Apostolico, Dept. of Information Engineering, University of Padova.

Seminars:

- 2007 - “From Conventional Biology to Computational Biology”, University of Verona.
- 2007 - “Some Results on Computational Biology”, *C. Offelli Award* for best young researcher, Dept. of Information Engineering, University of Padova.
- 2006 - “Mining Regularities: The Experience with Biological Data”, University of Trieste.
- 2005 - “Conservative Extraction of Over-Represented Extensible Motifs”, presented at the DIMACS Workshop on Detecting and Processing Regularities in High Throughput Biological Data, Rutgers University, New Jersey, US.
- 2005 - “Conservative Extraction of Over-Represented Extensible Motifs”, IBM T.J. Watson Research Center, New York, U.S..
- 2005 - “Pattern Discovery and Data Compression: Models, Methods and Tools”, Dept. of Information Engineering, University of Padova.
- 2004 - “From Motif Discovery to Off-line Data Compression”, research meeting within the “International School on Advanced BioMedicine and BioInformatics” Lipari Island, June 2004.

Specialized Courses and Conferences:

- “Information Beyond Shannon”, December 29-30, 2008, Venice, Italy.
- “7th European Conference on Computational Biology”, September 22-26, 2008, Cagliari, Italy.
- “Biological Networks: Evolution, Interaction and Computation”, June 14-21, 2008, Lipari Island.
- “Fifth IEEE International Workshop on High Performance Computational Biology”, HiCOMB 2006 is held in conjunction with the IEEE International Parallel and Distributed Processing Symposium, April 2006, Rhodes Island, Greece.
- “Tenth Annual International Conference on Research in Computational Molecular Biology”, RECOMB 2006, April 2-5, 2006, Venice, Italy.
- FIMA International Conference on “Models and Methods for Human Genomics”, Aya-Champoluc, Aosta Valley (Italy), January 23-27, 2006.

- “13th Annual International conference on Intelligent Systems for Molecular Biology (ISMB 2005)”, annual meeting of the International Society of Computational Biology, June 25-29, 2005, Detroit, Michigan, US.
- “DIMACS Workshop on Detecting and Processing Regularities in High Throughput Biological Data”, DIMACS Center, Rutgers University, New Jersey, US. June 20 - 22, 2005.
- “BITS 2005: the annual meeting of the bioinformatics Italian society”, Milan, March 17-19, 2005.
- “DIMACS Working Group on The Burrows - Wheeler Transform: Ten Years Later”, DIMACS Center, Rutgers University, New Jersey, US, August 19 - 20, 2004.
- “First International School on Advanced BioMedicine and BioInformatics”, Lipari Island, May 29 - June 5, 2004.
- “ASSIST Course”, Department of Informatics Univ. Pisa, May 5-7, 2004.
- “International Summer School on Grid Computing”, Vico Equense (NA), July 13-25, 2003, Director: Paul Messina, Argonne National Laboratory.
- Doctoral courses at the University di Padova: Writing for PhDs (Prof. S. Guth), Coding Theory (prof. F. Fagnani), Data Compression (prof. R. Giancarlo), Distributed Systems (prof. C. Ferrari), Machine Learning (prof. P. Frasconi).

Reviewer for:

International Journals:

Parallel Processing Letters (PPL), ACM Journal of Experimental Algorithmics (JEA), Theoretical Computer Science (TCS), Journal of Visual Communication and Image Representation (JVCI), BMC Algorithm for Molecular Biology (AMB), BMC Bioinformatics, Journal of Bioinformatics Research and Applications (IJBRA), ACM/IEEE Transaction in Computational Biology and Bioinformatics (TCBB), Journal of Computational Biology (JCB), Journal of Foundation of Computer Science (JFCS).

International Conferences:

European Conference on Parallel Computing (EuroPar), Workshop on Algorithms in Bioinformatics (WABI), Research in Computational Molecular Biology (RECOMB), The Asia-Pacific Bioinformatics Conference (APBC), Computational Systems Bioinformatics (CSB), Combinatorial Pattern Matching (CPM), International Workshop in Bioinformatics and Research Applications (IWBRA), IEEE International Conference on Data Mining (ICDM).

Awards and Fellowships:

- 2007 *C. Offelli Award* for best young researcher, Department of Information Engineering, University of Padova.

- 2005 International Society of Computational Biology Travel Fellowship to attend the “13th Annual International Conference on Intelligent Systems for Molecular Biology” (ISMB 2005), Detroit, USA.
- 2005 U.S. National Science Foundation Travel Fellowship to attend the workshop “DIMACS Workshop on Detecting and Processing Regularities in High Throughput Biological Data”, DIMACS Center, Rutgers University, New Jersey, USA.
- 2004 U.S. National Science Foundation Travel Fellowship to attend the workshop “DIMACS Working Group on The Burrows - Wheeler Transform: Ten Years Later”, DIMACS Center, Rutgers University, New Jersey, USA.

Patents:

Inventor of the following IBM patents:

- “Method and structure for lossy compression of continuous data with extensible motifs”, US Patent 7259701, 2007.
- “Methods and systems for conservative extraction of over-represented extensible motifs”, United States Patent Application 20070299615, filed.
- “Method and apparatus for detection consensus motifs in data sequences”, United States Patent Application 20080167850, filed.

PUBLICATIONS

International Journals:

1. A. Apostolico, M. Comin, L. Parida, "VARUN: Discovering Extensible Motifs under Saturation Constraints", accepted for publication in *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2009).
2. M. Comin, C. Guerra, G. Zanotti, "Mining Over-Represented 3DPatterns of Secondary Structures in Proteins", *Journal of Bioinformatics and Computational Biology*, Vol. 6 , No. 06 (2008), pp. 1067 - 1087.
3. G. Ciriello, M. Comin, C. Guerra, "Algorithmic Re-Structuring and Data Replication for Protein Structure Comparison on a GRID", *Future Generation Computer System* (23) Special Issue on Life Science Grids for Biomedicine and Bioinformatics, (2007), pp. 391-397.
4. A. Apostolico, M. Comin, L. Parida, "Mining, Compressing and Classifying with Extensible Motifs", *BioMed Central Algorithms for Molecular Biology*, Vol. 1:4, (2006).
5. A. Apostolico, M. Comin, L. Parida, "Bridging lossy and lossless compression by motif pattern discovery", (extended abstract), *Electronic Notes in Discrete Mathematics*, Vol. 21 (2005), pp. 219-225.
6. M. Comin, C. Guerra, G. Zanotti, "PROuST: a comparison method of three-dimensional structures of proteins using indexing techniques", *Journal of Computational Biology*, Mary Ann Liebert Inc. , Vol. 11, No. 6 (2004), pp. 1061-1072.
7. M. Comin, C. Ferrari, C. Guerra, "Grid deployment of bioinformatics applications: a case study in protein similarity determination", *Parallel Processing Letters*, World Scientific Publishing Company, Vol. 14 No. 2 (2004), pp. 163-176.

Book chapters and invited contributions:

8. M. Comin, C. Guerra, G. Zanotti, "Invariant geometric properties of secondary structure elements in proteins", to appear in "Biological Data Mining" edited by J. Chen and S. Lonardi, Chapman & Hall/CRC.
9. M. Comin, L. Parida, "Detection of Subtle Variations as Consensus Motifs". in "SAIL String Algorithms, Information and Learning" edited by R. Giancarlo and S. Lonardi, special issue of *Theoretical Computer Science*, Vol. 395 No. 2-3 (2008), pp. 158-170.
10. A. Apostolico, M. Comin, L. Parida, "Bridging Lossy and Lossless Compression by Motif Pattern Discovery", *General Theory of Information Transfer and Combinatorics*, *Lecture Notes in Computer Science*, Vol. 4123 (2006), pp. 793-813.
11. M. Cannataro, M. Comin, C. Ferrari, C. Guerra, A. Guzzo, P. Veltri, "Modelling a Protein Structure Comparison Application on the Grid using PROTEUS", SAG 2004, *Lecture Notes in Computer Science* Vol. 3458 (2005), pp. 75-85.

International Conferences:

12. M. Comin, C. Guerra, G. Zanotti, “Mining Over-Represented 3D Patterns of Secondary Structures in Proteins”, Proceedings of *7th International Workshop on Data Mining in Bioinformatics (BIOKDD '07)*, pp 19-26.
(accepted 25%)
13. M. Comin, L. Parida, “Subtle Motifs Discovery for Detection of DNA Regulatory Sites”, Proceeding of *Asian-Pacific Bioinformatics Conference, APBC 2007*, pp. 27-36.
(accepted 17%)
14. G. Ciriello, M. Comin, C. Guerra, “Application Re-Structuring and Data Management on a GRID Environment: a Case Study for Bioinformatics”, Proceedings of *IEEE International Workshop on High Performance Computational Biology*, 2006:9.
15. G. Ciriello and M. Comin, “High-Performance Protein Structure Comparison”, Proceedings of the *Tenth Annual International Conference on Research in Computational Molecular Biology*, RECOMB (2006).
16. A. Apostolico, M. Comin, L. Parida, “Conservative Extraction of Over-represented Extensible Motifs”, Proceeding of the 13th Annual International conference on Intelligent Systems for Molecular Biology (ISMB 2005) Supplement of *Bioinformatics* 2005: pp. 9-18.
(accepted 13%, impact factor 6.02)
17. A. Apostolico, M. Comin, L. Parida, “Off-line Compression by Extensible Motifs”, Proceedings of *IEEE Data Compression Conference*, Computer Society Press, pp. 450, (2005).
18. A. Apostolico, M. Comin, L. Parida, “Motifs in Ziv-Lempel-Welch Clef”, Proceedings of *IEEE Data Compression Conference*, Computer Society Press, pp. 72-81, (2004).
(accepted 20%)

Other Manuscripts:

19. M. Comin, “Motif Discovery and Data Compression: Models, Methods and Applications”. Ph.D Thesis, Ph.D Program in Computer Science, Department of Information Engineering, University of Padova (XIX cycle). (supervisor: Prof. Alberto Apostolico)
20. M. Comin, “Motivi Strutturali e Compressione Fuori Linea”. Laurea Thesis, Computer Science, Department of Information Engineering, University of Padova.