

SUDEEPA ROY

571 Moore GRW
3330 Walnut Street
Philadelphia, PA 19104
+1 215 573 2581

sudeepa@cis.upenn.edu
roy.sudeepa@gmail.com
<http://www.cis.upenn.edu/~sudeepa/>
+1 857 253 1568

Research Interests Data provenance; probabilistic databases; database privacy; algorithms with applications in databases; information extraction and data mining.

Education **University of Pennsylvania** 2007 - current

Philadelphia, PA, USA.
PhD student, Computer and Information Science
Dissertation research: Data Provenance and Applications
Advisors: Prof. Susan B. Davidson and Prof. Sanjeev Khanna
Current CGPA: 3.98/4.0.

Indian Institute of Technology 2004 - 2006

Kanpur, India.
Master of Technology in Computer Science and Engineering
Masters' thesis: On "Go With the Winners" Algorithm
Advisors: Prof. Manindra Agrawal and Prof. Somenath Biswas
CGPA: 10.0/10.0.

Jadavpur University 2000 - 2004

Kolkata, India.
Bachelor of Computer Science and Engineering
CGPA: 3.92/4.0.

Experience **IBM Research, Almaden** Summer 2010 and Summer 2011

San Jose, CA, USA.
Summer Internship
I worked in the Infrastructure for Intelligent Information Systems group on improving the output quality of an information extraction system.

University of Pennsylvania Fall 2008, Spring 2009, Spring 2009

Philadelphia, PA, USA.
Teaching Assistant
I worked as a teaching assistant for the undergraduate courses *Automata, Computability and Complexity* and *Introduction to Algorithms*.

Google India Pvt. Ltd. 2006 - 2007

Bangalore, India.
Software Engineer
I mainly worked on designing and building the frontend and backend of a spam reporting framework for the social network site Orkut.

Honeywell Technology Solutions Lab Summer 2005

Bangalore, India.
Summer Intern
I designed and built part of a tool that automatically translates Simulink models into the input language of a model checker NuSMV.

Publications

CONFERENCE PUBLICATIONS

Queries with Difference on Probabilistic Databases

Sanjeev Khanna, Sudeepa Roy and Val Tannen

International Conference on Very Large Data Bases (VLDB) 2011

Provenance Views for Module Privacy

Susan B. Davidson, Sanjeev Khanna, Tova Milo, Debmalya Panigrahi and Sudeepa Roy

Symposium on Principles of Database Systems (PODS) 2011

Faster Query Answering in Probabilistic Databases using Read-Once Functions

Sudeepa Roy, Vittorio Perduca and Val Tannen

International Conference on Database Theory (ICDT) 2011

An Optimal Labeling Scheme for Workflow Provenance Using Skeleton Labels

Zhuowei Bao, Susan Davidson, Sanjeev Khanna and Sudeepa Roy

ACM SIGMOD International Conference on Management of Data (SIGMOD) 2010

Optimizing User Views for Workflows

Olivier Biton, Susan Davidson, Sanjeev Khanna and Sudeepa Roy

International Conference on Database Theory (ICDT) 2009

STCON in Directed Unique-Path Graphs

Sampath Kannan, Sanjeev Khanna and Sudeepa Roy

Foundations of Software Technology and Theoretical Computer Science (FSTTCS) 2008

Automatic Translation of Simulink Models into Input Language of a Model Checker

Meenakshi B., Abhishek Bhatnagar and Sudeepa Roy

International Conference on Formal Engineering Methods (ICFEM), 2006

WORKSHOP PUBLICATIONS/VISION PAPERS/INVITED PAPERS

Hiding Data and Structure in Workflow Provenance

Susan Davidson, Zhuowei Bao and Sudeepa Roy

Invited paper, International Workshop Conference on Databases in Networked Information Systems (DNIS) 2011

On provenance and privacy

Susan Davidson, Sanjeev Khanna, Sudeepa Roy, Julia Stoyanovich, Val Tannen and Yi Chen

Invited paper, International Conference on Database Theory (ICDT) 2011

Enabling Privacy in Provenance-Aware Workflow Systems

Susan Davidson, Sanjeev Khanna, Sudeepa Roy, Julia Stoyanovich, Val Tannen, Yi Chen and Tova Milo

Vision Track, Conference on Innovative Data Systems Research (CIDR) 2011

Privacy Issues in Scientific Workflow Provenance

Susan Davidson, Sanjeev Khanna, Sudeepa Roy and Sarah Cohen Boulakia

International Workshop on Workflow Approaches to New Data-centric Science (WANDS) 2010

MANUSCRIPTS

Optimization Problems for Dictionary Refinement in Information Extraction

Sudeepa Roy, Laura Chiticariu, Vitaly Feldman, Frederick R. Reiss and Huaiyu Zhu

Submitted.

A Propagation Model for Provenance Views of Public/Private Workflows

Susan B. Davidson, Tova Milo and Sudeepa Roy

Submitted.

Patent **Automatic Translation of Simulink Models into Input Language of a Model Checker**
Meenakshi B., Abhishek Bhatnagar, Sudeepa Roy
US Patent 7698668

Skills C, C++, Java, Matlab, R

Awards and Honors Recipient of Google U.S./Canada PhD Fellowship in Structured Data Analysis, 2011.

Recipient of Doctoral Fellowship, University of Pennsylvania, 2007-2011.

Obtained certificate of appreciation from Indian Institute of Technology Kanpur for Academic Excellence in Computer Science and Engineering in the year 2004-05.

Secured All India Rank of 35 (99.90 percentile) in Graduate Aptitude Test in Engineering (GATE), 2004 in Computer Science and Engineering among 35,019 candidates.

Secured a rank of 46 in WBJEE (state-level engineering entrance examination) among approximately 46000 candidates.

Professional Service External reviewer for the conferences VLDB, PODS, SODA; and the journals Transactions on Knowledge and Data Engineering, Transactions on Algorithms, and SIAM Journal on Computing.

References Prof. Susan B. Davidson (advisor)
Professor
Dept. of Computer & Information Science, University of Pennsylvania
email - susan@cis.upenn.edu

Prof. Sanjeev Khanna (advisor)
Professor
Dept. of Computer & Information Science, University of Pennsylvania
email - sanjeev@cis.upenn.edu

Dr. Laura Chiticariu
Researcher
IBM Research, Almaden
email - chiti@us.ibm.com

Prof. Tova Milo
Professor
School of Computer Science, Tel Aviv University
email - milo@cs.tau.ac.il

Prof. Val Tannen
Professor
Dept. of Computer & Information Science, University of Pennsylvania
email - val@cis.upenn.edu