The Fundamental Limits of Data Privacy

Peter Kairouz
ECE Department
University of Illinois at Urbana-Champaign
Communication

- transfer of information from one point in space-time to the other
Wireless communication

- the fundamental limits of wireless communication are well understood
Unprecedented level of connectivity
We’re being watched!
Recent data privacy leaks
de-anonymizing Netflix data, identifying personal genomes
Global privacy model

Users → Data → Trusted Curator → Released info → Malicious Analyst

[Dwork 2006]
Global privacy model

Users → Data → Trusted Curator → Released info → Malicious Analyst

Other Movies You Might Enjoy

- Add Amelie
- Add Y Tu Mama También
- Add Guys and Balls
- Add Mostly Martha
- Add Only Human
- Add Russian Dolls

Eiken has been added to your Queue at position 2.
This movie is available now.
Move To Top Of My Queue

< Continue Browsing Visit your Queue >
Global privacy model

NiH
National Institutes of Health
Local privacy model

[Duchi 2012]
Local privacy model

Users → Malicious Analyst → Users

Google

facebook

Apple

Microsoft
Local privacy model

Users → Malicious Analyst → Users

Car rental
Economy 2 door sedan
Hertz rental car reservation
Name: Mr. John Smith
Booking Number: E12345678
Thu, 18 Apr, 2013 11:40
Fri, 26 Apr, 2013 21:50
Hertz San Diego
697 Harbor Dr, San Diego, CA 92101

- Get directions
- Manage reservation
- View email

Next Appointment
Agency Meeting
11:30 AM
Ninth Ave, New York, NY 10031

- Email guests

Flights
Delta Air Lines flight 8772
from DeltaAirLines@delta.com
Status: Scheduled / Fri, Nov 29, 2013
Depart San Francisco International
SFO 11:45 PM
Terminal 1

- Chromecast
  Shipped: 1 min ago
  1 item from
  Amazon.com
  Estimated arrival
  Tuesday, May 13

- Get directions
- Manage reservation
- View email

Restaurant Reservations
Broder
2508 SE Clinton St, Portland, OR 97202
Reservation in 1 hour
Travel time walking 45 minutes

- Get directions
- Manage reservation
- View email

Weather
SCATTERED
5mph
80%

TUE
THU
FRI

68° 67° 65° 57°
48° 44° 48° 46°

- Navigate / 57 mins via US - 101

Hotels
The Connaught Hotel
Carlos Place, Mayfair, London W1K 2AL, United Kingdom
Check in from 12:00pm today

- Call
- Hotel information
- View email

Friends' Birthdays
Conrad

- View email
Differential privacy

- $Q$ is a privacy mechanism
- privacy enforced by imposing **differential privacy** parametrized by $\varepsilon$
Differential privacy

$\varepsilon$ controls the level of privacy
- Large $\varepsilon$, low privacy
- Small $\varepsilon$, high privacy
Global Privacy Model
The Laplace mechanism

Laplace Mechanism
What would Shannon do?

- there is a **fundamental tradeoff** between **privacy** and **utility**
Data independent noise is optimal
Staircase mechanisms are optimal

Staircase Mechanism

\[ f_N \]
Staircase mechanisms are optimal

Differential privacy
C Dwork - Automata, languages and programming, 2006 - Springer
Abstract In 1977 Dalenius articulated a desideratum for statistical databases: nothing about an individual should be learnable from the database that cannot be learned without access to the database. We give a general impossibility result showing that a formalization of ...
Cited by 1744 Related articles All 22 versions Web of Science: 293 Cite Save

Differential privacy: A survey of results
C Dwork - Theory and applications of models of computation, 2008 - Springer
Abstract Over the past five years a new approach to privacy-preserving data analysis has born fruit [13, 18, 7, 19, 5, 37, 35, 8, 32]. This approach differs from much (but not all!) of the related literature in the statistics, databases, theory, and cryptography communities, in that ...
Cited by 749 Related articles All 24 versions Cite Save

Mechanism design via differential privacy
F McSherry, K Talwar - ... of Computer Science, 2007. FOCS'07. ..., 2007 - ieeexplore.ieee.org
Abstract We study the role that privacy-preserving algorithms, which prevent the leakage of specific information about participants, can play in the design of mechanisms for strategic agents, which must encourage players to honestly report information. Specifically, we ...
Cited by 573 Related articles All 24 versions Cite Save

Differential privacy via wavelet transforms
X Xiao, G Wang, J Gehrke - Knowledge and Data Engineering, ..., 2011 - ieeexplore.ieee.org
Abstract—Privacy preserving data publishing has attracted considerable research interest in recent years. Among the existing solutions, e-differential privacy provides the strongest privacy guarantee. Existing data publishing methods that achieve e-differential privacy, ...
Local Privacy Model
Local privacy

have you ever used illegal drugs?
Local privacy

Users → $Q$ → Malicious Analyst ← $Q$ ← Users

have you ever used illegal drugs?

say yes

answer truthfully

[Warner 1965]
What would Shannon do?

- there is a fundamental tradeoff between privacy and utility
Main result: binary data

- for binary data:

  lie w.p. \( \frac{1}{e^\epsilon + 1} \)

  say the truth w.p. \( \frac{e^\epsilon}{e^\epsilon + 1} \)

- optimal for all utilities obeying the data processing inequality

\[ JMLR \ 2015, \ NIPS 14 \]
Main result: general data

- for general k-ary data:

Binary Mechanism
Randomized Response
staircase mechanisms
all differentially private mechanisms

[NIPS 14, JMLR 15]
Randomized Response

\[ Y = \begin{cases} 
1 & \frac{e^\epsilon}{e^\epsilon + 3} \\
2 & \frac{e^\epsilon}{e^\epsilon + 3} \quad \frac{1}{e^\epsilon + 3} \\
3 & \frac{1}{e^\epsilon + 3} \quad \frac{e^\epsilon}{e^\epsilon + 3} \\
4 & \frac{e^\epsilon}{e^\epsilon + 3}
\end{cases} \]

- optimal in the low privacy regime
Binary mechanism

\[ Y = \begin{cases} 1 & \frac{e^\varepsilon}{e^\varepsilon + 1} \\ \frac{1}{e^\varepsilon + 1} & \end{cases} \]

\[ X = \begin{cases} 1 & \frac{1}{e^\varepsilon + 1} \\ \frac{e^\varepsilon}{e^\varepsilon + 1} & \end{cases} \]

- optimal in the **high privacy regime**
Acknowledgments

Sewoong Oh

Pramod Viswanath

Quan Geng