# Abhinav Sarje

abhinav.sarje@gmail.com • (515) 203 9030 • www.abhinavsarje.net

**Education**

**Iowa State University**, Ames, IA, USA — August 2005 – August 2010
Ph.D. in Computer Engineering (*advisor:* Dr. Srinivas Aluru) — *GPA:* 3.97/4.00
*Thesis:* Applications on emerging paradigms in parallel computing

**Indian Institute of Technology Guwahati**, Guwahati, India — August 2000 – May 2004
Bachelor of Technology in Computer Science & Engineering — *GPA:* 8.35/10.00

**Career Interests**

- High-Performance Computing
- Algorithms on Emerging Architectures
- Computational Biology and Scientific Computing
- Parallel Algorithms & Applications
- Cloud Computing Abstractions
- String and Text Algorithms

**Select Publications**

## Journal Articles

- **A. Sarje**, J. Zola and S. Aluru, **"Accelerating Pairwise Computations on the Cell Processors"**, IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 1, pp. 69–77, 2011.
- J. Zola, M. Aluru, **A. Sarje** and S. Aluru, **"Parallel Information Theory Based Construction of Genome-wide Gene Regulatory Networks"**, IEEE Transactions on Parallel and Distributed Systems, vol. 21, no 12, pp. 1721–1733, 2010.
- **A. Sarje** and S. Aluru, **"Parallel Genomic Alignments on the Cell Broadband Engine"**, IEEE Transactions on Parallel and Distributed Systems, vol. 20, no. 11, pp. 1600–1610, 2009.

[in preparation] **A. Sarje** and S. Aluru, **"An Abstract Parallel Framework for Computations on Trees"**, for IEEE Transactions on Parallel and Distributed Systems.

## Refereed Conference Publications

- **A. Sarje** and S. Aluru, **"A MapReduce Style Framework for Computations on Trees"**, in proc. of the 39th International Conference on Parallel Processing (ICPP), pp. 343–352, 2010.
- **A. Sarje**, J. Zola and S. Aluru, **"Constructing Gene Regulatory Networks on Clusters of Cell Processors"**, in proc. of the 38th International Conference on Parallel Processing (ICPP), pp. 108–115, 2009.
- **A. Sarje** and S. Aluru, **"Parallel Biological Sequence Alignments on the Cell Broadband Engine"**, in proc. of the 22nd IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 1–11, 2008.
- **A. Sarje**, A. Chawre and S. Nair, **"Reinforcement Learning of Player Agents in RoboCup Soccer Simulation"**, in proc. of the 4th IEEE International Conference on Hybrid Intelligent Systems (HIS), pp. 480–481, 2004.

[under review] **A. Sarje** and S. Aluru, **"All-Pairs Computations on Many-core Graphics Processors"**, in 40th International Conference on Parallel Processing (ICPP), 2011.

## Book Chapters

- **A. Sarje**, J. Zola and S. Aluru, **"Pairwise Computations on the Cell Processor with Applications in Computational Biology"**, in Scientific Computing with Multicore and Accelerators, Eds. J. Dongarra, D. A. Bader and J. Kurzak, Chapman & Hall/CRC Computational Science Series, December 2010.
- **A. Sarje** and S. Aluru, **"Parallel Algorithms for Alignments on the Cell B.E."**, in Bioinformatics: High Performance Parallel Computer Architectures, Ed. B. Schmidt, Taylor & Francis Group/CRC Embedded Multi-Core Systems Series, July 2010.

## Miscellaneous

- **A. Sarje**, **"Parallel Techniques for Efficient Pairwise Computations on Emerging Architectures"**, in Ph.D. Forum at the 24th International Parallel and Distributed Processing Symposium (IPDPS), 2010 (**TCPP best paper award**).
- **A. Sarje** and S. Aluru, **"A MapReduce Style Framework for Trees"**, Technical Report, Department of Electrical and Computer Engineering, Iowa State University, 2009.

## Work Experience

**Truveo**, AOL Inc. — Software Engineer
San Francisco, CA, U.S.A. — September 2010 – *present*
Part of the web video search company, dealing with an index of over 650 million videos on the web. Working on parallelization and efficiency improvement of the Truveo notification system, and task generation & management system for celebrity detection in videos based on audio/video analysis. Also, part of a new project initiative on developing parallel/distributed infrastructure for building large-scale web video associations.

**Iowa State University**, Computer Engineering Department — Research Assistant
Ames, IA, U.S.A. — August 2005 – August 2010
Worked on the development and implementation of parallel algorithms on emerging paradigms for various problems with applications in computational and systems biology, materials science, and fluid dynamics. This includes parallel algorithms for genomic alignments and scheduling of generic all-pairs computations on emerging multi- and many-core architectures, cloud computing abstractions for tree structures, and optimal parallel algorithms for the all-$k$-nearest neighbor problem, and multi-search on trees.

**Samsung India Software Operations** — Software Engineer
Bangalore, India — December 2004 – July 2005
Contributed to the design and development of traffic management algorithms for network traffic distributor (IP switch). Involved in the development and testing of IP Multimedia Subsystem (IMS) for large scale 3G wireless network deployment.

**Center for Development of Telematics** — Research Engineer
Bangalore, India — August 2004 – December 2004
Involved in the development and testing of network management software for telecommunications switches and servers. Worked on VoIP management project being initiated in the group.

**Linköping University**, Computational Biology Group, IFM — Project Student
Linköping, Sweden — May 2003 – July 2003
Worked on a survey of data clustering methods used for microarray data, and implementation of agglomerative hierarchical clustering and basic $k$-means clustering algorithms as libraries in Java to be used by the researchers as modules for microarray data analysis.

**Indian Institute of Technology Bombay**, Computer Science Department — Summer Intern
Mumbai, India — May 2002 – June 2002
Worked on fault testing of Xilinx FPGA demonstration boards using VHDL to design testbeds.

## Relevant Technical Skills

[*Expert/good knowledge of*]

- **Languages:** C++, C, Ruby, Unix shell scripting.
- **Parallel programming standards and paradigms:** MPI-2, NVIDIA CUDA, Intel TBB, IBM Cell SDK, OpenMP, Posix threads, Hadoop/MapReduce.
- **Parallel architectures:** Clusters, MPPs/Supercomputers, Many/Multi-cores and heterogeneous systems including Cell B.E. and GPGPUs.

[*Intermediate/basic knowledge of*]

- OpenCL, Python, Perl, Java, MIPS Assembly, XML/XHTML/HTML5, CSS, PHP, MySQL.

**Primary Research Work**

- **Accelerating pairwise computations on clusters of multi/many-cores.** Developed optimal parallel strategy for scheduling large number of all-pairs computations with various input decompositions, for clusters of Cell processors, extendible to multi-cores with the strategy being cache-aware. Implemented as an open source library using C++, MPI-2 and Cell SDK 3.1. We applied this to systems biology and materials science applications. We also developed an architecture-aware GPU-accelerated version of the library using CUDA with C++, and conducted an in-depth analysis of its performance, as well as compared to performance with the Cell version, and multi-core CPU versions implemented using OpenMP and Intel Threading Building Blocks.
- **Parallel genomic alignments on heterogeneous multi-cores.** Developed parallel algorithm for computing optimal pairwise sequence alignments (global/local, spliced and syntenic) on a multi-core processor while adhering to optimized memory transfers and linear memory usage in light of the limited local memories and caches on the computing cores of the processor. We demonstrated performance gain of its implementation with C and Cell SDK 3.0 on Cell B.E.
- **An abstract framework for trees on clouds.** Developed an abstract and generalized framework for computations and searches on tree structures, on the style of the MapReduce paradigm. The computational framework enables simplified writing of numerous applications through two user defined functions, while hiding all system details like the tree distribution and parallel algorithms from the user. Implemented as a generic programming library using C++ and MPI-2 on a large multi-core cluster.
- **Provably optimal parallel algorithm for $k$-nearest neighbors.** [Ongoing] Developing a distributed-memory message-passing model based optimal parallel algorithm for the all-$k$-nearest neighbors problem and related problems like the $k^{\text{th}}$-only nearest neighbor and foreign nearest neighbors.
- **Parallel algorithm for multi-search on trees.** [Ongoing] Developing a distributed-memory message-passing model based parallel algorithm for the *multi-search* problem on a search tree structure.

**Relevant Graduate-level Course-work**

| | |
|---|---|
| Parallel Algorithms and Programming | Distributed Algorithms |
| Parallel Algorithms for Scientific Applications | Theory of Computation |
| Fundamental Algorithms in Computational Biology | Computer Systems Architecture |
| Probabilistic Algorithms in Computer Engineering | Database Implementation |
| Algorithms in Computational Geometry | Principles of Compilers |

**Select Graduate-level Course-work Projects**

- **Parallel algorithm for single linkage clustering.** Developed a novel parallel algorithm for single-linkage hierarchical clustering of multi-dimensional data points for the distributed-memory message-passing model of computation with run-time complexity of $O(\frac{n^2}{p}\log n)$.
- **Theoretical analysis of distributed file systems.** Various distributed file systems were modeled using the distributed shared memory computation model. This was then used to theoretically analyze three prominent file systems – NFS, AFS and GFS, to classify them according to the data consistency models for distributed shared memory systems. Fault resiliency provided by these file systems was also theoretically analyzed.

**Professional Achievements, Affiliations & Activities**

- Best paper award in TCPP Ph.D. forum, held at IPDPS 2010.
- Research Excellence Award, Iowa State University, 2010.
- Research Assistantship, Iowa State University, 2005–2010.
- Member of the IEEE, and the IEEE Computer Society.
- External referee for various conferences, including IPDPS 2010, ICPP 2010, CF 2010, ICCS 2010, COCOON 2009, PPAM 2009, IC3 2009, HiPC 2008, CPM 2008.
- Merit certificate for securing 100% in Physics, CBSE Examination 2000.
- Merit certificates for First, Junior & Intermediate levels of National Mathematics Olympiad (India), 1997, 1998 & 1999, and National Science Olympiad (India), 1999.

**References**

*Available on request.*