

Curriculum vitae

Salvatore Orlando

November 2014

Contents

1	Personal information	2
2	Biographical sketch	2
2.1	Education and Work experience	3
3	Research activities	3
4	Teaching Activities	5
5	Scientific Supervision Activities	6
6	University Services	7
7	Coordination of / Participation in research projects	8
8	Committes of Scientific Events	9
9	Publications	13
	Journal Articles	13
	Refereed International Conference and Workshop Papers	14
	Edited Books and Journals	20
	Chapters in International Research Books	20
	Refereed National Conference Papers	21
	Other Works	22
	PhD Thesis	22

1 Personal information

Prof. Salvatore ORLANDO, PhD

Born in Cosenza, ITALY - June 14th, 1961

Associate Professor of Computer Science (INF/01 - Informatica)

Università Ca' Foscari Venezia

Dipartimento di Scienza Ambientali, Informatica e Statistica

Head of the BSc and MSc Programs in Computer Science

Phone: +39 041 2348428

Cell. Phone: +39 349 8625043

email: orlando@unive.it

URL: <http://www.dais.unive.it/~orlando>

Home Address: Via Raffaello Sanzio, 18/A - 56122 Pisa - Phone: +39 050 531330 -

Cell. phone: +39 349 8625043

2 Biographical sketch

Salvatore Orlando (Laurea/MSc (summa cum laude, 1985) and PhD (1991) in Computer Science, University of Pisa) was a post-doc fellow of the HP laboratories, and a post-doc fellow of the University of Pisa. In 1994 he joined as an assistant professor the Ca' Foscari University of Venice, where since 2000 he has been an associate professor. He is also a research associate at ISTI-CNR of Pisa. Since 2010 he has been appointed as Head of the BSc and MSc Programs in Computer Science of Ca' Foscari University. He is also the delegate of the Rector for national and international university orientation.

His research interests include the design of scalable algorithms for data/web mining problems, distributed and P2P systems for information retrieval, parallel/distributed systems and programming environments. In the last year he has supervised the research work of many doctoral students of the PhD Program in Computer Science of Ca' Foscari University, seven of which already completed their studies. In addition, he has coordinated the research activities of several post-docs.

He was a visiting scientist at the School of Computer Science of Carnegie Mellon University (USA), and at University of Paderborn and Heinz Nixdorf Institut (Germany).

Salvatore Orlando (Google h-index = 25) has published more than 110 papers in international journals, book chapters, and conference/workshop proceedings. He co-chaired conferences (EuroPVM/MPI, SASO), several conference tracks, and workshops. He has served on the program committees of international conferences, among which the premier conferences on data mining run by ACM, IEEE, SIAM (KDD, ICDM, SDM), and many others, such as Europar, ECML/PKDD, INFOSCALE, CCGRid, SASO.

Salvatore Orlando gave the tutorial "Mining query logs" at the international conferences European Conference on Information Retrieval (ECIR) 2009 and Web Intelligence (WI) 2009.

He was/has been the local coordinator of the Venetian research units involved in the PRIN project "GeoPKDD: Geographic Privacy-aware Knowledge Discovery and Delivery" (2004), and in the project "MOTUS: Mobility and Tourism in Urban Scenarios" (2009 - 2012) - Industria2015 - "Sustainable Mobility".

He participated in several Italian and European projects, among which we can mention the EU FP6 - STREP: "SAPIR: Search In Audio Visual Content Using Peer-to-peer IR" (2007-2009), and the EU FP6 NOE "CoreGRID: European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, GRID and Peer-to-Peer Technologies" (2005-2008).

2.1 Education and Work experience

- 1981-1985** Laurea degree in *Scienze dell'Informazione* (M.Sc. in Computer Science) - magna cum laude - University of Pisa.
- 1986** Consultant of a software house, Numerica Progetti s.r.l., for developing applications in the field of telecommunications, computer graphics, and multimedia data-bases.
- 1987** National military service.
- 1987-1990** Ph.D student of the Dept. of Computer Science of the University of Pisa. Received the Ph.D. in Computer Science with a dissertation entitled “*Exploiting Synchronous Models for Distributed Memory Multiprocessor Architectures*”, external referees: Prof. T. Hey (University of Southampton) e Dr. D. May. (Inmos Labs).
- 1991-1992** Fellow of the Hewlett Packard - Pisa Science Center. Joint project between the Dept. of Computer Science of the University of Pisa and the Hewlett Packard. Title of the research: *Development of hardware/software environments for massively parallel architectures.*
- 1993** Post-doctoral fellow of the Dept. of Computer Science of the Pisa University.
- 1994-2000** Assistant Professor - Università Ca' Foscari Venezia.
- 2000-Today** Associate Professor of Computer Science - Università Ca' Foscari Venezia.
Currently he is the Head of Studies in Computer Science (BSc and MSc) of Ca' Foscari University, is a member of the board (giunta) of “Dipartimento di Scienze Ambientali, Informatica e Statistica”, deputy head of the Department for academics affairs, and delegate of the Rector for national and international university orientation.

3 Research activities

The main topics covered by the research activities of S. Orlando are (i) Parallel and Distributed Systems and Languages, (ii) Data and Web Mining, and (iii) High Performance Information Retrieval and Web Search.

In the area of *Parallel Computing*, S. Orlando has focused his research activities on the so-called template/skeleton approach, which restricts the computational model seen by programmers and allows them to only use a small set of skeletons or high-level parallel language constructs. This restriction allows for applying both static and dynamic optimizations, which can be aggressive, but also specific for the target platforms by guaranteeing performance.

Nowadays this skeleton-based paradigm, in particular the composition of Map and Reduce skeletons, have become a very topical research subjects, due to the popular software frameworks by Google and Yahoo!, which can be used to compute large amount of data on a distributed/parallel cluster.

The research of S. Orlando in this field may be considered twofold. One is concerned with language aspects related to programming ease, hierarchical composition of constructs, structured features of languages. The other is concerned with the experimental studies of implementation templates for different architectures, and is based on “algorithmic forms” that have been proposed to solve important representative problems. All these features are summarized in the novel structured parallel language P³L, its compilers and programming environment [23, 22].

Notable results were obtained in the field of run-time support for data-parallel languages. In particular, S. Orlando devised novel run-time systems to compile unbalanced data parallel loops, which exploit partial data replication, dynamic scheduling, and prefetching techniques to minimize communication/synchronization overhead and evenly distribute the workload among processors [20, 19].

In addition, S. Orlando conceived innovative run-time layers to compile skeleton-based languages that allow programmers to mix control and data parallelism. More specifically, a library used to efficiently compile the coordination of data parallel tasks/modules [18, 17].

In the area of *Distributed Computing*, the issue of locating suitable resources in a dynamic Grid systems has been studied. A solution that takes in account the high variability of metadata characterizing Grid resources has been devised: it is based on a distributed discovery service implemented over a novel P2P network [71, 65, 66, 11].

In the field of data-intensive and scalable algorithms for *Data and Web Mining*, the research activities has regarded the study of scalable algorithm for the extraction of frequent patterns from transactional databases. Some devised algorithms – such as DCP [91] and DCI [86] – improved the state-of-art algorithms, often of orders of magnitude. In addition, they have been studied to be scalable, by optimizing the usage of main memory, and to dynamically adapt the algorithm behavior to the features of the datasets, such as size, density/sparsity, and data correlation. DCI is also the mining engine of ConQueSt [8], an integrated constraint-based system for mining and querying, which includes a specific query language and a graphical interface, and recently we studied its parallelization over a GPU [36].

DCI-closed [16] is an extremely efficient algorithm for the extraction of closed frequent patterns, which are lossless compressed representation of all frequent patterns. For this algorithm we also developed an out-of-core algorithm that scales very well when the datasets to mine are colossal [67], and an innovative multi-threaded parallel version [59].

Recently, still in the field of pattern extraction, we formalized the problem of discovering the Top-K patterns from binary datasets in presence of noise, as the minimization of a novel cost function. The problem of reducing the amount of patterns extracted by a mining algorithm is now considered particularly significant and urgent by the data mining community. According to the Minimum Description Length principle, the proposed cost function used by a greedy algorithm (PaNDa) favors the extraction of a succinct pattern sets that approximately describe the input data [40].

In the related field of *Data Warehousing*, a particular type of data has been analyzed: the sampled spatio-temporal trajectories of objects. Due to the current sensing and network technologies, we can nowadays accumulate huge trajectory datasets. Such datasets are produced as streams of trajectory observations and must be either quickly processed for real-time applications, like traffic control management, or carefully mined for complex, knowledge discovering tasks. To deal with such huge datasets, and extract some kind of actionable knowledge, data warehouses seems to be a natural choice. We devised a framework modeling a *trajectory data warehouse* (TDW) [6, 9, 2], aimed at storing significant aggregate information on trajectories of moving objects, which also offers visual OLAP operations for data analysis. The framework also includes a module that transforms and loads the TDW starting from streams of trajectory observations.

Finally, we have started an activity in social mining, studies correlation and predictive capability of several time series extracted from social networks [33, 31].

The activity of S. Orlando related to high performance *Information Retrieval and Web Search* is particularly important. The problem of realizing a high-throughput query engine for Web Search has been extensively studied, with a proposal of a software architectures that combine data and control parallelism [90]. The issue of improving the performance of Web Search has been faced also in the context of enhancing the index compression, since this implies a better exploitation of the memory hierarchies [76]. Recently, we studied how to selectively approximate query processing to increase the overall throughput of a distributed search engine [29, 32].

The query logs of web search engines (WSEs), and their power law features, have been extensively studied to devise a new strategy, namely SDC, for caching the query results of a WSE [15], and thus improving the overall query throughput. SDC exploits the knowledge about the queries submitted to the WSE in the past to enhance cache effectiveness. The caching technique also combines an adaptive prefetching strategy, which does not cause large additional loads on the WSE.

Always the study of query logs has allowed to devise innovative techniques to partitioning and partially replicating the inverted index of search engines [61].

Recently, still with the aim of improving the performance of WSEs, we have studied the issues in producing the snippets appearing in the result pages returned by a WSE. This generation process is very expensive. Thus we have proposed and evaluated a caching system based on the concept of supersnippet, that is the set of sentences in a document that are more likely to answer future queries, and that are dynamically built by exploiting the log of past queries and associated results [38].

Still a deep analysis of WSE query logs is at the basis of a novel technique, aimed at improving query suggestion. The idea is to understand if a combination of queries issued by a user are somehow related, and thus can be ascribed to the same Web-mediated task. We started this research activity with proposing a novel method to identify such task, with the aim of exploiting this knowledges to devise new recommendation mechanisms for queries on a task-based perspective [39, 34, 4].

We also investigated how to improve the performance of a content-based image retrieval (CBIR) system. We studied the query log of a real CBIR system available on the Web, where each query is a sample image, and the returned results are a set of images, whose visual contents are similar to the one of the query image. A novel similarity cache, which intercepts the stream of queries issued to the CBIR system, is proposed and evaluated. Unlike traditional caching, the proposed cache can manage not only exact hits, but also approximate ones, which are solved by similarity with respect to the result sets of past queries present in the cache [5, 48].

Finally, we have started a new activity on information extraction, related to named-entity linking [30, 28] and relation extraction [35].

4 Teaching Activities

Academic Year	MSc, BSc, PhD Courses at Università Ca' Foscari
1990-91	Set of lectures for the PhD course entitled "Parallel Architectures" – Università di Pisa
1995-96 and 96-97	Set of lectures for the courses: (a) Architettura degli Elaboratori (BSc), (b) Laboratorio di Architettura degli Elaboratori (BSc) – Ca' Foscari
1997-98, 98-99, and 99-2000	(a) Sistemi Distribuiti (BSc), (b) Laboratorio di Informatica: Architettura degli Elaboratori (BSc) – Ca' Foscari
2000-01	(a) Architettura degli Elaboratori (BSc), (b) Parallel Computing (PhD) – Ca' Foscari
2001-02	(a) Architettura degli Elaboratori A and B (BSc), (b) Calcolo Parallelo con Laboratorio (MSc) – Ca' Foscari
2002-03, 2003-04, 2004-05, 2005-06, 2006-07, 2007-08, and 2008-09	(a) Architettura degli Elaboratori, mod I and II (BSc), (b) Calcolo Parallelo con Laboratorio (MSc), (c) Data Mining (MSc) – Ca' Foscari
2009-10, 2010-11, 2011-12, 2012-13	(a) Architettura degli Elaboratori, mod II (BSc), (b) High Performance Computing (MSc), (c) Data and Web Mining (MSc) – Ca' Foscari
2013-14	(a) Architettura degli Elaboratori, mod II (BSc), (b) High Performance Computing (MSc) – Ca' Foscari

5 Scientific Supervision Activities

In the last years Prof. Orlando supervised many PhD students of the Doctoral Program in Computer Science at Ca' Foscari. In particular, Paolo Palmerini (XVI ciclo, PhD, 2003), Claudio Silvestri (XVIII cycle, PhD, 2005), Francesco Lelli (XIX cycle, PhD 2007), Matteo Mordacchini (XIX cycle, PhD 2007), Claudio Lucchese (XX cycle, PhD 2008), Fernando Josè Braz (XXI cycle, PhD 2009), Gabriele Tolomei (XXIII cycle, PhD 2011), Andrea Triossi (XXIV cycle, PhD 2012), Daniele Broccolo (XXV cycle, PhD 2014).

- Paolo Palmerini completed his studies with a dissertation entitled: *On performance of data mining: from algorithms to managing systems for data exploration.*
- Claudio Silvestri completed his studies with a dissertation entitled: *Distributed and Stream Data Mining Algorithms for Frequent Pattern Discovery.*
- Francesco Lelli completed his studies with a dissertation entitled: *Bringing Instruments to a Service-Oriented Interactive Grid.*
- Matteo Mordacchini completed his studies with a dissertation entitled: *Grid and Peer-to-Peer Resource Discovery Systems.*
- Claudio Lucchese completed his studies with a dissertation entitled: *High Performance Closed Frequent Itemsets Mining inspired by Emerging Computer Architectures.*
- Fernando Josè Braz completed his studies with a dissertation entitled: *Warehousing and Mining Aggregate Measures Over Trajectories of Moving Objects.*
- Gabriele Tolomei completed his studies with a dissertation entitled: *Enhancing Web Search User Experience: from Document Retrieval to Task Recommendation*
- Andrea Triossi completed his studies with a dissertation entitled: *Hardware Execution of Constraint Handling Rules.*
- Daniele Broccolo completed his studies with a dissertation entitled: *Query Log Based Techniques to Improve the Performance of a Web Search Engine.*

Currently S. Orlando is the supervisor of the following students of the Dottorato in Informatica at Università Ca' Foscari Venezia: Francesco Lettich (XXVII cycle).

Prof. Orlando was the supervisor of a fellows within the EU project Eureka HPCC-SEA (1998).

He also coordinated the research activities of two fellows within the national projects PRIN 2004 "GeoPKDD - Estrazione e presentazione di conoscenza geografica con tecniche di data mining che salvaguardano la privacy".

Finally, he coordinated the research activities of two fellows within the project: *Servizi avanzati per l'info mobilità e il turismo* (2009), funded by Fondo Sociale Europeo (FSE).

Prof. Orlando was the supervisor of numerous BSc and MSc theses of the Computer Science programs at Università Ca' Foscari Venezia, and of some other theses at Università di Pisa.

6 University Services

In the years 1996, 1997, and 1998 Prof. Orlando was a member of the governing board of CICSD (Centro Interdipartimentale di Calcolo scientifico e didattico) at Ca' Foscari.

Since 2000 he has been a member of the board of the doctoral program in Computer Science of Ca' Foscari.

Prof. Orlando was nominated Deputy Head of Dipartimento di Informatica at Ca' Foscari for the years 2008-10.

Since 2000 he is a member of the Teaching Committee of the BSc and MSc program in Computer Science at Ca' Foscari.

Since 2010, S. Orlando he is the *Head* of the Teaching Committee of the BSc and MSc programs in Computer Science at Ca' Foscari.

Since 2010, Prof. Orlando is a member of the governing board of CINI (National Inter-university Consortium for Informatics).

Since 2013, Prof. Orlando is delegate of the Rector for national and international university orientation.

Participation in Committees for University public competitions.

1999: Member of the competitive examination committee for the recruitment of 4 assistant professors at Università di Pisa.

2000: Member of the competitive examination committee for the admission to the PhD program in Computer Science at Ca' Foscari.

2000 and 2001: President of the examination committee for the recruitment of 3 computer technicians at Ca' Foscari.

2004: Member of the competitive examination committee for the recruitment of an assistant professor at Università di Bologna.

2004: Member of the committee for the competitive recruitment examination for an associate professor at Università di Cagliari.

2005: Member of the committee for the competitive recruitment examination for an assistant professor at Università di Bologna.

2008: Member of the final examination committee of Dottorato di Ricerca in Informatica - Università di Bologna.

2008: Member of the committee for the competitive examination for the admission to the PhD program in Computer Science at Ca' Foscari.

2008: Member of the final examination committee of Dottorato di Ricerca in Ingegneria dei Sistemi e Informatica - Università della Calabria.

2008: Member of the final examination committee of Dottorato di Ricerca in Informatica - Università di Pisa.

2009: Member of the final examination committee of Dottorato di Ricerca in Ingegneria dei Sistemi e Informatica - Università della Calabria.

2009: Member of the final examination committee of Dottorato di Ricerca in InInformatica - Università di Torino.

2010: President of the examination committee for accessing the Albo Professionale - Ingegneri dell'Informazione.

2011: Member of the final examination committee of Dottorato di Ricerca in InInformatica - Università di Bari.

2011: Member of the committee for the competitive examination for the admission to the PhD program in Computer Science at Ca' Foscari.

2012: Member of the final examination committee of Dottorato di Ricerca in in INFORMATICA and MATEMATICA E CALCOLO SCIENTIFICO at Università di Cagliari.

7 Coordination of / Participation in research projects

Coordination of research units

- Coordinator of the Ca' Foscari research units of the national project: *TETRis - Servizi innovativi Open Source su TETRA - PROGETTO PON01_00451*" (2011-13)
- Coordinator of the Ca' Foscari research units of the national project: *MOTUS - Mobility and Tourism in Urban Scenarios* (2009-1012), funded by Ministero dello Sviluppo Economico - Program "Industria 2015 - Mobilità Sostenibile".
- Coordinator of the project: *Servizi avanzati per l'info mobilità e il turismo* (2009), funded by Fondo Sociale Europeo (FSE).
- Coordinator of the Ca' Foscari research units of the national project PRIN2004 "GeoPKDD: GeographicPrivacy-aware Knowledge Discovery and Delivery" (2005-2006).
- Coordinator of a biennial research contract (2001-2003) between Tektronics S.p.A. and Ca' Foscari University, whose focus was the Data Mining analysis of telephonic traffic data.
- Coordinator of a research unit of the EU project *Eureka HPCC-SEA* at Consorzio Pisa Ricerca, aiming at studying and realizing the full integration of High Performance Fortran within a skeleton-based parallel language (SKIE).

Participation in research units

- Joint project between Hewlett-Packard Labs and Dipartimento di Informatica - Università di Pisa, entitled *Pisa Parallel Processing Project (P⁴)* on the topics: "Software Environments for Highly Parallel Systems".
- National project ASI-PQE2000 "Sviluppo di applicazioni di osservazione della terra mediante sistemi e tool di calcolo ad alte prestazioni"
- National project MURST (2002-2003): "Architetture Software per infrastrutture di rete ad accesso eterogeneo (SAHARA)".
- National strategic project: Grid Computing: tecnologie abilitanti e applicazioni per eScience (2002-2004).
- National strategic project: "ECD: Technologies and Services for Enhanced Contents Delivery" (2002-2004)
- National FIRB Project: "Grid.it: Piattaforme abilitanti per griglie computazionali a elevate prestazioni orientate a organizzazioni virtuali scalabili" (2003-2005).
- European Integrated Project (IP) - IST FP6: "NextGRID: The Next Generation Grid" (2004-2007).
- European Research Network - IST FP6: "CoreGRID: Foundations, Software Infrastructures and Applications for large scale distributed, Grid and Peer-to-Peer Technologies" (2004-2008).
- IST EU FP6 Project: "SAPIR - Search In Audio Visual Content Using Peer-to-peer Information Retrieval" (2007-2009).
- EU Coordination Action "MODAP (Mobility, Data Mining, and Privacy)" (2009-2012)
- EU ICT Policy Support programme: "ASSETS: Advanced Search Service and Enhanced Technological Solutions for the European Digital Library" (2010-12)

8 Committees of Scientific Events

Chair of Committees

- 2003** Co-Chair of the Program Committee of EuroPVM/MPI '03 – Venezia, Italia – September 29 - Ottobre 2, 2003.
- 2010** Co-Chair of the topic "Parallel and Distributed Databases, Data Mining, and Knowledge Discovery" of Europar 2004.
- 2005** Program Co-Chair of the 8th High Performance and Distributed Mining Workshop (HPDM'05) - SIAM Conf. on Data Mining - 2005.
- 2008** Workshop Chair of the 2nd Int. Conf. on Scalable Information Systems (INFOSCALE'08).
- 2008** Special Sessions & Workshops Co-Chair of the 2008 High Performance Computing and Simulation Conference (HPCS 2008).
- 2008** Local Chair of the 2nd IEEE Int.l Conf. on Self-Adaptive and Self-Organizing Systems (SASO 2008).
- 2009** Awards Co-Chair of the 2008 High Performance Computing & Simulation Conference (HPCS 2009)
- 2010** Program Co-Chair of SDIA 2010 - Scalable Data Intensive Applications - Special session of PDP 2010.
- 2010** Program Co-Chair of the topic "Parallel and Distributed Data Management" of Europar 2010.
- 2010** Awards Co-Chair of the High Performance Computing & Simulation Conference (HPCS 2010)
- 2011** Program Global Chair of the topic "Parallel and Distributed Data Management" of Europar 2011.
- 2011** Awards Co-Chair of the High Performance Computing & Simulation Conference (HPCS 2011)
- 2012** Awards Co-Chair of the High Performance Computing & Simulation Conference (HPCS 2012)
- 2013** Awards Co-Chair of the High Performance Computing & Simulation Conference (HPCS 2013)

Member of Program Committees

- 1998** Second European Conference on Parallel and Distributed Systems (Euro-PDS'98) – Vienna, Austria – July 1-3, 1998
- 2001** International Conference ParCo2001 – Parallel Computing 2001 – Napoli, Italia – September 4-7, 2001.
- 2003** 6th HPDM (High-performance and distributed datamining) Workshop, in conjunction with 3rd SIAM Conference on Data Mining, May 2003, San Francisco, CA.
- 2004** 7th HPDM (High-performance and distributed datamining) Workshop, in conjunction with 4th SIAM Conference on Data Mining, April, 2004, Orlando, Florida.
- 2004** International Conference on Computational Science 2004 (ICCS 2004) - Krakov, Poland, June 2004.
- 2004** Europar 2004 - Pisa, Italy, August/September 2004.
- 2004** 1st Int. Workshop on Grid and P2P Computing Impacts on Large Scale Heterogeneous Distributed Database Systems (GLOBE'04) - Zaragoza, Spain.

- 2004** EuroPVM/MPI '04 – Budapest, Hungary, September 19-22, 2004.
- 2005** 2005 ACM Symposium on Applied Computing (SAC 2005) - Special Track on Data Mining.
- 2005** International Conference on Computational Science 2005 (ICCS 2005).
- 2005** European Grid Conference 2005 (EGC 2005).
- 2005** IEEE International Conference on Information Technology: Coding and Computing (ITCC 2005) - Track on Next-Generation Web and Grid Systems.
- 2005** 2nd Int. Workshop on Grid and P2P Computing Impacts on Large Scale Heterogeneous Distributed Database Systems (GLOBE'05).
- 2005** Euro PVM/MPI 2005 Conference.
- 2005** SIGIR 2005 – Workshop on Heterogeneous and Distributed Information Retrieval.
- 2005** International Conference on High Performance Computing and Communications (HPCC 05).
- 2006** ACM Symposium on Applied Computing (SAC 2006) - Special Track on Data Mining.
- 2006** IASTED Conf. on Parallel and Distributed Computing and Networks - PDCN 2006.
- 2006** SIAM Conference on Data Mining (SDM'06).
- 2006** Int. Conf. on Computational Science and its Applications (ICCSA 2006).
- 2006** 9th High Performance and Distributed Mining Workshop (HPDM'06).
- 2006** Int. Conference on Computational Science 2006 (ICCS 2006).
- 2006** Euro PVM/MPI 2006 Conference.
- 2006** Int. Conference on High Performance Computing and Communications (HPCC 06).
- 2006** 2nd Int. Workshop on Grid and P2P Computing Impacts on Large Scale Heterogeneous Distributed Database Systems (GLOBE'06).
- 2006** ECML PKDD 2006 Workshop on Parallel Data Mining.
- 2006** Workshop on Inform. Retrieval in Peer-to-Peer Networks, ACM 15th CIKM 2006.
- 2007** IASTED Int. Conf. on Parallel and Distributed Computing and Networks - PDCN 2007.
- 2007** ACM Symposium on Applied Computing (SAC 2007) - Special Track on Data Mining.
- 2007** SIAM Conference on Data Mining (SDM'07).
- 2007** 11th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2007).
- 2007** IEEE 21st Int.l Conference on Advanced Information Networking and Applications (AINA-07).
- 2007** Int.l Conference on Computational Science 2007 (ICCS 2007).
- 2007** Int. Conf. on Computational Science and its Applications (ICCSA 2007).
- 2007** Int.l Workshop on High Performance Data Mining and Applications (HPMDA'07).
- 2007** 4th Int. Workshop on Grid and P2P Computing Impacts on Large Scale Heterogeneous Distributed Database Systems (GLOBE'07).
- 2007** Euro PVM/MPI 2007 Conference.
- 2008** IASTED Int. Conf. on Parallel and Distributed Computing and Networks - PDCN 2008.
- 2008** 2008 ACM Symposium on Applied Computing (SAC 2008) - Tracks on "Data Mining" and "Engineering Large-Scale Distributed Systems".
- 2008** 2008 SIAM Conference on Data Mining (SDM'08).

- 2008** 12th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2008).
- 2008** 8th IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2008).
- 2008** Euro PVM/MPI 2008 Conference.
- 2008** 2nd IEEE Int.l Conf. on Self-Adaptive and Self-Organizing Systems (SASO 2008).
- 2008** IASTED Int. Conf. on Parallel and Distributed Computing and Networks - PDCN 2008.
- 2008** European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2008).
- 2008** Int.l Conf. on Computational Science 2008 (ICCS 2008).
- 2008** 1st Int.l Conf. on Data Management in Grid and P2P Systems (Globe'2008).
- 2008** 10th IEEE Int.al Conf. on High Performance Computing and Communications (HPCC-08).
- 2008** 16th Italian Symposium on Advanced Database Systems (SEBD 2008).
- 2008** 1st Int.l Conf. on Data Management in Grid and P2P Systems (Globe'2008).
- 2008** 10th International Workshop on High Performance Data Mining" at IEEE ICDM 2008.
- 2009** IASTED Int. Conf. on Parallel and Distributed Computing and Networks - PDCN 2009.
- 2009** 13th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2009).
- 2009** 15th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (KDD 2009).
- 2009** 11th IEEE Int.l Conf. on High Performance Computing and Communication (HPCC 2009).
- 2009** 2nd Int.l Conf. on Data Management in Grid and P2P Systems (Globe'2009).
- 2009** IEEE Int.l Conf. on Data Mining (ICDM 2009).
- 2009** ACM Symposium on Applied Computing (SAC 2009) - Special Track on Data Mining.
- 2010** 14th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2010).
- 2010** 4th IEEE Int.l Conf. on Self-Adaptive and Self-Organizing Systems (SASO 2010).
- 2010** 10th IEEE Int.l Conf. on Data Mining (ICDM 2010).
- 2011** PDP 2011 - The 19th Euromicro Int.l Conf. on Parallel, Distributed and Network-Based Computing
- 2011** 15th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2011).
- 2011** ACM Symposium on Applied Computing (SAC 2011) - Special Track on Data Mining.
- 2011** 2nd Italian Information Retrieval (IIR 2011) Workshop.
- 2011** ACM Symposium on Applied Computing (SAC 2011) - Special Track on Data Mining.
- 2011** 11th IEEE Int.l Conf. on Data Mining (ICDM 2011).
- 2012** European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2012).

- 2012** 21st ACM International Conference on Information and Knowledge Management (CIKM 2012).
- 2012** 12th IEEE Int.l Conf. on Data Mining (ICDM 2012).
- 2013** ACM Symposium on Applied Computing (SAC 2013) - Special Track on Data Mining.
- 2013** Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2013).
- 2013** 21st Italian Symposium on Advanced Database Systems (SEBD 2013)
- 2013** 22nd ACM International Conference on Information and Knowledge Management (CIKM 2013)
- 2013** IEEE Int.l Conf. on Data Mining (ICDM 2013).
- 2013** IEEE Int.l Conf. on Big Data (IEEE BigData 2013).

9 Publications

Journal Articles

- [1] C. Lucchese, S. Orlando, and R. Perego. A unifying framework for mining approximate top-k binary patterns. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 26(12): 2900-2913 (2014).
- [2] L. Leonardi, S. Orlando, A. Raffaetà, A. Roncato, C. Silvestri, G. Andrienko, and N. Andrienko. A general framework for trajectory data warehousing and visual OLAP. *Geoinformatica*, Springer, 18(2): 273-312 (2014), doi: 10.1007/s10707-013-0181-3.
- [3] F. Giummolè, S. Orlando, and G. Tolomei. A Study on Microblog and Search Engine User Behaviors: How Twitter Trending Topics Help Predict Google Hot Queries. *HUMAN, ASE Science Journal*, Vol. 2, No. 3, pp. 195-209, 2013, ISBN: 978-1-62561-001-0 (Special issue - Best papers of the IEEE/ASE SocialCom2013 Conference).
- [4] C. Lucchese, S. Orlando, R. Perego, F. Silvestri, G. Tolomei. Discovering Tasks from Search Engine Query Logs. *ACM TOIS - Transactions on Information Systems*, 2013, Vol. 31, No. 3, pp. 1-43, doi: 10.1145/2493175.2493179, ISSN: 1046-8188. Selected by ACM Computing Reviews as a notable article in computing in 2013.
- [5] F. Falchi, C. Lucchese, S. Orlando, R. Perego, F. Rabitti. Similarity Caching in Large-Scale Image Retrieval, *Information Processing and Management*, Elsevier, vol. 48, pp. 803-818, 2012, ISSN: 0306-4573.
- [6] A. Raffaetà, L. Leonardi, G. Marketos, G. Andrienko, N. Andrienko, E. Frentzos, N. Giatrakos, S. Orlando, N. Pelekis, A. Roncato, C. Silvestri. Visual Mobility Analysis using T-Warehouse, *International Journal of Data Warehousing and Mining*, IGI Global, ISSN: 1548-3924, Volume 7, Number 1, Pages 1-27, 2011.
- [7] C. Lucchese, C. Mastroianni, S. Orlando, D. Talia. Mining@home: Towards a Public Resource Computing Framework for Distributed Data Mining, *Concurrency & Computation: Practice and Experience*, Wiley, ISSN: 1532-0634, Volume 22, Issue 5, 2010, Pages 658-682.
- [8] F. Bonchi, F. Giannotti, C. Lucchese, S. Orlando, R. Perego, R. Trasarti. A Constraint-based Querying System for Exploratory Pattern Discovery, *Information Systems*, Elsevier, Volume 34, Issue 1, March 2009, pages: 3-27.
- [9] S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, C. Silvestri. Trajectory Data Warehouses: Design and Implementation Issues, *Journal of Computing Science and Engineering (JCSE)*. KIISE Publ., Vol. 1, No. 2, pp. 240-261, Dicembre 2007 (Special issue - Best papers of the 9th Int.l Conf. on Data Warehousing and Knowledge Discovery (DaWaK '07))
- [10] F. Lelli, E. Frizziero, M. Gulmini, G. Maron, S. Orlando, A. Petrucci, S. Squizzato. The many faces of the integration of instruments and the grid, *Int. Journal of Web and Grid Services (IJWGS)*, ISSN: 1741-1106, 3(3). pp. 239-266, 2007, Inderscience Publishers.
- [11] M. Marzolla, S. Orlando, M. Mordacchini. P2P Systems for Resource Discovery in a Dynamic Grid, *Parallel Computing Journal*, 33, pp. 339-358, 2007, Elsevier Science.
- [12] C. Silvestri, and S. Orlando. Approximate Mining of Frequent Patterns on Streams, *Int. Journal of Intelligent Data Analysis*, IOS Press, Vol. 11, No. 1, pp.49-73, 2007, ISSN 1088-467X.
- [13] F. Lelli, S. Orlando, G. Maron, S. Pinter. Bringing instruments into a Grid: An empiric approach. *WSEAS Transactions in Computers*, vol. 6, 2007, p. 153-159, ISSN: 1109-2750.

- [14] F. Silvestri, D. Puppini, D. Laforenza. Toward a Search Architecture for Software Components, *Concurrency & Computation: Practice & Experience*, ISSN: 1532-0634, Vol. 18, pp. 1317-1331, Wiley, 2006.
- [15] T. Fagni, R. Perego, F. Silvestri, and S. Orlando. Boosting the Performance of Web Search Engines: Caching and Prefetching Query Results by Exploiting Historical Usage Data. *ACM TOIS - Transactions on Information Systems*, vol. 24, pp. 51-78.
- [16] C. Lucchese, S. Orlando, and R. Perego. Fast and Memory Efficient Mining of Frequent Closed Itemsets, *IEEE Transactions on Knowledge and Data Engineering*, January 2006 (Vol. 18, No. 1), pp. 21-36.
- [17] S. Orlando, P. Palmerini, and R. Perego. Mixed Data and Task Parallelism with HPF and PVM. *CLUSTER COMPUTING: The Journal of Networks, Software and Applications*, Baltzer Science Publishers, Vol.3, n.3, 2000, pp. 201-213.
- [18] S. Orlando and R. Perego. *COLT_{HPF}*: a Run-Time Support for the High-Level Coordination of HPF Tasks. *Concurrency: Practice and Experience*, ISSN: 1532-0634, Wiley, Vol. 11(8), pp. 407-434, 1999.
- [19] S. Orlando and R. Perego. SUPPLE: an Efficient Run-Time Support for Non-Uniform Parallel Loops. *Journal of Systems Architecture*, Special Issue on New Trends in Programming and Execution Models for Parallel Architectures, Heterogeneously Distributed Systems and Mobile Computing, Elsevier Science publisher, Vol. 45, pp. 1323-1343, 1999.
- [20] S. Orlando and R. Perego. A Comparison of Implementation Strategies for Non-Uniform Data Parallel Computations. *Journal of Parallel and Distributed Computing*, Academic Press, vol. 52, n. 2, August 1998, pp. 132-149.
- [21] S. Orlando and R. Perego. Exploiting Partial Replication in Unbalanced Parallel Loops Scheduling on Multicomputers. *Microprocessing and Microprogramming - The EUROMICRO Journal*, Elsevier Science publisher, 41, 1996.
- [22] B. Bacci, M. Danelutto, S. Orlando, S. Pelagatti, and M. Vanneschi. *P³L*: a Structured High-level Parallel Language and its Structured Support. *Concurrency: Practice and Experience*, Wiley, Vol. 7(3), pp. 225-255, 1995.
- [23] M. Danelutto, R. Di Meglio, S. Orlando, S. Pelagatti, and M. Vanneschi. A Methodology for the Development and the Support of Massively Parallel Programs. *Future Generation Computer Systems J.*, 8:205-220, 1992. Also published as selected paper in *Programming Languages for Parallel Processing*, IEEE Computer Society Press, D. B. Skillicorn and D. Talia Ed., 319-334, 1994.

Refereed International Conference and Workshop Papers

- [24] Salvatore Trani, Diego Ceccarelli, Claudio Lucchese, Salvatore Orlando, Raffaele Perego. Manual Annotation of Semi-Structured Documents for Entity-Linking. *CIKM 2014: 2075-2077* (Demo paper).
- [25] Diego Ceccarelli, Claudio Lucchese, Salvatore Orlando, Raffaele Perego, Salvatore Trani. Dexter 2.0 - an Open Source Tool for Semantically Enriching Data. *International Semantic Web Conference (Posters & Demos) 2014: 417-420*.
- [26] Stefano Calzavara, Gabriele Tolomei, Michele Bugliesi, Salvatore Orlando. Quite a mess in my cookie jar!: leveraging machine learning to protect web authentication. *WWW Conference 2014: 189-200*.

- [27] Claudio Silvestri, Francesco Lettich, Salvatore Orlando, and Christian S. Jensen. GPU-based Computing of Repeated Range Queries over Moving Objects. 22nd Euromicro Int.l Conf. on Parallel, Distributed and Network-Based Processing (PDP 2014), pages 640-647.
- [28] D. Ceccarelli, C. Lucchese, S. Orlando, R. Perego, and S. Trani. Learning relatedness measures for entity linking. Proc. of. ACM International Conference on Information and Knowledge Management, CIKM 2013, pages 139-148.
- [29] D. Broccolo, C. Macdonald, S. Orlando, I. Ounis, R. Perego, F. Silvestri, and N. Tonellotto. Load-Sensitive Selective Pruning for Distributed Search. Proc. of. ACM International Conference on Information and Knowledge Management, CIKM 2013, pages 379-388.
- [30] D. Ceccarelli, C. Lucchese, S. Orlando, R. Perego, and S. Trani. Dexter: an Open Source Framework for Entity Linking. Proc. of the CIKM 2013 Workshop on Exploiting Semantic Annotation (ESAIR), 2013, pages 17-20.
- [31] G. Tolomei, S. Orlando, D. Ceccarelli, C. Lucchese. Twitter Anticipates Bursts of Requests for Wikipedia Articles. Proc. of the CIKM 2013 Workshop on Data-driven User Behavioral Modelling and Mining from Social Media (DUBMOD@CIKM), 2013, pages 5-8.
- [32] D. Broccolo, C. Macdonald, S. Orlando, I. Ounis, R. Perego, F. Silvestri, and N. Tonellotto. Query Processing in Highly-Loaded Search Engines". Proc. of SPIRE 2013, LNCS, Springer, ISBN 978-3-319-02431-8, pages 49-55.
- [33] F. Giummolè, S. Orlando, and G. Tolomei. Trending Topics on Twitter Improve the Prediction of Google Hot Queries. Proc. of 2013 ASE/IEEE Int.l Conference on Social Computing (Socialcom 2013) - selected among the best conference papers for a journal version, pp. 39-44, 2013.
- [34] C. Lucchese, S. Orlando, R. Perego, F. Silvestri, and G. Tolomei. Modeling and Predicting the Task-by-Task Behavior of Search Engine Users. Proc. of 10th Int.l Conference on Open Research Areas in Information Retrieval, OAIR 2013, RIAO series. pages 77-84.
- [35] S. Orlando, F. Pizzolon, and G. Tolomei. SEED: A Framework for Extracting Social Events from Press News. Proc. of the 2nd Int.l Workshop on Web of Linked Entities (WOLE 2013), WWW2013 Workshop. WWW (Companion Volume), pages 1285-1294.
- [36] C. Silvestri, S. Orlando. Exploiting GPUs in Frequent Itemset Mining. 20th Euromicro Int.l Conf. on Parallel, Distributed and Network-Based Processing (PDP 2012). Garching, Germany, 15-17 Feb, 2012, p. 416-425, IEEE Computer Society, ISBN: 9780769546339.
- [37] A. Triossi, S. Orlando, A. Raffaetà, T. Fruehwirth. Compiling CHR to parallel hardware. Proc. of the 14th Symposium on Principles and Practice of Declarative Programming. p. 173-184, ACM, ISBN: 9781450315227, 2012.
- [38] D. Ceccarelli, C. Lucchese, S. Orlando, R. Perego, F. Silvestri. Caching query-biased snippets for efficient retrieval Proc. of 14th Int.l Conf. on Extending Database Technologies (EDBT 2011), Upsala. Sweden, March 22-24, 2011.
- [39] C. Lucchese, S. Orlando, R. Perego, F. Silvestri, G. Tolomei. Detecting Task-based Query Sessions using Collaborative Knowledge In Proc. of 4th ACM Int.l Conference on Web Search and Data Mining (WSDM 2011), pages 277-286, 2011.
- [40] C. Lucchese, S. Orlando, R. Perego. Mining Top-K Patterns from Binary Datasets in presence of Noise. In Proc. of the 10th SIAM Int.l Conf. on Data Mining (SDM'10), April 2010.
- [41] G. Tolomei, S. Orlando, F. Silvestri. Towards a task-based search and recommender systems ICDE Workshops, 2010, pages 333-336
- [42] C. Lucchese, S. Orlando, R. Perego. Generative Pattern Model for Mining Binary Datasets. In Proc. of the 2010 ACM Symposium on Applied Computing (SAC '10), Track on Data Mining.

- [43] L. Leonardi, G. Marketos, E. Frentzos, N. Giatrakos, S. Orlando, N. Pelekis, A. Raffaetà, A. Roncato, C. Silvestri. Y. Theodoridis. T-Warehouse: Visual OLAP Analysis on Trajectory Data. Proc. of ACM ICDE 2010, Demo paper.
- [44] C. Lucchese, S. Orlando, R. Perego, F. Silvestri, G. Tolomei. Detecting Task-Based Query Sessions Using Collaborative Knowledge Web Intelligence/IAT Workshops 2010, pages 128-131.
- [45] A. Triossi, S. Orlando, A. Raffaetà, F. Raiser, T. Frühwirth. Constraint-based hardware synthesis. Proc. of 24th Workshop on (Constraint) Logic Programming, 2010.
- [46] M. Mordacchini, P. Dazzi, G. Tolomei, R. Baraglia, S. Orlando, F. Silvestri. Challenges in Designing an Interest-based Distributed Aggregation of Users in P2P Systems. Int.l Conf. on Ultra Modern Telecommunications 2009.
- [47] S. Orlando, F. Silvestri. Mining Query Logs. ECIR 2009, LNCS, Volume 5478/2009, pages 814-817, Springer.
- [48] F. Falchi, C. Lucchese, S. Orlando, R. Perego, F. Rabitti. Caching Content-based Queries for Robust and Efficient Image Retrieval. Proc. of 12th Int.l Conf. on Extending Database Technologies (EDBT 2009), pages 780-790, 2009, ACM.
- [49] L. Leonardi, S. Orlando, A. Raffaetà, A. Roncato, C. Silvestri. Frequent Spatio-Temporal Patterns in Trajectory Data Warehouses. In Proc. of the 2009 ACM Symposium on Applied Computing (SAC '09), Track on Data Mining, pages 1433-1440, 2009.
- [50] F. Falchi, C. Lucchese, S. Orlando, R. Perego, F. Rabitti. A metric cache for similarity search. Proc. of the 2008 ACM CIKM workshop on Large-Scale distributed systems for information retrieval (LSDS-IR) 2008, pages 43-50, ACM.
- [51] F.Silvestri, R. Baraglia, C. Lucchese, S. Orlando, R. Perego. (Query) History Teaches Everything, Including the Future". Paper invited to LA-Web 2008, pp. 12-22, IEEE CS press, October 28th-30th, Vila Velha, Espirito Santo, Brazil, 2008.
- [52] G. DaCosta, S. Orlando, M. D. Dikaiakos. Multi-set DHT for interval queries on dynamic data. 1st Int.l Conf. on Data Management in Grid and P2P Systems (Globe'2008), LNCS, Springer.
- [53] C. Gennaro, M. Mordacchini, S. Orlando, F. Rabitti Processing Complex Similarity Queries in Peer-to-Peer Networks. Proc. of The 23rd Annual ACM Symposium on Applied Computing (SAC '08), 2008, Track on "Engineering Large-Scale Distributed Systems"
- [54] F. Braz, S. Orlando, R. ORSINI, A. Raffaetà, A. Roncato, C. Silvestri. Designing a Spatio-Temporal Trajectory Data Warehouse. In Proc. of XXII Brazilian Symposium on Databases, 2007.
- [55] C. Gennaro, M. Mordacchini, S. Orlando, F. Rabitti MRoute: A Peer-to-Peer Routing Index for Similarity Search in Metric Spaces. Proc. of the 5th Int. Workshop on Databases, Information Systems and Peer-to-Peer Computing (DBISP2P 2007) in conjunction with VLDB 2007. 2007.
- [56] F. Lelli, G. Maron, S. Orlando Towards Response Time Estimation in Web Services. IEEE International Conference on Web Services (ICWS 2007), pp. 1138-1139, IEEE, ISBN: 9780769529240, 2007.
- [57] F. Lelli, G. Maron, S. Orlando. Client Side Estimation of a Remote Service Execution. 15th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2007), October 24-26, 2007, Istanbul, Turkey.
- [58] G. Da Costa, M. D. Dikaiakos, S. Orlando. Nine months in the life of EGEE: a look from the South. 15th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2007), October 24-26, 2007, Istanbul, Turkey. pages 281-287.

- [59] C. Lucchese, S. Orlando, R. Perego. Parallel Mining of Frequent Closed Patterns: Harnessing Modern Computer Architectures. Proc. of the IEEE Int.l Conf. on Data Mining 2007 (ICDM 07). pp 242–251, 2007.
- [60] S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, C. Silvestri. Spatio-Temporal Aggregations in Trajectory Data Warehouses. 9th Int.l Conf. on Data Warehousing and Knowledge Discovery (DaWaK '07), LNCS 4654, Springer, 2007, pp. 66–77.
- [61] C. Lucchese, S.Orlando, R. Perego, F.Silvestri. Mining Query Logs to Optimize Index Partitioning in Parallel Web Search Engines. Proc. of 2nd Int. Conf. on Scalable Information Systems (INFOSCALE'07), 2007.
- [62] F. Braz, S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, C. Silvestri. Approximate Aggregations in Trajectory Data Warehouses. Proc. ICDE Workshop on Spatio-Temporal Data Mining 2007 (STDM07), IEEE Press.
- [63] R. Baraglia, C. Lucchese, S. Orlando, M. Serranò, F. Silvestri. A Privacy Preserving Web Recommender System. Proc. of The 21th Annual ACM Symposium on Applied Computing (SAC '06), 2006, pp. 559-563.
- [64] F. Bonchi, F. Giannotti, C. Lucchese, S. Orlando, R. Perego, R. Trasarti. CONQUEST: a Constraint-based Querying System for Exploratory Pattern Discovery. Proc. of ACM ICDE 2006, Demo paper, pp. 159.
- [65] M. Marzolla, M. Mordacchini, S. Orlando. A P2P Resource Discovery System Based on a Forest of Trees 17th Int. Workshop on Database and Expert Systems Applications (DEXA'06) - GLOBE06, pp. 261-265, IEEE press.
- [66] M. Marzolla, M. Mordacchini, S. Orlando. Tree Vector Indexes: Efficient Range Queries for Dynamic Content on Peer-to-Peer Networks. Proc. of the 14th Euromicro Conference on Parallel, Distributed and Network based Processing (PDP 2006), IEEE Press, pp. 457–464, 2006.
- [67] C. Lucchese, S. Orlando, R. Perego. Mining frequent closed itemsets out-of-core. Proc. of the SIAM Conference on Data Mining (SDM '06), pp. 417-427, 2006.
- [68] E. Frizziero, M. Gulmini, F. Lelli, G. Maron, A. Oh, S. Orlando, A. Petrucci, S. Squizzato, S. Traldi. Instrument Element: A New Grid component that Enables the Control of Remote Instrumentation. 6th IEEE Int. Symposium on Cluster Computing and the Grid Workshops (CCGRIDW'06) - Workshop on Grid Testbeds, 2006.
- [69] G. Maron, F. Lelli, S. Orlando. Improving the Performance of XML Based Technologies by Caching and Reusing Information. IEEE Int.l Conference on Web Services (ICWS 2006), pp. 689–700, 2006.
- [70] C. Lucchese, S. Orlando, R. Perego. Distributed Mining of Frequent Closed Itemsets: Some Preliminary Results. Proc. of the 8th Int. Workshop on High Performance and Distributed Mining (HPDM '05), in conjunction with SIAM Data Mining Conf. (SDM'05). April, 2005.
- [71] M. Marzolla, M. Mordacchini, S. Orlando. Resource Discovery in a Dynamic Grid Environment. 16th Int. Workshop on Database and Expert Systems Applications (DEXA'05) - GLOBE05 - pp. 356-360, IEEE press.
- [72] C. Silvestri, S. Orlando. Approximate Mining of Frequent Patterns on Streams. 2nd Int. ECML-PKDD Workshop on Knowledge Discovery from Data Streams, 2005.
- [73] C. Silvestri, S. Orlando. Distributed Approximate Mining of Frequent Patterns. Proc. of the 2005 ACM Symposium on Applied Computing, SAC 2005, special track on Data Mining.
- [74] C. Lucchese, S. Orlando, R. Perego. DCI_{Closed} : A Fast and Memory Efficient Algorithm to Mine Frequent Closed Itemsets. Proceedings of the Workshop on Frequent Itemset Mining Implementations (FIMI'04), in conjunction with ICDM'04, November 1-4, 2004. Brighton, UK.

- [75] D. Puppini, F. Silvestri, D. Laforenza, S. Orlando. A Search Engine for Software Components. Proc. of the 2004 IEEE/WIC/ACM Int. Conf. on Web Intelligence (WI 2004). Beijing, China. September 20-24, 2004. pp. 495-498.
- [76] F. Silvestri, S. Orlando, R. Perego. Assigning Identifiers to Documents to Enhance the Clustering Property of Fulltext Indexes. Proceedings of the 27th Annual Int.l ACM SIGIR Conf. on Research and Development in Information Retrieval, pp. 305–312, 2004.
- [77] C. Silvestri, S. Orlando. Distributed association mining: an approximate method. Proceedings of the 7th Int. Workshop on High Performance Distributed Mining (HPDM'04), in conjunction with 4th International SIAM Conference on Data Mining, Orlando, Florida.
- [78] F. Silvestri, S. Orlando, R. Perego. WINGS: a Parallel Indexer for Web Contents. Proc. of International Conference on Computational Science 2004 (ICCS'04) - Krakov, Poland. LNCS, Springer, to appear.
- [79] P. Palmerini, S. Orlando, R. Perego. Statistical Properties of Transactional Databases. Proceedings of the 2004 ACM Symposium on Applied Computing, SAC 2004, special track on Data Mining, pp. 515–519, March 14-17, 2004, Nicosia, Cyprus.
- [80] S. Orlando, C. Silvestri, R. Perego. A new algorithm for gap constrained sequence mining. Proceedings of the 2004 ACM Symposium on Applied Computing, SAC 2004, special track on Data Mining, pp. 540–547, March 14-17, 2004, Nicosia, Cyprus. ISBN:1-58113-812-1.
- [81] F. Silvestri, S. Orlando, R. Perego. Assigning Document Identifiers to Enhance Compressibility of Web Indexes. Proceedings of the 2004 ACM Symposium on Applied Computing, SAC 2004, special track on Data Mining, pp. 600-605, March 14-17, 2004, Nicosia, Cyprus.
- [82] C. Lucchese, S. Orlando, P. Palmerini, R. Perego, F. Silvestri. kDCI: a Multi-Strategy Algorithm for Mining Frequent Sets. Proceedings of the Workshop on Frequent Itemset Mining Implementations (FIMI'03), in conjunction with ICDM'03. 19 November 2003, Melbourne, Florida, USA.
- [83] T. Fagni, S. Orlando, P. Palmerini, R. Perego, F. Silvestri. A Hybrid Strategy for Caching Web Search Engine Results. Poster paper Twelfth International World Wide Web Conference (WWW2003), May 20-24, 2003, Budapest, Hungary.
- [84] S. Orlando, R. Perego, C. Silvestri. CCSM: an Efficient Algorithm for Constrained Sequence Mining. Proc. of the 6th Int.l Workshop on High Performance Data Mining: Pervasive and Data Stream Mining, in conjunction with Third International SIAM Conference on Data Mining (SDM'03), May 1-3, 2003, San Francisco, CA.
- [85] R. Baraglia, M. Danelutto, D. Laforenza, S. Orlando, P. Palmerini, P. Pesciullesi, R. Perego, M. Vanneschi. AssistConf: a Grid configuration tool for the Assist parallel programming environment. Proc. of the 11th Euromicro Conference on Parallel Distributed and Network based Processing (PDP2003) Genoa, Italy, February 5-7, 2003. IEEE CS Press, pp. 193-200.
- [86] S. Orlando, P. Palmerini, R. Perego, F. Silvestri. Adaptive and Resource-Aware Mining of Frequent Sets. Proc. of the 2002 IEEE Int. Conf. on Data Mining (ICDM 2002), Maebashi City, Japan, December 9 - 12, 2002. IEEE CS Press, pp. 338–345.
- [87] S. Orlando, P. Palmerini, R. Perego, F. Silvestri. Scheduling High Performance Data Mining Tasks on a Data Grid Environment. Proc. of Int. Conf. Euro-Par 2002, 27-30 August 2002, Paderborn, Germany. Lecture Notes In Computer Science 2400 - Springer, pp. 375-384.
- [88] S. Orlando, P. Palmerini, R. Perego, F. Silvestri. An Efficient Parallel and Distributed Algorithm for Counting Frequent Sets. Proc. of Int. Conf. VECPAR 2002, June 2002, Porto, Portugal. Appeared as Selected Paper in the LNCS 2565, Springer, pp. 197-204.

- [89] S. Orlando, P. Palmerini, R. Perego, F. Silvestri. A Scalable Multi-Strategy Algorithm for Counting Frequent Sets. Proc. of the 5th Workshop on High Performance Data Mining, April 2002, Washington, USA.
- [90] S. Orlando, R. Perego, F. Silvestri. Design of a Parallel and Distributed Web Search Engine. 2001 Parallel Computing Conference (ParCo 2001), 4-7 September 2001. Imperial College Press, pp. 197–204.
- [91] S. Orlando, P. Palmerini, R. Perego, Enhancing the Apriori Algorithm for Frequent Set Counting. Proceedings of DaWak 2001, September 5-7, 2001, Munich, Germany, Lecture Notes In Computer Science 2114, Springer, 2001, pp. 71–82.
- [92] C. Gennaro, S. Orlando, and R. Perego. Integrating HPF in a Skeleton Based Parallel Language. Proceedings of the 9th Euromicro Workshop on Parallel and Distributed Processing, PDP2001, Mantova, Italy February 7th-9th, 2001, IEEE CS Press, pp. 195–202.
- [93] S. Ciarpaglini, L. Folchi, S. Orlando, S. Pelagatti, and R. Perego. Integrating Task and Data Parallelism with taskHPF. In *Proc. of PDPTA'2000 - 2000 International Conference on Parallel and Distributed Processing Techniques and Applications*, Las Vegas, Nevada, USA, June 26-29, 2000, pp. 2485–2491.
- [94] S. Orlando, P. Palmerini, and R. Perego. Coordinating HPF programs to mix task and data parallelism. In *Proc. of the 2000 ACM Symposium on Applied Computing (SAC'2000) - Special Track on Coordination Models, Languages and Application*, Villa Olmo, Como, Italy, March 19-21, 2000, ACM Press, pp. 240–247.
- [95] R. Baraglia, D. Laforenza, S. Orlando, P. Palmerini, and R. Perego. Implementation Issues in the Design of I/O Intensive Data Mining Applications on Clusters of Workstations. In *Proc. of IPDPS-2000 - Workshop on High Perf. Data Mining*, Cancun, Mexico, 2000, LNCS 1800 - Springer Verlag, pp. 350-357.
- [96] S. Orlando and R. Perego. Scheduling Data-Parallel Computations on Heterogeneous and Time-Shared Environments. In *Proc. of EUROPAR'98*, Southampton, UK, September 1998, LNCS 1470 - Springer Verlag - Pag. 356-366.
- [97] S. Orlando and R. Perego. A Coordination Layer for Exploiting Task Parallelism with HPF. In *Proc. of the Fourth ACM Workshop on Languages, Compilers, and Run-time Systems for Scalable Computers (LCR98)*, Carnegie Mellon University, Pittsburgh, PA, USA, May 1998, LNCS 1511 - Springer Verlag - Pag. 386–393.
- [98] S. Orlando and R. Perego. An MPI-Based Run-Time Support to Coordinate HPF Tasks. In *Proc. of the EuroPVM/MPI98 Conference*, Liverpool, UK, September 1998, LNCS 1497 - Springer Verlag - Pag 289–296.
- [99] S. Orlando and R. Perego. A Support for Non-Uniform Parallel Loops and its Application to a Flame Simulation Code. In *Proc. of Irregular '97 - Fourth Int. Symposium on Solving Irregularly Structured Problems in Parallel*, LNCS 1253, Springer Verlag, pages 186–197, June 1997. Paderborn (Germany).
- [100] S. Orlando and R. Perego. A template for non-uniform parallel loops based on dynamic scheduling and prefetching techniques. In *Proc. of the 1996 ACM Int. Conf. on Supercomputing*, May 1996. Philadelphia-PA.
- [101] S. Orlando and R. Perego. A Load Balancing Strategy for Iterated Parallel Loop Scheduling. In *Proc. of the HPCN Europe 96 - Int. Conf. and Exhibition*, pages 978–979. LNCS 1067 - Springer Verlag, April 1996. Bruxelles.

- [102] S. Orlando and R. Perego. An Efficient Template for the Highly Parallel Implementation of Non-Uniform Loops. In *Proc of the 2nd EUROPEAN SCHOOL OF COMPUTER SCIENCE – Parallel Progr. Environ. for High Perf. Computing*, pages 31–34, April 1996. Alpe d’Huez - France.
- [103] B. Bacci, M. Danelutto, S. Pelagatti, S. Orlando, and M. Vanneschi. Summarizing an experiment in parallel programming language design. In *Proc. of the HPCN Europe 95 – Int. Conf. and Exhibition*, pages 7–13, LNCS 919 - Springer Verlag, May 1995.
- [104] S. Orlando and R. Perego. Scheduling Non-Uniform Parallel Loops on Highly Parallel Multicomputers. In *Proc. of the IEEE Int. Conf. on High Performance Computing – HiPC95*, pages 547–553. MacGraw Hill, December 1995. New Delhi, India.
- [105] B. Bacci, M. Danelutto, S. Pelagatti, S. Orlando, and M. Vanneschi. Unbalanced Computations onto a Transputer Grid. In *Proceedings of The 1994 Transputer Research and Application Conference*, pages 268–282, October 1994. Athens, Georgia, USA.
- [106] S. Orlando and M. Vanneschi. Compiling Geometric Paradigms through Local Communications. In *Proc. of IFIP Int. Conf. on Applications in Parallel and Distributed Computing*, pages 175–184, Caracas, Venezuela, April 1994.
- [107] B. Bacci, M. Danelutto, S. Orlando, S. Pelagatti, and M. Vanneschi. Efficient Compilation of Structured Parallel Programs for Distributed Memory MIMD Machines. In *Proc. of Int. Conf. ParCo93 - Parallel Computing 93*, pages 565–568, 1994.
- [108] S. Orlando and M. Vanneschi. A Synchronous Model for Static Optimization of Multicomputer Programs. In *Proc. of the IFIP Conf. on Progr. Env. for Parallel Computing*, pages 193–214, Edinburg, Scotland, April 1992. IFIP Transactions Series, North-Holland.
- [109] F. Baiardi and S. Orlando. Strategies for a Massively Parallel Implementation of Simulated Annealing. In *Proc. of Int. Conf. PARLE ’89*, pages 273–287, Eindhoven, The Netherlands, June 1989. LNCS 366 Springer-Verlag.

Edited Books and Journals

- [110] J. Dongarra, D. Laforenza, S. Orlando (Eds.). *Recent Advances in Parallel Virtual Machine (PVM) and Message Passing Interface (MPI) - 10th European PVM/MPI Users’ Group Meeting*, Venice, Italy, September 29 - October 2, 2003, Proceedings. Lecture Notes in Computer Science, 2840, Springer.
- [111] S. Orlando, D. Laforenza (Eds.). Selected Papers from the EUROPVM/MPI 2003 Conference, Venice, Italy, 29 September-2 October 2003. *Int. Journal of High Performance Computing Applications*, Spring 2005, Volume 19, No. 1, Sage Publications.
- [112] D. B. Skillicorn, A. Hameurlain, P. Watson, S. Orlando. *Topic 5: Parallel and Distributed Databases, Data Mining and Knowledge Discovery*. Euro-Par 2004, pp. 346, LNCS 3149, Springer.
- [113] R. Sakellariou, S. Orlando, J.-L. Larriba-Pey, S. Parthasarathy, D. Zeinalipour-Yazti. *Parallel and Distributed Data Management*. Euro-Par 2010, pp. 316, LNCS 6271/2010, Springer.
- [114] S. Orlando, G. Antoniu, A. Ghoting, M. S. Pérez-Hernández. Introduction to Topic 5: Parallel and Distributed Data Management. Euro-Par 2011: pp. 351-352, LNCS 6852/2011, Springer.

Chapters in International Research Books

- [115] R. Baraglia, C. Lucchese, S. Orlando, R. Perego, F. Silvestri. Preserving Privacy in Web Recommender Systems. In *Privacy Aware Knowledge Discovery: Novel Applications and New techniques*, F. Bonchi and E. Ferrari eds., pages 369–391, Taylor and Francis publisher, 2010.
- [116] S. Orlando, A. Raffaetà, A. Roncato, C. Silvestri. Warehousing and Mining Streams of Mobile Object Observations. In *Intelligent Techniques for Warehousing and Mining Sensor Network Data*, Alfredo Cuzzocrea eds., IGI Global publication, ISBN: 978-1-60566-328-9, 2010.
- [117] D. Barbalace, C. Lucchese, C. Mastroianni, S. Orlando, D. Talia. Mining@home: Public Resource Computing for Distributed Data Mining. In *From Grids to Service and Pervasive Computing*, 2008, pp. 217–227, SPRINGER, ISBN: 9780387094540
- [118] F. Bonchi, F. Giannotti, C. Lucchese, S. Orlando, R. Perego, R. Trasarti. On Interactive Pattern Mining from Relational Databases. *Knowledge Discovery in Inductive Databases*, Revised Selected and Invited Papers. LNCS 4747, pp. 42–62, 2007, Springer.
- [119] D. Laforenza, C. Lucchese, S. Orlando, R. Perego, D. Puppini, F. Silvestri. On the Value of Query Logs for Modern Information Retrieval. In A. Soro, G. Paddeu, G. Armano. *Distributed Agent-based Retrieval Tools*. Polimetrica International Scientific Publisher, 2006.
- [120] C. Lucchese, S. Orlando, R. Perego, C. Silvestri. Mining Frequent Closed Itemsets from Distributed repositories. In D. Talia, A. Bilas, M. Dikaiakos. *Knowledge and Data Management in Grids*. vol. 3 of CoreGrid Series, Spinger, 2007.
- [121] R. Baraglia, M. Danelutto, T. Fagni, D. Laforenza, S. Orlando, A. Paccosi, N. Tonello, M. Vanneschi, C. Zoccolo. HPC Application Execution on Grids. Getov, Laforenza, Reinefeld eds., *Future Generation Grids*, CoreGRID Series, Vol 3, pp. 263–282, Springer, 2006. ISBN 0-387-27935-0.
- [122] D. Puppini, F. Silvestri, D. Laforenza, S. Orlando. Toward GRIDLE: A Way to Build Grid Applications Searching Through an Ecosystem of Components. In Jose C. Cunha and Omer F. Rana eds., *Grid Computing: Software Environments and Tools*, Springer Verlag, 2006, ISBN: 1-85233-998-52004.
- [123] B. Bacci, B. Cantalupo, M. Danelutto, S. Orlando, D. Pasetto, S. Pelagatti, M. Vanneschi. An environment for structured parallel programming. In Grandinetti, Kowalick eds., *Advances in High Performance Computing*, pages 219–234, Kluwer Academic Publishers, 1997.
- [124] S. Orlando. Compile-time Parallelizing Techniques. In G. Balbo and M. Vanneschi, editors, *General Purpose Parallel Computers: Architectures, Programming Environments, and Tools*, chapter 8, pages 209–250. ETS Pisa, 1995.

Refereed National Conference Papers

- [125] L. Leonardi, S. Orlando, A. Raffaetà, A. Roncato, C. Silvestri. Visual OLAP analysis on network constrained trajectories. Proc. of the 18th Italian Symposium on Advanced Database Systems (SEBD 2010), Rimini, Italy, June 20-23.
- [126] F. Falchi, C. Lucchese, S. Orlando, R. Perego, Fausto Rabitti. Caching Algorithms for Similarity Search. Proc. of the 17th Italian Symposium on Advanced Database Systems (SEBD 2009).
- [127] C. Gennaro, M. Mordacchini, S. Orlando, F. Rabitti. A Scalable Distributed Data Structure for Multi-Feature Similarity Search. Proc. of the 16th Italian Symposium on Advanced Database Systems (SEBD 2008), Mondello (PA), Italy, June 22-25.

- [128] S. Orlando, A. Raffaetà, A. Roncato, C. Silvestri. Trajectory Data Warehouses: Storing and Aggregating Frequent ST Patterns. Proc. of the 16th Italian Symposium on Advanced Database Systems (SEBD 2008), Mondello (PA), Italy, June 22-25.
- [129] S. Orlando, R. Orsini, A. Raffaetà, A. Roncato, C. Silvestri. Trajectory Data Warehouses: Design Issues and Use Cases. Proc. of the 15th Italian Symposium on Advanced Database Systems (SEBD 2007), pp. 208–219, June 17-20, 2007, Torre Canne - Fasano (BR)
- [130] S. Orlando, R. Perego, F. Silvestri. Assigning Document Identifiers to Enhance Compressibility of Fulltext Indices. Proc. of the 12th Italian Symposium on Advanced Database Systems (SEBD 2004), pp. 222–229.
- [131] P. Criscione, R. Scopigno, and S. Orlando. Metodologia di programmazione parallela P^3L : un'applicazione a problemi di geometria computazionale. In *Proceedings of the Annual AICA Conference*, pages 221–229, September 1995. Chia, Italy.
- [132] B. Bacci, M. Danelutto, S. Pelagatti, S. Orlando, and M. Vanneschi. Methodologies and tools for massively parallel programming: the P^3L approach. In *Proceedings of the Annual AICA Conference*, pages 471–484, September 1994. Palermo, Italy.
- [133] R. Bisiani and S. Orlando. Supporting Parallel Programming with Distributed Artificial Intelligence Techniques. In *Proceedings of the Annual AICA Conference*, pages 441–448, September 1994. Palermo, Italy.
- [134] S. Antonelli, M. Danelutto, R. Di Meglio, S. Orlando, and S. Pelagatti. Una metodologia di programmazione per macchine a parallelismo massiccio - applicazione al problema del ray tracing. In *Atti del Workshop su Elaborazione Parallela: Ricerca ed Applicazioni*, pages 139–156, Roma, Italy, 1992.
- [135] S. Orlando and R. Perego. Un Modello per il Controllo di Transazioni Concorrenti in Ambiente Distribuito. In *Atti del Congresso Annuale AICA*, pages 71–84, Trieste, Italy, October 1989.

Other Works

- [136] R. Baraglia, D. Laforenza, S. Orlando, P. Palmerini, R. Perego. Design of Efficient Input/Output Intensive Data Mining Applications. ERCIM News, no. 44, 2000, a publication of the European Community in Information Technology.
- [137] R. Perego, S. Orlando. Task Parallelism in an HPF Framework. ERCIM News, no. 34, 1998, a publication of the European Community in Information Technology.
- [138] S. Orlando. Combining Template-based Task Parallelism and HPF Data Parallelism. Schloss Dagstuhl – Invited Seminar on “High Level Parallel Programming: Applicability, Analysis and Performance”, April 1999.
- [139] S. Orlando. Load Balancing techniques for Non-uniform Data-Parallel Computations. Schloss Dagstuhl – Invited Seminar on “Theory and Practice of Higher-Order Parallel Programming”, February 1997.
- [140] S. Orlando. La sicurezza nelle transazioni elettroniche. In L. Pilotti ed., *La comunicazione in rete per le PMI*, chapter 6, pages 183–206. Il Sole 24 Ore Libri, December, 1996.
- [141] S. Orlando, R. Perego. A Template for the Implementation of Non-Uniform Loops on Massively Parallel Machines. In *Science and Supercomputing at CINECA, 1995 Report*, G. Erbacci and M. Voli (Eds.), 1996, pages 192–199.

PhD Thesis

- [142] S. Orlando. *Exploiting Synchronous Models for Distributed Memory Multiprocessor Architectures*.
PhD thesis, Dipartimento di Informatica, Università di Pisa - Italy, March 1991. TD-3/91.

Prof. Salvatore Orlando

A handwritten signature in dark ink, appearing to read 'Salvatore Orlando', written in a cursive style.