Ashwin Kumar.V. Machanavajjhala

Cornell University
Department of Computer Science
Ithaca, NY 14853

Cell: (607) 351-0477
Email: mvnak@cs.cornell.edu
http://www.cs.cornell.edu/∼mvnak

Research Interests

Data Privacy, and its relation to Web-Scale Data Management, and Cryptography.

Education

Ph.D., Computer Science, **Cornell University, Ithaca, NY, USA** **Expected 2008**
Thesis Advisor: Prof. Johannes Gehrke

B.Tech., Computer Science, **Indian Institute of Technology - Madras, Chennai, India** **2002**
Thesis Advisor: Prof. C. Pandu Rangan

Honors and Awards

- The poster, *Beyond k-Anonymity: New Scheme for Privacy Preserving Data Publishing*, was judged the **best visionary poster**, in DB/IR Day, April 2005.
- Graduated with the 2^{nd} highest CGPA amongst 400 students in the class of 2002, IIT-Madras.
- Awarded the Biswajit Sain Endowment Prize for securing the highest CGPA in 1999-2000.
- Awarded the Academic Distinction Certificate for securing the 2^{nd} highest CGPA in 1998-1999.
- Ranked 6^{th} amongst about 100,000 candidates in the Joint Entrance Exams to the IITs, 1998.
- Amongst the top 35 students in India screened for the International Physics Olympiad, 1998.

Research Experience

Thesis Work, Cornell University **2004 - current**
Title: Privacy in Data Publishing.
Demonstrated using simple attacks that k-anonymous data, where each individual in the data is indistinguishable from at least $k-1$ other individuals, is not private. Proposed **ℓ-Diversity**, a practical and efficiently implementable condition that provably guarantees privacy against adversaries with limited amounts of background knowledge.

Initiated a formal study of adversarial background knowledge expressed as propositional formulae. Defined basic units of knowledge that can express any adversarial knowledge, and helped propose efficient algorithms to compute the ℓ basic units of knowledge that lead to the maximum disclosure of sensitive information from the published data.

Performed the first formal privacy analysis of OnTheMap, a real world application that plots commute patterns of workers on the U.S. map to study workforce indicators, in collaboration with Prof. John Abowd. Identified an open problem that applying existing provably private anonymization techniques to sparse data greatly diminishes the utility of the resulting data. When applied to the sparse OnTheMap data, these algorithms led to the generation of spurious commute patterns. Proposed algorithms for publishing useful commute patterns while preserving the privacy of worker residences, despite the sparsity of data.

Research Intern, Yahoo! Research **Summer 2007**
Work in submission at SIGMOD 2008 (double-blind review). Filed two patents as a result of this work.

Research Intern, IBM Almaden Research Labs **Summer 2005**
Title: Efficient Privacy Enforcement
Studied the tractability of enforcing complex privacy policies, and identified large classes of efficiently enforceable privacy policies using a connection to the problem of query containment. Joint work with Dr. Rakesh Agrawal, Dr. Ramakrishnan Srikant, Dr. Alexandre Evfimievsky, and Prof. Johannes Gehrke.

Research Assistant, Cornell University **2002-2004**
Title: PEPPER: Peer-to-Peer Data Storage and Indexing
Studied load balancing in peer-to-peer data management systems that support range queries with Prof. Johannes Gehrke and Prof. Jayavel Shanmugasundaram. Proposed new load balance algorithms that provably bounded the load imbalance in the system below $2+\epsilon$, beating the previous bound of 4.6. Also contributed to the design and implementation of *P-Ring*, a new P2P system for storage and indexing data, which has been deployed on PlanetLab, and is currently being commercialized by ATC, NY.

Undergraduate Research, I.I.T-Madras **2001-2002**
Title: Secure Communication on Arbitrary Networks
Studied secure communication across a network containing malicious routers. Characterized the possibility of secure communication in terms of the connectivity of the underlying network under the non-threshold adversary model, and proposed efficient secure communication protocols on networks that allow secure communication.

Intern, Infosys Technologies Limited, India **Summer 2001**
Title: Software Security Assurance
Proposed a security assurance model for off the shelf components using Aspect Oriented Programming and Proof Carrying Codes.

Teaching Experience

Cornell Mentoring Program, **2003, 2004, 2006**
Advised an incoming graduate student on selecting courses and projects that would potentially suit their research interests.
CS 632: Advanced Database Systems, Teaching Assistant, **Spring 2005**
Taught two lectures on data privacy and online aggregation in addition to grading responsibilities.
Privacy Reading Group, Organizer, **Fall 2004**
Headed a weekly reading group on Data Privacy that attracted other graduate students to this field of research.
CS 482: Algorithms, Teaching Assistant, **Spring 2003**
Duties involved grading weekly assignments and holding office hours.
ECE/CS 314: Computer Organization, Teaching Assistant, **Fall 2002**
Gave weekly lectures on material not covered in class in addition to grading and advising responsibilities. Assisted the students with a project where they were required to implement a fully functional MIPS architecture at the gate level.
ID101: Engineering Design, Teaching Assistant, **Fall 1999**
Responsible for counseling a group of 20 students who were given a set of open-ended design projects. Met with the faculty to assess student performance, and help design assignments and course material.

Professional Services

Professional Affiliations. Student member of ACM and IEEE.
Reviewer. TODS (2007), TKDE (2007), and JCSS (2006).
External Reviewer. SIGMOD (2005, 2006), PODS (2005), VLDB (2004, 2005, 2006, 2007), KDD (2004, 2006), STOC (2007), and EDBT (2006).

Relevant Graduate Level Courses

Advanced Database Systems, Reasoning about Uncertainty, Probability and Measure Theory, Cryptography, System Security, Game Theory, Random Graphs and Spectral Methods, Approximation Algorithms, Theory of Computation, Analysis of Algorithms, and Advanced Programming Languages.

Publications

Journal.

J-1. *Ashwin Machanavajjhala*, Daniel Kifer, Johannes Gehrke, Muthuramakrishnan Venkitasubramaniam, "ℓ-Diversity: Privacy beyond k-Anonymity", In *TKDD*, vol 1, no. 1, 2007.

J-2. *M. V. N. Ashwin Kumar*, Arun K. Singh, Ramesh Babu, "A Security Assurance Framework for Component Based Software Development", In *Informatica*, vol 25, no. 4, 2001.

Conference.

C-1. *Ashwin Machanavajjhala*, Daniel Kifer, John Abowd, Johannes Gehrke, Lars Vilhuber, "Privacy: Theory meets Practice on the Map", In *Proc. ICDE 2008*.

C-2. Adina Crainiceanu, Prakash Linga, *Ashwin Machanavajjhala*, Johannes Gehrke, Jayavel Shanmugasundaram, "P-Ring: An Efficient and Robust P2P Range Index Structure", In *Proc. SIGMOD 2007*.

C-3. David Martin, Daniel Kifer, *Ashwin Machanavajjhala*, Johannes Gehrke, Joseph Halpern, "Worst Case Background Knowledge", In *Proc. ICDE 2007*.

C-4. *Ashwin Machanavajjhala*, Johannes Gehrke, "On the Efficiency of Checking Perfect Privacy", In *Proc. ACM PODS 2006*.

C-5. *Ashwin Machanavajjhala*, Johannes Gehrke, Daniel Kifer, Muthuramakrishnan Venkitasubramaniam, "ℓ-Diversity: Privacy beyond k-Anonymity", In *Proc. ICDE 2006*.

C-6. *M. V. N. Ashwin Kumar*, Pranava. R. Goundan, K. Srinathan, C. Pandu Rangan, "On Perfectly Secure Communication over Arbitrary Networks", In *Proc. ACM PODC 2002*.

C-7. K. Srinathan, *M. V. N. Ashwin Kumar*, C. Pandu Rangan, "Asynchronous Secure Communication tolerating Mixed Adversaries", In *Proc. ASIACRYPT 2002*, LNCS, Springer Verlag.

C-8. *M. V. N. Ashwin Kumar*, K. Srinathan, C. Pandu Rangan, "Asynchronous Perfectly Secure Computation tolerating Generalized Adversaries", In *Proc. ACISP 2002*, LNCS, Springer Verlag.

C-9. Pranava. R. Goundan, K. Srinathan, *M. V. N. Ashwin Kumar*, R. Nandakumar, C. Pandu Rangan, "Theory of Equal-Flows in Networks", In *Proc. COCOON-2002*, LNCS, Springer Verlag.

Workshop, Posters and Demos.

O-1. Muthuramakrishnan Venkitasubramaniam, *Ashwin Machanavajjhala*, David Martin, Johannes Gehrke, "Trusted CVS", In *ICDE Workshops - STD3S 2006*.

O-2. *Ashwin Machanavajjhala*, Daniel Kifer, Johannes Gehrke, "Beyond k-Anonymity: New Schemes for Privacy Preserving Data Publishing", (Poster Paper) **Best Visionary Poster** In DB/IR Day, April, 2005.

O-3. Adina Crainiceanu, Prakash Linga, *Ashwin Machanavajjhala*, Johannes Gehrke, Jayavel Shanmugasundaram, "An Indexing Framework for P2P Systems", (Demo Paper) In *Proc. ACM SIGMOD 2004*.

O-4. Adina Crainiceanu, Prakash Linga, *Ashwin Machanavajjhala*, Johannes Gehrke, Jayavel Shanmugasundaram, "A Storage and Indexing Framework for P2P Systems", (Poster Paper) In *Proc. WWW 2004*.

In Submission: Three papers for *SIGMOD 2008* (titles suppressed due to double-blind reviews).

In Preparation: *Ashwin Machanavajjhala*, Johannes Gehrke, "Randomization Methods for ensuring Data Privacy", Encyclopedia of Database Systems, Springer, 2008.

Patents Filed

P-1. Erik Vee, Minos Garofalakis, Jayavel Shanmugasundaram, *Ashwin Machanavajjhala*, (title suppressed), US Patent (Y02628US00) filed by Yahoo!.

P-2. Srujana Merugu, Philip Bohannon, Pedro deRose, *Ashwin Machanavajjhala*, (title suppressed), US Patent (Y02634US00) filed by Yahoo!.

References

Prof. Johannes Gehrke
Associate Professor
Department of Computer Science
Cornell University
Ithaca NY 14853
Email: johannes@cs.cornell.edu

Dr. Jayavel Shanmugasundaram
Senior Research Scientist
Yahoo! Research
Santa Clara CA 95054
Email: jaishan@yahoo-inc.com

Prof. John M. Abowd
Professor
Department of Labor Economics
Cornell University, ILR School
Ithaca NY 14853
Email: john.abowd@cornell.edu

Dr. Raghu Ramakrishnan
Chief Scientist for Audience, and
Research Fellow
Yahoo! Research
Santa Clara CA 95054
Email: ramakris@yahoo-inc.com

Dr. Ramakrishnan Srikant
Senior Staff Research Scientist
Google
Mountain View CA 94043
Email: srikant@google.com