CSE 410 Fall 2025
Privacy-Enhancing Technologies

Marina Blanton

Department of Computer Science and Engineering
University at Buffalo

Lecture 9: Protecting Data during Computation

Computing with Private Data

Computing on private data can be accomplished if the data does not leave the premises of where it was collected

- analysis of one's medical history collected by a doctor's office
- internally evaluating a company's performance
- computing a student's GPA

Computing with Private Data

Computing on private data can be accomplished if the data does not leave the premises of where it was collected

- analysis of one's medical history collected by a doctor's office
- internally evaluating a company's performance
- computing a student's GPA

This holds even if the data is a subject of legal restrictions

- legal restrictions often control data sharing
- they can also regulate protection of data at rest

Computing with Private Data

Computing on private data can be accomplished if the data does not leave the premises of where it was collected

- analysis of one's medical history collected by a doctor's office
- internally evaluating a company's performance
- computing a student's GPA

This holds even if the data is a subject of legal restrictions

- legal restrictions often control data sharing
- they can also regulate protection of data at rest

Privacy-preserving computation (or secure computation) refers to computing on private data across different trust domains

Computing with Private Data

Computing across different trust domains can take many forms

- evaluating predisposition to a genetic disease
  - a patient has a DNA, a service provider has genetic pattern

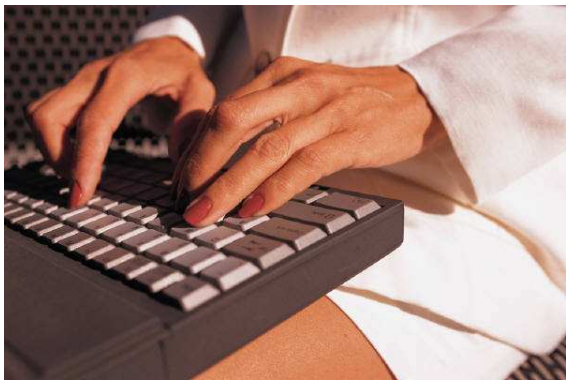
- determining the best treatment for a rare condition by multiple hospitals
  - hospitals cannot easily share patients' data

- computation outsourcing to a cloud provider
  - offloading an expensive task, e.g.., pharmaceutical evaluation of new drug

The Famous Millionaires Problem

- Alice and Bob would like to determine who is richer without revealing their worth to each other

Alice
private x



$\longrightarrow$
$\longleftarrow$
$\longrightarrow$
$\longleftarrow$
$\longrightarrow$

Bob
private y

output $x < y$

Secure Computation

Secure multi-party computation refers to the ability of multiple individuals to evaluate a function on their respective private inputs without disclosing them

- Each participant holds their own private data
- Participants jointly perform the computation on cryptographically protected private data
- Private data can only leave the owner after applying proper cryptographic protection
- Only the outcome is revealed to the intended output recipients

Secure Computation

- This is very different from traditional ways of computing:
  - obtain access to private data and promise to comply with data usage requirements
  - or private data is accessible to software, but the software doesn't let users to "see" the data

Secure Computation

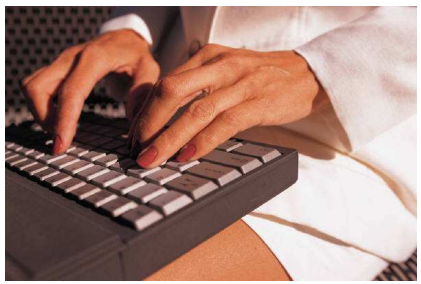
- This is very different from traditional ways of computing:
  - obtain access to private data and promise to comply with data usage requirements
  - or private data is accessible to software, but the software doesn't let users to "see" the data
- Contrast this with secure computation:
  - nothing about private data is recoverable by others throughout the computation
  - observed information is the same as if the computation was performed by a trusted third party

Secure Computation

The security expectations are as if a trusted third party performed the computation and handed the result to the participants

$b = (x < y)$

Alice



input x, output b

Bob

input y, output b

Secure Computation

Secure computation solutions can be categorized into 3 major types:

- those based on homomorphic encryption
- those based on secret sharing
- those based on garbled circuit evaluation

Homomorphic Encryption

Homomorphic encryption is a special type of encryption that, given ciphertexts, permits computation on the underlying plaintexts

$$\mathsf{Enc}_k(m_1) \otimes \mathsf{Enc}_k(m_2) = \mathsf{Enc}_k(m_1 \oplus m_2)$$

Contrast this with conventional encryption we previously considered

- changes to a ciphertext often garble the data

Additional properties of homomorphic encryption restrict the type of arithmetic that can be used

Homomorphic Encryption

Different types of homomorphic encryption are known:

- partially homomorphic encryption
- fully homomorphic encryption

Partially homomorphic encryption

- supports a single operation on ciphertexts

Homomorphic Encryption

Different types of homomorphic encryption are known:

- partially homomorphic encryption
- fully homomorphic encryption

Partially homomorphic encryption

- supports a single operation on ciphertexts
- additively homomorphic encryption
  $\mathsf{Enc}_k(m_1) \cdot \mathsf{Enc}_k(m_2) = \mathsf{Enc}_k(m_1 + m_2)$
- multiplicatively homomorphic encryption
  $\mathsf{Enc}_k(m_1) \cdot \mathsf{Enc}_k(m_2) = \mathsf{Enc}_k(m_1 \cdot m_2)$
- intuition

Homomorphic Encryption

Fully homomorphic encryption (FHE)

- supports two operations on ciphertexts: addition and multiplication
- allows for any functionality to be evaluated on encrypted data
- this representation is called an arithmetic circuit

Homomorphic encryption enables computation on encrypted data and results in efficient protocols for certain problems

- it is well suited for secure computation outsourcing

Examples

Secret Sharing

Another way to compute on private values is by splitting them into multiple shares

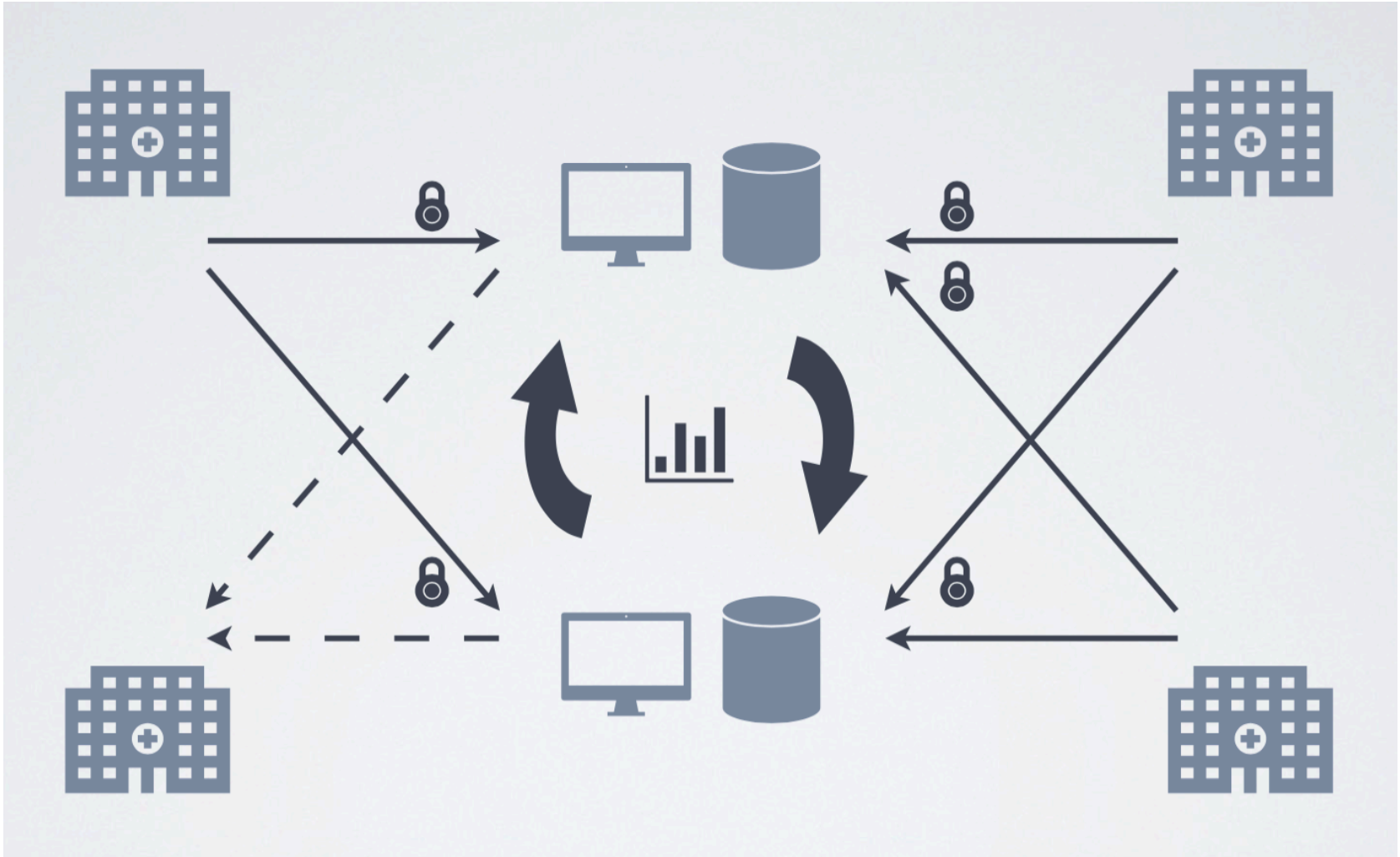
- this is called secret sharing
- given value s, generate shares s_1, s_2, ..., s_n
- each s_i is stored in a different place and doesn't reveal the secret
- access to enough shares allows for s to be reconstructed, but individual shares don't reveal anything
- specifically, (n, t) threshold secret sharing means
  - a secret s is divided into n shares
  - access to $\leq t$ shares reveals nothing about s
  - access to t shares allows for s to be reconstructed

Computation Using Secret Sharing

A number of participants would like to perform joint computation on their private inputs

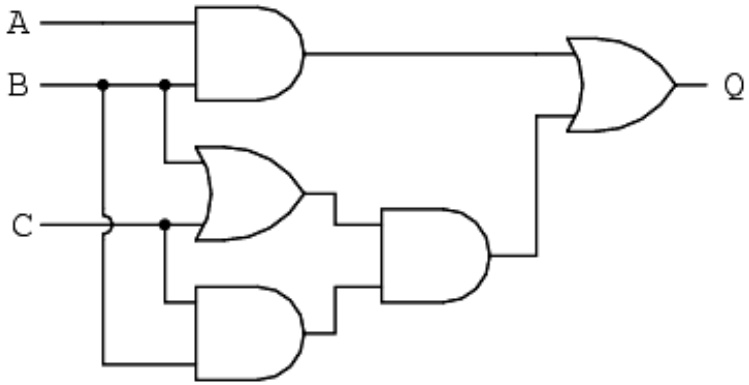
- secret sharing: each input owner creates shares of its private inputs and communicates a share to each party running the computation
- secure computation: computation parties evaluate the function on shares one operation at a time
- once the result is computed, shares are communicated to the output recipients
- output reconstruction: each output recipient reconstruct its output from the received shares

Computation Using Secret Sharing



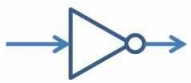
Garbled Circuit Evaluation

With garbled circuit evaluation, the computation is represented as Boolean circuit



NOT

x	F
0	1
1	0



AND

x	y	F
0	0	0
0	1	0
1	0	0
1	1	1

OR

x	y	F
0	0	0
0	1	1
1	0	1
1	1	1

XOR

x	y	F
0	0	0
0	1	1
1	0	1
1	1	0

Garbled Circuit Evaluation

The computation is performed by two parties:

- one party plays the role of a circuit garbler
- the other party is the circuit evaluator

The circuit garbler associates a random label with each value of each wire

The circuit evaluator evaluates the circuit on private inputs without knowing the meaning of values it handles

The idea is to decouple evaluation from its meaning

Secure Multi-Party Computation

Today is prime time for secure computation

- speed and abilities of secure computation techniques have improved dramatically

Secure Multi-Party Computation

Today is prime time for secure computation

- speed and abilities of secure computation techniques have improved dramatically
- a variety of tools have been developed

Secure Multi-Party Computation

Today is prime time for secure computation

- speed and abilities of secure computation techniques have improved dramatically
- a variety of tools have been developed
- a number of companies now offer this as a product

Real-World Uses

Danish sugar beet auction

- country-wide sealed bid auction that cryptographically protected farmer's bids

Tax fraud detection in Estonia

- citizens pay taxes on goods acquired abroad
- the government can compute taxes without access to purchases

Correlation between student employment and college graduation

- employment information was extracted from tax records
- it was combined with university records

Real-World Uses

City of Boston gender wage gap study

- analysis of wages by gender and race in the Greater Boston area
- significantly larger participation the second year

Effectiveness of Google ads

- Google knows on what ads a user clicked
- the store knows how much a user spent at the store
- this allows for determining effectiveness of ads

Privacy-Preserving Data Analytics

Data collection and analysis are all around us

- this allows for personalized medicine, targeted advertisement, ...
- this also brings unintended data leakage and disclosure

Privacy-Preserving Data Analytics

Data collection and analysis are all around us

- this allows for personalized medicine, targeted advertisement, ...
- this also brings unintended data leakage and disclosure

Privacy-preserving data mining and machine learning are receiving a significant amount of attention

- data can be distributed across multiple entities or organizations
- a model can be trained by one party, but queried by others
- the task of building a model can be outsourced

Summary

Computing on private data requires protection when the data resides in different trust domains

Secure multi-party computation is making rapid progress and is entering our lives

- it is an additional mechanism to keep our data protected
- it allows for our data to be used without disclosing it