Dr. Roberto Yus
https://robertoyus.com/

Data Privacy

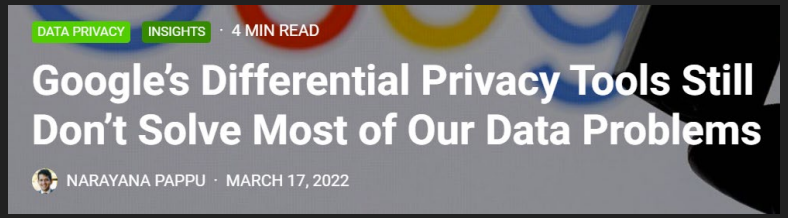
CMSC 491/691

L08 – Secure Multi-Party Computation

Icons from https://thenounproject.com/

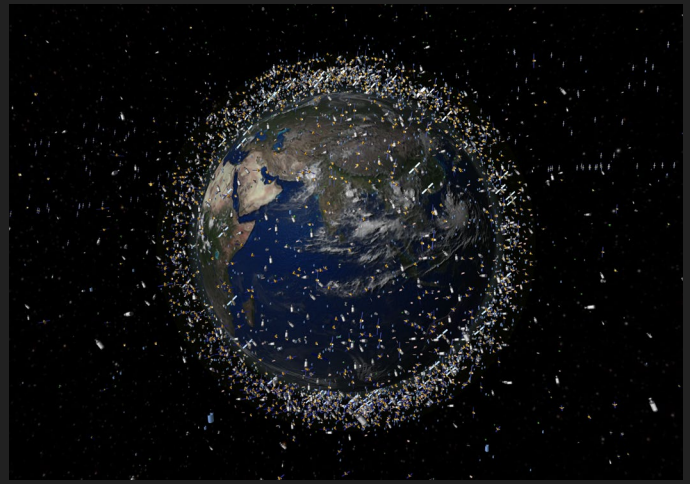
Previously on…

- Differential Privacy is current state of the art for privacy protection
- Privacy parameter (ε) to adjust the tradeoff between the level of privacy loss and data quality
- Not for all privacy problems!
  - Statistical releases
  - Works well with large amounts of data



In the news!

Privacy and Collaboration



Parties: International satellite operators

Desired output: potential collisions

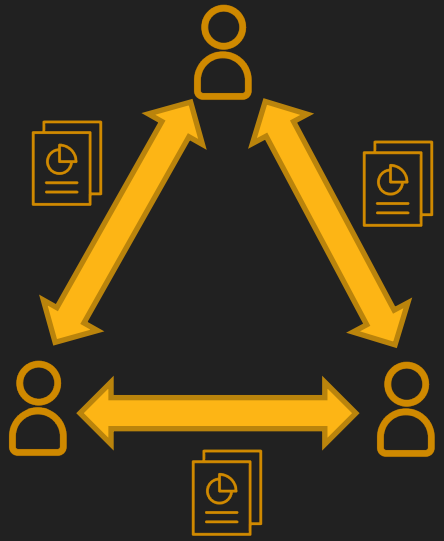
Private information: location, maneuver schedule

Other Examples

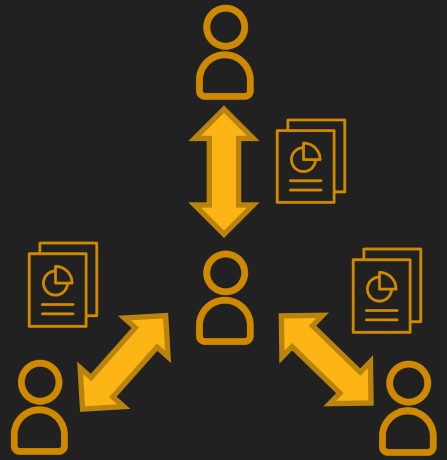
- Elections
  - N parties, each vote “yes” or “no”
  - Goal: determine whether the majority voted “Yes”, but no voter should learn how other people voted
- Auctions
  - Each bidder makes an offer
    - Offer should be committing! (can’t change it later)
  - Goal: determine whose offer won without revealing losing offers
- Distributed data mining
  - Two companies want to compare their datasets without revealing them
    - For example, compute the intersection of two lists of names
- Database privacy
  - Evaluate a query on the database without revealing the query to the database owner
  - Evaluate a statistical query on the database without revealing the values of individual entries
  - Many variations

Current Approaches to Collaboration

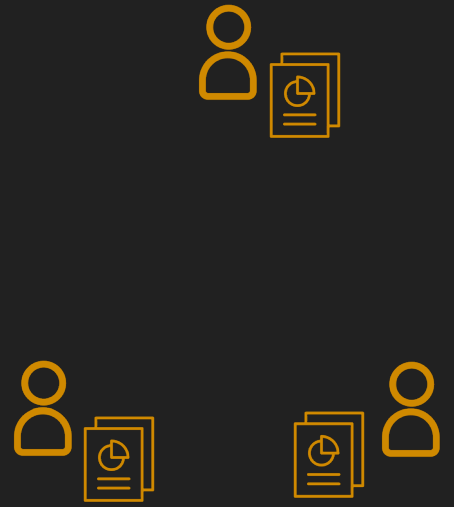
Share with each other



Share with external party



Don’t share



Collaboration requires giving data to trusted parties, accepting security and privacy risks

Secure Multi-Party Computation (MPC)

- Goal: replace trusted party with technology
- Requirements
  - **Correctness:** everyone learns correct result of computation
  - **Privacy/security:** no one learns anything beyond result
- **MPC provides correctness and security without trusted party**
  - For any computation
  - For any number of parties

Canetti, Ran, et al. "Adaptively secure multi-party computation." Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. 1996.

MPC: The First 40 Years

Shamir secret sharing

GMW

BGW

1980s: *Existence*	1990s: *Adolescence*	2000s: *Idealism*	2010s: *Pragmatism*

Yao's garbled circuits

Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

MPC: The First 40 Years

Shamir secret sharing

GMW

BGW

Beaver triples

Packed SS

1980s: *Existence*	1990s: *Adolescence*	2000s: *Idealism*	2010s: *Pragmatism*

Yao's garbled circuits

point & permute

row reduction

Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

MPC: The First 40 Years

Shamir secret sharing

GMW

BGW

Beaver triples

Packed SS

Homomorphic secret sharing

1980s: *Existence*	1990s: *Adolescence*	2000s: *Idealism*	2010s: *Pragmatism*

Yao's garbled circuits

point & permute

row reduction

OT extension

free XOR

Fairplay

Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

MPC: The First 40 Years

Shamir secret sharing GMW BGW Beaver triples Packed SS Homomorphic secret sharing Homomorphic Enc and MACs × via OT

1980s: *Existence*	1990s: *Adolescence*	2000s: *Idealism*	2010s: *Pragmatism*

Yao's garbled circuits

point & permute

row reduction

OT extension

free XOR

Fairplay

fleXOR

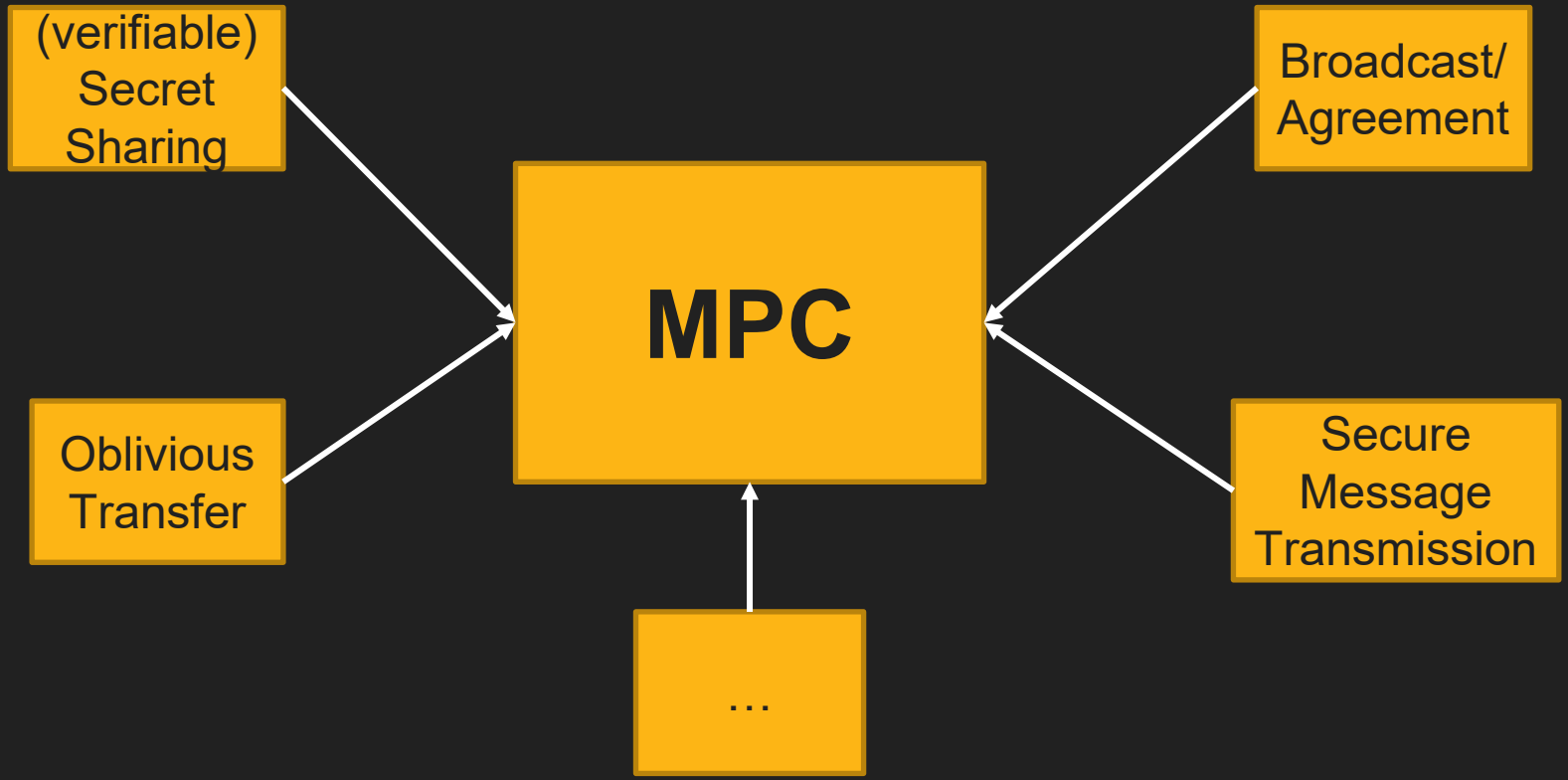
half gates

Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

Adversarial Models

- Some participants may be **dishonest** (corrupt)
  - If all were honest, we would not need secure multi-party computation
- **Semi-honest** (aka passive; honest-but-curious)
  - Follows protocol, but tries to learn more from received messages than they would learn in the ideal model
- **Malicious**
  - Deviates from the protocol in arbitrary ways, lies about his inputs, may quit at any point

Building Blocks of MPC



How MPC Works

Jane $6200

John $5800

Bobby $7300

Jack $5100

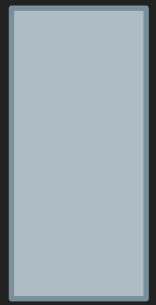
What's the average salary?

(6200 + 5800 + 7300 + 5100) / 4
= $6100 per person

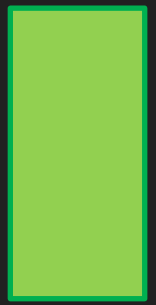
We want to know the avg salary,
but **we don't want anybody to know our salary**

How MPC Works

Jane $6200

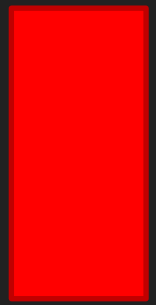


John $5800



Bobby $7300

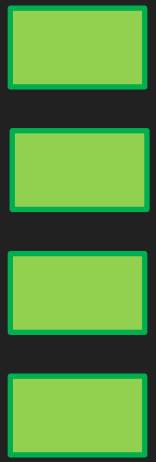
Jack $5100



How MPC Works

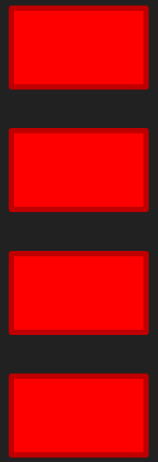
Jane $6200

John $5800

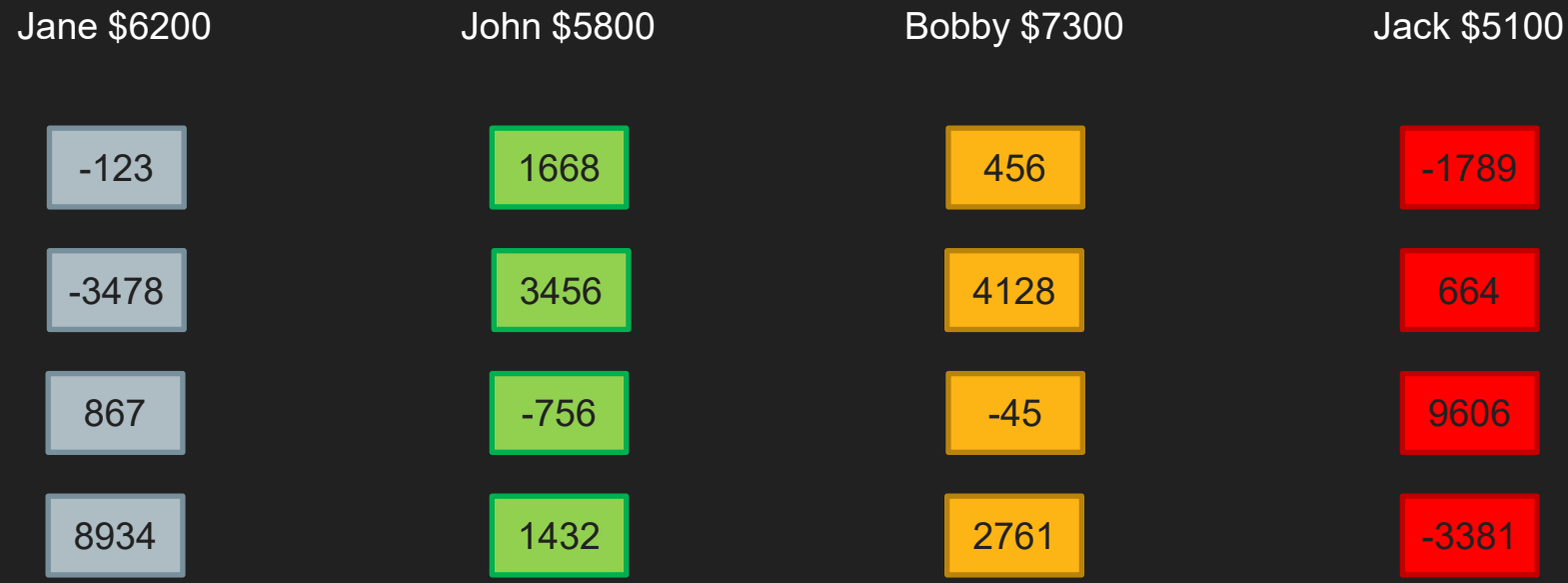


Bobby $7300

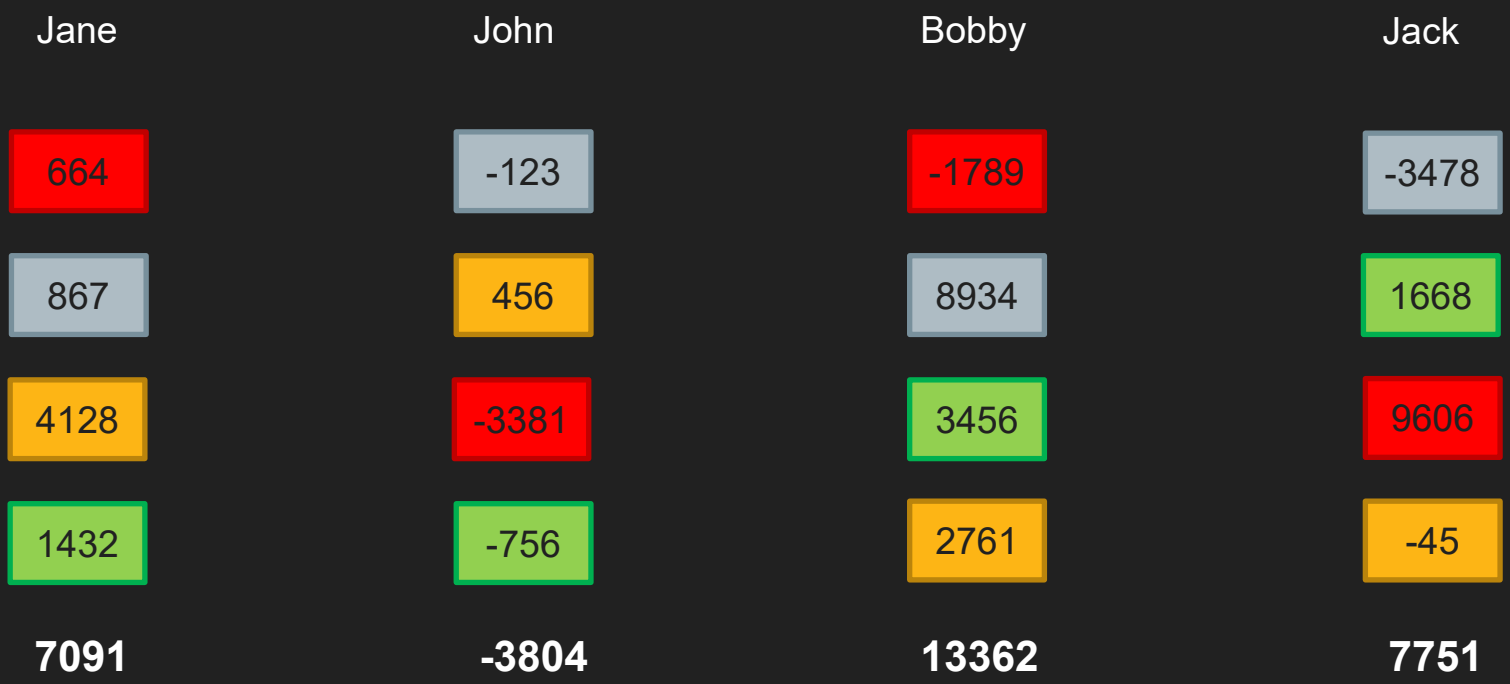
Jack $5100



How MPC Works



How MPC Works



How MPC Works

Jane $6200 7081

John $5800 -3804

Bobby $7300 13362

Jack $5100 7751

Raw data amount: 6200 + 5800 + 7300 + 5100
= 24400

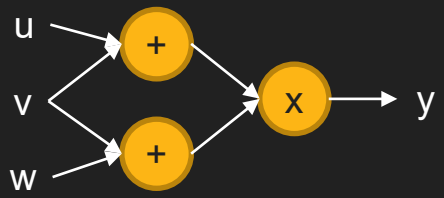
Random data amount: 7081 – 3804 + 13362 + 7751
= 24400

Average: 24400 / 4 = 6100 per person

No one learns anything beyond the result of computation

MPC for Any Function

MPC for Arithmetic Computation



MPC for add, multiply primitives over integers can securely compute any function!

MPC for Boolean Computation

MPC for XOR, AND primitives over bits can securely compute any function!

MPC can securely compute any function using arithmetic or Boolean primitives

Why Secret-Sharing?

- Encryption techniques are **computationally secure**
  - A powerful adversary can break the encryption technique
    - Google, with sufficient computational capabilities, broke SHA-1 (https://shattered.io/)

- **Information-theoretical security**
  - Secure regardless of the computational power of an adversary
  - Quantum secure

The Concept of Secret Sharing

(n, t) LOCKED BOX REPRESENTATION

A secret s

The Concept of Secret Sharing

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

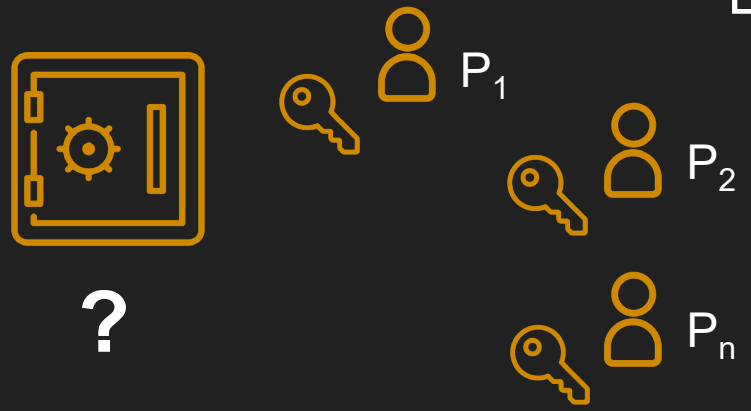
Any t parties cannot open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

Ex: t = 1



Any t parties cannot open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

?

Ex: t = 1

Any t parties cannot open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

?

Ex: t = 1

Any t parties cannot open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box

Ex: t = 1

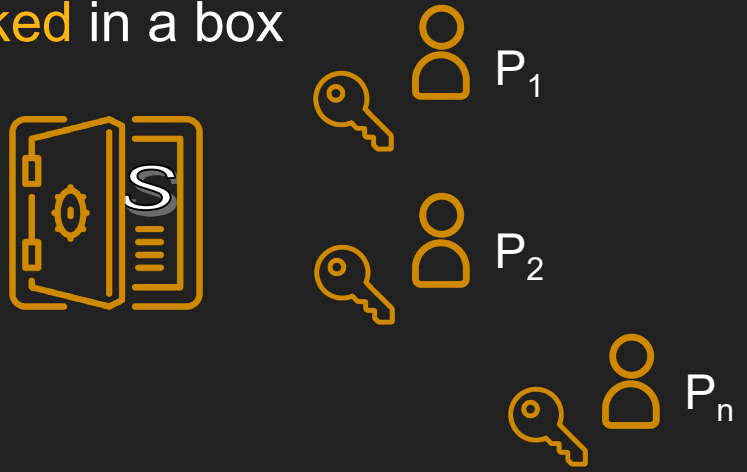
Any t parties cannot open the box

Any (t + 1) parties can open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box



Ex: t = 1

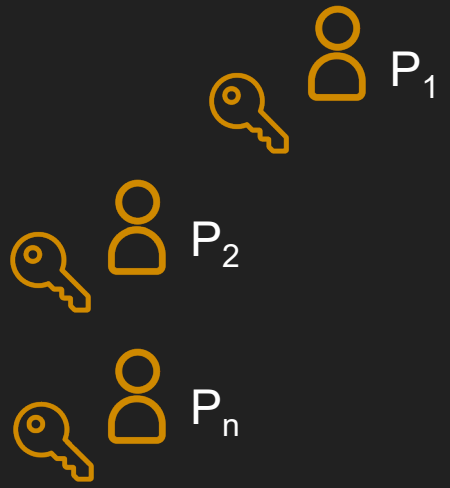
Any t parties cannot open the box

Any (t + 1) parties can open the box

Secret Sharing: Properties

(n, t) LOCKED BOX REPRESENTATION

A secret s locked in a box



Ex: t = 1

Any t parties cannot open the box

Any (t + 1) parties can open the box

Shamir's Secret-Sharing (SSS)

Secret-Share Creation

e.g., under the assumption that no server will collude

Secret S

Share 1 (s_1)

Share 2 (s_2)

Share 3 (s_3)

Share 4 (s_4)

S1

S2

S3

S4

Each server **cannot** learn the secret **S**

Secret Owner

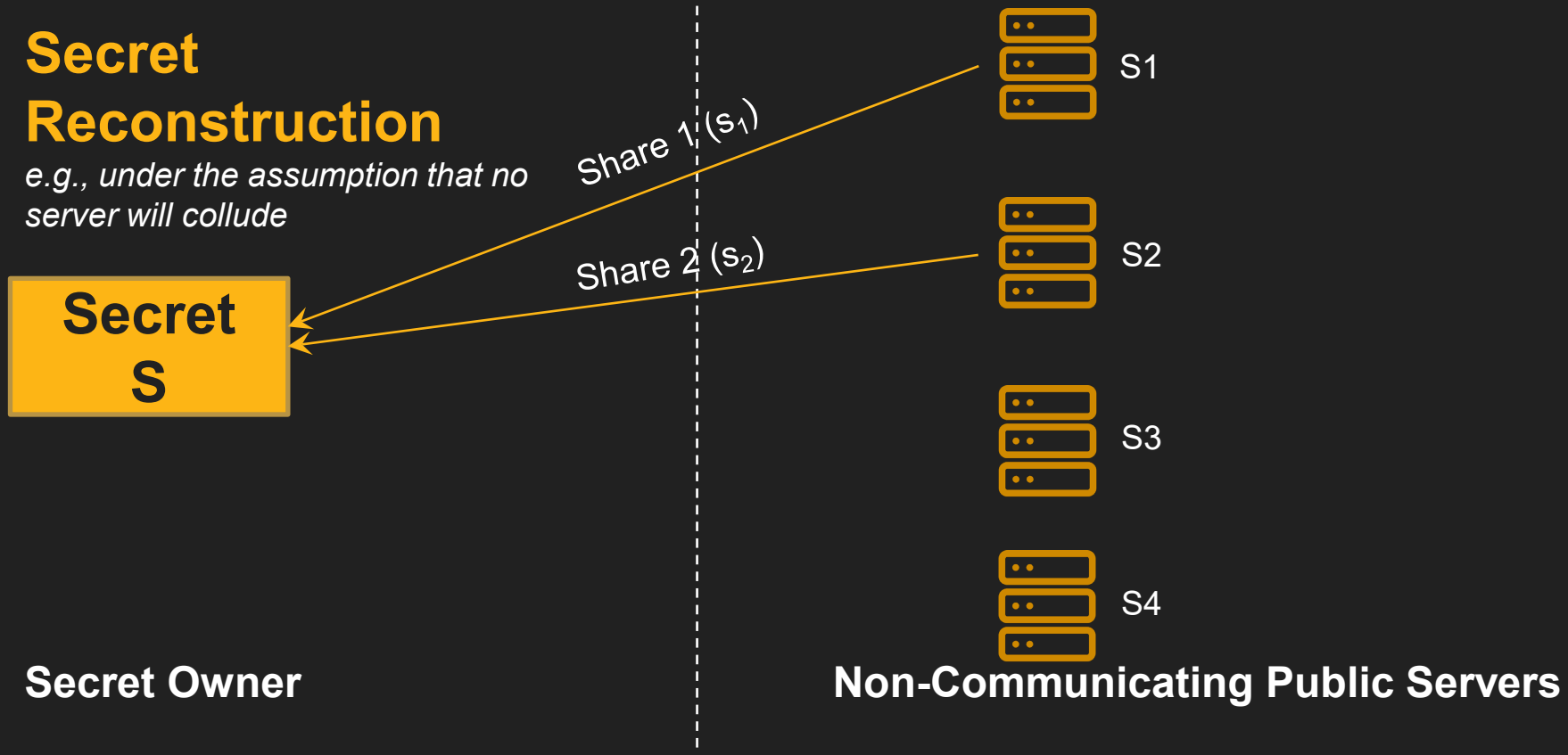
Non-Communicating Public Servers

Adi Shamir. "How to share a secret." Communications of the ACM 22, no. 11 (1979): 612-613.

Shamir's Secret-Sharing (SSS)

Secret Reconstruction

e.g., under the assumption that no server will collude

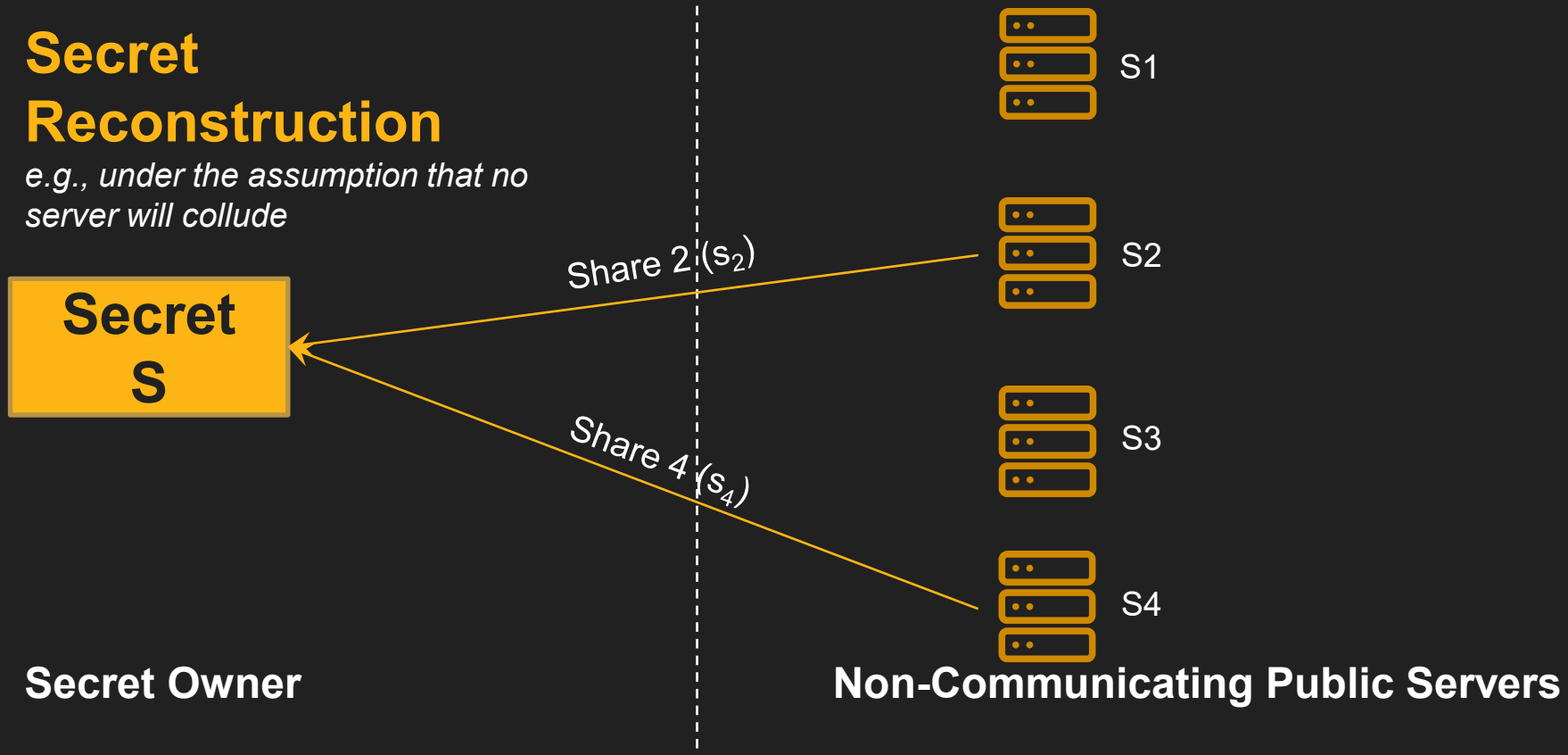


Adi Shamir. "How to share a secret." Communications of the ACM 22, no. 11 (1979): 612-613.

Shamir's Secret-Sharing (SSS)

Secret Reconstruction

e.g., under the assumption that no server will collude

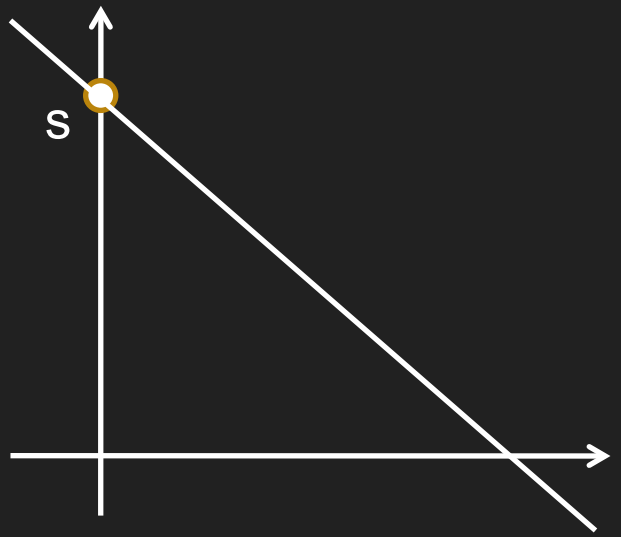


Adi Shamir. "How to share a secret." Communications of the ACM 22, no. 11 (1979): 612-613.

SSS Intuition

n = 3 and t = 1

f(0) = 5

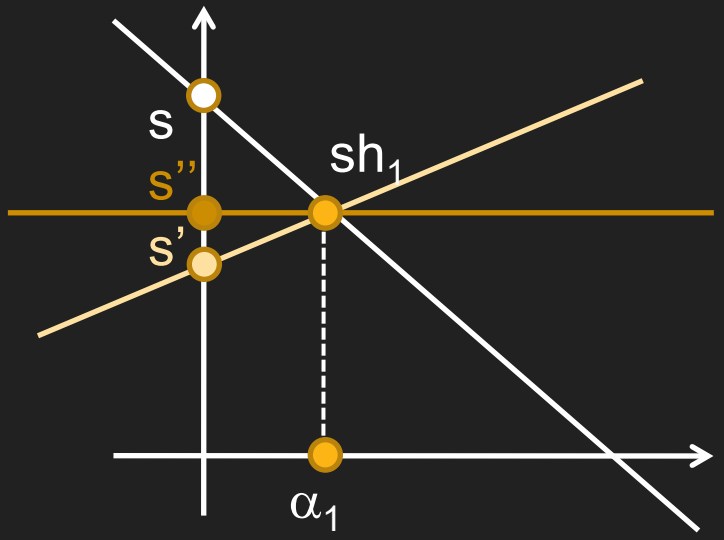


f(x) = 5 – 1x

SSS Intuition

n = 3 and t = 1

f(0) = 5



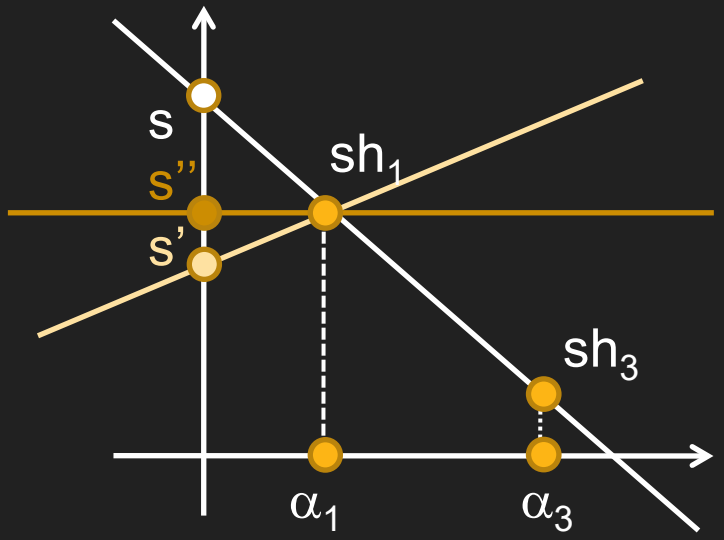
f(x) = 5 – 1x

Only 1 share → all possible straight lines over the field

SSS Intuition

n = 3 and t = 1

f(0) = 5



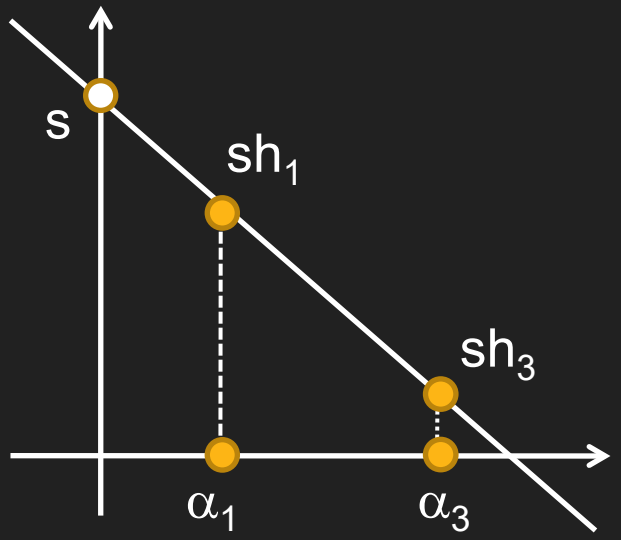
f(x) = 5 – 1x

Only 1 share → all possible straight lines over the field

SSS Intuition

n = 3 and t = 1

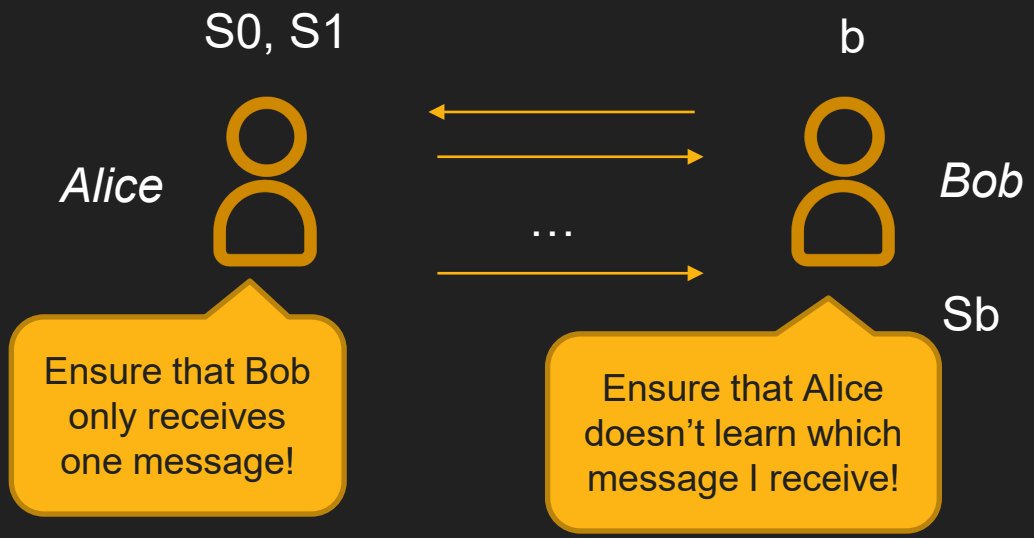
f(0) = 5



f(x) = 5 – 1x

Any set of 2 shares → original straight line and the secret

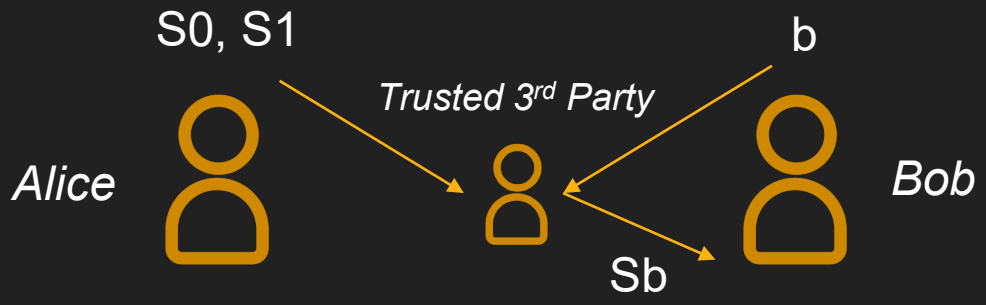
Oblivious Transfer (OT)



1-out-of-2 Oblivious Transfer

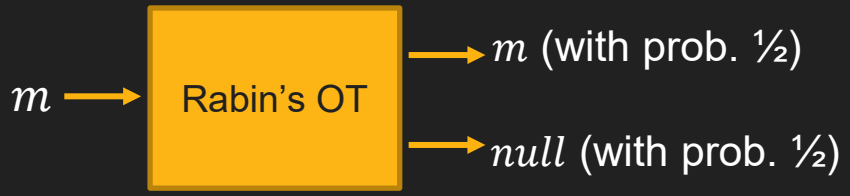
Michael O. Rabin. "How to exchange secrets by oblivious transfer" Technical Report TR-81, Aiken Computation Laboratory, Harvard University, 1981

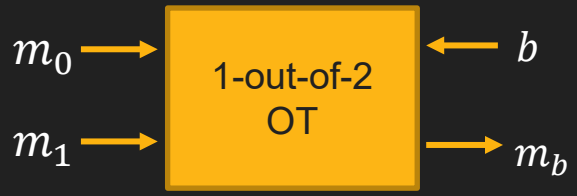
Oblivious Transfer (OT)

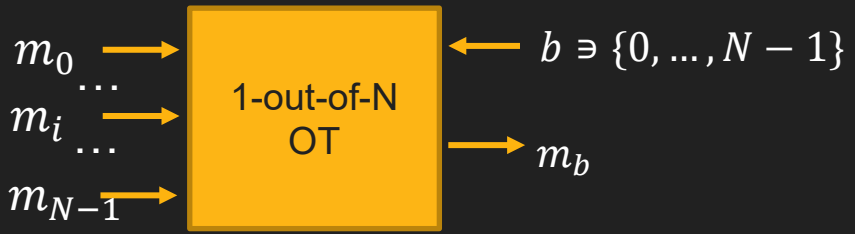


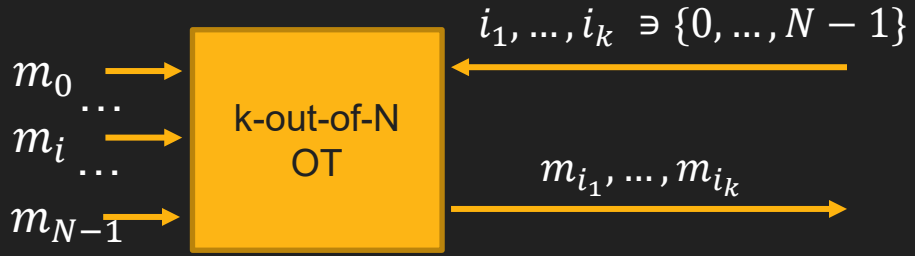
"Learn no more than what they would if they were interacting with a trusted third party"

OT Variants

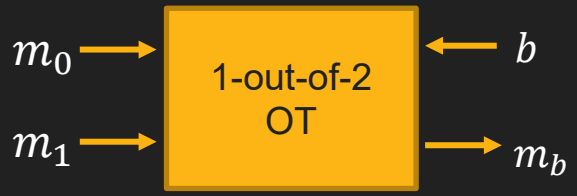

m
Rabin's OT
m (with prob. ½)
$null$ (with prob. ½)


m_0
m_1
1-out-of-2 OT
b
m_b


m_0
m_i
m_{N-1}
1-out-of-N OT
$b \ni \{0, \dots, N-1\}$
m_b


m_0
m_i
m_{N-1}
k-out-of-N OT
$i_1, \dots, i_k \ni \{0, \dots, N-1\}$
$m_{i_1}, \dots, m_{i_k}$

Using OT to Compute Operations



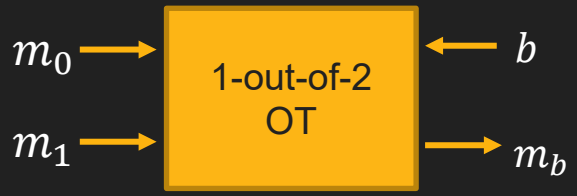
$$m_b = (1 \oplus b)m_0 \oplus bm_1$$

A	B	A ⊕ B
0	0	0
0	1	1
1	0	1
1	1	0

If $b = 0 \rightarrow m_b = m_0 \oplus 0 = m_0$

If $b = 1 \rightarrow m_b = 0 \oplus m_1 = m_1$

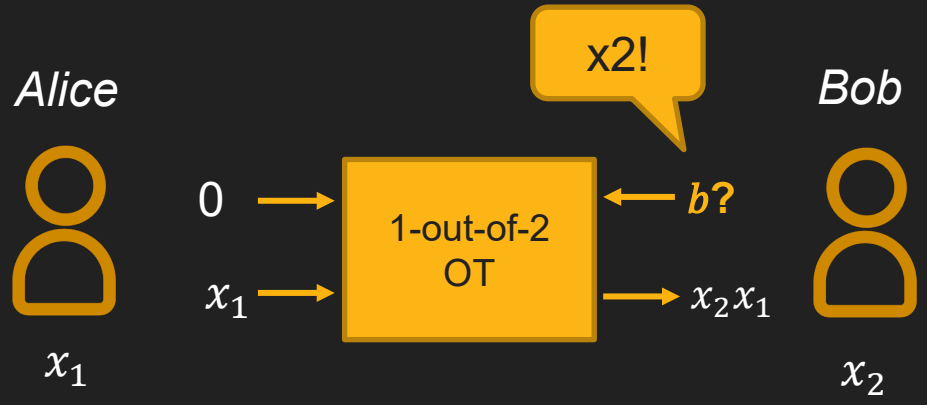
Using OT to Compute Operations



$$m_b = (1 \oplus b)m_0 \oplus bm_1$$

A	B	A ⊕ B
0	0	0
0	1	1
1	0	1
1	1	0

$$m_b = (1 \oplus x_2)0 \oplus x_2x_1 = x_2x_1$$



Compute AND operator!

How is MPC Deployed in Practice?

Input Parties

...

Computing Parties

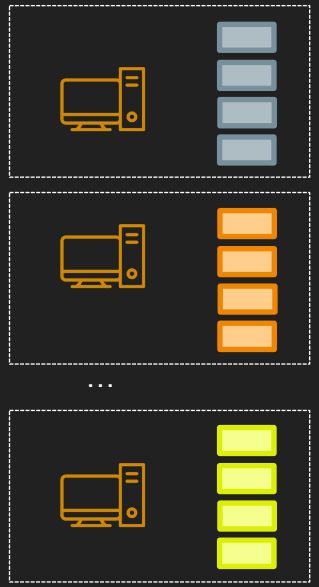
Result Parties

How is MPC Deployed in Practice?

Input Parties

Computing Parties

Result Parties

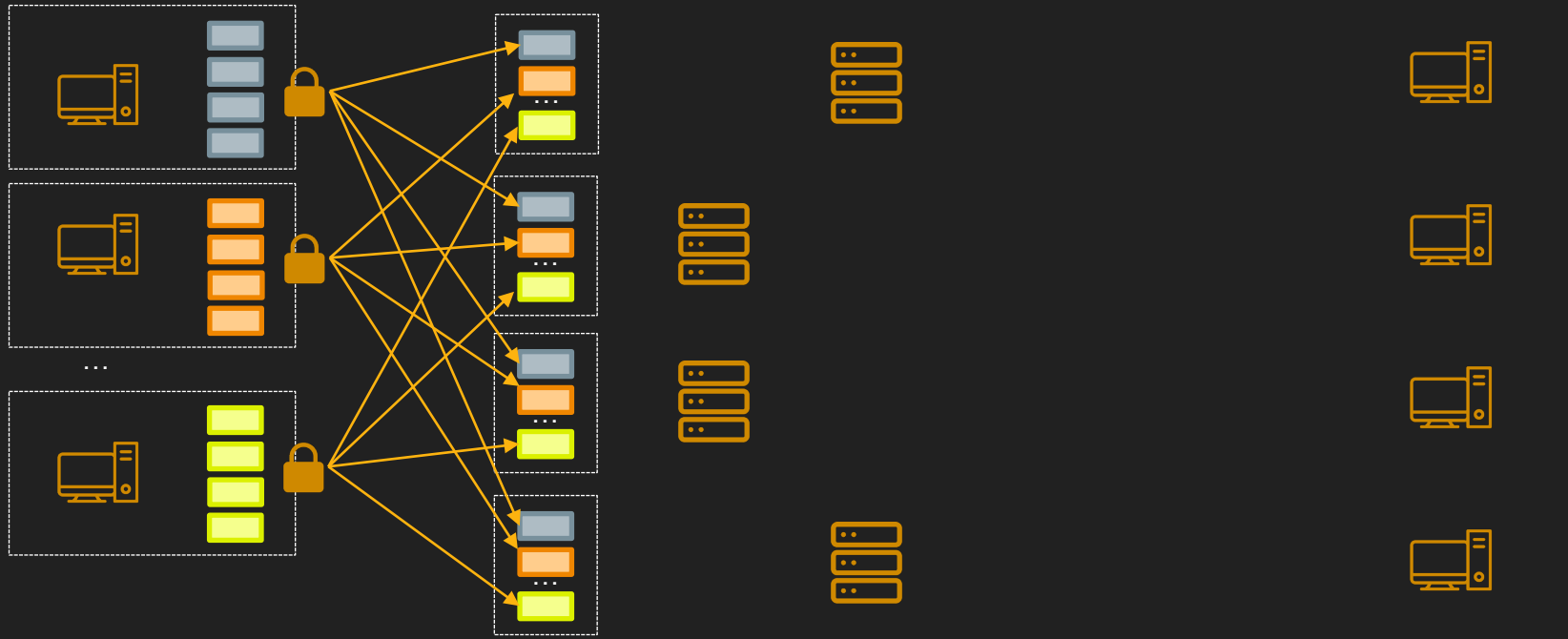


How is MPC Deployed in Practice?

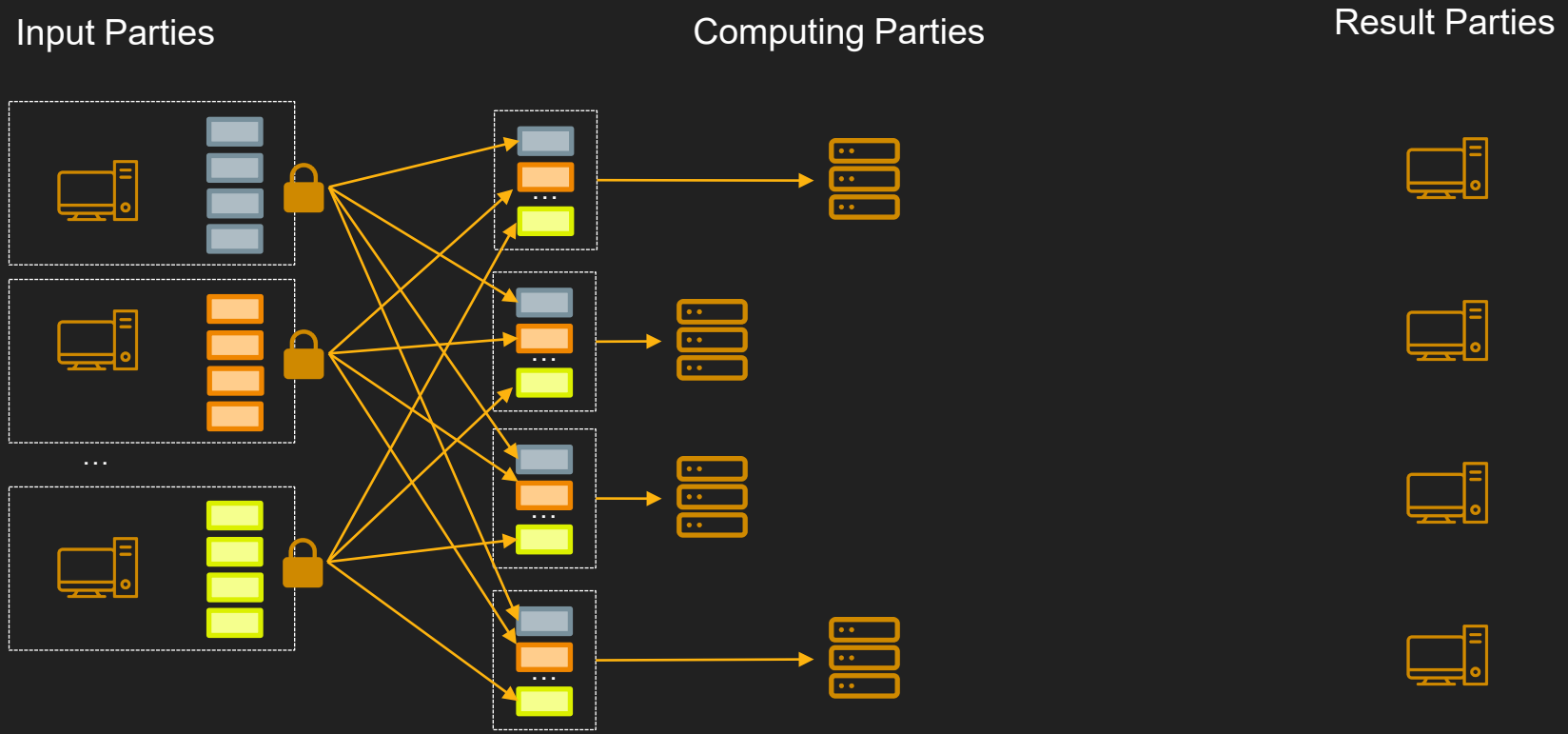
Input Parties

Computing Parties

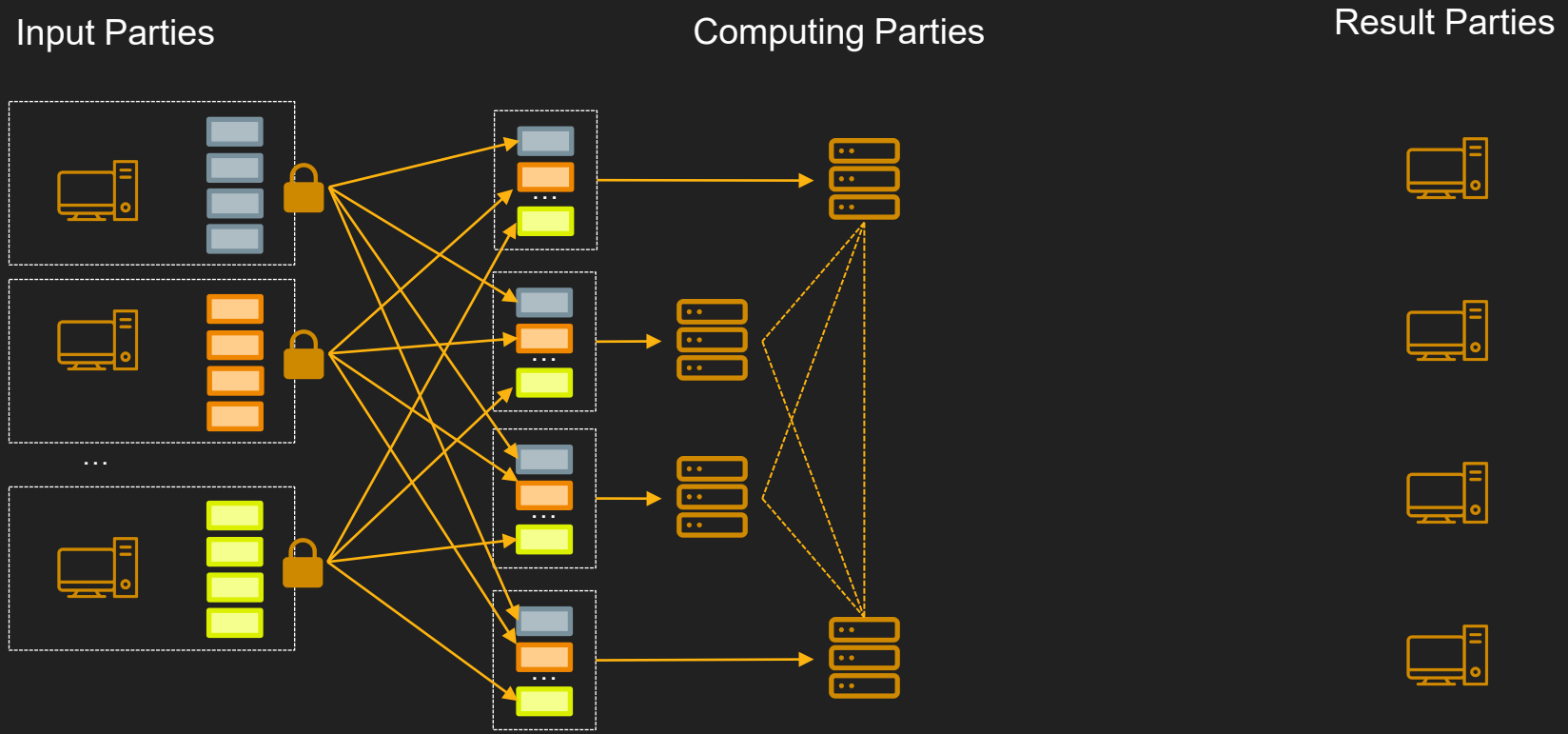
Result Parties



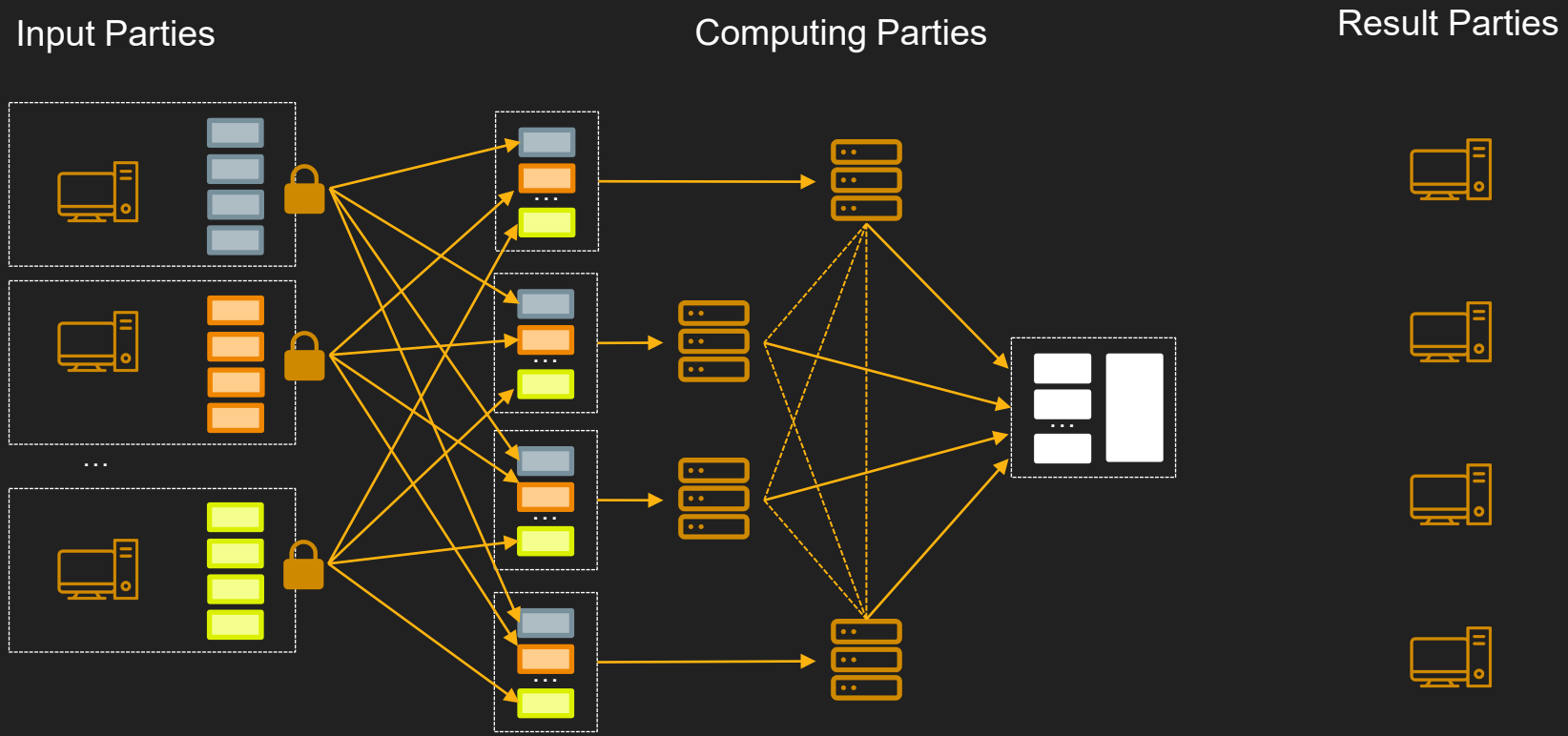
How is MPC Deployed in Practice?



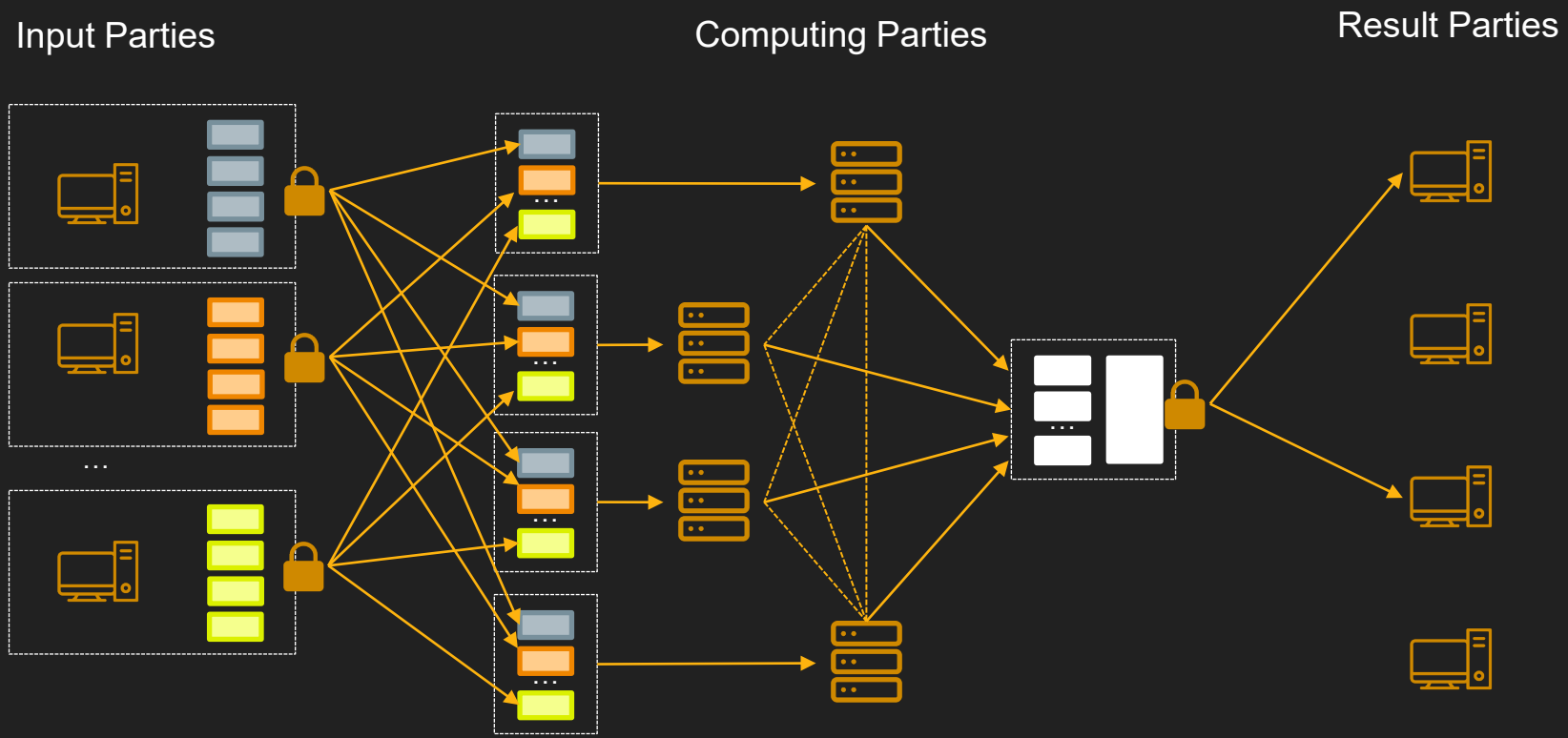
How is MPC Deployed in Practice?



How is MPC Deployed in Practice?

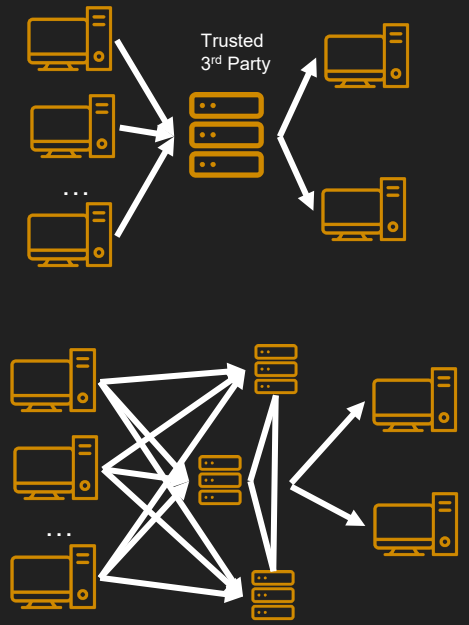


How is MPC Deployed in Practice?

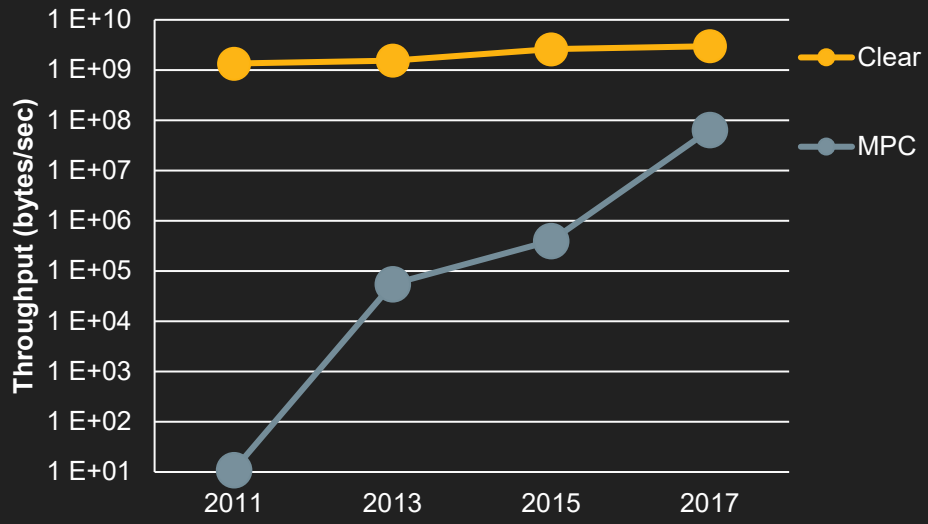


MPC Challenges

- Communication overheads!



- High Computational Cost!

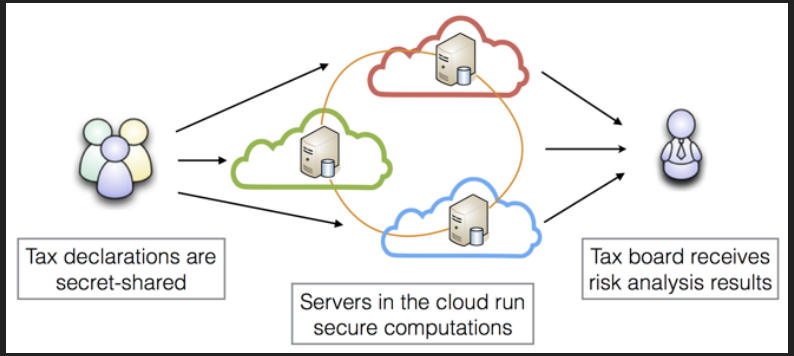


Source: Mayank Varia, Boston University, "A Survey of MPC Offerings"

MPC in Use!

Tax Fraud

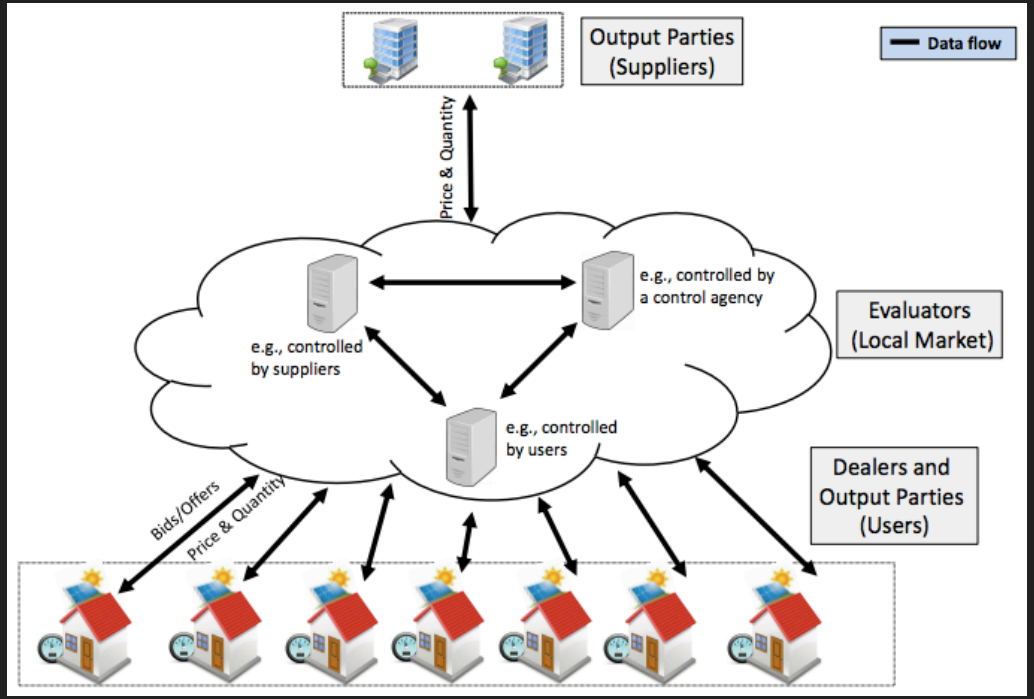
- ITL economic benchmarks
  - Collection of Estonian companies
  - Aggregate economic indicators: profit, # employees, salaries



- VAT tax revenue
  - Worked with Estonian Tax and Customs Board
  - Test if Company A's VAT credit == Company B's VAT reported

https://sharemind.cyber.ee/tax-vat-fraud

Electricity Markets

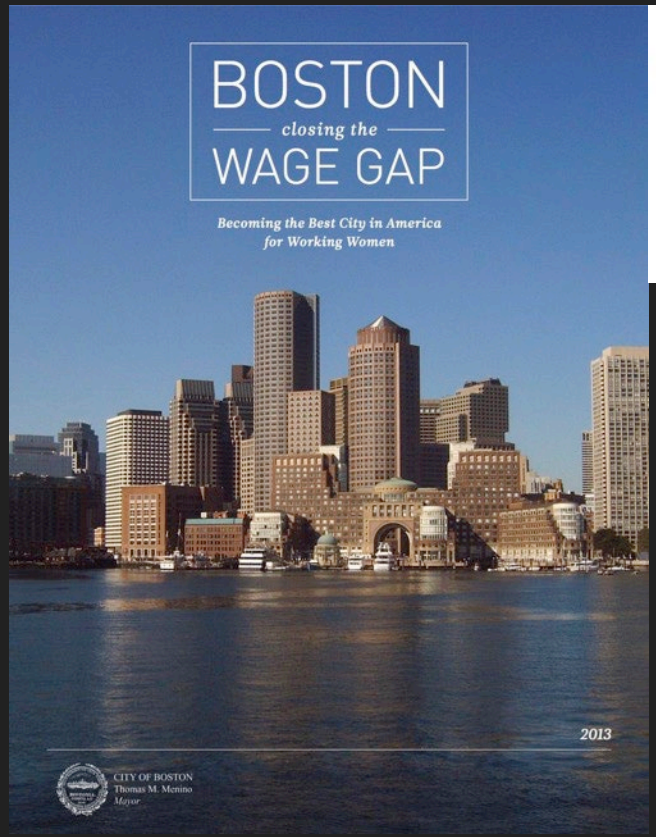


Energy trading with smart meters

- Handles 2500 bids in ~5 min
- Auction run every 30 min

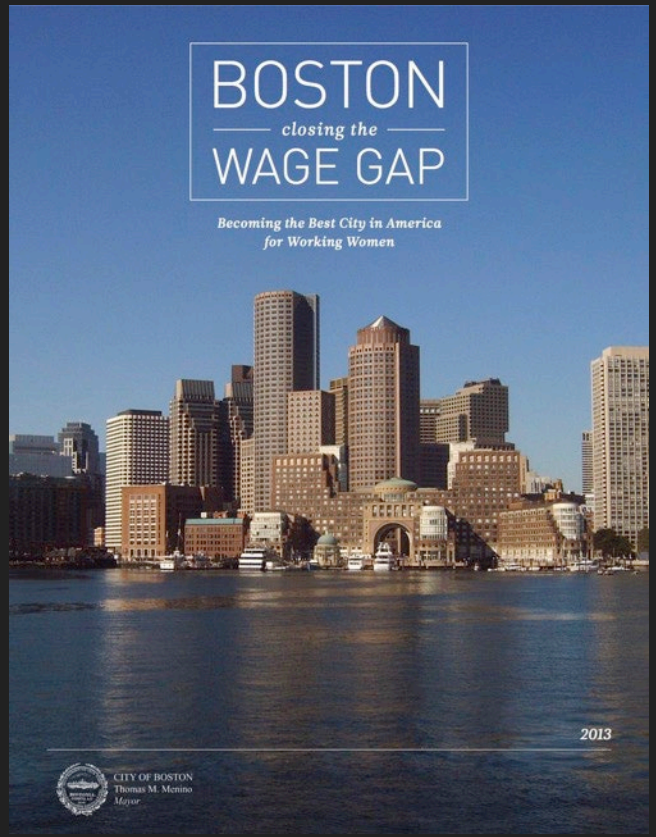
Abidin, Aly, Cleemput, and Mustafa, An MPC-based Privacy-Preserving Protocol for a Local Electricity Trading Market

Public Good (Wage Disparity)



https://thebwwc.org/mpc

Public Good (Wage Disparity)

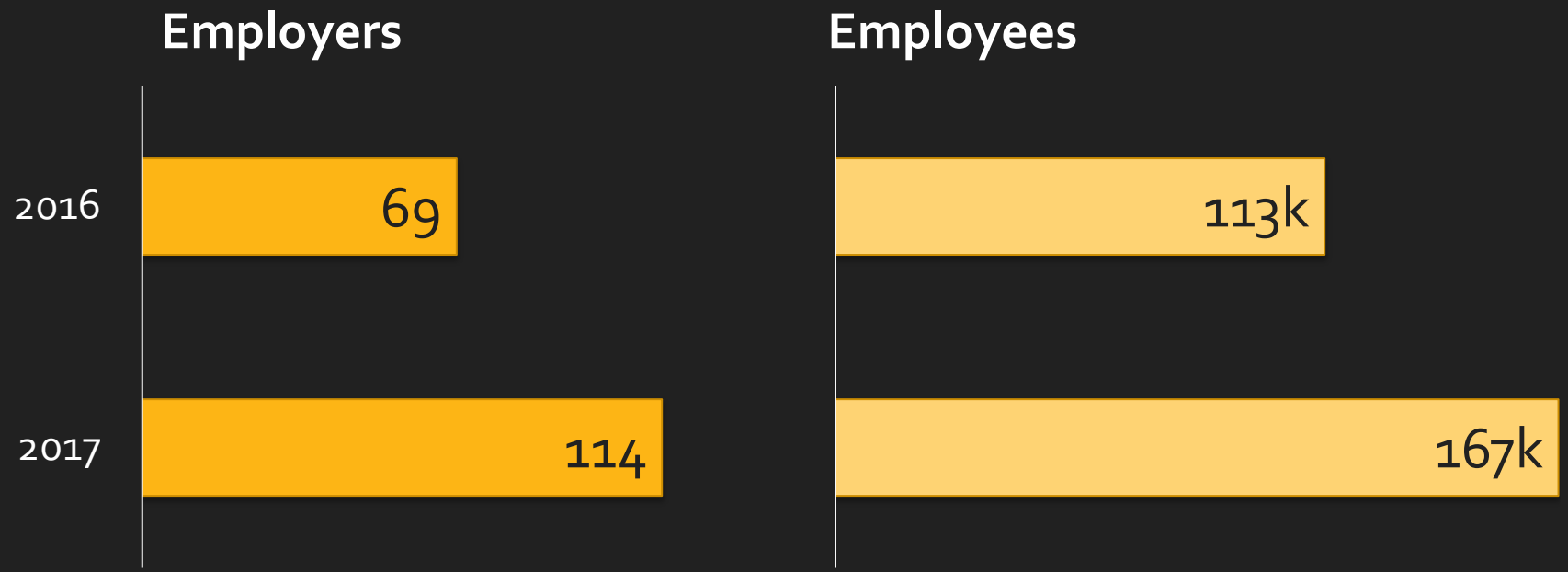


Goal 3: Evaluating Success

Employers agree to contribute data to a report ***compiled by a third party*** on the Compact's success to date. ***Employer-level data would not be identified*** in the report.

https://thebwwc.org/mpc

Public Good (Wage Disparity)

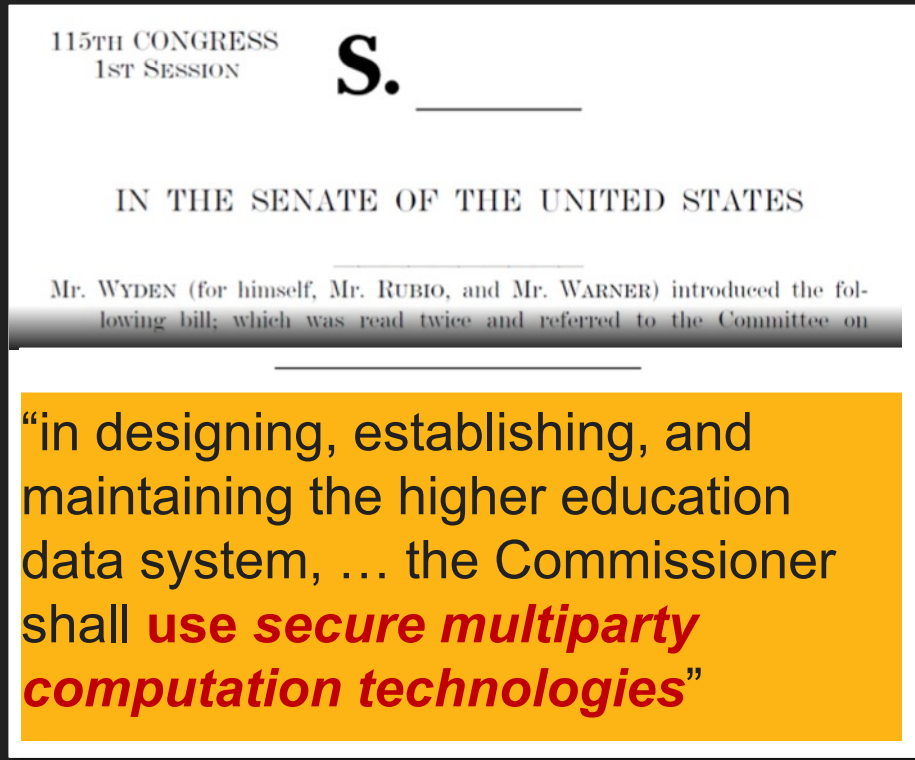


https://thebwwc.org/mpc

Bestavros, Azer, Andrei Lapets, and Mayank Varia. "User-centric distributed solutions for privacy-preserving analytics." Communications of the ACM 60.2 (2017): 37-39.

“Student Right to Know Before You Go” Bill

- Empower prospective college students to make more informed decisions
- Measure annual earnings and accumulated debt of recent graduates



115TH CONGRESS
1ST SESSION

S. _____

IN THE SENATE OF THE UNITED STATES

Mr. WYDEN (for himself, Mr. RUBIO, and Mr. WARNER) introduced the following bill; which was read twice and referred to the Committee on

_____

“in designing, establishing, and maintaining the higher education data system, … the Commissioner shall **use *secure multiparty computation technologies***”

https://www.congress.gov/bill/115th-congress/house-bill/4479

“Student Right to Know Before You Go” Bill

MAY 9, 2013

Earlier Version — Introduced

This activity took place on a related bill, S. 915 (113th).

MAY 5, 2015

Earlier Version — Introduced

This activity took place on a related bill, S. 1195 (114th).

NOV 29, 2017

Earlier Version — Introduced

This activity took place on a related bill, S. 2169 (115th).

MAR 6, 2019

Introduced

Bills and resolutions are referred to committees which debate the bill before possibly sending it on to the whole chamber.

Read Text »

MAR 30, 2022

Reintroduced Bill — Introduced

This activity took place on a related bill, S. 3952.

https://www.govtrack.us/congress/bills/116/s681

Conclusions

- MPC provides a mechanism to promote collaboration
- Goal: prevent other parties from learning about shared data
- MPC maintains data usability and “privacy” (more like “confidentiality”)
  - Not the differential privacy definition of privacy! Attacks are still possible…
- High computational and communication costs!
- Assumptions about maliciousness of participants

Group Activity

- Think about your group project
- Do you need to collaborate to learn something?
  - What data would you need to share?
  - Who would be the collaboration parties?
  - What would you want to learn?
  - What would you want to protect?