COMPUTER SCIENCE DEPARTMENT • 5000 FORBES AVENUE • PITTSBURGH PA 15213 • USA
HTTP://WWW.CS.CMU.EDU/~SHUCHI • E-MAIL **SHUCHI@CS.CMU.EDU** • PHONE 412 897 7443

SHUCHI CHAWLA

RESEARCH INTERESTS

Combinatorial Optimization, Approximation Algorithms, Game Theory, Database Privacy, Machine Learning Theory and Complexity Theory.

EDUCATION

2000-present **Carnegie Mellon University**, Pittsburgh, PA, USA.
Ph.D. in Computer Science
Advisor: Prof. Avrim Blum
Dissertation: Approximation Algorithms for Path-Planning and Partitioning Problems in Graphs

1996-2000 **Indian Institute of Technology**, New Delhi, INDIA.
B.Tech. in Computer Science and Engineering
Advisor: Prof. Huzur Saran
Senior Thesis: QoS based Scheduling in BLUETOOTH

ACADEMIC HONORS

2004-2005 Awarded *IBM Ph.D. Fellowship*

2000 Awarded *Carnegie Mellon Graduate Fellowship*

2000 Ranked 2nd in the graduating class in IIT, Delhi.

2000-1997 Awarded *Institute Merit Award* in semesters 1 through 8 by IIT, Delhi.

1996 Ranked 32nd in All India Joint Entrance Examination (JEE-1996) for admission to the Indian Institutes of Technology.

1995 Gold medalist at the *Regional Mathematics Olympiad*, Delhi, India.

1994 Recipient of the *National Talent Search Examination (NTSE) Scholarship* awarded by National Council for Educational Research and Training (NCERT), India.

RESEARCH EXPERIENCE

August, 2004 **Microsoft Research, Silicon Valley Campus, CA.**
Mentor: Cynthia Dwork
Project: Database Privacy

May – July, 2004 **IBM Almaden Research Center, CA.**
Mentor: Ravi Kumar
Project: Hardness of approximation of Sparsest Cut

May – Aug, 2003 **Microsoft Research, Silicon Valley Campus, CA.**
Mentor: Cynthia Dwork
Project: Database Privacy and Rank Aggregation

May – Sept, 1999 **IBM India Research Lab, New Delhi, India.**
Mentors: Rajeev Shorey and Alok Aggarwal
Project: Pricing on the Internet

REFREED PUBLICATIONS

1. **Embeddings of Negative-type Metrics and Improved Approximations to Sparsest Cut**, Shuchi Chawla, Anupam Gupta and Harald Räcke, *Symposium on Discrete Algorithms (SODA)*, 2005.

2. **Toward Privacy In Public Databases**, Shuchi Chawla, Cynthia Dwork, Frank McSherry, Adam Smith and Hoeteck Wee, *Theory of Cryptography Conference (TCC)*, 2005.

3. **Approximation Algorithms for Deadline-TSP and Vehicle Routing with Time-Windows**, Nikhil Bansal, Avrim Blum, Shuchi Chawla and Adam Meyerson, *ACM Symposium on the Theory of Computation (STOC)*, 2004.

4. **Worst-case Payoffs of a Location Game**, Shuchi Chawla, Uday Rajan, R. Ravi and Amitabh Sinha, short paper (poster) in *ACM Conference on Electronic Commerce (EC)*, 2004.

5. **Approximation Algorithms for Orienteering, Discounted Reward Collection, and Other Variants of Prize Collecting TSP**, Avrim Blum, Shuchi Chawla, David Karger, Terran Lane, Maria Minkoff and Adam Meyerson, *44th Symposium on the Foundations of Computer Science (FOCS)*, 2003.

6. **Scheduling for Flow-Time with Admission Control, or, How to Manage your ``To-do'' List**, Nikhil Bansal, Avrim Blum, Shuchi Chawla and Kedar Dhamdhere, *European Symposium on Algorithms (ESA)*, 2003.

7. **Profit Maximizing Mechanisms for the Extended Multicasting Game**, Shuchi Chawla, David Kitchin, Uday Rajan, Ramamoorthi Ravi and Amitabh Sinha, short paper (poster) in *ACM Conference on Electronic Commerce (EC)*, 2003.

8. **Online Oblivious Routing**, Nikhil Bansal, Avrim Blum, Shuchi Chawla and Adam Meyerson, *Symposium on Parallel Algorithms and Architecture (SPAA)*, 2003.

9. **On the Scaling of Congestion in the Internet Graph**, Aditya Akella, Shuchi Chawla, Arvind Kannan and Srinivasan Seshan, *ACM SIGCOMM Computer Communication Review*, Special Issue on the Science of Network Design (2004), 34(3):43—55.
   Preliminary version appeared in *Principles of Distributed Computing (PODC)*, 2003.

10. **Correlation Clustering**, Nikhil Bansal, Avrim Blum and Shuchi Chawla, *Machine Learning*, invited to the Special Issue on Clustering (2004), 56(1-3):89—113.
    Preliminary version appeared in *43rd Symposium on the Foundations of Computer Science (FOCS)*, 2002.

11. **Static Optimality and Dynamic Search Optimality in Lists and Trees**, Avrim Blum, Shuchi Chawla and Adam Kalai, *Algorithmica,* invited to the Special Issue on Online Algorithms (2003), 36: 249—260.

Preliminary version appeared in *Symposium on Discrete Algorithms (SODA)*, 2002.

12. **Learning from Labeled and Unlabeled Data using Graph Mincut**, Avrim Blum and Shuchi Chawla, *International Conference on Machine Learning (ICML)*, 2001.

13. **QoS based Scheduling for Incorporating Variable Rate Coded Voice in BLUETOOTH**, Shuchi Chawla, Huzur Saran and Mitali Singh, *International Conference on Communications (ICC)*, 2001.

SELECTED OTHER TECHNICAL WRITINGS

14. **On the Utility of Privacy-Preserving Histograms,** Shuchi Chawla, Cynthia Dwork, Frank McSherry and Kunal Talwar, submitted to Principles of Database Systems (PODS), 2005.

15. **On the Hardness of Approximating Sparsest Cut and Multicut**, Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani and D. Sivakumar, submitted to Computational Complexity (CCC), 2005.

16. **Bayesian Optimal No-deficit Mechanism Design**, Shuchi Chawla, Jason Hartline, Uday Rajan and R. Ravi. Technical report CMU-CS-04-153, August 2003.

17. **Approximation Algorithms for Path-planning and Clustering Problems on Graphs**, Shuchi Chawla, Thesis Proposal, Computer Science Department, Carnegie Mellon University, February 2004.

18. **Mechanisms for Internet Routing: A Study**, Aditya Akella, Shuchi Chawla and Srini Seshan. Technical report CMU-CS-02-163, July 2002.

19. **QoS based Scheduling in BLUETOOTH**, Shuchi Chawla, B. Tech. Thesis, IIT, Delhi, May 2000.

TEACHING EXPERIENCE

Fall, 2004 — **Randomized Algorithms**
Graduate level, CMU 15-859 Lecturer: Prof. Anupam Gupta
I helped Prof. Gupta design the course and pick the topics to be covered. I taught about half of the lectures, designed homeworks, and helped with the grading.

Spring, 2003 — **Algorithms and Applications**
Senior level, CMU 15-499 Lecturer: Prof. Guy Blelloch
I taught a few lectures, held office hours, designed homeworks and exams, and graded them.

Spring, 2001 — **Great Theoretical Ideas in Computer Science**
Freshman level, CMU 15-251 Lecturer: Prof. Steven Rudich
I taught weekly recitations, held office hours, designed homeworks and exams, and graded them.

INVITED TALKS

- **Approximation Algorithms for Path-Planning Problems**
  - University of Washington, Seattle — Nov, 2003
  - IBM T. J. Watson Research Center — Mar, 2004
  - Princeton University — Mar, 2004
  - Cornell University — Apr, 2004
- **From Idiosyncratic to Stereotypical: Toward Privacy in Public Databases**
  - Microsoft Research, Redmond — Nov, 2003
  - DIMACS Workshop on Privacy-Preserving Data Mining — Mar, 2004

PROFESSIONAL SERVICES

- Member of the **CMU SCS Speakers Club**, 2002-2005.
- Member of the **CMU CS Doctoral Review Committee**, 2004-2005.
- Served as an external reviewer for Information Processing Letters, INFORMS Journal on Computing, Internet Mathematics, Journal of Computer and System Sciences, Machine Learning, Networks, Theoretical Computer Science, FOCS, STOC, SODA, Approx and ESA.

PERSONAL INFORMATION

Nationality	Indian
Visa Status	F-1

REFERENCES

Prof. Avrim Blum
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
avrim@cs.cmu.edu
Ph. 412-268-6452

Prof. Anupam Gupta
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
anupamg@cs.cmu.edu
Ph. 412-268-7127

Prof. R. Ravi
Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213
ravi@cmu.edu
Ph. 412-268-3694

Dr. Cynthia Dwork
Microsoft Research
Silicon Valley Campus
Mountain View, CA 94043
dwork@microsoft.com
Ph. 650-693-3701

Prof. Moses Charikar
Dept. of Computer Science
Princeton University
Princeton, NJ 08544
moses@cs.princeton.edu
Ph. 609-258-7477

Approximation Algorithms for Orienteering and Discounted-Reward TSP

Avrim Blum, Shuchi Chawla, David Karger, Terran Lane, Maria Minkoff and Adam Meyerson
44th Symposium on the Foundations of Computer Science, 2003.

In this paper, we give the first constant-factor approximation algorithm for the rooted *Orienteering* problem. Given a graph with lengths on edges and prizes (rewards) on nodes, and a start node *s*, the goal in this problem is to find a path that maximizes the reward collected, subject to a hard limit on the total length of the path. We also give a constant factor approximation for a new problem motivated by robot navigation, that we call Discounted-Reward TSP. In this problem, instead of a length limit we are given a discount factor γ, and the goal is to maximize total discounted reward collected, where reward for a node reached at time t is discounted by γ^t. Our algorithm for Orienteering is the first approximation for the problem, solving a long-standing open question in Operations Research.

Embeddings of Negative-type Metrics and An Improved Approximation to Generalized Sparsest Cut

Shuchi Chawla, Anupam Gupta and Harald Racke
ACM SIAM Symposium on Discrete Algorithms, 2005.

In this paper, we study the metrics of *negative type*, which are metrics *(V,d)* such that $\sqrt{d}$ is a Euclidean metric; these metrics are thus also known as ``ℓ_2-squared'' metrics. We show how to embed *n*-point negative-type metrics into Euclidean space ℓ_2 with distortion $O(\log^{3/4} n)$, improving on a classic $O(\log n)$ distortion embedding due to Bourgain (1985). This embedding result, in turn, implies an $O(\log^{3/4} k)$-approximation algorithm for the Sparsest Cut problem with non-uniform demands. This result also implies that *n*-point subsets of the Manhattan metric embed into the Euclidean metric with distortion $O(\log^{3/4} n)$.

On the Hardness of Approximating Multicut and Sparsest Cut

Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani and D. Sivakumar
Submitted to *Computational Complexity (CCC), 2005*.

We consider the following *Multicut* problem – given a weighted undirected graph, with *k* pairs of vertices called demand pairs, find the smallest weight set of edges, the removal of which disconnects all the demand pairs. This problem is approximable within $O(\log k)$ using a linear programming relaxation. We prove that assuming the so-called *Unique Games conjecture* of Khot (2002), it is NP-hard to approximate this problem to within an $\Omega(\log\log k)$ factor. Our result extends to several related problems including *Sparsest Cut* and *Min-2CNF≡ Deletion*. Prior to our work, Multicut and Min-2NF≡ Deletion were known to be hard to approximate only within very small constant factors. Ours is the first hardness of approximation result for the Sparsest Cut problem.

Correlation Clustering

Nikhil Bansal, Avrim Blum and Shuchi Chawla
Machine Learning, Special Issue on Clustering (2004), 56(1-3):89—113.

> We introduce the following clustering problem. We are given a complete graph on n vertices (items), where each edge (u,v) is labeled either + or - depending on whether u and v have been deemed to be similar or different. The goal is to produce a partition of the vertices that agrees as much as possible with the edge labels. This formulation is motivated from a document clustering problem in which one has a pairwise similarity function f learned from past data, and the goal is to partition the current set of documents in a way that correlates with f as much as possible. An interesting feature of this formulation is that, unlike for traditional clustering measures, here one does not need to specify the number of clusters as a separate parameter. Our main results are a constant factor approximation for minimizing disagreements with edge labels and a PTAS for maximizing agreements with the labels.

Toward Privacy In Public Databases

Shuchi Chawla, Cynthia Dwork, Frank McSherry, Adam Smith and Hoeteck Wee
Theory of Cryptography Conference, 2005.

> We begin a theoretical study of the *census problem*. Informally, in a census, individual respondents give private information to a trusted party (the census bureau), who publishes a *sanitized* version of the data. There are two fundamentally conflicting requirements: *privacy* for the respondents and *utility* of the sanitized data. Unlike in secure function evaluation, here privacy is paramount; what cannot be learnt safely should not be learnt. An important contribution of this work is a definition of privacy for statistical databases, together with a method for describing and comparing the privacy offered by specific sanitization techniques. We obtain several privacy results using two different sanitization techniques, and two utility results involving clustering. Of these, the most interesting is a histogram-based summarization of data that preserves privacy with a high probability.

Bayesian Optimal No-deficit Mechanism Design

Shuchi Chawla, Jason Hartline, Uday Rajan and R. Ravi
Technical report CMU-CS-04-153, August 2003.

> We consider the problem of designing an auction in a Bayesian setting that gives the maximum expected profit to the auctioneer. Myerson (1981) reduces this problem to that of computing the *efficient allocation* on the bidders' *virtual valuations*, simple functions of the bidders' value distributions. We review this technique for arbitrary single parameter agent problems and further consider the problem of merging the worst case *no-deficit* condition with this average case *Bayesian expected profit maximization* – we maximize profit subject to the constraint that the auctioneer *never* makes a loss. We show that among ex post incentive compatible *no-deficit* mechanisms, the Myerson mechanism is optimal for *supermodular* costs, Myerson merged with a simple "thresholding" mechanism is optimal for *all-or-nothing* costs, and neither mechanism is optimal for general *submodular* costs. We also consider the computational hardness of the problem and Bayesian incentive compatible mechanisms.