

## Areas of Interest

Information Retrieval, Machine Learning, Natural Language Processing, Language Modeling, Data Mining, Design and Analysis of Algorithms, Graph Theory, Stochastic Modeling, Fault Tolerance.

## Education

### PhD (Computer Science)

2002-2007

Advisor: John H. Reif

Dissertation: *Theory and Experiments in Self-Assembly and Nanorobotics***Duke University**

Durham, NC

### Bachelor of Technology (Computer Science)

1998-2002

**IIT (Indian Institute of Technology)**

New Delhi, India

## Work Experience

### Software Developer

Dec'07-present

**Microsoft Corporation**

Redmond, WA

I am working in Bing Search at Microsoft. As a member of Dynamic Ranking team in Bing Search Relevance group, I work in the areas of Information Retrieval, Data Mining, Machine Learning and Algorithm Design. We deal with huge data sets on a daily basis, and solve design and implementation problems on large scale systems. The efficiency in terms of time and space is primary consideration due to the scale of the system. In particular, my first project dealt with the impact of proximity of query words in documents on the search relevance, and to leverage it in order to improve search relevance. Information Retrieval and Machine Learning techniques were applied to improve the search relevance, and algorithmic and programming skills were used in design and implementation of various components. My second project deals with spell-correction in the query words. This project also has a deeply ingrained algorithmic and machine learning component. The work requires intensive team work and extends to collaborations across different teams. I use C++ for writing production code, SQL and C# for data mining, and sometimes Perl for analytical purpose.

### Research Assistant

May'03-Dec'07

**Dept. of Computer Science, Duke University,**

Durham, NC

I worked in the areas of Algorithms, Graph-theory, Complexity theory, Markov chains, and Stochastic modeling with the application in Self-Assembly and Nanorobotics. One of my projects was design and implementation of a framework for modeling nanorobotical devices. It involved Randomized algorithms, Stochastic modeling, Collision detection, efficient String matching and Graph algorithms. Design and implementation of efficient data structure was also critical to time-space optimization of this computation intensive system. In another of my project, I worked on various algorithmic design problems in the area of fault-tolerance in self-assembly. In addition to these, I also worked on Markov Chains, Graph-theoretic and Complexity-theoretic problems in the settings of self-assembly.

### Teaching Assistant

Jan'03-Dec'06

**Dept. of Computer Science, Duke University**

Durham, NC

Numerical Analysis (Fall 07), Design and Analysis of Algorithms (Fall 06), Randomized Algorithms (Spring 06), Design and Analysis of Algorithms (Fall04), Numerical Analysis (Fall03), Mathematical Foundations of Computer Science(Spring03)

### Intern

May-July'01

**Max Planck Institut fur Informatik,**

Saarbrucken, Germany

Implemented algorithm for single-source shortest-paths on arbitrary directed graphs in linear average-case time.

## Selected Publications

1. Sudheer Sahu, John Reif, "Capabilities and Limits of Compact Error Resilience Methods for Algorithmic Self-Assembly", *Algorithmica*, Springer New York, 2008.

2. Urmi Majumder, Sudheer Sahu, John Reif, "Stochastic Analysis of Reversible Self-Assembly", *Journal of Computational and Theoretical Nanoscience, American Scientific Publishers*, 5(7):1289-1305, 2008
3. Sudheer Sahu, Bei Wang, John Reif, "A Framework for Modeling DNA Based Molecular Systems", *Journal of Computational and Theoretical Nanoscience, American Scientific Publishers*, 5(11):2124-2134, 2008
4. John Reif, Sudheer Sahu, Peng Yin, "The Complexity of Graph Self-Assembly in Accretive Systems and Self-Destructible Systems", *LNCS, Springer Berlin/Heidelberg Publisher*, 3892:257-274, 2006.
5. Peng Yin, Sudheer Sahu, John Reif, "Autonomous DNA Cellular Automata", *DNA11, Ontario, Canada, June 6-9, 2005*.
6. Sudheer Sahu, Peng Yin, John Reif, "A Self-Assembly Model of DNA Tiles with Time Dependent Glue Strength", *DNA11, Ontario, Canada, June 6-9, 2005*.
7. John Reif, Sudheer Sahu, Peng Yin, "Compact Error-Resilient Computational Tiling Assemblies", *DNA10, Milano, Italy, June 7-10, 2004*.
8. Peng Yin, Andrew Turberfield, Sudheer Sahu, John Reif, "Design of an Autonomous DNA Nanomechanical Device Capable of Universal Computation", *DNA10, Milano, Italy, June 7-10, 2004*.

## Technical Reports

1. Sudheer Sahu, Deepak Ajwani, "Cache-Efficient Fast Fourier Transform", *B.Tech Project, IIT Delhi, 2002*.
2. Sudheer Sahu, "Single-source shortest-paths on arbitrary directed graphs in linear average-case time", *Max Planck Institut fur Informatik, July 2001*.
3. Sudheer Sahu, Anshul Sawant, "Terrain Visualisation for 3D Geometric Terrains", *Mini-Project, IIT Delhi, 2001*.

## Selected Course Projects

1. "Clustering Error-Prone Gene Expression Data". (*Computational Functional Genomics*, Spring 03)
2. "Genome Rearrangement with Gene Families". (*Algorithms in Computational Biology*, Fall 02)
3. "Accurate Static Branch Prediction using Value Range Propagation". (*Computer Architecture*, Fall 02)
4. "Hardware Sorter". (*Digital Hardware Design*, Spring 00)
5. "Design and simulation of processor based on PowerPC architecture with extended instruction set" (*Computer Architecture*, Fall 99)

## Graduate Courses:

**Theory and Algorithms:** *Design and Analysis of Algorithms, Complexity Theory, Randomized Algorithms, Approximation Algorithms, Numerical Analysis, Algorithms in Computational Biology, Computational Functional Genomics, Numerical Artificial Intelligence* **Mathematics:** *Applied Stochastic Processes, Topology, Combinatorics* **Systems and Architecture:** *Advanced Computer Architecture*

## Undergraduate Courses:

**Theory and Algorithms:** *Data Structures, Analysis and Design of Algorithms, Theory of Computation, Discrete Structures, Logic for Computer Science, Programming Languages, Numerical and Scientific computing.* **Mathematics:** *Analysis and Differential Equation, Complex Analysis and Vector Spaces, Optimization, Probability and Stochastic Processes.* **Graphics:** *Computer Graphics, Digital Image Analysis.* **Systems and Architecture:** *Computer Architecture, Digital Hardware Design, Operating Systems, File Structures and Information Systems, Computer Networks, DBMS.*

## Professional Services

- **Assistant Organizer:** *Conference on Foundations of Nanoscience: Self-assembled Architectures and Devices (FNANO)* in 2004, 2005, 2006 and 2007.

- **Assistant Organizer:** *NSF workshop: Emerging Opportunities of Nanoscience to Energy Conversion and Storage*, Nov 21-23, 2005, Washington DC.
- **Referee to Conferences:** Tenth International Meeting on DNA Computing (DNA10) 2004, Eighth International Conference on Knowledge Based Intelligent Information and Engineering Systems(KES) 2004, European Symposium on Algorithms (ESA) 2006.

## Academic Achievements

- National Talent Search Scholarship from NCERT (National Council for Educational Research and Training) in 1996
- Rank 1 (among approx. 75,000 candidates) in MP-PET, Rank 2 (among approx. 110,000 candidates) in DCE, and Rank 90 (among approx. 150,000 candidates) in IIT-JEE in 1998

## Computer Skills

Languages:	C, C++, Perl
Platforms:	Unix, Linux, Windows
Other tools/skills:	HTML, XML, Matlab