Privacy Research

A categorization

- Three types of privacy papers:
  - Attack:
    - “A system that was previously thought to be private/secure has an unexpected privacy vulnerability”
    - Includes papers improving/expanding on such a vulnerability
  - Design:
    - “We create a system to do something in a private way”
    - Biggest category of papers
  - Human:
    - “How do humans interact with privacy-enhancing/compromising technologies?”
    - Focuses on surveys/studies

PETS 2022.1 Attack papers

“A system that was previously thought to be private/secure has an unexpected privacy vulnerability”

Voice recording

Stylometry

Comprehensive measurement of web tracking

Bitcoin usage

Membership inference attacks

Tor Guard discovery

Data storage on mobile devices

PETS 2022.1 Design papers

"We create a system to do something in a private way"

Tor load balancing

Differential privacy (query: partition selection)

Facial recognition

Differential privacy (data collection)

Differential privacy (deep learning)

Zero-knowledge

Collaborative learning

Automated data protection policy verification

Power analysis of machine learning

Differential privacy for blockchains

Blockchain

FairSwap

SMPC histogram

Range-searchable symmetric encryption

Low-latency SMPC

Private set intersection

PETS 2022.1 Human papers

"How do humans interact with privacy-enhancing/compromising technologies?"

Covid certificates

IFTTT applets

Gmail's confidential mode

Social network privacy options

Tracking defenses

CCPA website compliance

Current directions not covered

- Machine learning & Privacy
  - Can we perform federated learning in a private manner?
    - The trained model doesn’t rely on any one person’s data too much
    - This can also be good against adversaries
  - Are trained models vulnerable to adversarial learning?
    - i.e. Does there exist a small perturbation that changes the class?
    - If the model were privacy-protecting, this would not happen
  - Does machine learning still work if the data is privatized?
    - What happens if we try to learn on k-anonymized data?

Current directions not covered

- IoTs
  - How are IoTs vulnerable?
    - Side channel attacks (traffic analysis, power analysis, etc.) to compromise privacy
    - Security issues (hijacking, confidentiality)
  - How do humans interact with IoTs?
  - How do app developers collect data?

Dealing with the exam

- “How can multiple choice questions be this hard?”
  - Often one or two wrong choices will “sound” correct, using familiar terminology in an incorrect way
- Understand each component of a design and how they work together
  - “What would happen if I removed or changed this?”
  - What are the underlying vulnerabilities to be solved?
- Questions will usually test “why”, not “what”
- Most advanced questions: combining two unrelated concepts together