Cryptography & Data Privacy
Research in the NSRC

Adam Smith

Assistant Professor
Computer Science and Engineering
Algorithms & Complexity

• Research on the theoretical foundations of computer science
  ➢ Algorithm design
  ➢ Complexity theory and lower bounds
  ➢ Cryptography and information theory
  ➢ Combinatorics and discrete mathematics

• Collaboration with other research groups
  ➢ Give: abstractions, modeling, applications of algorithmic techniques
  ➢ Get: new theoretical, mathematical challenges
Algorithms & Complexity: Faculty

- **Piotr Berman**  
  ➢ Combinatorial optimization, computational biology

- **Martin Fürer**  
  ➢ Complexity theory, combinatorics

- **Sean Hallgren (Sep. ‘07)**  
  ➢ Quantum computing, computational complexity

- **Sofya Raskhodnikova (Jan. ‘07)**  
  ➢ Sublinear algorithms, complexity, data privacy

- **Adam Smith (Jan. ‘07)**  
  ➢ Cryptography, data privacy, quantum information
Cryptography & Data Privacy

• Foundations of cryptographic protocols
  - Efficient Protocols for Multi-party Computations
  - Impossibility Results and Stronger Protocols for Deniable Authentication
    [Dodis, Katz, S., Walfish, in progress]

• Key Extraction from Noisy Secrets
  - biometrics, voiceprint [Eurocrypt 04/05, STOC 05, Crypto 2006]

• Quantum cryptography
  - Understanding how recent technology impacts security and deniability
    [STOC 02, FOCS 02, Eurocrypt 05, STOC 06]

• Privacy in Statistical Databases
Cryptography & Data Privacy

• Foundations of cryptographic protocols
  ➢ Efficient Protocols for Multi-party Computations
  ➢ Impossibility Results and Stronger Protocols for Deniable Authentication
    [Dodis, Katz, S., Walfish, in progress]

• Key Extraction from Noisy Secrets
  ➢ biometrics, voiceprint [Eurocrypt 04/05, STOC 05, Crypto 2006]

• Quantum cryptography
  ➢ Understanding how recent technology impacts security and deniability
    [STOC 02, FOCS 02, Eurocrypt 05, STOC 06]

• Privacy in Statistical Databases
Privacy in Statistical Databases

Large collections of personal information
- census data
- medical/public health data
- social networks
- recommendation systems
- trace data: search records, etc
- intrusion-detection systems
Privacy in Statistical Databases

- **Two conflicting goals**
  - **Utility**: Users can extract “global” statistics
  - **Privacy**: Individual information stays hidden
Our Work

• Unify approaches from disparate fields
  ➢ statistics, data mining, database theory, cryptography,...

• Rigorous formulations of “privacy”
  ➢ Want provable guarantees that sensitive info. is not leaked
  ➢ Should be secure against arbitrary side information

• New protocols / techniques [TCC’06, STOC’07,...]

• New attacks [Ganta and S., in progress]

• Supported by
New Attacks: “Composition” Attacks

- **Example:** two hospitals serve overlapping populations
  - What if they independently release “anonymized” statistics?
- **Composition attack:** Combine independent releases
  - popular schemes leak lots of information [Ganta and S.]
  - Certain categories of schemes doomed to vulnerability
  - Not necessary...
New Protocols \[\text{TCC '06, STOC '07}\]

- **New notion:** Differential privacy
  - Roughly: any single individuals’ data does not affect the release significantly
- **Robust against very strong attacks**
  - Correlation with arbitrary outside data collections
  - Composition attacks
- **Practical**
  - Common data mining algorithms can be modified to be D.P.
  - Apply current statistical methodology almost “as is”
Summary

• **Foundations of Cryptography**
  - Efficient protocols
  - Basic (im)possibility questions

• **Privacy in Statistical Databases**
  - New protocols, new attacks

• **In progress (we’re new!):**
  - Integration, collaboration with other SIIS projects
Thank you

SIIS Lab:

http://siis.cse.psu.edu

A & C group:

http://www.cse.psu.edu/theory

Me:

http://www.cse.psu.edu/~asmith