Pursuing Privacy in Recommender Systems

The View of Users and Researchers from Regulations to Applications

The 15th ACM Conference on Recommender Systems

Amsterdam, The Netherlands

30th September 2021

The team

Tommaso Di Noia
Politecnico di Bari

Yashar Deldjoo
Politecnico di Bari

Claudio Pomo
Politecnico di Bari

Antonio Ferrara
Politecnico di Bari

Fedelucio Narducci
Politecnico di Bari

Vito Walter Anelli
Politecnico di Bari

Luca Belli
Twitter

Table of contents

Big Data, Recommender Systems and Privacy

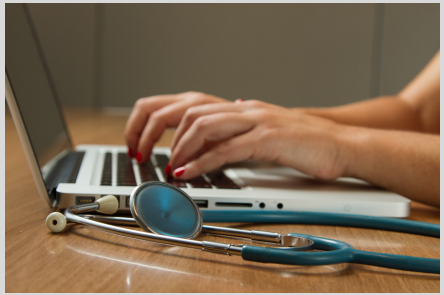
Privacy and Utility
in the Era of Regulations

«Every day, we create 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few. THIS DATA IS BIG DATA.»

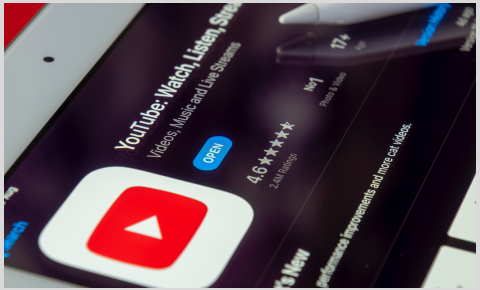
March 19, 2012

Big Data is the fuel of AI and RecSys

Healthcare



Entertainment



Smart Cities

Home Automation

Virtual Assistants

Banking

This is not a secret

Forbes

Nov 26, 2019, 11:40am EST | 34,050 views

Data Privacy Will Be The Most Important Issue In The Next Decade

Mary Meehan Contributor

Consumer Tech

I write about consumer insights and foresights to drive innovation.

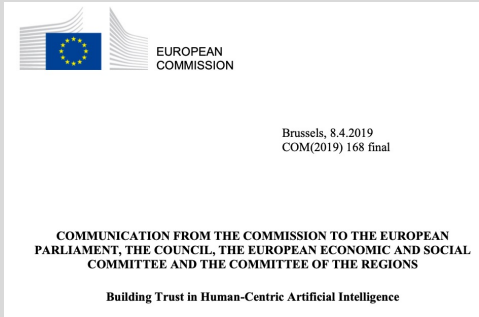
Listen to article 7 minutes

Building Trust in Human-Centric Artificial Intelligence

Privacy and data governance

- Privacy and data protection must be guaranteed at all stages of the AI system's life cycle.
- To allow individuals to trust the data processing, it must be ensured that they have full control over their own data, and that data concerning them will not be used to harm or discriminate against them.

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58496


EUROPEAN COMMISSION

Brussels, 8.4.2019
COM(2019) 168 final

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

Building Trust in Human-Centric Artificial Intelligence

The European General Data Protection Regulation (GDPR)

Protection and Control of individuals' personal data

A lot of parts involved: data subject, data controller, data processor...

GDPR FINES

If your data is breached

report it in

72

hours

face a fine up to

20M €

GDPR INDIVIDUAL'S RIGHTS

Make individuals aware of their rights and ensure mechanisms are in place to act on these

Examples

Right to access (Are my data collected and processed? Why? How long?)

Right to rectification, erasure, and restrict processing

RecSys: User-centric Data

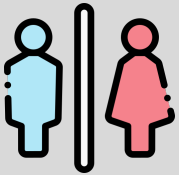
Medical records

User's check-in

Genomic data

User's preference

Search logs



Sensitive attribute
(e.g., gender)

The privacy-personalization trade-off

- The quality of the recommendations is correlated with the amount, richness, and freshness of the underlying user modeling data
- The same factors drive the severity of the privacy risk

Friedman A., Knijnenburg B.P., Vanhecke K., Martens L., Berkovsky S. (2015) Privacy Aspects of Recommender Systems. In: Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook (2nd edition). Springer, Boston, MA.

Privacy risks

- Direct access to data
  - Unsolicited data collection
  - Sharing data with third parties
  - Unsolicited access by employes

Friedman A., Knijnenburg B.P., Vanhecke K., Martens L., Berkovsky S. (2015) Privacy Aspects of Recommender Systems. In: Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook (2nd edition). Springer, Boston, MA.

A Face is Exposed for AOL searcher No. 4417749: the case of Thelma Arnold

- In August 2006, AOL released anonymized search logs
  657K users, 20M queries over three months

User 4417749

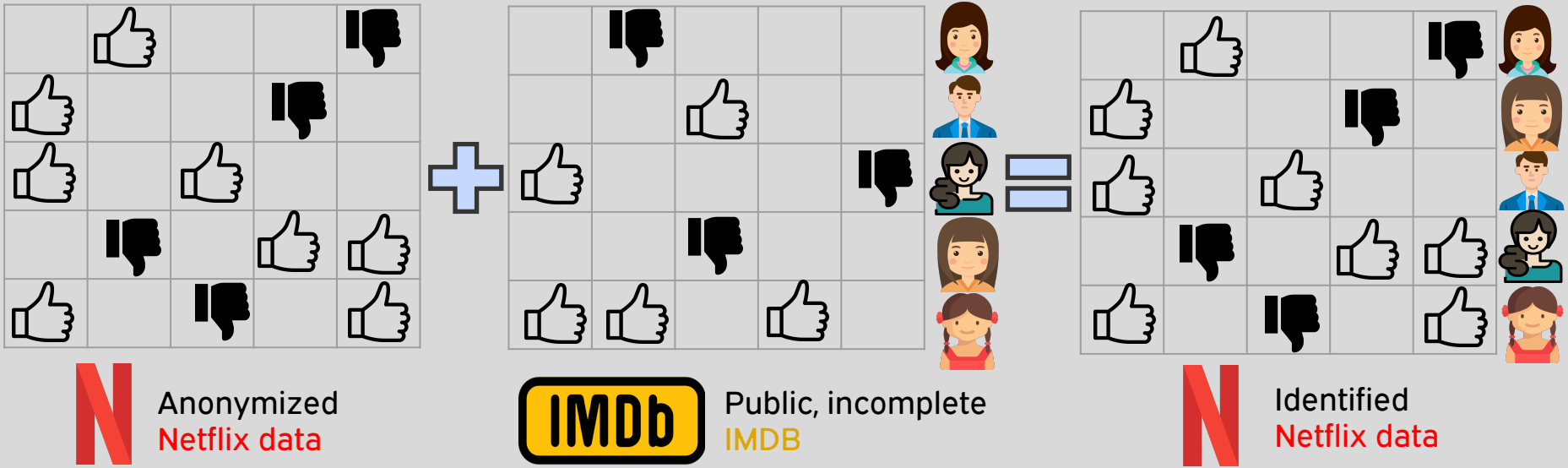
- "numb fingers"
- "60 single men"
- "dog that urinates on everything"
- "Landscape in Liburn, Ga"
- Several people names with last name Arnold
- Home sold in shadow lake subdivision

Netflix Prize

October 2006: Netflix announces Netflix Prize

- 10% of their users
- Average 200 ratings per user

Netflix Prize



Privacy risks

- Inference from User Preference Data
  - Exposure of sensitive information
  - Targeted Advertising
  - Discrimination

Friedman A., Knijnenburg B.P., Vanhecke K., Martens L., Berkovsky S. (2015) Privacy Aspects of Recommender Systems. In: Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook (2nd edition). Springer, Boston, MA.

Privacy risks

- Risks Imposed by other System Users
  - In collaborative approaches, users are compared with each other
  - Create fake profiles to identify other users' preferences
  - By observing changes in item-to-item collaborative systems an attacker may infer preferences of a target user

Friedman A., Knijnenburg B.P., Vanhecke K., Martens L., Berkovsky S. (2015) Privacy Aspects of Recommender Systems. In: Ricci F., Rokach L., Shapira B. (eds) Recommender Systems Handbook (2nd edition). Springer, Boston, MA.

Risks imposed by external entities

- Privacy breaches
  - Netflix
  - Facebook
  - Amazon
  - LinkedIn
  - Marriot

2

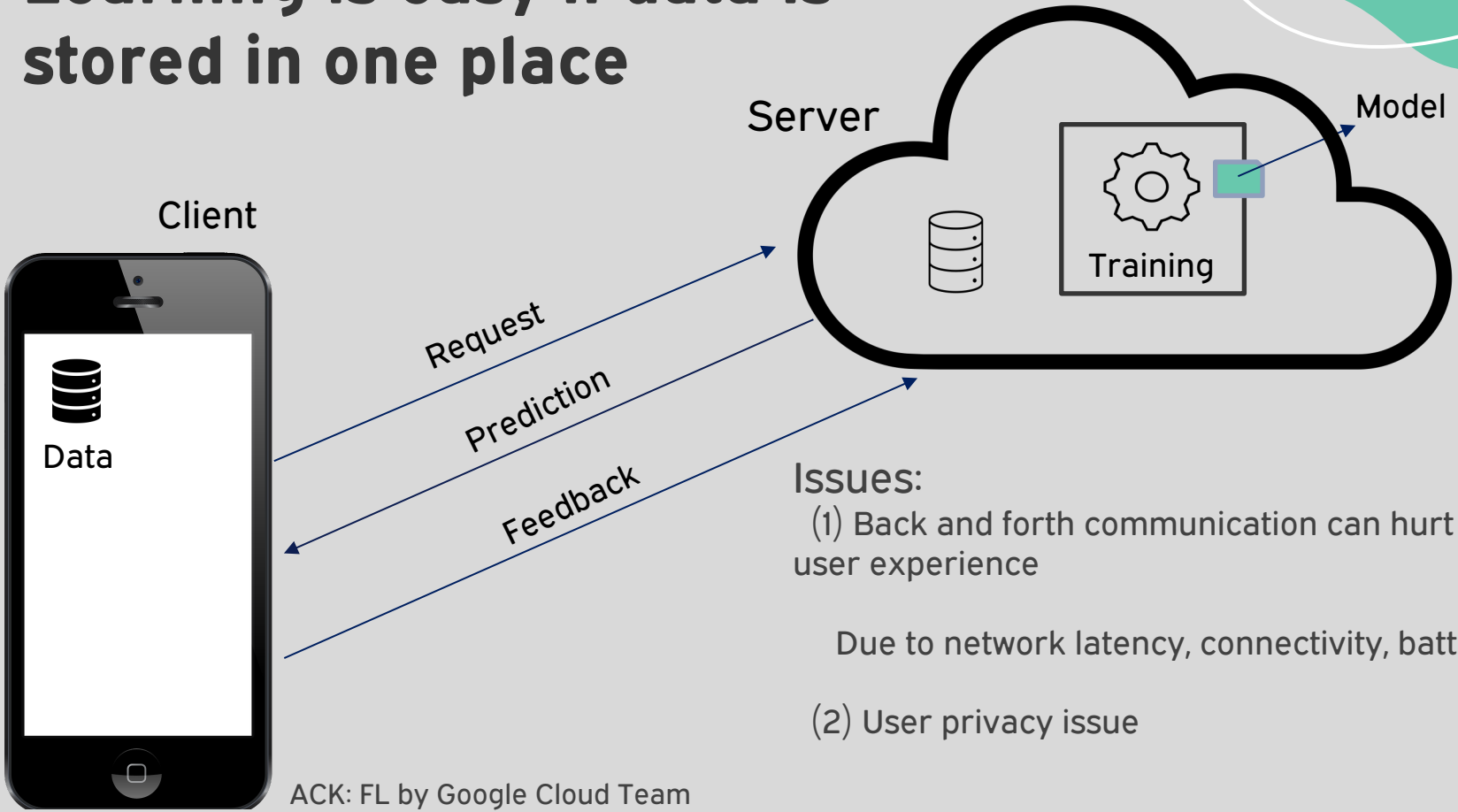
Privacy-Oriented Recommender Systems

Learning Paradigms
and Threats

What Privacy-Preserving Machine Learning tries to Protect

1. **Input training data**
2. **Output** predicted **labels**
3. **Model information**, including parameters, architecture, and loss function
4. **Identifiable information**, such as which site a record comes from

Learning is easy if data is stored in one place



Issues:

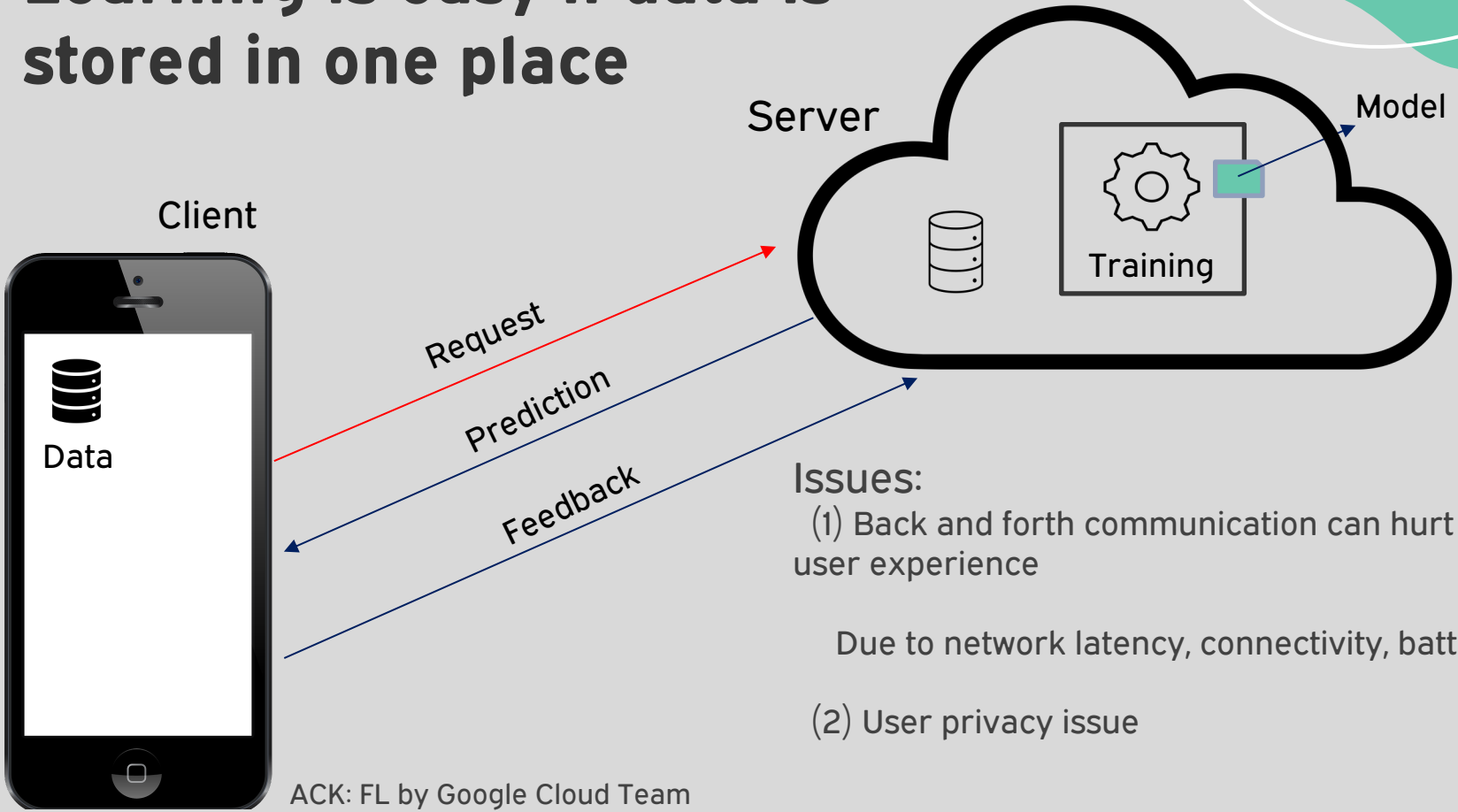
(1) Back and forth communication can hurt the user experience

Due to network latency, connectivity, battery life

(2) User privacy issue

ACK: FL by Google Cloud Team

Learning is easy if data is stored in one place



Issues:

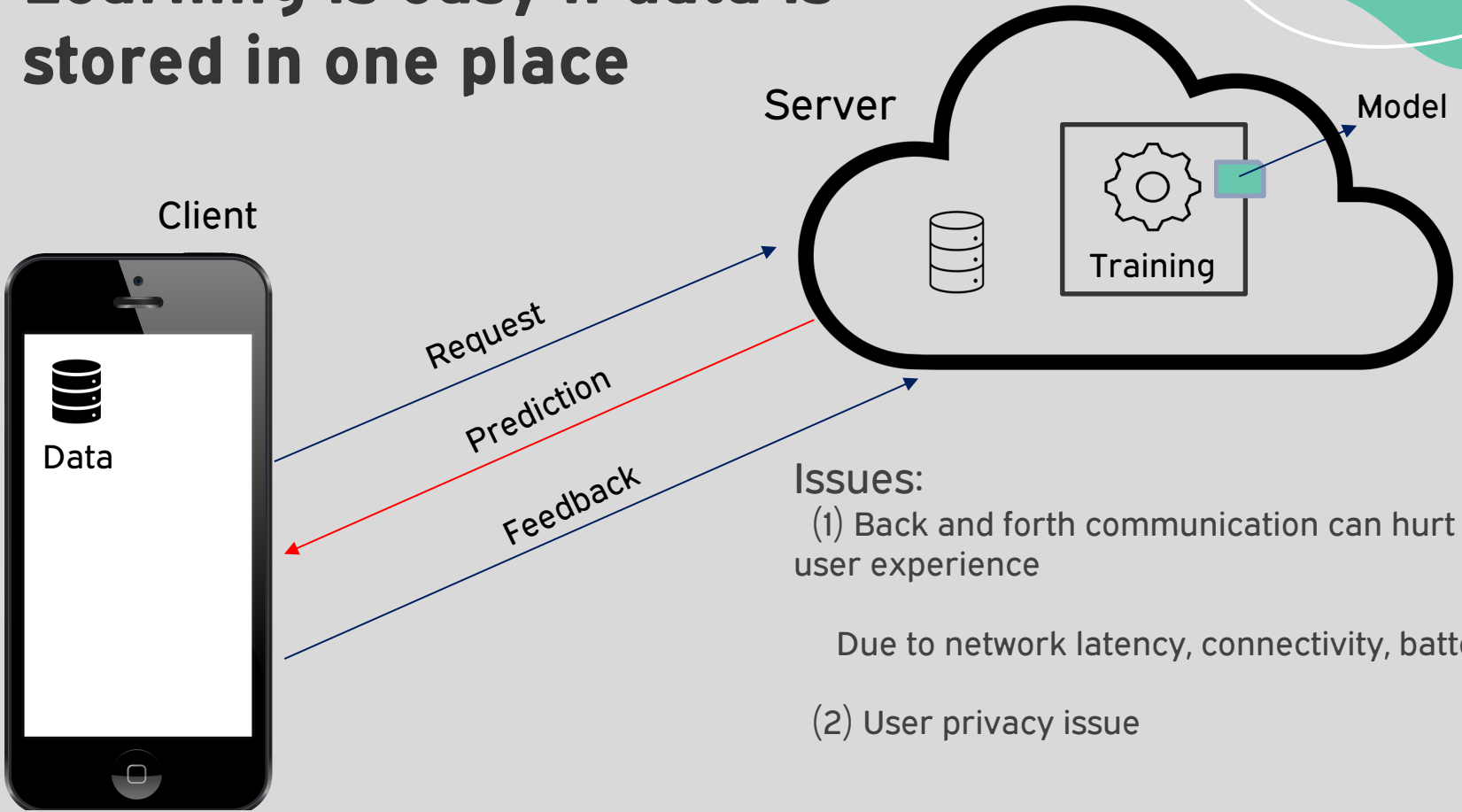
(1) Back and forth communication can hurt the user experience

Due to network latency, connectivity, battery life

(2) User privacy issue

ACK: FL by Google Cloud Team

Learning is easy if data is stored in one place



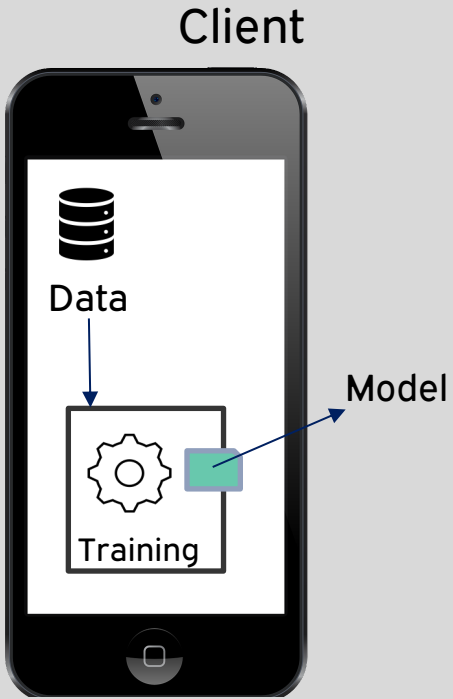
Issues:

(1) Back and forth communication can hurt the user experience

Due to network latency, connectivity, battery life

(2) User privacy issue

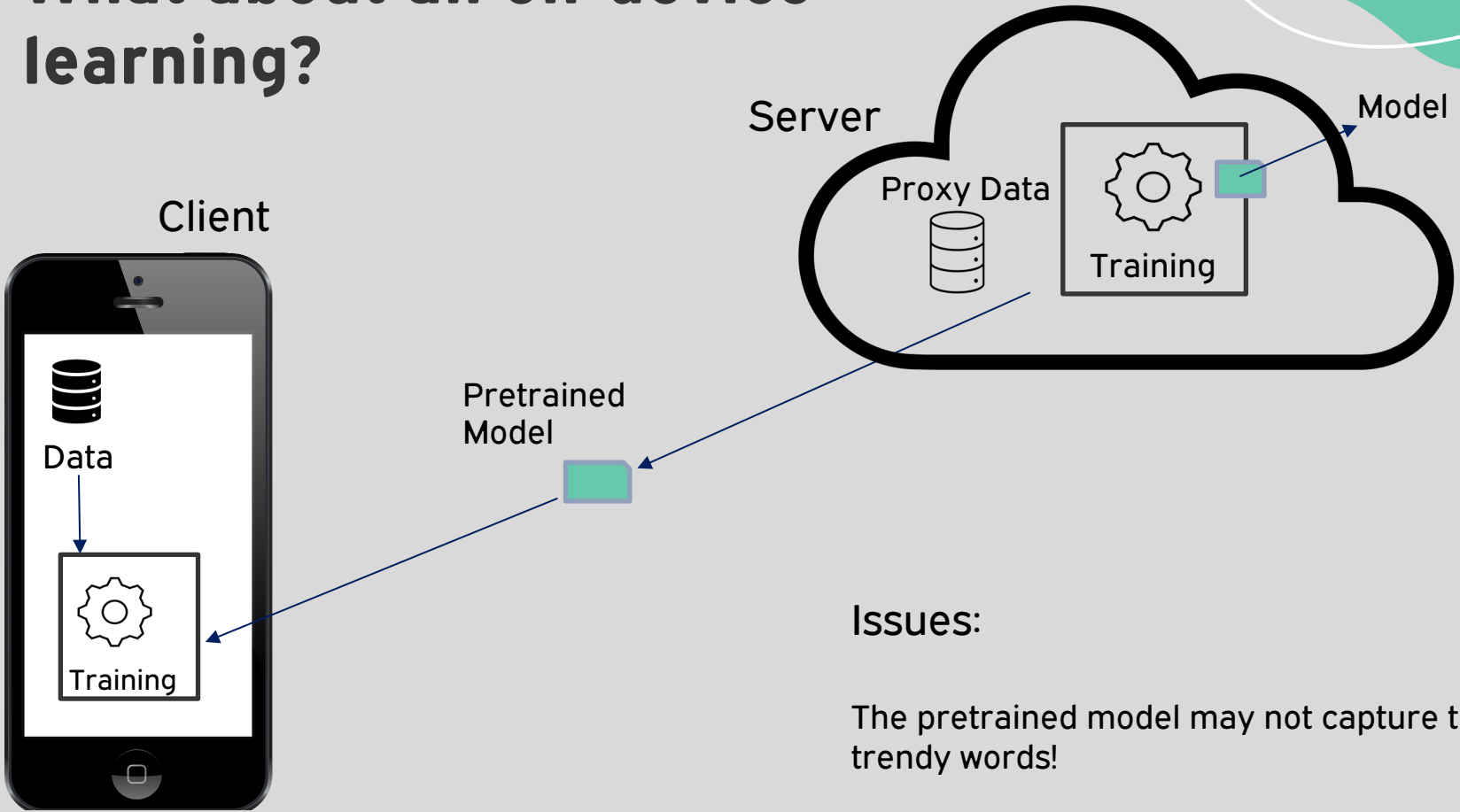
What about all on-device learning?



Issues:

Each individual device may not have enough data to build a good model!

What about all on-device learning?



Issues:

The pretrained model may not capture the trendy words!

What about all on-device learning?

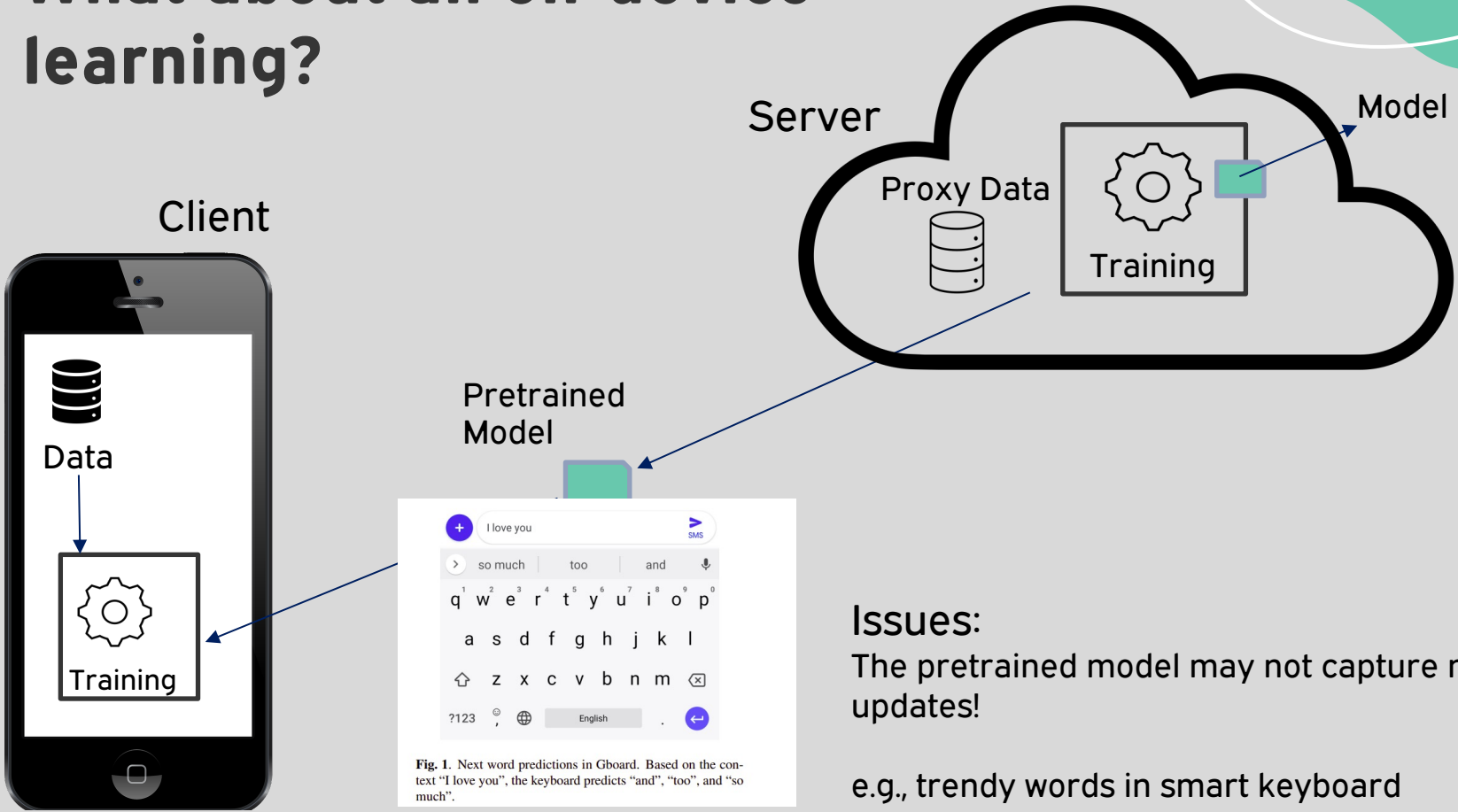


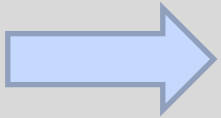
Fig. 1. Next word predictions in Gboard. Based on the context "I love you", the keyboard predicts "and", "too", and "so much".

Issues:

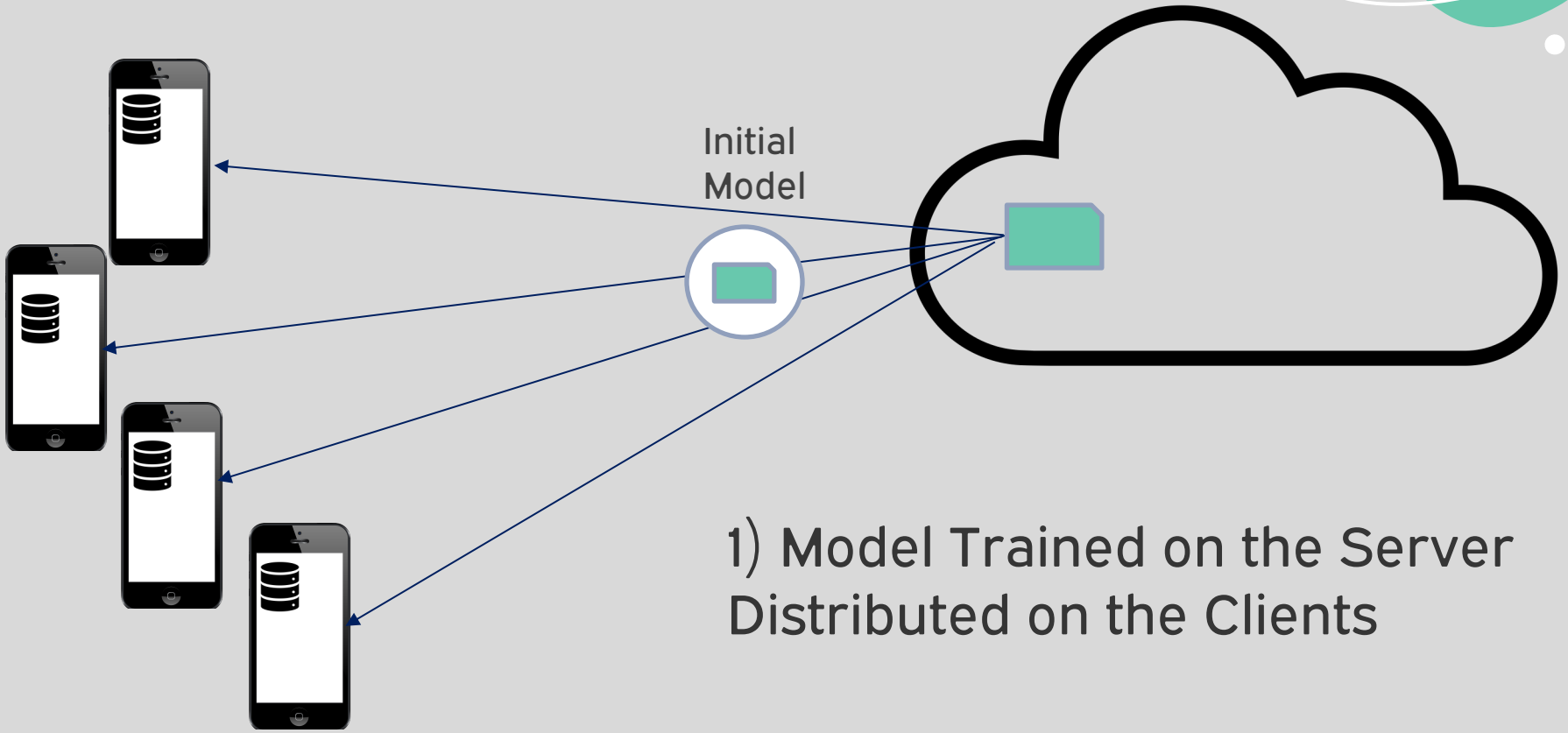
The pretrained model may not capture new updates!

e.g., trendy words in smart keyboard

How to Keep the Goodness of

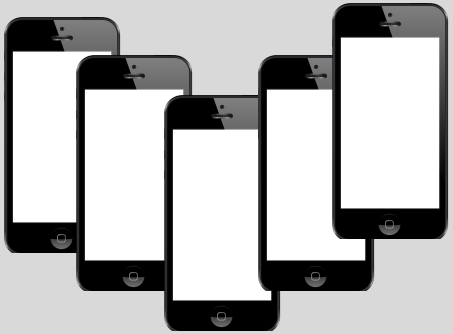
Decentralized Data  Users' Privacy  Federated Learning

Federated Learning



1) Model Trained on the Server
Distributed on the Clients

Federated Learning



Available

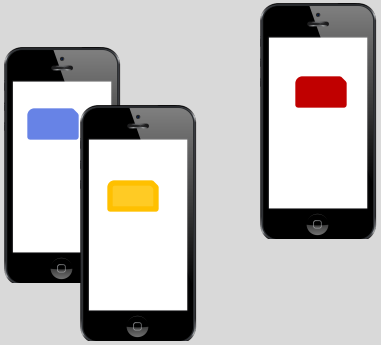
Not
Available

2) Find Available Devices

Federated Learning

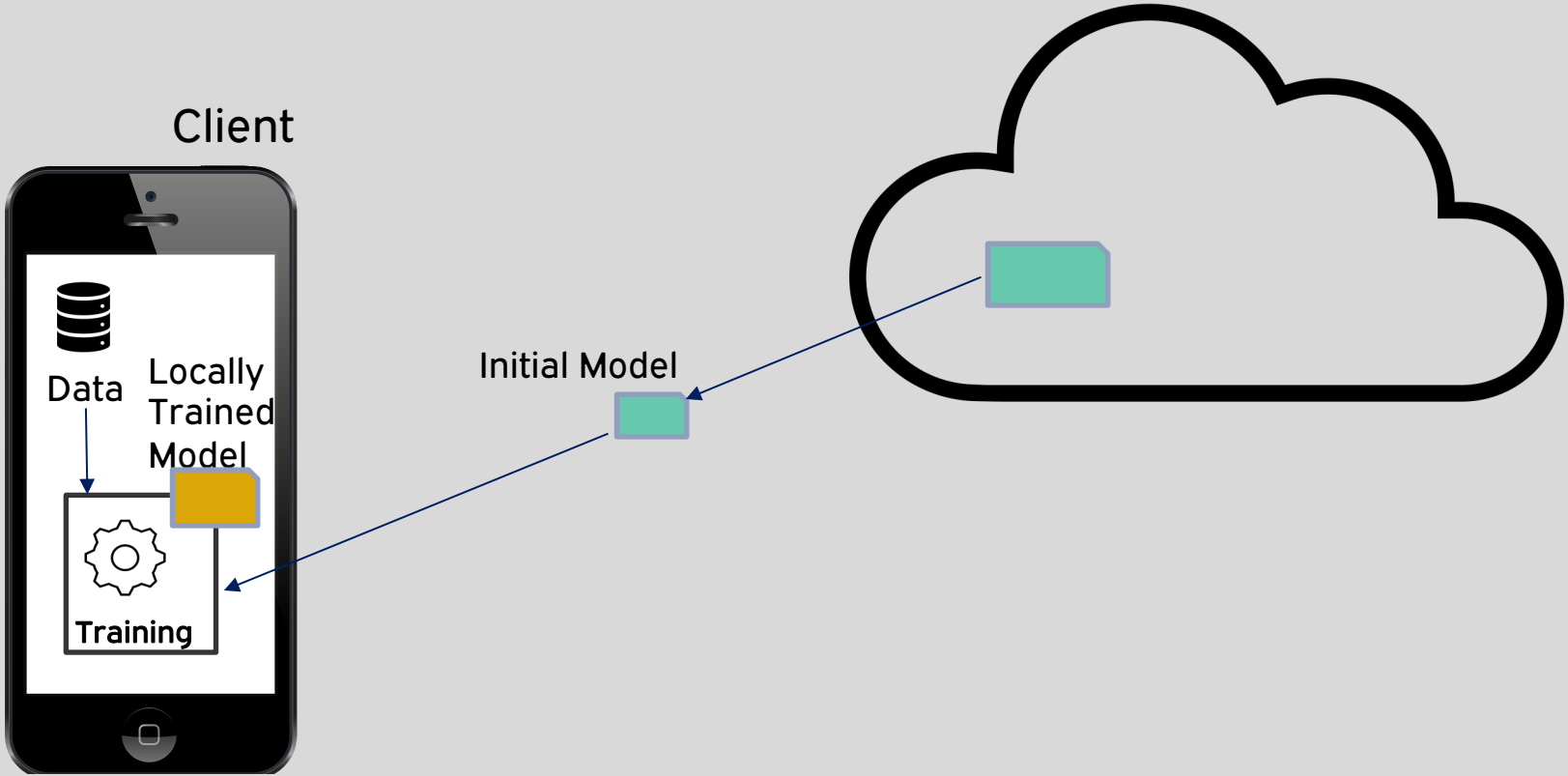
Available

3) Find Suitable Devices

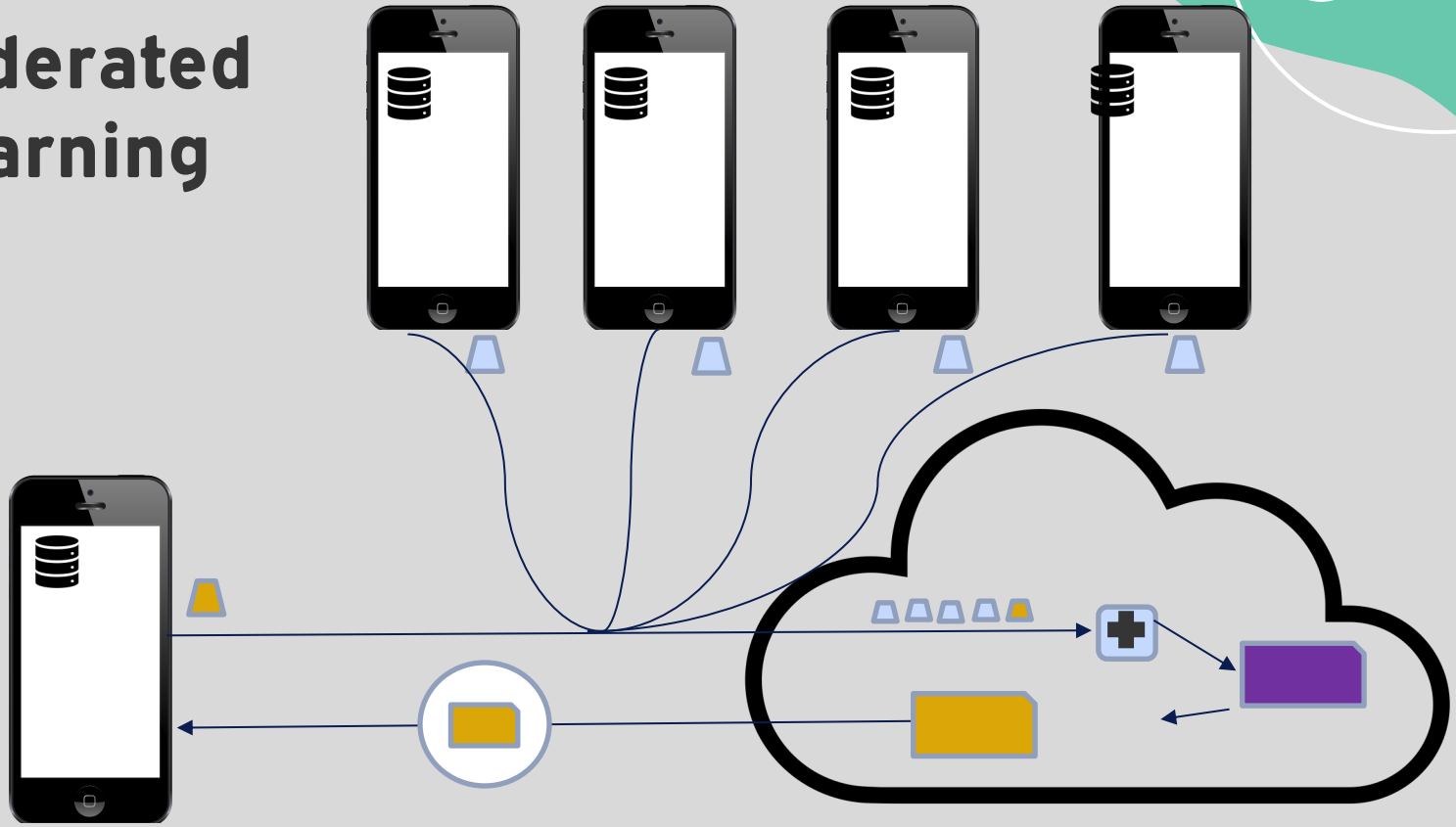


Suitable
Available

Federated Learning: How does it work?

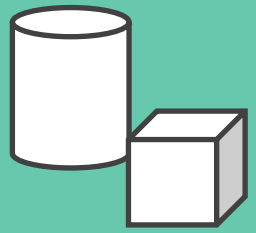


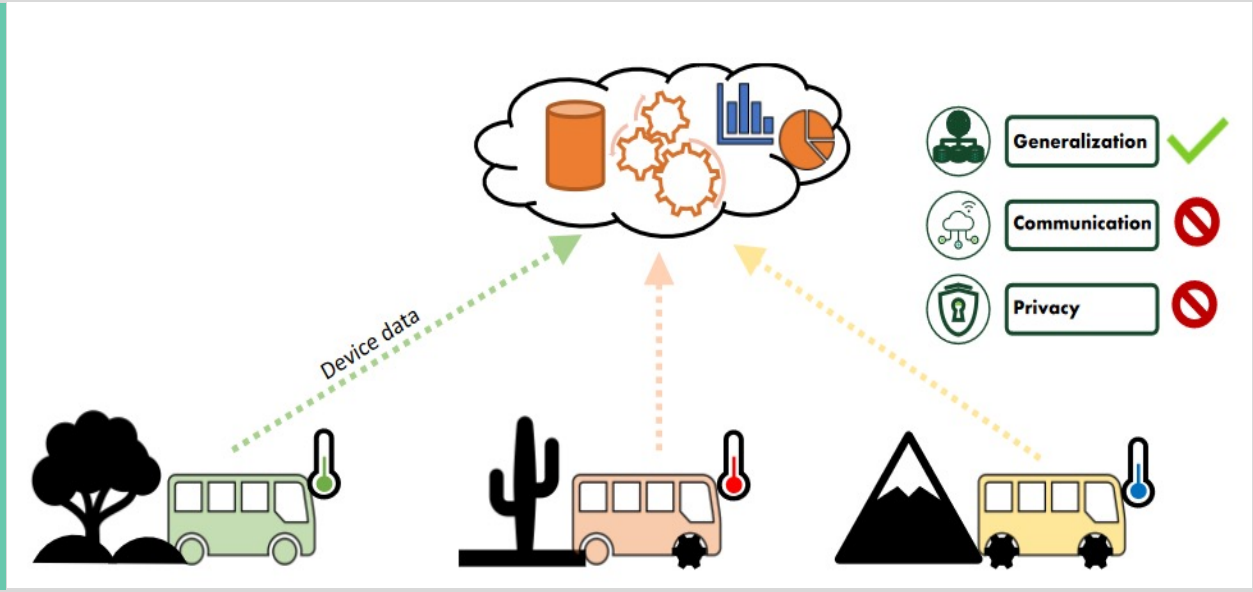
Federated Learning



Learning paradigms

Centralized Learning



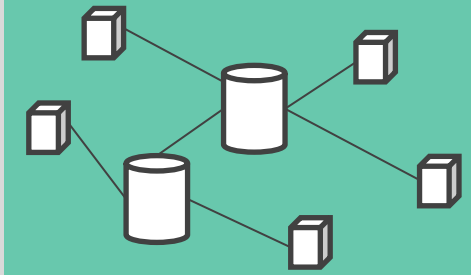


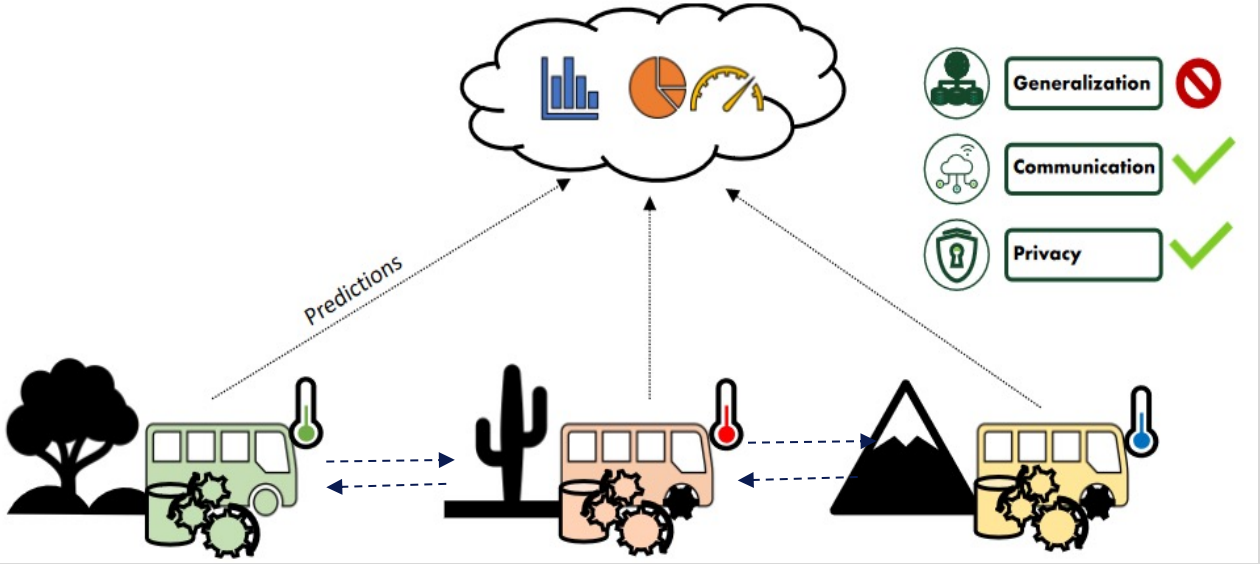
Centralized learning: Challenges

Connectivity	Data must be transmitted over a stable connection
Bandwidth	Bandwith is limited, and the amount of data may exceed what's reasonable to transfer to the cloud.
Latency	Real-time applications, e.g. automation, requires very low latency.
Connectivity	Sensitive operational data must remain on site

Learning paradigms

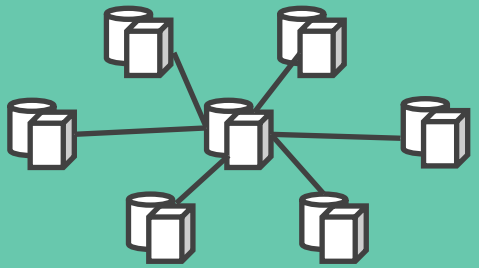
Decentralized Learning

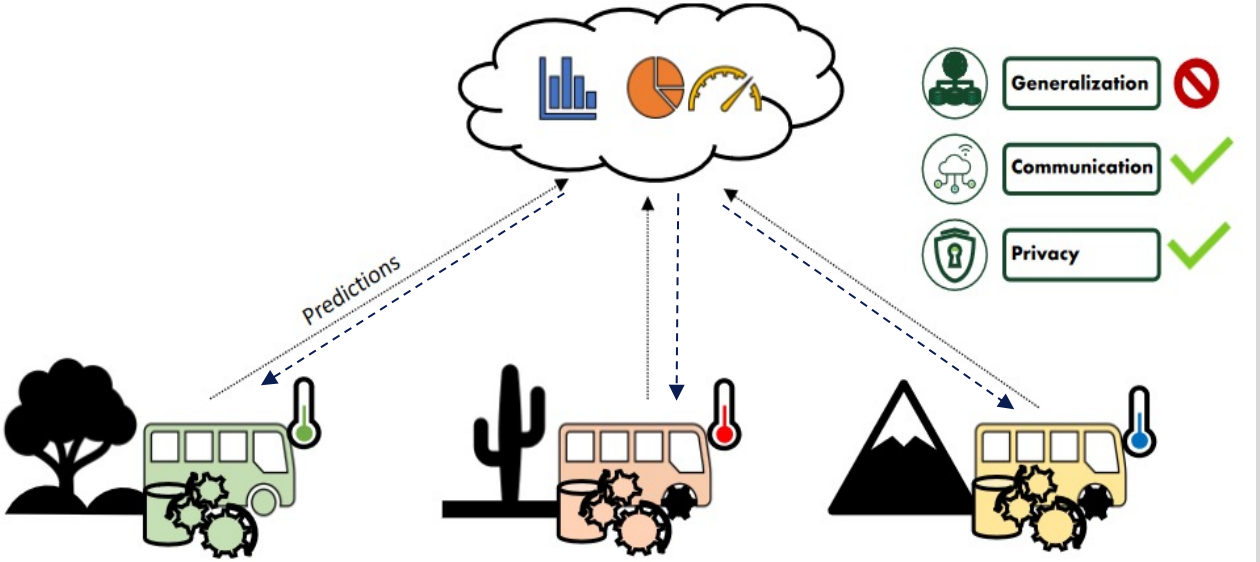




Learning paradigms

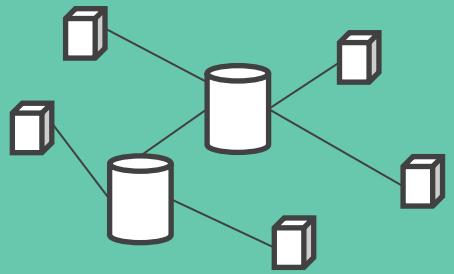
Distributed Learning

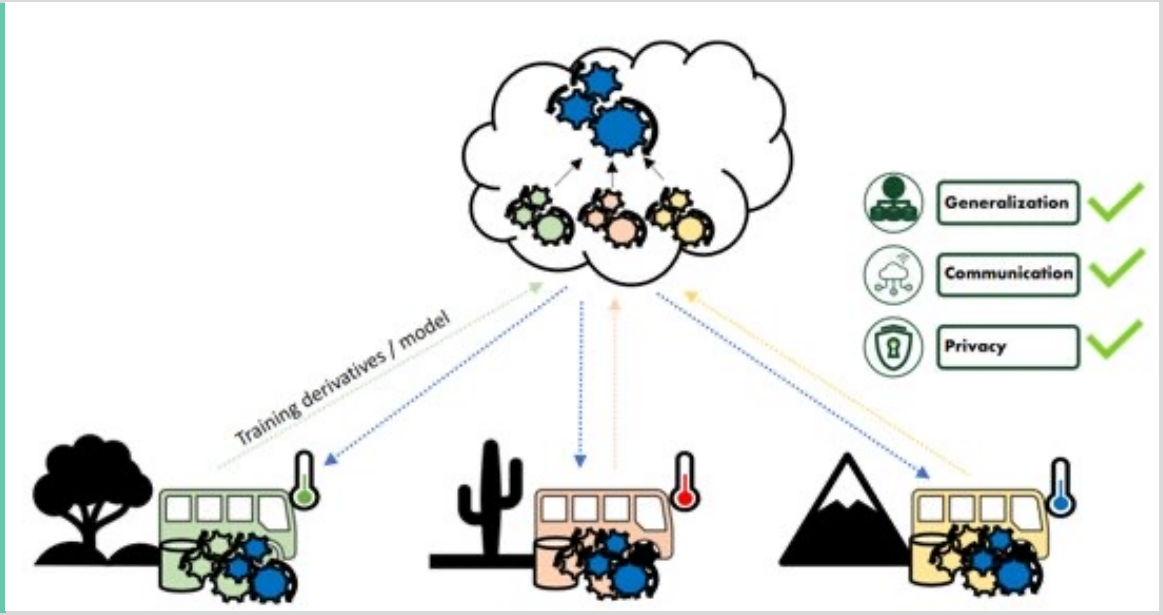




Learning paradigms

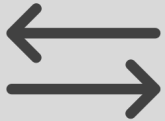
Federated Learning





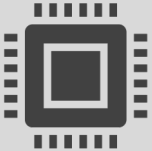
Federated learning: Advantages

Data privacy/security	Data pool not required for the model, and are remained on user's devices



Data diversity and Model Liability	FL facilitates access to heterogeneous data. Reduces legal liability of the model.

Real time continual learning	Model are constantly improved using client data with no need to aggregate data for continual learning



Hardware/Bandwidth efficiency	FL models do not need complex central server to analyze data/Does not require uploading large amount of data

Attack and Threat models

1. **Targets**
   Data vs Model

2. **Knowledge**
   White-box vs Black-box

3. **Methods**
   Model extraction vs
   Encoding Information

Attack methods

Model Extraction Attack
Are you cheating me?

Feature Estimation Attack
How are your data?

Membership Inference Attack
Do these features belong to you?

Model Memorization Attack
Are you a malicious ML Provider?

Model Extraction Attack

Knowledge
Black-Box

Attack methods
- Shadow Training
- Meta-model
- Hyperparameter-stealing
- Reverse Engineering

System type
Centralized/Federated

Model Extraction Attack

Shadow Training

The adversary's goal is to use as few queries as possible to f in order to efficiently compute an approximation f' that closely matches f

Meta-model

The metamodel not only learns to infer model characteristics from query outputs based on a static set of inputs, but it also looks for query inputs that are tailored to extract more data from the target models.

Feature Estimation Attack

Knowledge
Black-Box/White-Box

Attack methods
- Model Inversion
- Shadow Training
- GAN
- Power side-channel attack

System type
Centralized/Distributed

Feature Estimation Attack

Model Inversion

Adversary learns certain features $x_i \in x^*$ or statistical properties sach as $avg(x^*)$ of the training dataset. Given white-box access to f and auxiliary information side for a patient instance attempts to infer the patient's sensible data

Power side-channel attack

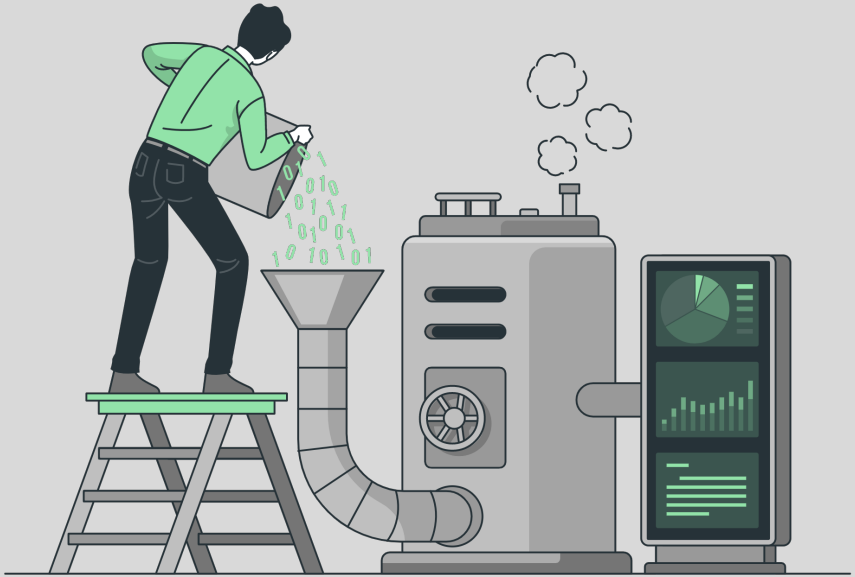
"The power side-channel leakage can be exploited to recover the secret keys in cryptographic devices". The attack on an FPGA-based convolutional neural network accelerator aims to recover the input image from the collected power traces without knowing the detailed parameters in the neural network.

Membership Inference Attack

Knowledge
Black-Box

Attack methods
- Shadow Training
- GAN
- Gradient-based
- Unsupervised binary classification

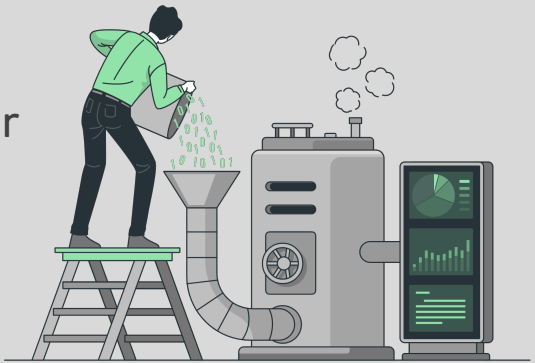
System type
Centralized/Distributed



Membership Inference Attack

Gradient-based

The attack is achieved by analyzing periodic updates to the shared model during training. The reason that this attack is effective is that the gradients in neural networks are based on features, thus observations of the participants' gradient updates can be used to infer the feature values.



GAN

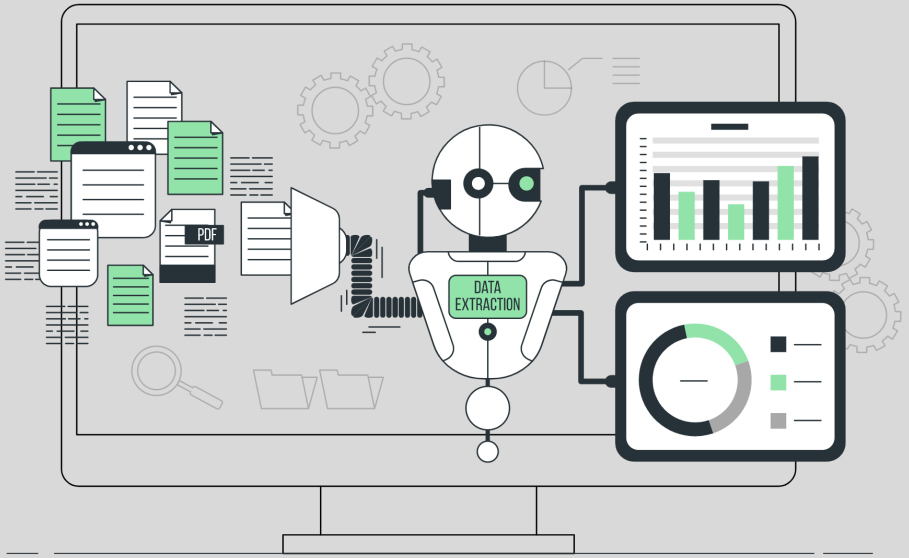
The proposed attack framework exploits GAN with a multi-task discriminator, which simultaneously discriminates category, reality and client identity of input samples, and doing so recovers user-specific private data

Model Memorization Attack

Knowledge
Black-Box/White-Box

Attack methods
- Encoding
- Reconstruction
- Hyperparameter-stealing
- Reverse Engineering

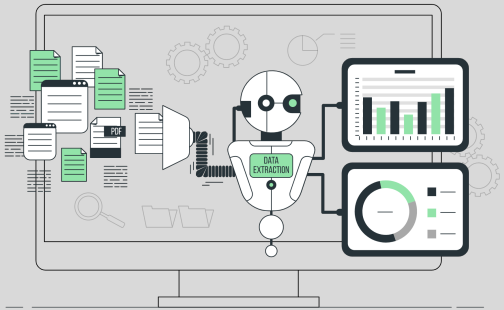
System type
Centralized/Federated



Model Memorization Attack

Encoding

- LSB encoding: the adversary can encode the “training dataset in the least significant (lower) bits of the model parameters.”
- Correlated value encoding: the adversary can “gradually encode information while training model parameters.”
- Sign encoding: similar to correlated value encoding, the adversary can use “the sign of model parameters to interpret as bit strings"



Why should I trust you?

Why

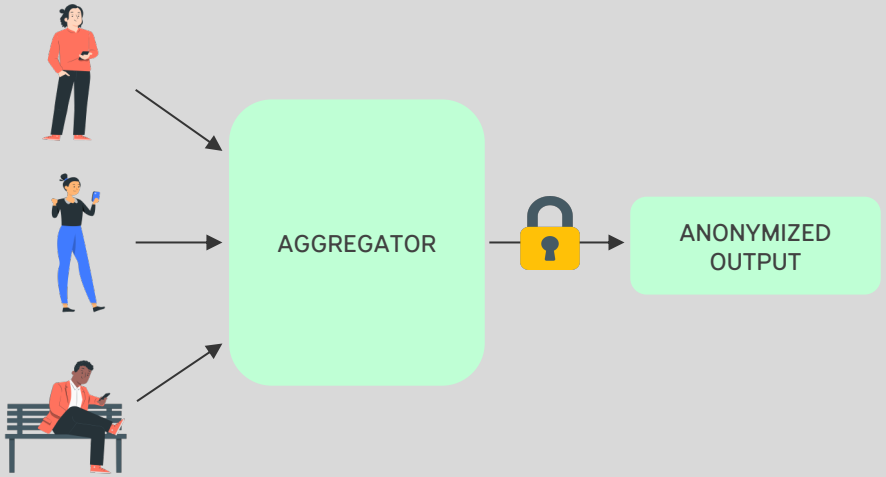
- Aggregator does not trust in plainly releasing both the output and the features of the model

What

- Model Extraction Attack
- Feature Estimation Attack

How

- (Central) Differential Privacy

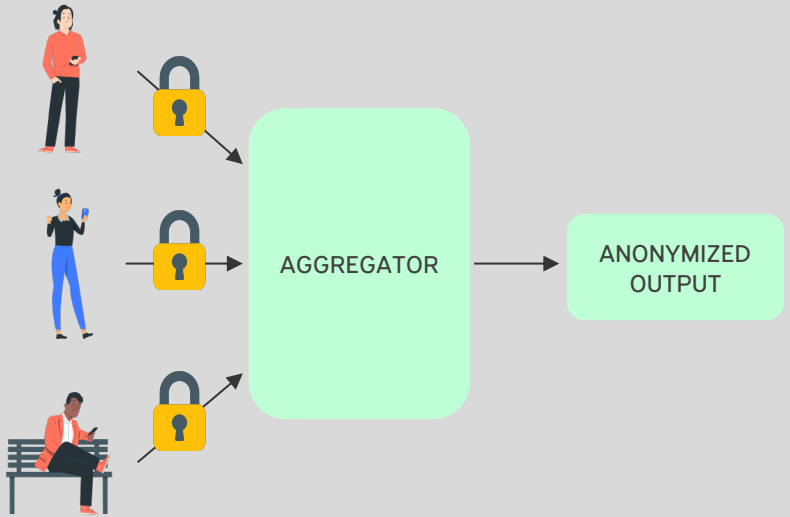


Why should I trust you?

Why
- The aggregator could be malicious
- Network layer is a potential point of failure

What
- Membership Memorization Attack

How
- Encryption
- Secure Multi-Party computation
- (Local) Differential Privacy



Techniques for Privacy-Preserving ML

From Differential Privacy
to Cryptography

Why Differential Privacy?

We want to learn nothing about individuals but still learn useful information about a population

How to analyze data while preserving privacy (e.g., census, epidemic tracking, etc.)?

This is a very old problem in literature!

1

De-identified data are not so secure

Actually, they are not completely de-identified

2

Releasing just statistics is still non-private

Overly accurate estimates of too many statistics is non-private

Why Differential Privacy?

1

De-identified data are not so secure

- AOL search database
- Netflix ratings

Get an anonymous dataset of telephone records
Correlate it with a merchant's telephone database

Google could easily deanonymize a database of online purchases

Why Differential Privacy?

2

Releasing just statistics is still non-private

Which is the number of faculty members in the university who have hearth disease?

Which is the number of faculty members in the university who have hearth disease and are not the President?

Now we know whether the President has hearth disease or not

What can we learn?

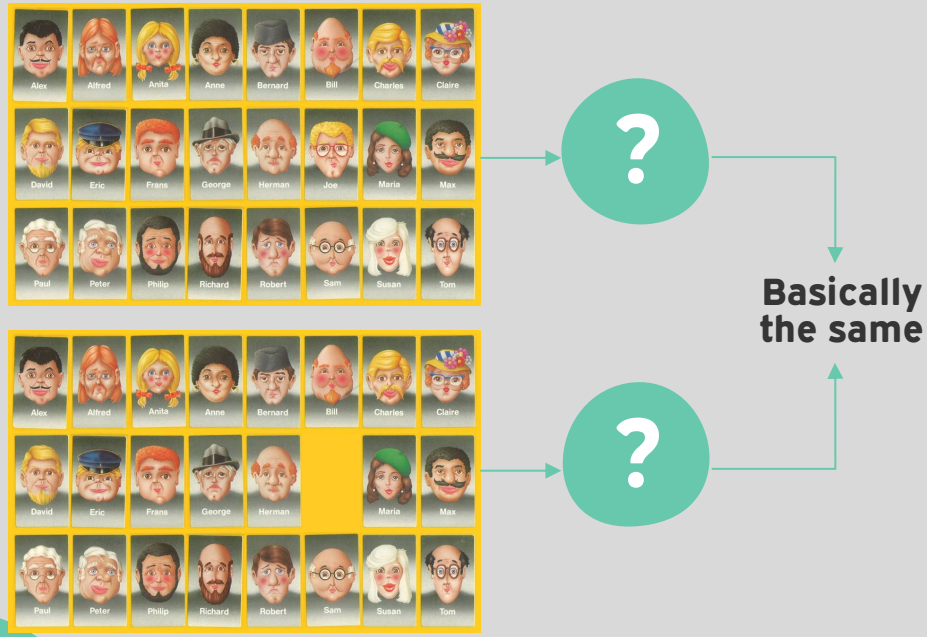
We cannot learn anything about Tommaso

If I learn that (almost) all humans have one left foot and one right foot, I am learning something new about him

? **Am I compromising Tommaso's privacy?**

No, we would have learned the same thing if Tommaso had been replaced by another random member of the population

What can we learn?



How to distinguish "learning about a population" from "learning about an individual"?

The outcome of any analysis is essentially equally likely, independent of whether any individual joins, or refrain from joining the dataset

Differential Privacy

$\mathcal{X}$ and $\mathcal{Y}$ are adjacent datasets
($\mathcal{Y}$ is equal to $\mathcal{X}$ but also contains one more example)

$\mathcal{M}$ is a randomized mechanism over a dataset

M gives ε-differential privacy if for all pairs of datasets $\mathcal{X}$ and $\mathcal{Y}$ and all events S

$$Pr[\mathcal{M}(\mathcal{X}) \in S] \leq e^{\varepsilon} Pr[\mathcal{M}(\mathcal{Y}) \in S]$$

- If $\varepsilon = 0$, we have no probability loss, and an attacker cannot distinguish the two datasets
- With current and future side information and with postprocessing, the probability ratio should hold

[Dwork, McSherry, Nissim and Smith, 2006]

A toy example

—Have you engaged in drug traffic in the past week?

The respondents:

- Flip a coin
- If tails, then respond thruthfully
- If heads, then flip a second coin and respond "Yes" if heads and "No" if tails

Randomized response mechanism

$$\frac{\Pr[\text{Response} = \text{Yes} | \text{Truth} = \text{Yes}]}{\Pr[\text{Response} = \text{Yes} | \text{Truth} = \text{No}]} = \frac{3/4}{1/4} = \frac{\Pr[\text{Response} = \text{No} | \text{Truth} = \text{No}]}{\Pr[\text{Response} = \text{No} | \text{Truth} = \text{Yes}]} = 3$$

Differential Privacy

Two or more differentially private mechanism can be combined
The composition theorem bounds cumulative privacy loss over multiple analyses

The combination of k algorithms, each ε-DP, still guarantees $k\varepsilon$-DP
In general, the ε's add up (basic composition theorem)

Combining more differentially private algorithms,
we can carry out complex differentially private analyses

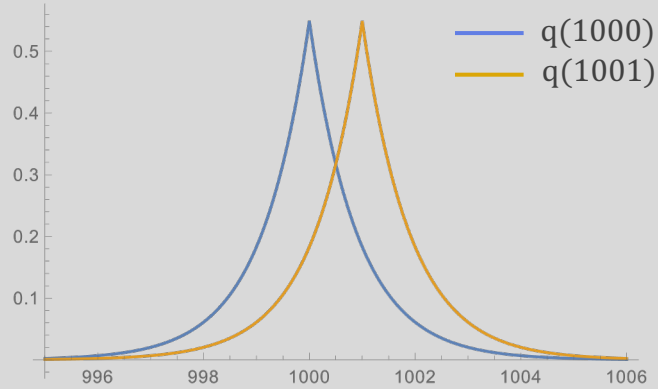
Answering multiple queries, we cannot avoid that the strength of our privacy guarantee will degrade with repeated use (e.g., convergence to the true value)

[Dwork, McSherry, Nissim and Smith, 2006]

Differential Privacy in practice

How many users in D have green eyes?

The attacker wants to know if the real count is 1000 or 1001



We publish a noisy version of the actual value

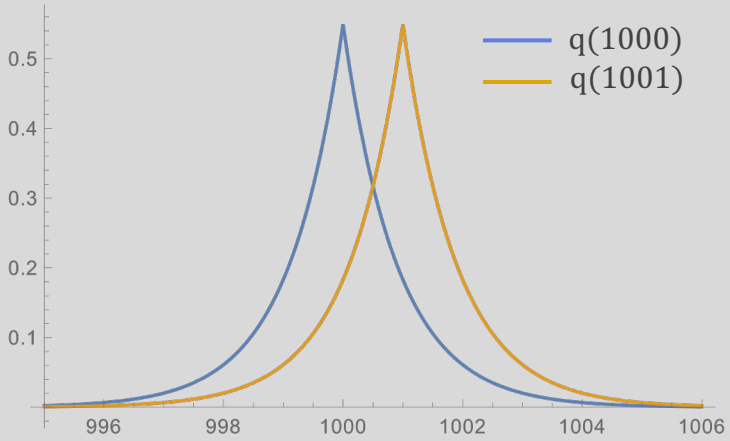
Instead of sending $q(D)$, we send $q(D) + \mathrm{Lap}\left(\frac{1}{\varepsilon}\right)$

We publish 1003

The hypothesis $q(D) = 1001$ is more likely from the point of view of the attacker

ACK: Damien Desfontaines

Differential Privacy in practice



$$Lap\left(x\middle|\frac{1}{\varepsilon}\right) = \frac{\varepsilon}{2}exp(-\varepsilon|x|)$$

$$\text{Density}(y|q(1000)) \propto exp\left(-\varepsilon\middle|y - f\left(q(1000)\right)\middle|\right)$$

$$Pr[\mathcal{M}(\mathcal{X}) \in S] \leq e^{\varepsilon} Pr[\mathcal{M}(\mathcal{Y}) \in S]$$

Differential Privacy requires the ratio between $\text{Density}(y|q(1001))$ and $\text{Density}(y|q(1000))$ to be limited between e^{ε} and $e^{-\varepsilon}$

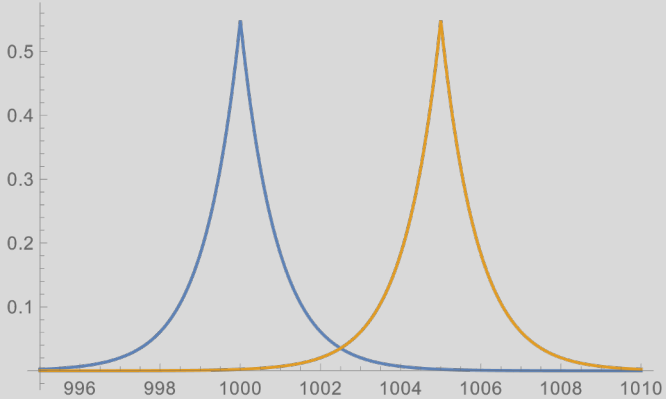
ACK: Damien Desfontaines

Differential Privacy in practice

?

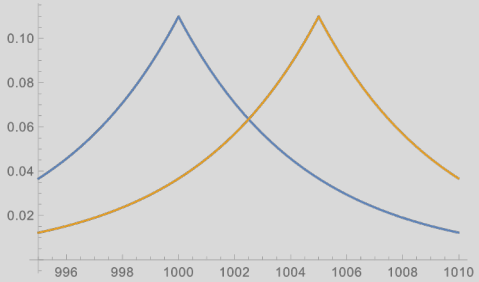
How many complaints have been sent?

The attacker has uncertainty about just one user in the dataset and wants to know how many complaints she sent



A user have sent 5 complaints

Using a noise $\mathrm{Lap}\left(\frac{1}{\varepsilon}\right)$, we would have the two curves at a large distance $(5\varepsilon\text{-DP})$



We should take into account the maximum contribution of a user

Using a noise $\mathrm{Lap}\left(\frac{5}{\varepsilon}\right)$, we would obtain ε-DP

ACK: Damien Desfontaines

Sensitivity in Differential Privacy

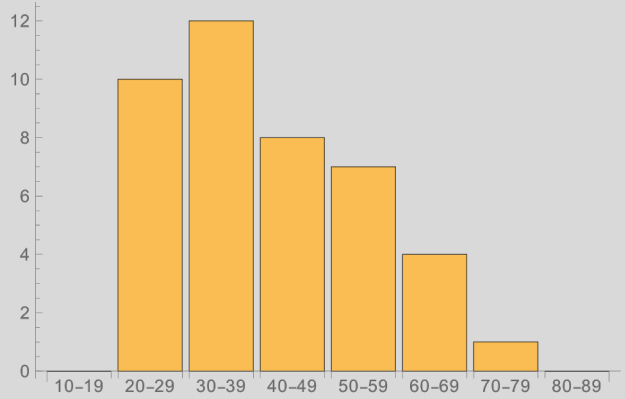
We need to know the **maximum contribution** of each user/element

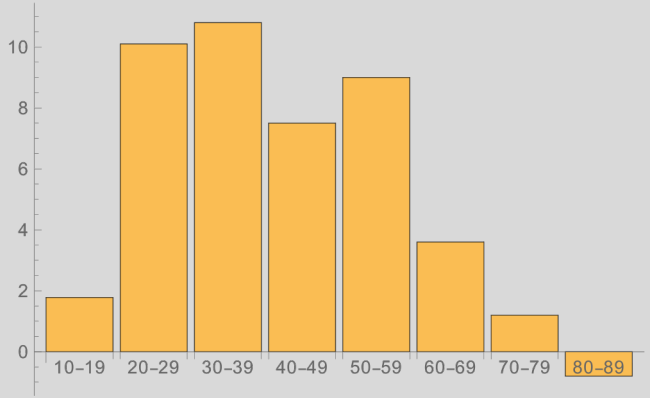
Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two **adjacent datasets**
The sensitivity Δf of a function f is
$\Delta f = max||f(\mathcal{D}_1) - f(\mathcal{D}_2)||$

Changing **one entry** in the dataset, the output of the function changes at most by Δf

In some cases, we need to **clamp the values to limit the sensitivity**, e.g., when the answer is a sum or an average
(but be careful to avoid bias in data)

Sensitivity in Differential Privacy





Release the number of users in each age range

Applying $\mathrm{Lap}\left(\frac{1}{\epsilon}\right)$ noise to each count,
the released histogram is ϵ-DP

ACK: Damien Desfontaines

Composition in Differential Privacy

Releasing multiple statistics that can be all affected by a single data point requires to "compose" the DP algorithms

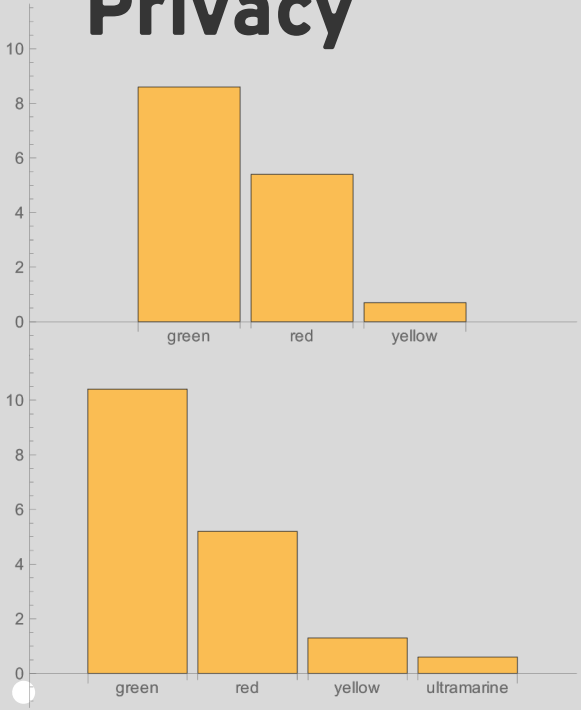
We ask the system how many users:

- are older than 35
- have visited a certain webpage
- live in the Netherlands

To make the final answer ε-DP, each answer should be $\frac{\varepsilon}{3}$-DP (noise x3)!

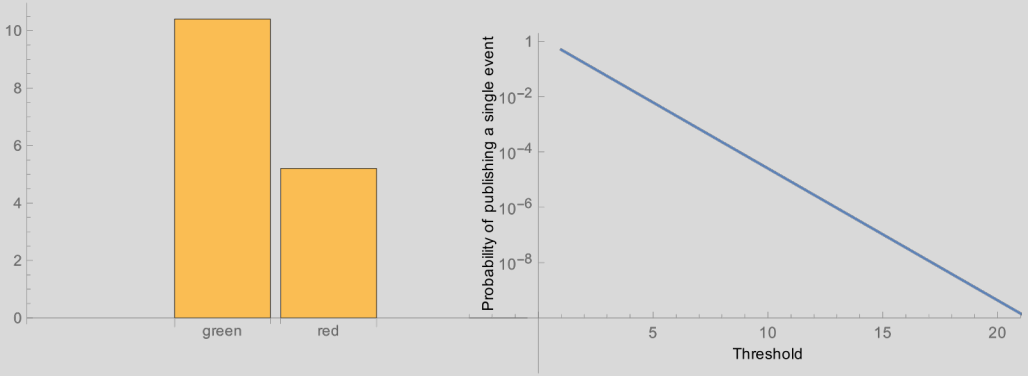
The sum of the ε_i of all the answers corresponds to the final privacy budget

(Almost) Differential Privacy



A collection of answers to an open-ended question can reveal which is the real dataset (distinguishing event)

Applying noise and threshold to drop rare categories, we reduce the odds of a distinguishing event



ACK: Damien Desfontaines

(Almost) Differential Privacy

We end up with the definition of (ε, δ)-DP

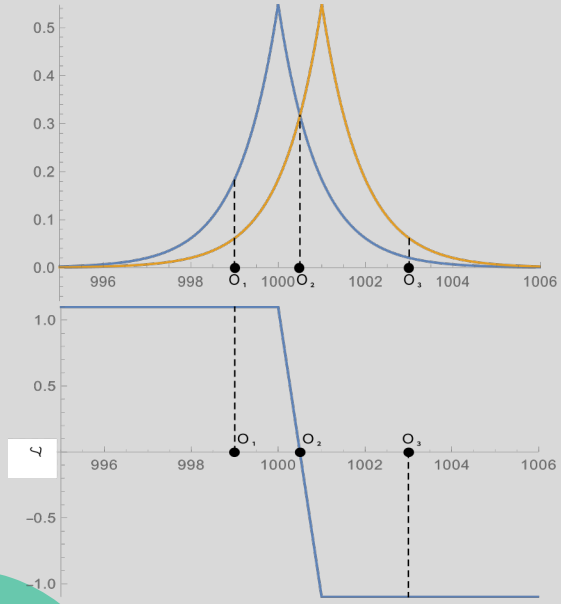
$$P(A(\mathcal{D}_1) \in S) \leq e^{\varepsilon} P(A(\mathcal{D}_2) \in S) + \delta$$

δ captures the odds that something goes wrong

δ is a cryptographically small probability,
smaller than the chance of an asteroid flying into the earth,
it's the probability of a disaster

Privacy Loss

$$\mathcal{L}(O) = ln\left(\frac{P(A(\mathcal{D}_1) = O)}{P(A(\mathcal{D}_2) = O)}\right)$$



The Privacy Loss random variable measures the actual value of ε given the output O, i.e. the knowledge gain of the attacker

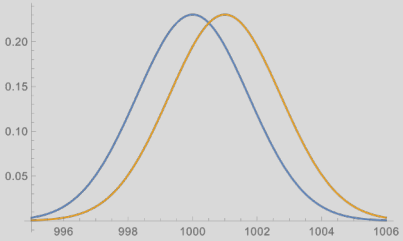
- With O_1, we increase by three times the knowledge of the attacker
- With O_3, we decrease by three times the knowledge of the attacker
- With O_2, the knowledge does not change

The limited loss proves that Laplace noise actually gives ε-DP

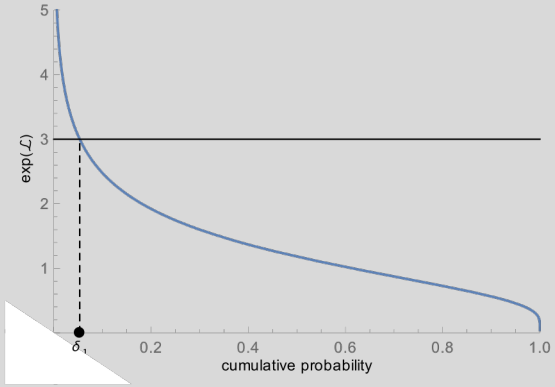
ACK: Damien Desfontaines

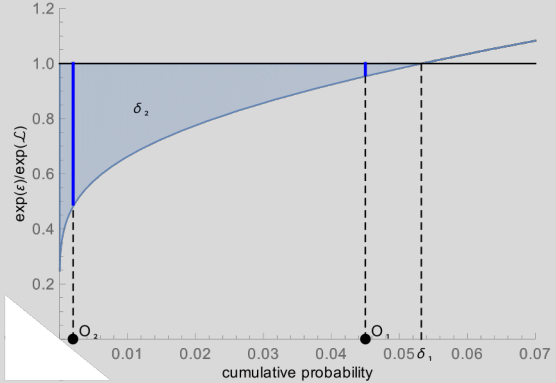
Privacy Loss

$$\mathcal{L}(O) = ln\left(\frac{P(A(\mathcal{D}_1) = O)}{P(A(\mathcal{D}_2) = O)}\right)$$



What happens if A applies a gaussian noise?





There's a probability for the loss to go above our ε threshold: that's the "terrible event"

Not all the terrible events are so terrible!
The blue area (losses weighted by their probabilities) represents the parameter δ

ACK: Damien Desfontaines

Why Gaussian noise?

User	Item
001	War and Peace
001	Ulysses
002	Ulysses
003	Pride and Prejudice

We want to release the number of users who bought each book (the catalog has 1,000 books), e.g. in a vector of 1,000 elements

Users can affect a single element once, but can affect all of them

We want to protect the presence of a user

We could add $\mathrm{Lap}\left(\frac{1000}{\varepsilon}\right)$ to each element of the vector

Intuitively, the sensitivity of the function producing the vector is 1,000

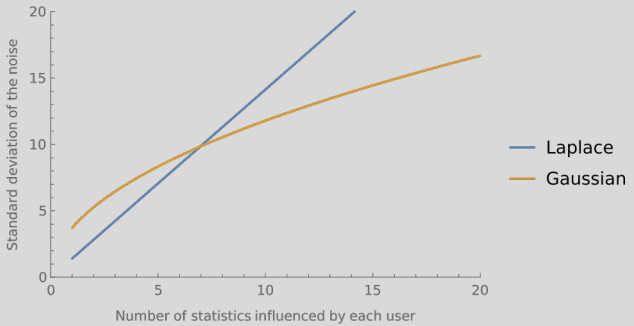
Why Gaussian noise?

So far we intuitively used the Manhattan distance to estimate the sensitivity (L1 sensitivity)

$$\Delta(f) = \max_{\mathcal{D}_1, \mathcal{D}_2} \sum_{i=1}^{1000} |f_i(\mathcal{D}_1) - f_i(\mathcal{D}_2)|$$

Gaussian noise needs to be scaled by Euclidean distance (L2 sensitivity)!

$$\Delta(f) = \sqrt{\max_{\mathcal{D}_1, \mathcal{D}_2} \sum_{i=1}^{1000} |f_i(\mathcal{D}_1) - f_i(\mathcal{D}_2)|}$$



e.g., $\varepsilon = 1, \delta = 10^{-5}$

The L2 sensitivity grows much more slowly than the L1 sensitivity!

ACK: Damien Desfontaines

The Exponential Mechanism

Sometimes we have queries whose answers:

- are categorical
- are real numbers where noise would destroy its value (return a precise value)
- are returned by functions with very high sensitivity

- Exponential mechanism selects the "best" element
- Utility function of database/output pairs with global sensitivity Δu

  $u\colon \mathbb{N}^{|\mathcal{D}|} \times \mathcal{R} \to \mathbb{R}$
- Outputs each element of $\mathcal{R}$ with probability proportional to $exp(\varepsilon u(d, r)/2\Delta u)$
- Guarantees ε-differential privacy

A little bit of practice

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
A		X	X				X
B			X	X		X	
C				X			
D					X	X	X
E		X	X	X	X	X	X
F			X	X			
G			X	X		X	
H	X	X	X		X		

How to pick a good day?

for each day d:

    count[d] = # people avail. on d

return argmax(count)

No privacy!

A little bit of practice

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
A		X	X				X
B			X	X		X	
C				X			
D					X	X	X
E		X	X	X	X	X	X
F			X	X			
G			X	X		X	
H	X	X	X		X		

How to pick a good day?

```
for each day d:
    count[d] = DP(# people avail. on d)
return argmax(count)
```

Use Laplacian or Gaussian mechanism

A little bit of practice

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
A		X	X				X
B			X	X		X	
C				X			
D					X	X	X
E		X	X	X	X	X	X
F			X	X			
G			X	X		X	
H	X	X	X		X		

How to pick a good day?

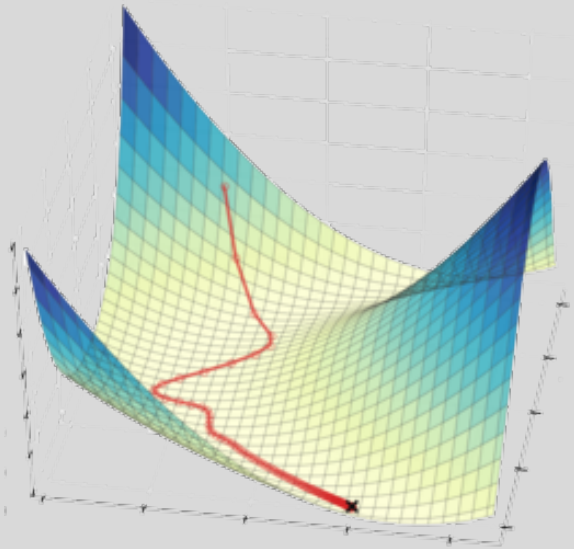
for each day d:

```
    count[d] = # people avail. on d
return DP(argmax(count))
```

Use Exponential mechanism

Mon	Tue	Wed	Thu	Fri	Sat	Sun

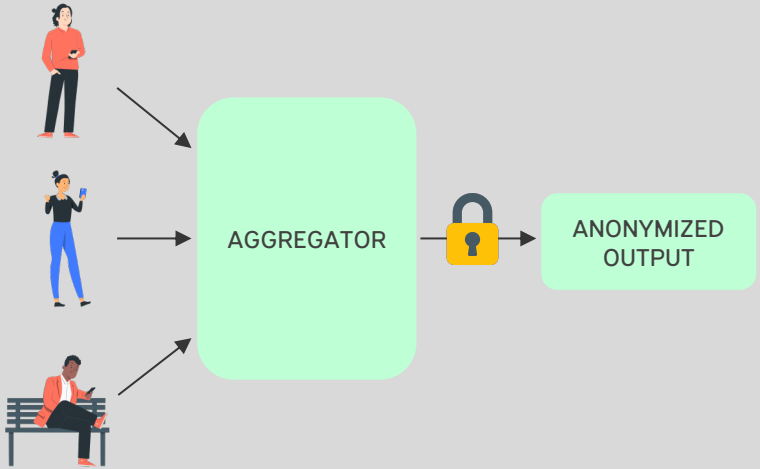
What about DP in machine learning?



- Just adding noise to the output doesn't work
  Non-convex functions are very sensitive (too much noise needed)
- Hard to characterize the dependence of final parameters on data
- Take the standard algorithm and add appropriate noise during training

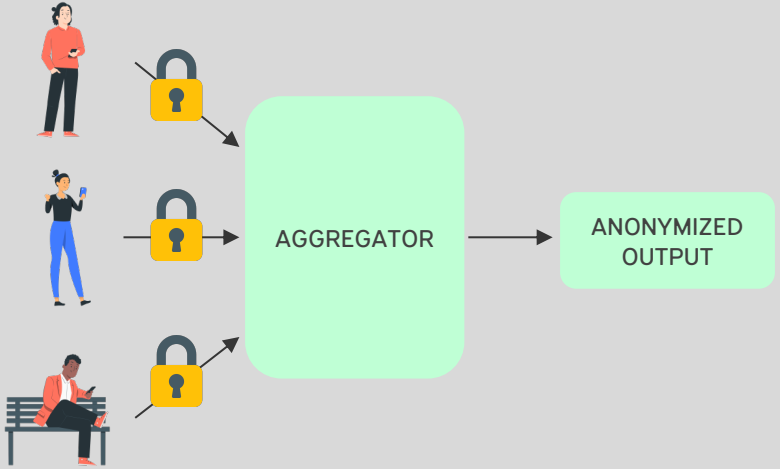
Hands-on for an example

Central DP vs. Local DP



Central Differential Privacy

Higher accuracy
Trusted aggregator
(e.g., US Census)



Local Differential Privacy

Higher noise
No trust required
(e.g., Google RAPPOR, Apple Emojis)

Differential Privacy in short

Strong privacy guarantees

No longer needed attack modeling

Quantifiable privacy loss

Composable mechanisms

Useful for analyzing any algorithm

Secure Multi-Party Computation

Compute a function jointly while keeping the inputs secret

$$y_1, \ldots, y_n = f(x_1, \ldots, x_n)$$

Some co-workers want to compute their average salary without revealing each other (or to another party) their individual salaries

- Computation distributed across multiple parties with no one seeing the other parties' data
- Can perform training and inference on encrypted data
- Generates computational and communication overhead

Secure Multi-Party Computation

(Additive) Secret Sharing

Distribute random pieces of a secret (shares) among parties

A secret share is a piece of incomplete information about the initial secret value

	Alice	Bob	Caroline
Alice (10 beers)	5 beers	3 beers	2 beers
Bob (20 beers)	-8 beers	10 beers	18 beers
Caroline (30 beers)	0 beers	35 beers	-5 beers
	-3 beers	**48 beers**	**15 beers**

← **Average?**
20 beers 😙

Secure Multi-Party Computation

Additive Secret Sharing

We can split a secret into N shares and keep it hidden as long as at most N-1 shareholders collaborate

We can sum shares of different secrets between them or sum and multiply any non-encrypted number (homomorphic addition)

Shamir's Secret Sharing

Control the minimum number R of shares needed to reconstruct a secret S (e.g. if some parties drop out)

We sample N points from a polynomial f of degree R with the condition that $f(0) = S$
Then, we can reconstruct S by knowing just R points

(Partially) supports multiplication

Secure Multi-Party Computation

SPDZ Protocol

Use of additive scheme for linear operations

Introduces multiplication:
- we need a honest crypto provider
- it shifts part of the computation in an offline preprocessing phase

The online phase (e.g., training and prediction of a model) has a speed up

Provides security for malicious parties

Additive schemes are adequate for integers: how to use them in ML?
Fixed Precision Encoding

Homomorphic Encryption

What happens if we have a limited number of participants (i.e. SMPC does not apply)?

Homomorphic Encryption allows meaningful calculations on encrypted data

- Only one party needed to encrypt and decrypt own data
- Can perform operations directly on encrypted data (without interactions)
  The result is equivalent to performing analogous operations without encryption!
- Computationally expensive
- Allows a little set of calculations

Homomorphic Encryption

$$\forall m_1, m_2 \in \mathcal{M}, \quad Enc(m_1 \odot_{\mathcal{M}} m_2) \leftarrow Enc(m_1) \odot_{\mathcal{C}} Enc(m_2)$$

In Paillier [1999], if $\odot_{\mathcal{C}}$ is a multiplication of cyphertexts we obtain the encryption of the sum

Partially Homomorphic Encryption
Can reach additive homomorphism or multiplicative homomorphism

Somewhat Homomorphic Encryption
You can apply operations for a limited number of times, since noise is used

Fully Homomorphic Encryption
Allows unlimited number of additions and multiplications over cyphertexts
(Performs a highly costly operation called bootstrap)

The final trade-off

Developing private ML needs a artful balance of efficiency-accuracy-privacy

HE and SMPC often replaceable
- HE: Little interaction and expensive computation
- SMPC: Cheap computation and significant amount of interaction

SMPC replaces computation with interaction, offering better practical performance

DP replaces accuracy with efficiency
If the coordinator is trusted, send plain data to preserve more accuracy

4

Privacy-Preserving Recommender Systems

Trending Research and
Open Challenges

Trending Researches

Years

2018-2021

Domain

PoI, Fitness, News, Movies, Web browsing, Online Shopping

Successive Point-of-Interest Recommendation With Local Differential Privacy

JONG SEON KIM[1], JONG WOOK KIM[2], (Member, IEEE), AND YON DOHN CHUNG[1], (Member, IEEE)
[1]Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea
[2]Department of Computer Science, Sangmyung University, Seoul 03016, South Korea

Years: 2021

Domain: PoI

Paradigm: Collaborative

Venue: IEEE Access

Successive Point-of-Interest Recommendation With Local Differential Privacy

JONG SEON KIM[1], JONG WOOK KIM[2], (Member, IEEE), AND YON DOHN CHUNG[1], (Member, IEEE)
[1]Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea
[2]Department of Computer Science, Sangmyung University, Seoul 03016, South Korea

Problem

- Only four location points are sufficient to identify most people

Task

- Given a user u and her check-in history at time T, we recommend a set of POIs that the user u is likely to visit at a time T+1

Paradigm

- Collaborative filtering

Successive Point-of-Interest Recommendation With Local Differential Privacy

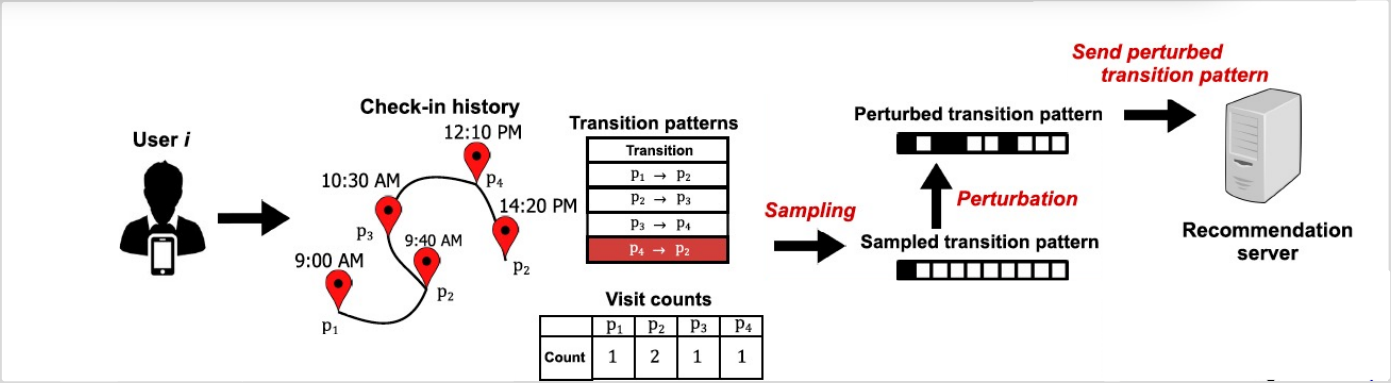
JONG SEON KIM[1], JONG WOOK KIM[2], (Member, IEEE), AND YON DOHN CHUNG[1], (Member, IEEE)
[1]Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea
[2]Department of Computer Science, Sangmyung University, Seoul 03016, South Korea

Solution

- Local Differential Privacy (DP), SPIREL
- DP - trusted data curator collects original data from user
- Local version of DP (LDP) each user perturbs her data
- Original data never leaves the device of the users

Successive Point-of-Interest Recommendation With Local Differential Privacy

JONG SEON KIM[1], JONG WOOK KIM[2], (Member, IEEE), AND YON DOHN CHUNG[1], (Member, IEEE)
[1]Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea
[2]Department of Computer Science, Sangmyung University, Seoul 03016, South Korea



- Each participant independently samples a single transition pattern between two consecutive POIs
- Perturbs the sampled transition pattern in their local device
- Submits the perturbed transition pattern to the server

Successive Point-of-Interest Recommendation With Local Differential Privacy

JONG SEON KIM[1], JONG WOOK KIM[2], (Member, IEEE),
AND YON DOHN CHUNG[1], (Member, IEEE)
[1]Department of Computer Science and Engineering, Korea University, Seoul 02841, South Korea
[2]Department of Computer Science, Sangmyung University, Seoul 03016, South Korea

Evaluation

- Gowalla, Taxi, Yelp, and Foursquare

Metrics

- HR, MRR

Results

- SPIREL provides significantly better POI recommendation performance than the existing privacy-preserving recommendation methods (based on DP)

Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, *Member, IEEE* and Chi-Yin Chow, *Senior Member, IEEE*

Years: 2021

Domain: PoI

Paradigms: Collaborative

Venue: User Modeling and User Adaptive Interaction

Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, *Member, IEEE* and Chi-Yin Chow, *Senior Member, IEEE*

Solution (twofold contribution)

- additive Markov chain (accuracy)
  - all visited locations in the check-in history, instead of only using latest visited location adopted by the first-order Markov chain

Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, *Member, IEEE* and Chi-Yin Chow, *Senior Member, IEEE*

Solution (twofold contribution)

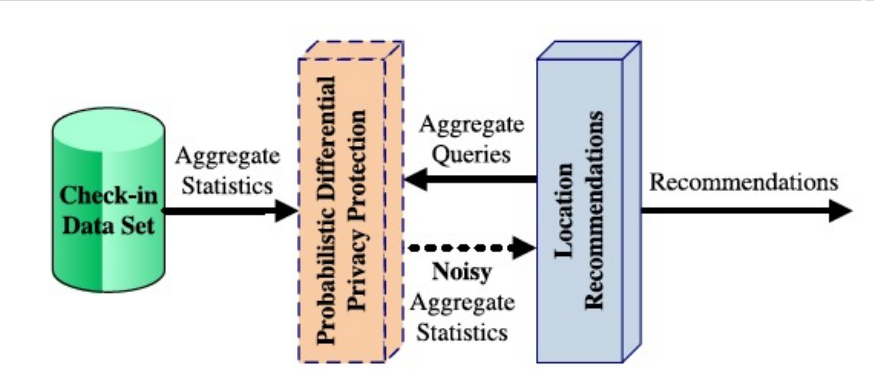
- Relaxed differential privacy (DP)
  - (ε, δ)-probabilistic DP

  δ is the maximum probability of a breach of ε-differential privacy

Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, *Member, IEEE* and Chi-Yin Chow, *Senior Member, IEEE*

- Extracts **aggregate statistics** from the check-in data set, **injects random noise** into the aggregate statistics
- Receives **aggregate queries** from the module of location recommendations
- Answers with the **noisy aggregate statistics**



Enabling Probabilistic Differential Privacy Protection for Location Recommendations

Jia-Dong Zhang, *Member, IEEE* and Chi-Yin Chow, *Senior Member, IEEE*

Evaluation

- Foursquare, Gowalla, and Brightkite

Metrics

- Accuracy, Efficiency (running time)

Results

- Extensive experimental results show that the model achieves high recommendation efficiency and accuracy, and strict location privacy

A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

Years: 2020

Domain: Fitness Trackers

Venue: User Modeling and User-Adapted Interaction

A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

Problem

- Support users in defining their privacy settings

How

- PDM (Personal datamanager) is e an intermediary between the user, her devices through their dedicated third parties, and other third parties (i.e., fourth parties) that want access to her data

A recommendation approach for user privacy preferences in the fitness domain

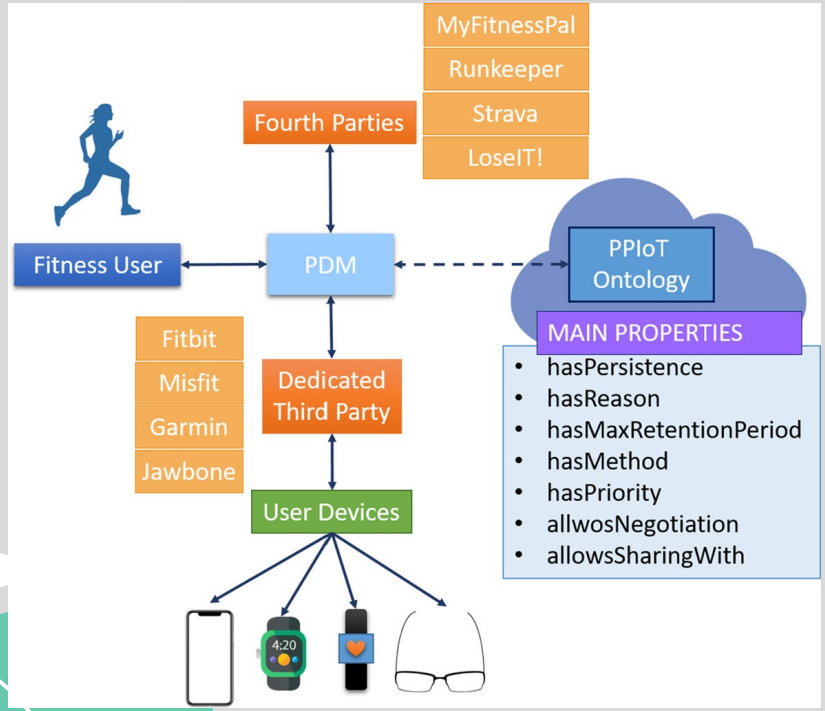
Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

Contributions

- Formally definition of user's privacy preferences
- Definition of privacy-profiles
- Prediction of privacy profiles
- Privacy profile recommendations

A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

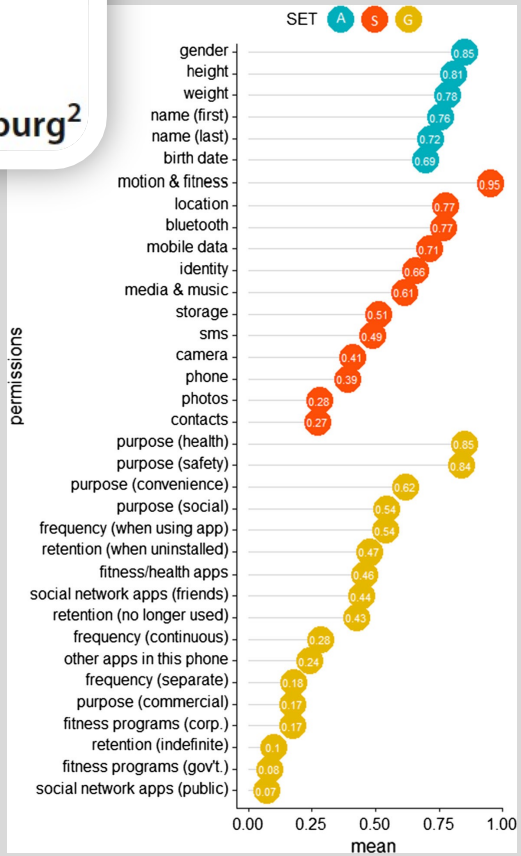


- PPIoT ontology fills a gap in privacy preference description for the IoT by combining and extending existing ontologies for privacy preference and for IoT
- It addresses the GDPR requirements

A recommendation approach for user privacy preferences in the fitness domain

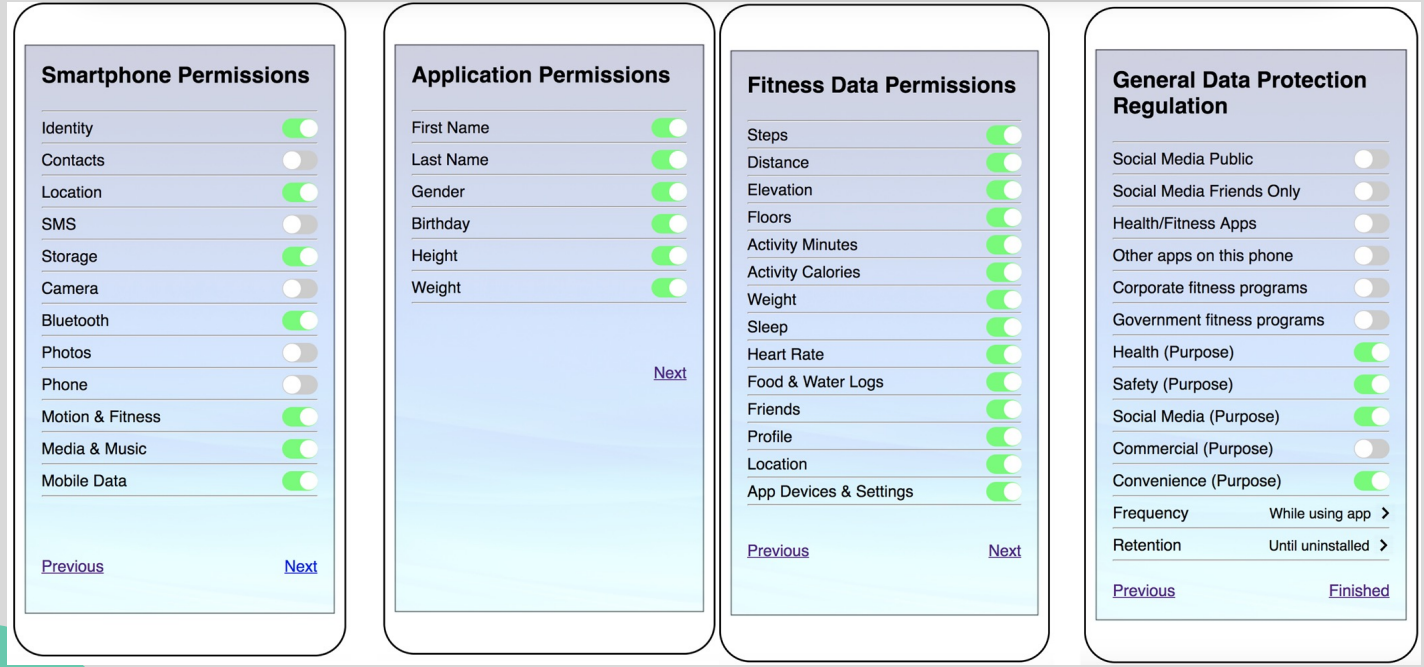
Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

- Demographics have a high disclosure rate
- Participants are more likely to allow motion, location, Bluetooth, and mobile data
- Photographs or contacts is granted much less often
- Open to data collection for health and safety



A recommendation approach for user privacy preferences in the fitness domain

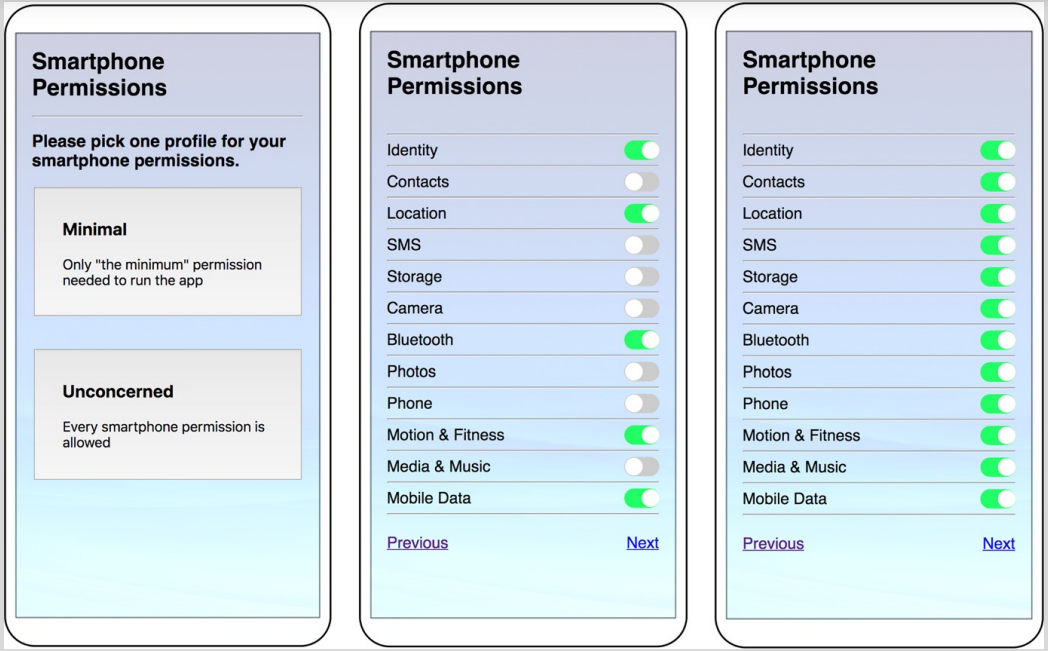
Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]



Smart Settings

A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]



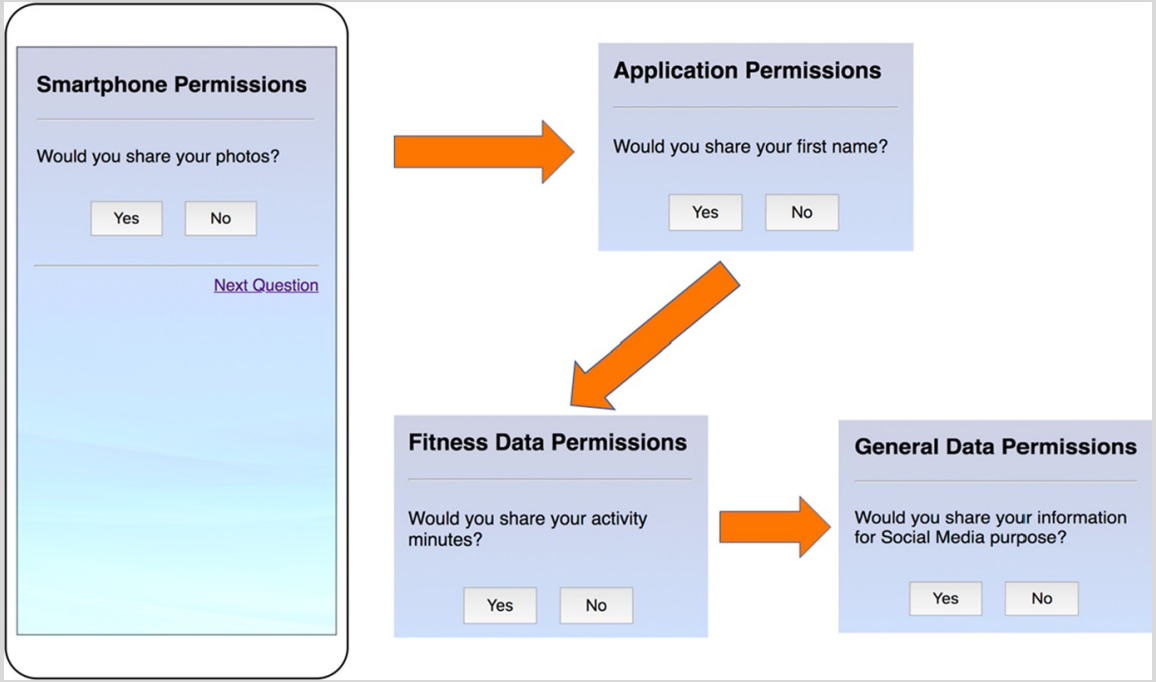
S set subprofiles

The «minimal» subprofile

The «unconcerned» subprofile

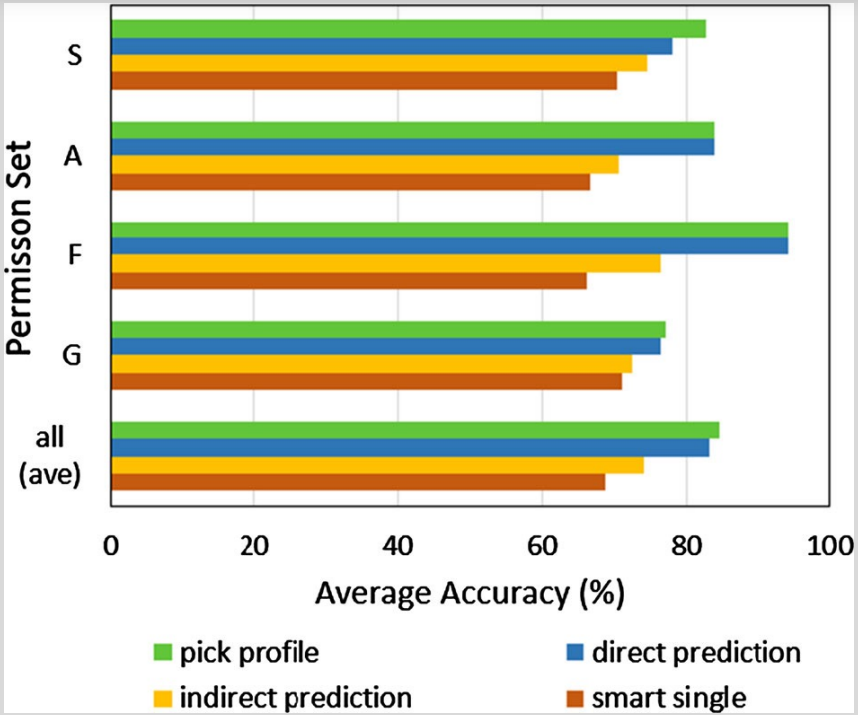
A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]



A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]



A recommendation approach for user privacy preferences in the fitness domain

Odnan Ref Sanchez[1] · Ilaria Torre[1] · Yangyang He[2] · Bart P. Knijnenburg[2]

Lesson learned

Users care more about “WHO” will receive that data rather than “WHAT” data are shared specifically

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

Years: 2020

Domain: Movies, Web Browsing

Paradigms: Content-based

Venue: Expert Systems With Application

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

Problem

- Privacy
  - service providers hold a database with information about all of their users
- Partial view
  - provider can rely only on data that were collected by its service itself
- The authors criticize the two main approaches:
  - perturbation (lower-quality recommendations)
  - cryptography (high communication and computational costs)

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

Solution

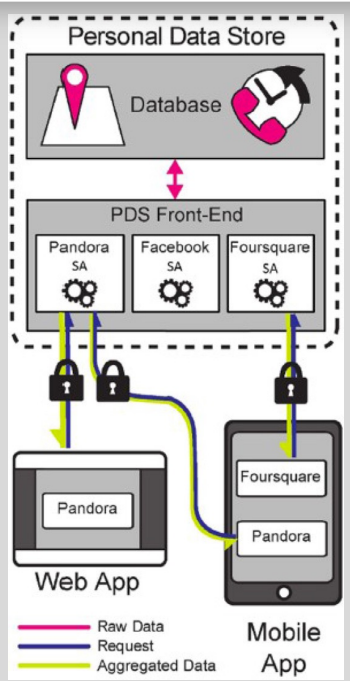
- a novel architecture for recommender systems based on Open Personal Data Store (openPDS)

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

- The OpenDS Architecture
  - storing personal data in a privacy preserving way
  - user's data stored in a central location, managed and controlled by the user



A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

The proposed idea

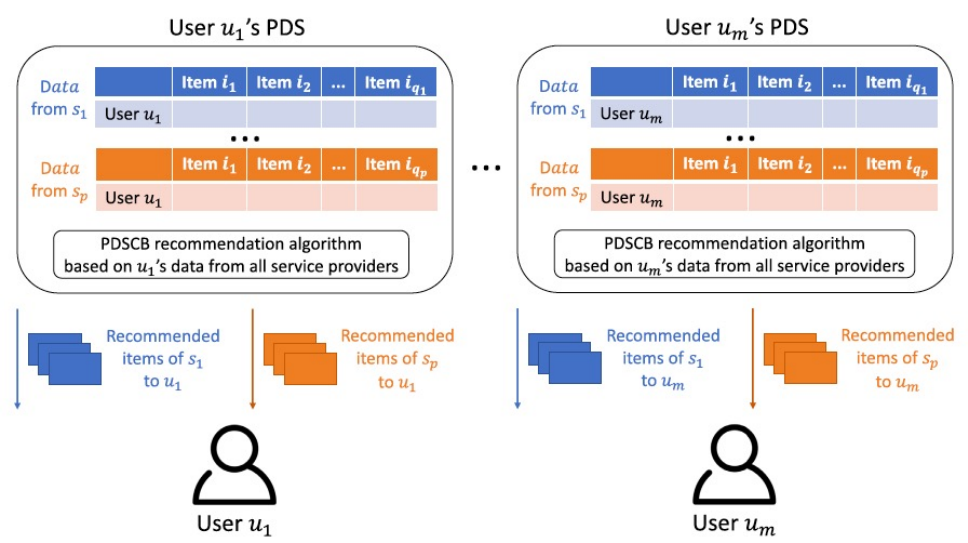
- a new recommendation algorithm more suitable to the PDS case, named PDS-inspired content-based (PDSCB)
- shifting from a collaborative filtering recommender system to a content-based one
- exploiting the user's data from multiple service providers, enabled by openPD

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

The proposed idea



A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

Evaluation

- movies and web browsing
- compare performance with popular non-privacy-aware collaborative-filtering algorithms

A personal data store approach for recommender systems: enhancing privacy without sacrificing accuracy

Itzik Mazeh, Erez Shmueli*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, Israel

Results

- the proposed privacy enhancing model is able to perform as well as or better than the common collaborative filtering alternatives
- improves its performance using additional service providers' data

Privacy-Preserving News Recommendation Model Learning

Tao Qi[1], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1] and Xing Xie[2]
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100080, China
{taoqi.qt,wuchuhan15}@gmail.com yfhuang@tsinghua.edu.cn
{fangzwu,xing.xie}@microsoft.com

Years: 2020

Domain: News

Paradigms: Collaborative with Federated Learning

Venue: EMNLP 2020

Privacy-Preserving News Recommendation Model Learning

Tao Qi[1], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1] and Xing Xie[2]
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100080, China
{taoqi.qt,wuchuhan15}@gmail.com yfhuang@tsinghua.edu.cn
{fangzwu,xing.xie}@microsoft.com

Problem

Users' behaviors on news websites and Apps are privacy-sensitive, the leakage of which may bring catastrophic consequences

Privacy-Preserving News Recommendation Model Learning

Tao Qi[1], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1] and Xing Xie[2]
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100080, China
{taoqi.qt,wuchuhan15}@gmail.com yfhuang@tsinghua.edu.cn
{fangzwu,xing.xie}@microsoft.com

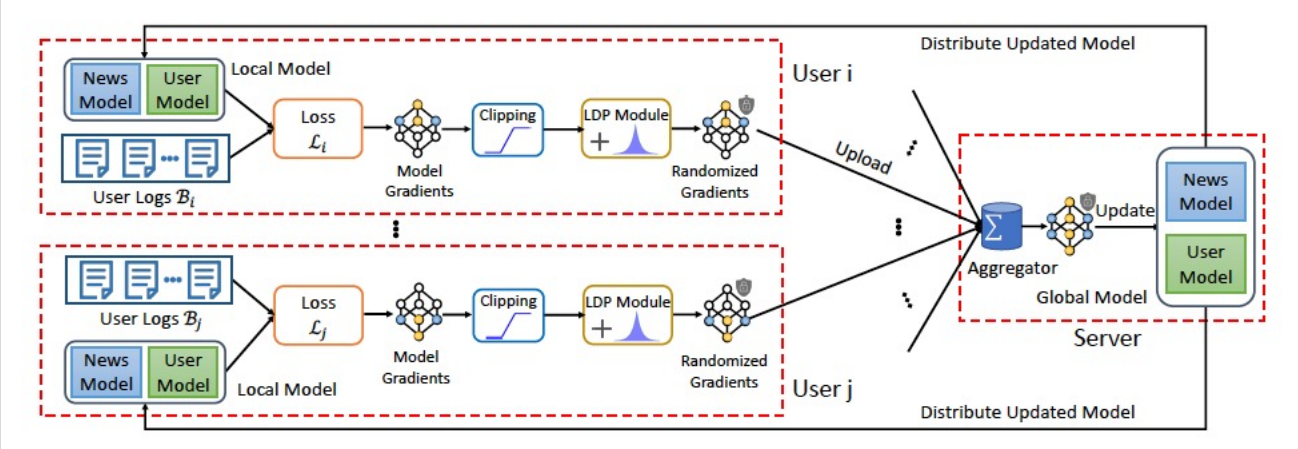
Solution

- A privacy-preserving method for news recommendation model training based on federated learning, where the user behavior data is locally stored on user devices

Privacy-Preserving News Recommendation Model Learning

Tao Qi[1], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1] and Xing Xie[2]
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100080, China
{taoqi.qt,wuchuhan15}@gmail.com yfhuang@tsinghua.edu.cn
{fangzwu,xing.xie}@microsoft.com

Solution



Privacy-Preserving News Recommendation Model Learning

Tao Qi[1], Fangzhao Wu[2], Chuhan Wu[1], Yongfeng Huang[1] and Xing Xie[2]
[1]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[2]Microsoft Research Asia, Beijing 100080, China
{taoqi.qt,wuchuhan15}@gmail.com yfhuang@tsinghua.edu.cn
{fangzwu,xing.xie}@microsoft.com

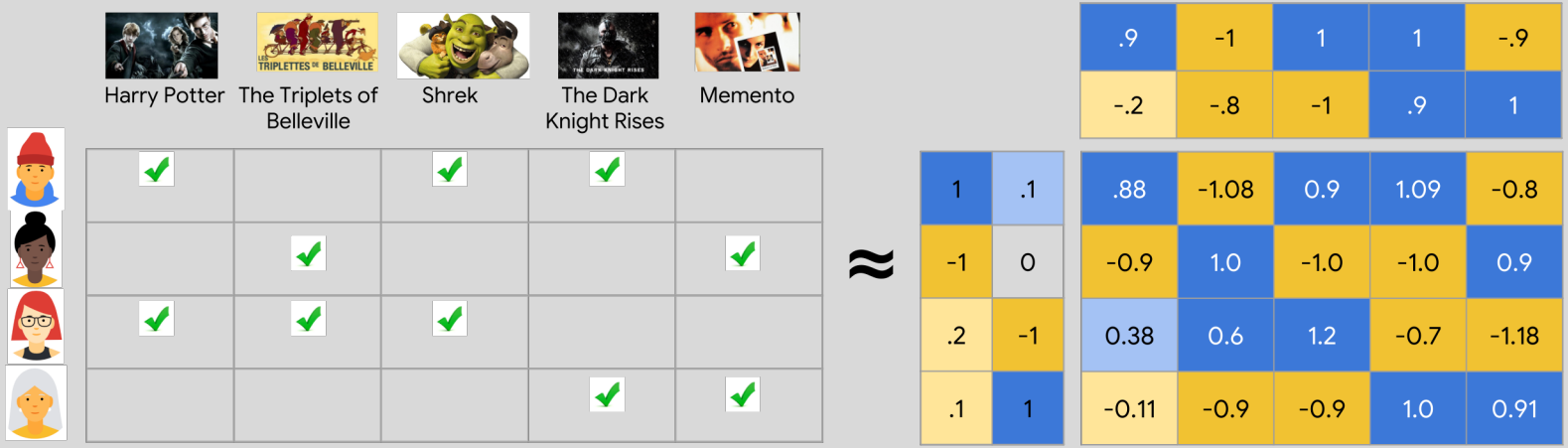
Evaluation

- Adressa, MSN-News (from Microsoft)

Results

- The proposed method can achieve comparable performance with SOTA news recommendation methods (FM, DFM, etc.), and meanwhile can better protect user privacy

Privacy-Preserving Recommender Systems: focus on Matrix Factorization



https://developers.google.com/machine-learning/recommendation/collaborative/matrix

Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

Year: 2018

Venue: AAAI

Domain: Point-of-Interest

Paradigm: Decentralized Learning

Core ideas

Which user should be communicated under DMF framework?

- Nearby User Communication

How far should users communicate with their neighbors?

- Random walk based decentralized training technique

What information should users communicate with each other?

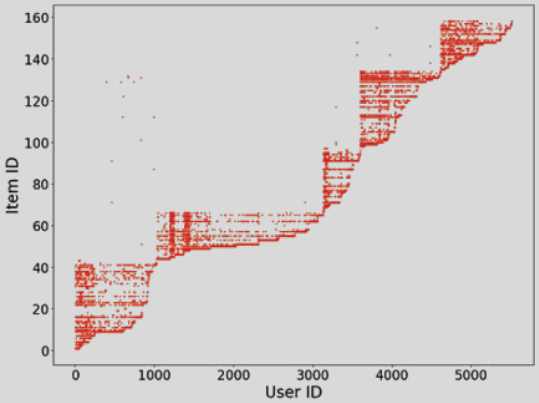
- common and personal embeddings

Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

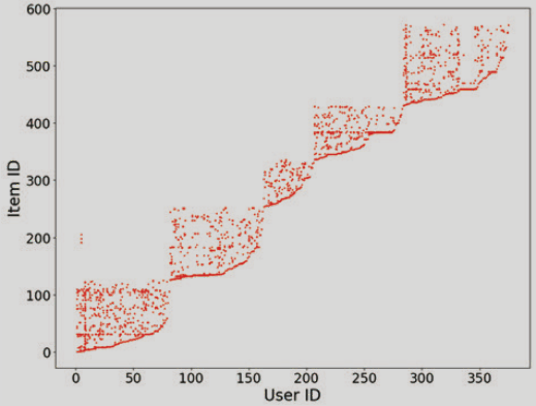
Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

Nearby Users

Foursquare



Alipay



Maintaining all the low rank matrices (and sensitive rating data) for training makes no sense!

Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

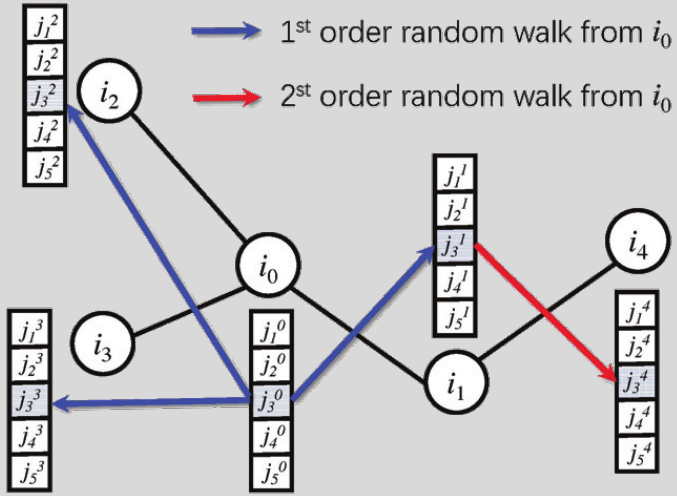
Decentralized Random walk

Adjacency matrix representing the communication graph among users

- how far should users communicate with their neighbors becomes the second challenging question.

Trade-off between decentralization and communication/computation cost:

- the further the communication is, the more users can collaborate
- the more communication and computation need to be done
- **Solution: random walk theory**



Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

Nearby User Communication:
What information should users communicate each other?

Goal: Protection of the value of the rating

For each user, the j-th item latent factor can be decomposed as follows:

$$v_j^i = p_j + q_j^i$$

Where p_j is a global latent factor, while q_j^i is a personal (local) latent factor.

Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

Nearby User Communication:
What information should users communicate each other?

Since p_j is also saved on one's device, each user has its own p_j^i

Therefore, inspired by [Yan et al. 2013] they propose to send the gradient of the loss w.r.t. p_j^i to each of one's neighbor to collaboratively learn p_j

[Yan et al. 2013] Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties. TKDE 25(11):2483–2493.

Privacy Preserving Point-of-Interest Recommendation Using Decentralized Matrix Factorization

Chaochao Chen, Ziqi Liu, Peilin Zhao,* Jun Zhou, Xiaolong Li
AI Department, Ant Financial Services Group, China
{chaochao.ccc, ziqiliu, peilin.zpl, jun.zhoujun, xl.li}@antfin.com

Meta Matrix Factorization for Federated Rating Predictions

Yujie Lin
Shandong University
Qingdao, China
yu.jie.lin@outlook.com

Pengjie Ren*
University of Amsterdam
Amsterdam, The Netherlands
p.ren@uva.nl

Zhumin Chen
Shandong University
Qingdao, China
chenzhumin@sdu.edu.cn

Zhaochun Ren
Shandong University
Qingdao, China
zhaochun.ren@sdu.edu.cn

Dongxiao Yu*
Shandong University
Qingdao, China
dxyu@sdu.edu.cn

Jun Ma
Shandong University
Qingdao, China
majun@sdu.edu.cn

Maarten de Rijke
University of Amsterdam & Ahold Delhaize
Amsterdam, The Netherlands
m.derijke@uva.nl

Xiuzhen Cheng
Shandong University
Qingdao, China
xzcheng@sdu.edu.cn

Year: 2020

Venue: SIGIR

Domain: Movies, Multidomain (Douban, Ciao)

Paradigm: Federated Learning

Core ideas

- Collaborative Memory
- Meta Recommender
- Rise-dimensional generation

Meta Matrix Factorization for Federated Rating Predictions

Yujie Lin
Shandong University
Qingdao, China
yu.jie.lin@outlook.com

Pengjie Ren*
University of Amsterdam
Amsterdam, The Netherlands
p.ren@uva.nl

Zhumin Chen
Shandong University
Qingdao, China
chenzhumin@sdu.edu.cn

Zhaochun Ren
Shandong University
Qingdao, China
zhaochun.ren@sdu.edu.cn

Dongxiao Yu*
Shandong University
Qingdao, China
dxyu@sdu.edu.cn

Jun Ma
Shandong University
Qingdao, China
majun@sdu.edu.cn

Maarten de Rijke
University of Amsterdam & Ahold Delhaize
Amsterdam, The Netherlands
m.derijke@uva.nl

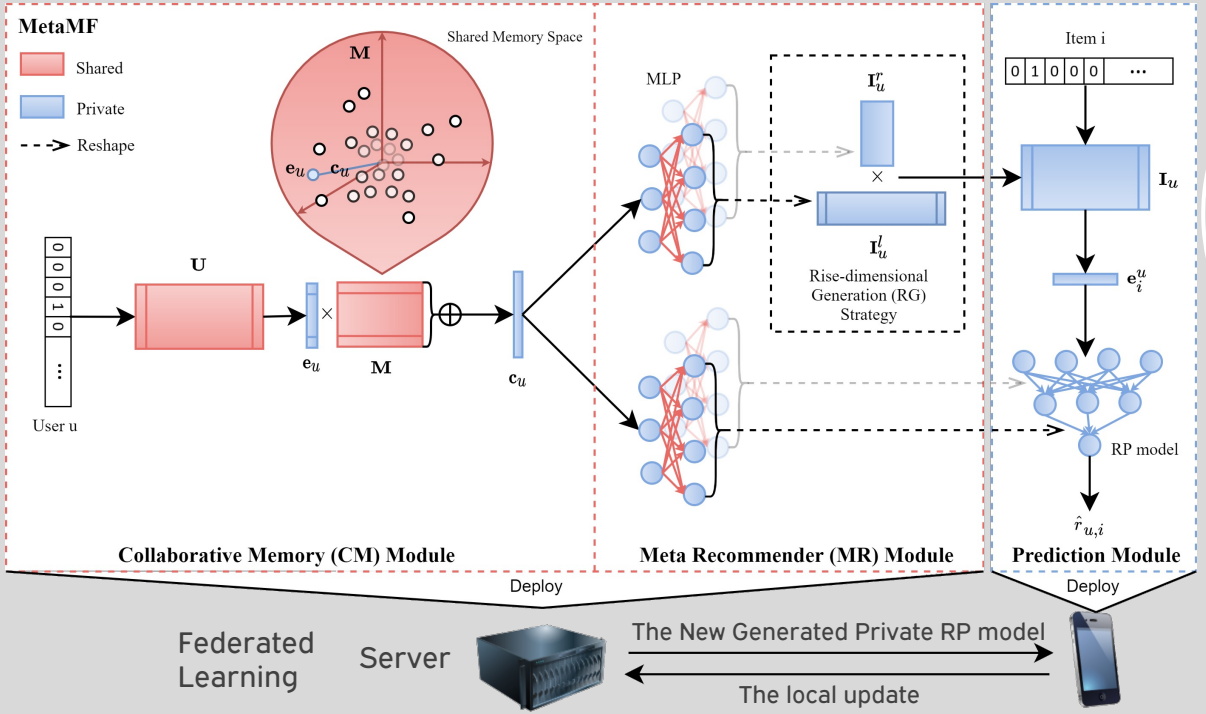
Xiuzhen Cheng
Shandong University
Qingdao, China
xzcheng@sdu.edu.cn

Model Overview - Collaborative Memory

Collaborative Memory

Meta Recommender

Prediction Module



Model Overview - Collaborative Memory

Collaborative Memory

User embedding:

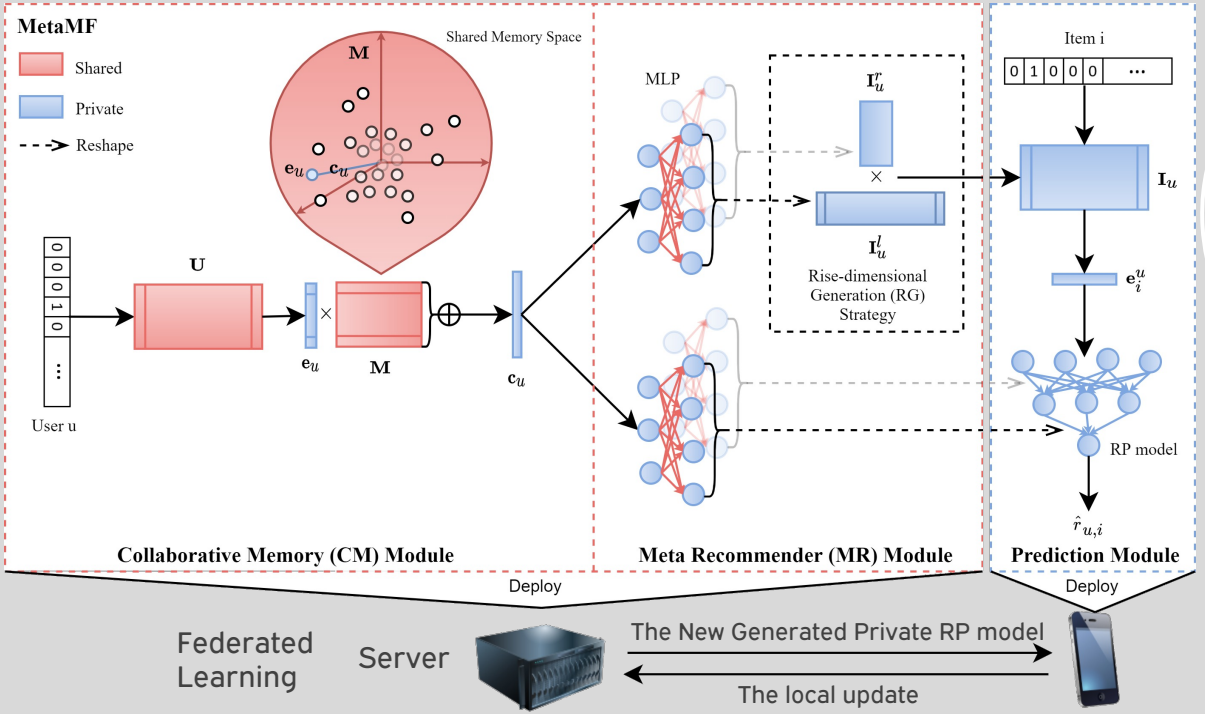
$$e_u \in R^{d_u}$$

Shared memory Matrix:

$$M \in R^{d_u \times m}$$

Collaborative vector:

$$c_u = \sum_i M(i,:)e_u(i)$$



Model Overview – Meta Recommender

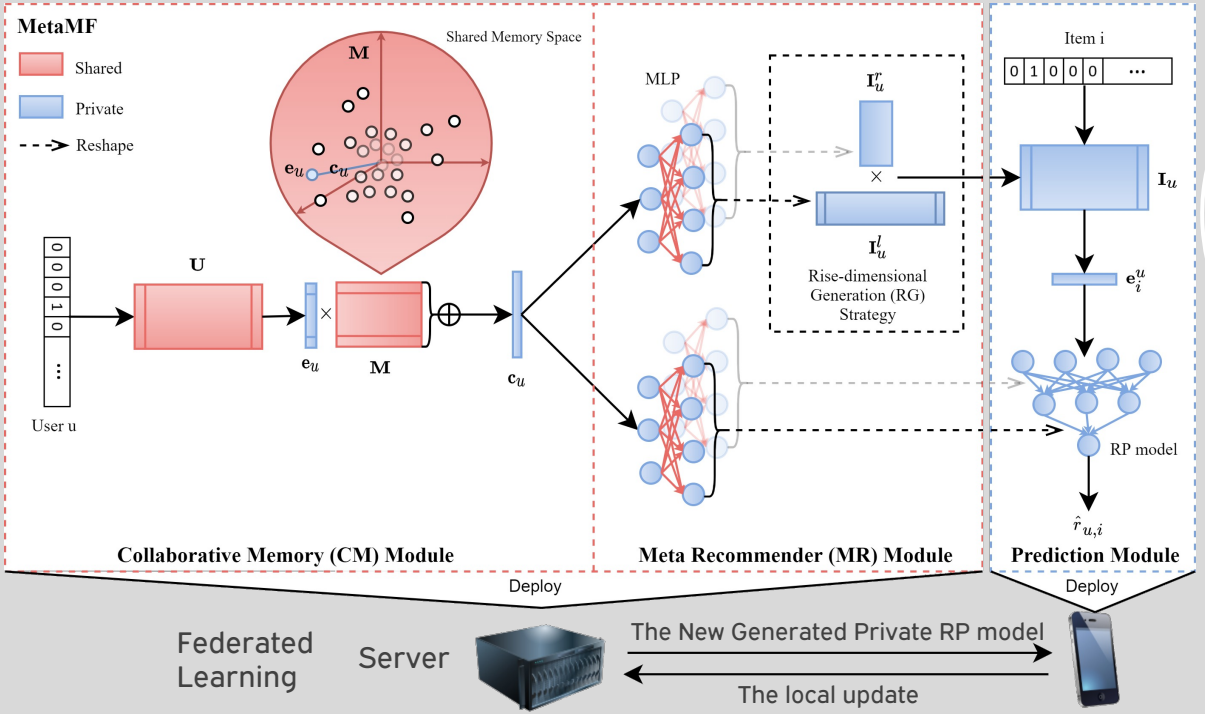
Meta Recommender

Private item embedding Matrix:
$I_u \in R^{d_i \times n}$

Low-dimensional item embedding Matrix:
$I_u^l \in R^{s \times n}$

Rise-dimensional item embedding Matrix:
$I_u^r \in R^{d_i \times s}$

MLPs to learn these matrices

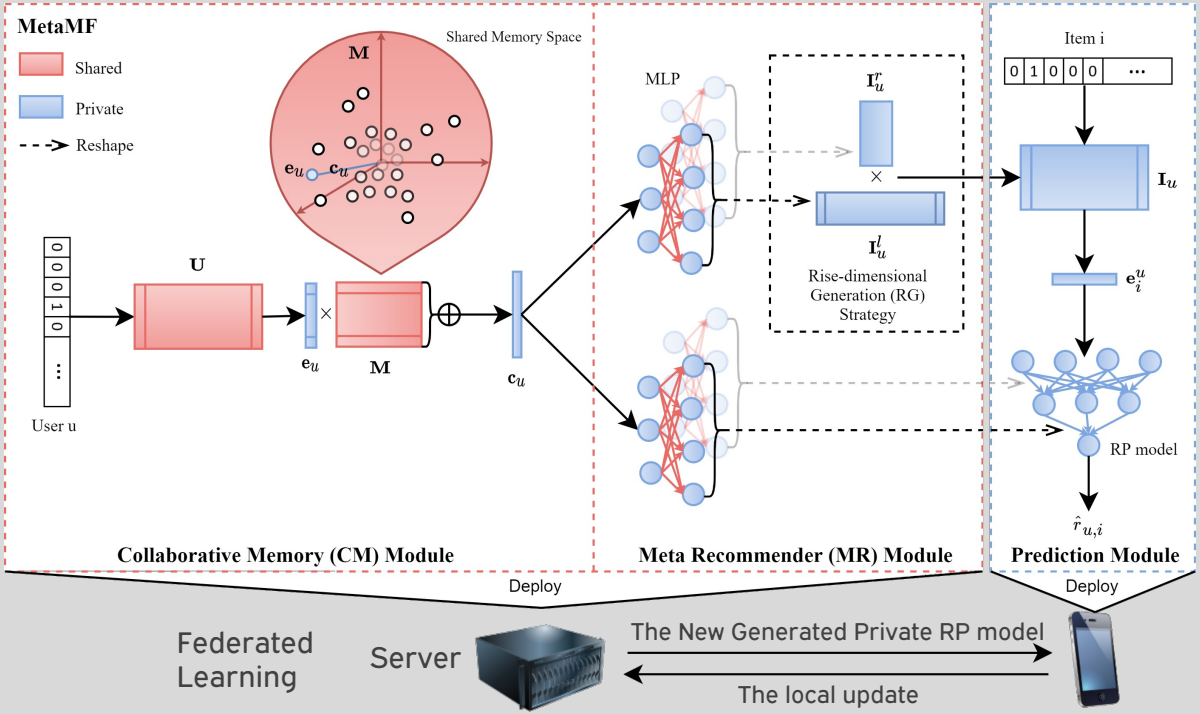


Model Overview – Prediction Module

The prediction module estimates the user's rating for a given item i using the **generated item embedding matrix I_u and RP model**

RP model:

- Private Rating Prediction Model, generated for each user
- MLP as predicting model



Secure Federated Matrix Factorization

Di Chai*, Leye Wang*, Kai Chen, and Qiang Yang, *Fellow, IEEE*

* *equal contribution, ranked alphabetically*

Year: 2020

Venue: IEEE Intelligent Systems

Domain: Movies

Paradigm: Federated Learning

Core ideas

- Gradient exchange does not preserve privacy
- Introduction of homomorphic encryption and FedMF Federated Matrix Factorization

Secure Federated Matrix Factorization

Di Chai*, Leye Wang*, Kai Chen, and Qiang Yang, *Fellow, IEEE*

** equal contribution, ranked alphabetically*

Gradients Leak Information

Previous works argued that gradients' exchange is sufficient to keep user's ratings private.

That is wrong! The authors derive the following equations:

$$\frac{G_{jk}^{t}}{u_{ik}^{t}} - \frac{G_{jk}^{t+1}}{u_{ik}^{t} + \frac{\alpha_k}{u_{ik}^{t}}} = \frac{u_{ik}^{t}}{G_{jk}^{t}}\beta_j + \frac{G_{jk}^{t}}{u_{ik}^{t}}\gamma_j$$

$$r_{ij} = \frac{G_{jk}^{t}}{u_{ik}^{t}} + \sum_{m=1}^{D} u_{im}^{t}\, v_{jm}^{t}$$

where **v** are item factors, **G** are gradients, and **Alpha, Beta**, and **Gamma** can be computed using item latent factors and gradients.

Therefore, we can compute u (user latent factor), and then the rating!

Secure Federated Matrix Factorization

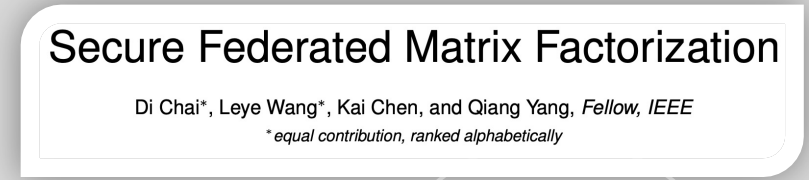
Di Chai*, Leye Wang*, Kai Chen, and Qiang Yang, *Fellow, IEEE*

*equal contribution, ranked alphabetically

Homomorphic Encryption and FedMF

Homomorphic encryption has been already introduced. Here we are interested in how to build FedMF (Federated Matrix Factorization) using it

1. **The server encrypts** item profile using public key, getting the ciphertext CV. From now on, the latest CV is prepared for all users' download.
2. **Each user downloads** the latest CV from the server, and decrypts it using secret key, getting the plain- text of V. **V is used to perform local update and compute the gradient G**. Then **G is encrypted** using public key, getting ciphertext CV . Then a TLS/SSL secure channel is built, CV is sent back to the server via this secure channel.
3. After receiving a user's encrypted gradient, **the server updates the item profile** and prepares the next ciphertext.
4. Step 2 and 3 are iteratively executed until convergence.



Major achievements

Security against Server: only ciphertext is sent to the server in FedMF. **So no bit of information will be leaked** to the server as long as homomorphic encryption system is in operation.

No Accuracy Decline: FedMF is accuracy **equivalent to** the **user-level distributed matrix factorization**. This is because the parameter updating process is the same as the distributed matrix factorization if the homomorphic encryption part is removed.

Secure Federated Matrix Factorization

Di Chai*, Leye Wang*, Kai Chen, and Qiang Yang, *Fellow, IEEE*

**equal contribution, ranked alphabetically*

Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]
[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Year: 2019

Venue: IJCAI

Domain: Movies

Paradigm: Centralized Learning

Core ideas

- Prediction-only model
- Fine-tuning almost infeasible when applying Homomorphic encryption
- Quantization method to boost encrypted operations

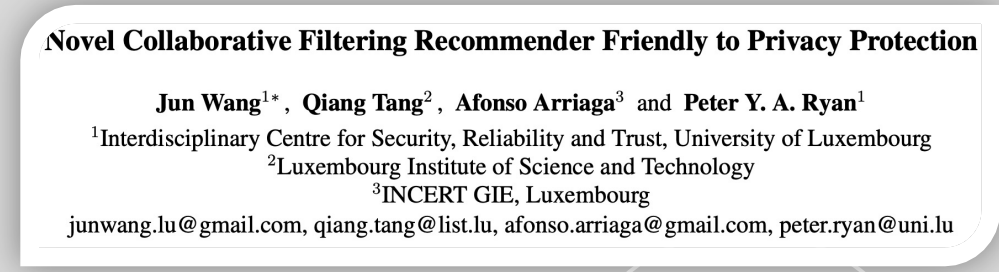
Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]
[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Privacy-preserving CryptoRec Protocols

Fast-Mode Prediction Protocol

- Pre-trained model
- Three-step protocol:
  - The **Client encrypts rating vector** (and sends it to the server)
  - The Server exploits it to **create encrypted predictions and send them back to the client**
  - The **Client decrypts** and use the predictions



Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]

[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Privacy-preserving CryptoRec Protocols

Accurate-Mode Prediction Protocol (full)

- Pre-trained model + fine tuning
- Three-step protocol:
  - The **Client encrypts** rating vector and indication vector (and sends it to the server)
  - The **Server exploits them to update** the model (the new model will be an encrypted model), create **encrypted predictions and send them back to the client**
  - The **Client decrypts** and use the predictions

Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]
[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Privacy-preserving CryptoRec Protocols

Accurate-Mode Prediction Protocol (full)

Limitation:

- **prohibitive cost** in ciphertext-only computation and corresponding storage
- **Too expensive to respond to a single query** while the accuracy improvement is limited.

Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]

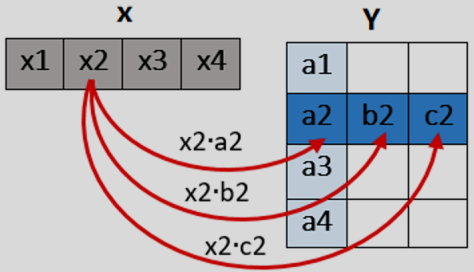
[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Privacy-preserving CryptoRec Protocols

Solution: Sparse-Quantization-Reuse Method

Goal: minimize the number of multiplications on encrypted data.

- **Sparsify a pre-learned CryptoRec model** (by removing parameters which don't contribute to final predictions).
- **Remove** all **weights** whose absolute values fall **below** a **threshold**
- **Quantize the weights** to enforce more weights to share the same values ([Han et al., 2015]), classify each feature matrix into 512 clusters
- **Reuse the shared multiplicative results** if possible



[Han et al., 2015] Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding

Novel Collaborative Filtering Recommender Friendly to Privacy Protection

Jun Wang[1*], **Qiang Tang**[2], **Afonso Arriaga**[3] and **Peter Y. A. Ryan**[1]

[1]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg
[2]Luxembourg Institute of Science and Technology
[3]INCERT GIE, Luxembourg
junwang.lu@gmail.com, qiang.tang@list.lu, afonso.arriaga@gmail.com, peter.ryan@uni.lu

Differentially Private Matrix Factorization*

Jingyu Hua, Chang Xia, Sheng Zhong
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology, Nanjing University, China
huajingyu@nju.edu.cn, changxia656569@gmail.com, zhongsheng@nju.edu.cn

Year: 2015

Venue: IJCAI

Domain: Movies

Paradigm: Federated Learning

Core ideas

- Trusted Recommender - never disclose users' privacy

- Untrusted Recommender - central recommender is considered to be non-trusted. Users can no longer feel free to send their raw ratings of items to the recommender.

Differentially Private Matrix Factorization*

Jingyu Hua, Chang Xia, Sheng Zhong
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology, Nanjing University, China
huajingyu@nju.edu.cn, changxia656569@gmail.com, zhongsheng@nju.edu.cn

Trusted Recommender Scenario

The Recommender wants to learn and publish item embeddings matrix V satisfying ε-differential privacy

Goal: Matrix U must be kept secret, otherwise the attacker can predict a specific user's ratings of all the items as long as she knows which vector in U corresponds to this user.

Solution: objective perturbation method, which is first proposed by [Chaudhuri and Monteleoni, 2009]

Steps:

- Matrix U is pre-computed, then it is condered as a constant in the loss minimization formula.
- Matrix V is the variable to update
- Add noise vector to perturb the objective function

[Chaudhuri and Monteleoni, 2009] Privacy-preserving logistic regression. In Advances in Neural Information Processing Systems, pages 289–296, 2009

Differentially Private Matrix Factorization*

Jingyu Hua, Chang Xia, Sheng Zhong
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology, Nanjing University, China
huajingyu@nju.edu.cn, changxia656569@gmail.com, zhongsheng@nju.edu.cn

Untrusted Recommender Scenario

Similar to modern Federated recommenders with Differential Privacy

One possibility may be to leverage existing secure multi-party computing methods based on cryptographic techniques, **BUT**..

- These methods rely on **too many heavyweight cryptographic operations** (far from practical)

The objective perturbation (previous slide) guarantees that publishing final V does not breach ε-differential privacy. Nevertheless, **the untrusted recommender WILL OBTAIN final V and gradients! But how?**

Differentially Private Matrix Factorization*

Jingyu Hua, Chang Xia, Sheng Zhong
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology, Nanjing University, China
huajingyu@nju.edu.cn, changxia656569@gmail.com, zhongsheng@nju.edu.cn

Untrusted Recommender Scenario

Just a sec! We did mention that **Final V and gradients contain noise**!

Yes… BUT…

iteration after iteration, the recommender can easily eliminate their effect, named difference attack (some similarities with «Secure Federated Matrix Factorization» proof)

Solution:

- **add additional noises** that are **generated by the user** and that are unknown to the recommender

Have we heard about that before?

Differentially Private Matrix Factorization*

Jingyu Hua, Chang Xia, Sheng Zhong
State Key Laboratory for Novel Software Technology,
Department of Computer Science and Technology, Nanjing University, China
huajingyu@nju.edu.cn, changxia656569@gmail.com, zhongsheng@nju.edu.cn

Jo-DPMF: Differentially private matrix factorization learning through joint optimization

Feng Zhang [a,b,*], Victor E. Lee [c], Kim-Kwang Raymond Choo [d,e,a]

[a] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[b] Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China
[c] GraphSQL Inc., Mountain View, CA 94043, USA
[d] Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA
[e] Information Assurance Research Group, University of South Australia, Adelaide, SA 5095, Australia

Year: 2018

Venue: Information Sciences

Domain: Movies

Paradigm: Federated Learning

Core ideas

- Differential privacy with Laplace noise in the objective function
- K-coRating Matrix

Jo-DPMF: Differentially private matrix factorization learning through joint optimization

Feng Zhang[a,b,*], Victor E. Lee[c], Kim-Kwang Raymond Choo[d,e,a]

[a] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[b] Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China
[c] GraphSQL Inc., Mountain View, CA 94043, USA
[d] Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA
[e] Information Assurance Research Group, University of South Australia, Adelaide, SA 5095, Australia

Differential privacy with Laplace noise

The objective function is integrated with a **carefully designed Laplace noise** factor

The authors provide a theorem to prove that the solution (the factorized matrices) of the **optimization problem constrained by the perturbed objective function satisfy differential privacy**

Jo-DPMF: Differentially private matrix factorization learning through joint optimization

Feng Zhang[a,b,*], Victor E. Lee[c], Kim-Kwang Raymond Choo[d,e,a]

[a] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[b] Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China
[c] GraphSQL Inc., Mountain View, CA 94043, USA
[d] Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA
[e] Information Assurance Research Group, University of South Australia, Adelaide, SA 5095, Australia

K-coRating

$$u[i] = \begin{cases} 0 & \text{if } r_{u,i} \text{ is NULL} \\ 1 & \text{if } r_{u,i} \text{ is NOT NULL} \end{cases}$$

k-coRated privacy: A rating matrix M satisfies k-coRated privacy if every user u ∈ U has at least (k - 1)-coRated Equivalent users

- k-coRating tries to **split the users into groups**, forcing each group to satisfy k-coRated privacy.
- To make each group satisfy k-coRated privacy, some NULL cells have to be filled up with some data.
- In [Zhang et al. 2014], **trust derivation** and **Pearson correlation similarity** are shown as the best methods to generate the filling data.
- Here, trust derivation is adopted

[Zhang et al. 2014] k-coRating: filling up data to obtain privacy and utility, in: Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14), pp. 320–327

Jo-DPMF: Differentially private matrix factorization learning through joint optimization

Feng Zhang[a,b,*], Victor E. Lee[c], Kim-Kwang Raymond Choo[d,e,a]

[a] School of Computer Science, China University of Geosciences, Wuhan 430074, China
[b] Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430074, China
[c] GraphSQL Inc., Mountain View, CA 94043, USA
[d] Department of Information Systems and Cyber Security, University of Texas at San Antonio, San Antonio, TX 78249, USA
[e] Information Assurance Research Group, University of South Australia, Adelaide, SA 5095, Australia

Differentially private graph-link analysis based social recommendation

Taolin Guo*, Junzhou Luo, Kai Dong, Ming Yang

School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, PR China

Year: 2018

Venue: Information Sciences

Domain: Social networks

Paradigm: Federated Learning

Core ideas

Apply Differential privacy to a completely novel domain: social recommendation

Differentially private graph-link analysis based social recommendation

Taolin Guo*, Junzhou Luo, Kai Dong, Ming Yang

School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, PR China

What is the problem with differential privacy and social recommendation?

[Machanavajjhala et al., 2011] suggest that

- differential privacy in a graph-link analysis based social recommendation requires that **modifying an arbitrary social link** **should have a negligible effect** on the recommendations to any target node.
- They argue that there is an **inherent high sensitivity in social recommendation** since the presence or absence of a link in a social network affects the recommendations of multiple nodes, result in very poor recommendation accuracy even to ensure an unreasonable privacy guarantee.

[Machanavajjhala et al., 2011] Personalized social recommendations - accurate or private? 4 (7) (2011) 440–450.

Differentially private graph-link analysis based social recommendation

Taolin Guo*, Junzhou Luo, Kai Dong, Ming Yang

School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, PR China

Is there a problem with differential privacy and social recommendation?

- Sensitivity captures the **maximum extent to the output of a function** that can be affected **by any single input**.

In Machanavajjhala's work, the output of a recommendation function is a single node. However, the **output of a recommendation function used to compute sensitivity is a utility vector**, which contains the utilities of all nodes be recommended to a target node.

As a result, **Machanavajjhala** defined **sensitivity captures** no longer the maximum affection to the output of all potential recommendations, but **the maximum sum of this affection**. This lead to greatly **overestimate** the value of sensitivity.

Solution: Define and compute the sensitivity of the utility function!

Differentially private graph-link analysis based social recommendation

Taolin Guo*, Junzhou Luo, Kai Dong, Ming Yang

School of Computer Science and Engineering, Southeast University, Nanjing, Jiangsu 211189, PR China

5

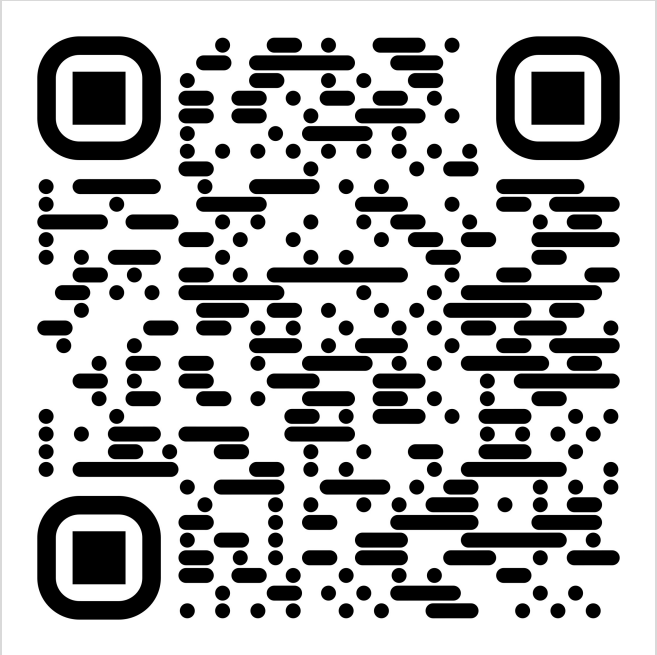
Preserving Privacy at Scale

The Case of Twitter

6

Hands-On Session

Privacy-Preserving Techniques for Recommender Systems



https://bit.ly/3CYyeuJ

Thank you!

Questions?