Midterm

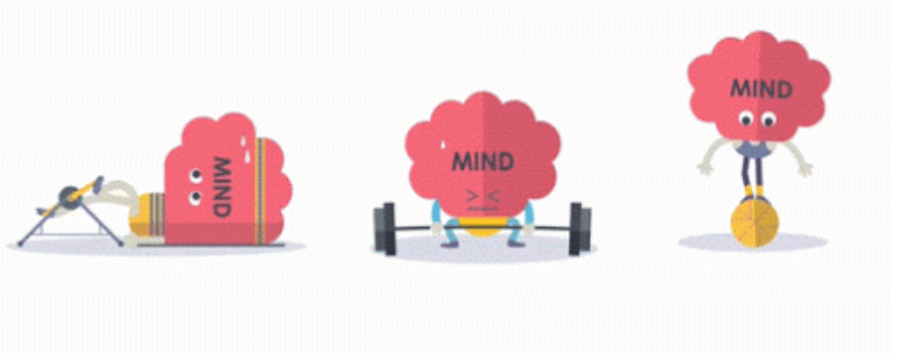
Written questions – NO COMPUTER

Questions can be about the class or the exercises

Mix of Multi-answer questions and short questions
Short questions MUST be answered in at MOST 3 lines

This room and 2 other rooms. We will announce on Moodle

Let’s exercise your privacy brain



True or False?

- Protecting Netflix ratings from other users using privacy settings can be seen as SOCIAL privacy

YES, the adversary are other users, the goal is not to surprise the user regarding what others will see

- Using TLS to encrypt your communications with the email server can be seen as ANTI-SURVEILLANCE privacy

Nope, TLS provides encryption against third parties, but anti-surveillance would also consider the server adversarial

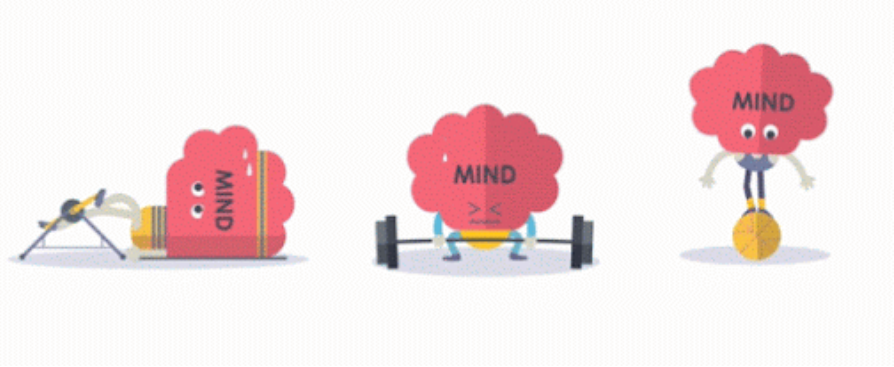
- Privacy as CONTROL aims to hide as much information as possible from the provider

Nope, privacy as control aims at providing users with ways to decide where the information flows. At no point considers the provider as a problematic flow

- A privacy technology that requires more than one entity to decrypt data follows the CONFIDENTIALITY paradigm

YES, this approach ensures that a minimal amount of information has to be leaked to individual entities

Let’s exercise your privacy brain



Remember the privacy properties:
pseudonymity, anonymity, unlinkability, unobservability, plausible deniability

What property/properties Alice enjoys

- if she uses one-time credentials to users login in a website (imagine a random number that can only be shown once, like a ticket to the cinema)
  - with respect to the website? **Anonymity, Unlinkability**
  - with respect to the credential issuer? **If the credential is obtained in person: none. If the credential is obtained online, pseudonymity.**

- regarding email receivers, if every time she wants to send an email she ask one of our friends to send it for us
  - if each friend sends at most one email - **Anonymity**
  - if friends can send more than one email - **Pseudonymity**

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

SPRING
SECURITY AND PRIVACY ENGINEERING LABORATORY

Information Security and Privacy (COM-402)

Part 5: Privacy enhancing technologies

Data anonymization

Carmela Troncoso

SPRING Lab

carmela.troncoso@epfl.ch

Some slides/ideas adapted from: JP Hubaux, Bryan Ford, Vitaly Shmatikov

PETs for Data anonymization

Scenario:

You have a set of data that contains personal data and you would like to anonymize it to:

- not be subject to data protection while processing
- make it public for profit
- make it public for researchers

Goal:

Produce a dataset that **preserves the utility** of the original dataset **without leaking information** about individuals. *This process is known as "*database sanitization*"*

REMEMBER: ANONYMITY IS ABOUT DECOUPLING IDENTITY AND ACTION!

To achieve anonymity we must **decouple** users from their attributes

✘ Let's make users pseudonymous!

✘ Let's remove identities!

Some attributes are quasi-identifiers!

✘ Let's remove some attributes!

Medical Data

QID			SA
Zipcode	Age	Sex	Disease
47677	29	*	Ovarian Cancer
47602	22	*	Ovarian Cancer
47678	27	*	Prostate Cancer
47905	43	*	Flu
47909	52	*	Heart Disease
47906	47	*	Heart Disease

Voter registration data

Name	Zipcode	Age	Sex
Alice	47677	29	F
Bob	47983	65	M
Carol	47677	22	F
Dan	47532	23	M
Ellen	46789	43	F

To achieve anonymity we must **decouple** users from their attributes

✘ Let's make users pseudonymous!

✘ Let's remove identities!

Some attributes are quasi-identifiers!

✘ Let's remove some attributes!

Impossible to know what will be a QID

Medical Data

QID			SA
Zipcode	Age	Sex	Disease
47677	29	*	Ovarian Cancer
47602	22	*	Ovarian Cancer
47678	27	*	Prostate Cancer
47905	43	*	Flu
47909	52	*	Heart Disease
47906	47	*	Heart Disease

Voter registration data

Name	Zipcode	Age	Sex
Alice	47677	29	F
Bob	47983	65	M
Carol	47677	22	F
Dan	47532	23	M
Ellen	46789	43	F

Caucasian	HIV+	Flu
Asian	HIV-	Flu
Asian	HIV+	Herpes
Caucasian	HIV-	Acne
Caucasian	HIV-	Herpes
Caucasian	HIV-	Acne

Anonymization: k-anonymity

Key Attribute / Identifier | **Quasi-identifier** | **Sensitive attribute**

name	gender	zipcode	problem
John	Male	1012	Cancer
Zoey	Female	1013	Flu
Nathan	Male	1016	Heart Disease
Lucas	Male	1015	Heart Disease
Sam	Female	1003	Flu
Max	Male	1012	Flu
Mathias	Male	1014	HIV+
Sarah	Female	1012	Herpes
Julia	Female	1012	Flu

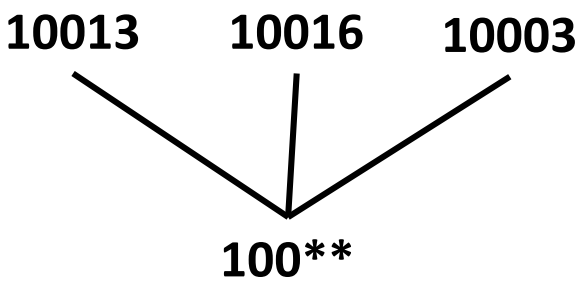
Anonymization: k-anonymity

- Each person contained in the database cannot be distinguished from at least k-1 other individuals whose information also appears in the released database.

***Generalization**: replace attributes with less specific, but semantically consistent values*

EXAMPLE

zipcode



name	gender	zipcode	problem
	Male	1012	Cancer
	Female	100**	Flu
	Male	100**	Heart Disease
	Male	100**	Heart Disease
	Female	100**	Flu
	Male	1012	Flu
	Male	100**	HIV+
	Female	1012	Herpes
	Female	1012	Flu

k=2

What is the rationale?

name	gender	zipcode	Favourite color
John	Male	1012	Blue
Zoey	Female	1013	Red
Nathan	Male	1016	Red
Lucas	Male	1015	Black
Sam	Female	1003	Yellow
Max	Male	1012	Red
Mathias	Male	1014	Black
Sarah	Female	1012	Blue
Julia	Female	1012	Red

Who has cancer, John or Max???

name	gender	zipcode	problem
	Male	1012	Cancer
	Female	100**	Flu
	Male	100**	Heart Disease
	Male	100**	Heart Disease
	Female	100**	Flu
	Male	1012	Flu
	Male	100**	HIV+
	Female	1012	Herpes
	Female	1012	Flu

k=2

What is the rationale?

name	gender	zipcode	Favourite color
John	Male	1012	Blue
Zoey	Female	1013	Red
Nathan	Male	1016	Red
Lucas	Male	1015	Black
Sam	Female	1003	Yellow
Max	Male	1012	Red
Mathias	Male	1014	Black
Sarah	Female	1012	Blue
Julia	Female	1012	Red

Does John have Cancer or Flu?

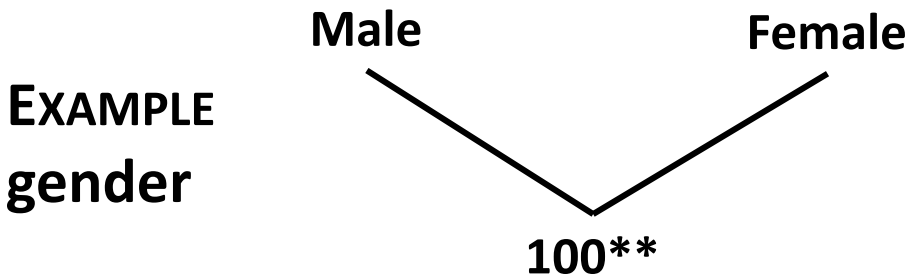
name	gender	zipcode	problem
	Male	1012	Cancer
	Female	100**	Flu
	Male	100**	Heart Disease
	Male	100**	Heart Disease
	Female	100**	Flu
	Male	1012	Flu
	Male	100**	HIV+
	Female	1012	Herpes
	Female	1012	Flu

k=2

Anonymization: k-anonymity

- To improve anonymity, identifying attributes can be ***suppressed***

(note that suppression is the ultimate generalization!)



name	gender	zipcode	problem
	*	1012	Cancer
	*	100*	Flu
	*	100*	Heart Disease
	*	100*	Heart Disease
	*	100*	Flu
	*	1012	Cancer
	*	100*	HIV+
	*	1012	Herpes
	*	1012	Flu

k=4

This is 3-anonymous, any problem? (think about the rationale)

Homogeneity attack

Bob	
Zipcode	***Age***
47678	27

Background knowledge attack

Carl	
Zipcode	***Age***
47673	36

A 3-anonymous patient table

Zipcode	Age	Disease
476**	2*	Heart Disease
476**	2*	Heart Disease
476**	2*	Heart Disease
4790*	≥40	Flu
4790*	≥40	Heart Disease
4790*	≥40	Cancer
476**	3*	Heart Disease
476**	3*	Cancer
476**	3*	Cancer

If you have background, e.g., “heart diseases are very unlikely in populations of 30 year old”

It is highly likely that Carl has cancer!!

ℓ-Diversity

- An equivalence class has *l*-diversity if there are at least ℓ well-represented values for the sensitive attribute.
- A database has ℓ-diversity if every equivalence class has ℓ–diversity.

Zipcode	Age	Salary	Disease
476**	2*	20K	Gastric Ulcer
476**	2*	30K	Gastritis
476**	2*	40K	Stomach Cancer
4790*	≥40	50K	Gastritis
4790*	≥40	100K	Flu
4790*	≥40	70K	Bronchitis
476**	3*	60K	Bronchitis
476**	3*	80K	Pneumonia
476**	3*	90K	Stomach Cancer

A 3-diverse hospital records dataset

ℓ-Diversity: problems

Similarity attack

Bob	
Zip	***Age***
47678	27

A 3-diverse patient table

Zipcode	Age	Salary	Disease
476**	2*	20K	Gastric Ulcer
476**	2*	30K	Gastritis
476**	2*	40K	Stomach Cancer
4790*	≥40	50K	Gastritis
4790*	≥40	100K	Flu
4790*	≥40	70K	Bronchitis
476**	3*	60K	Bronchitis
476**	3*	80K	Pneumonia
476**	3*	90K	Stomach Cancer

Conclusion

1. Bob's salary is in [20k,40k], which is relatively low
2. Bob has some stomach-related disease

l-diversity does not consider semantics of sensitive values!

ℓ-Diversity: problems

Q1: 423, >50**
Q2: 423, <60**

Original dataset

…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Flu

99% have cancer

Anonymization A

Q1	Flu
Q1	Flu
Q1	Cancer
Q1	Flu
Q1	Cancer
Q1	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer

Anonymization B

Q1	Flu
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Flu
Q2	Flu

l-diversity does not consider distribution of semantic values!

ℓ-Diversity: problems

Q1: 423, >50**
Q2: 423, <60**

Original dataset

…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Flu

99% have cancer

Anonymization A

Q1	Flu
Q1	Flu
Q1	Cancer
Q1	Flu
Q1	Cancer
Q1	Cancer
Q2	Cancer
Q2	Cancer
Q2	
Q2	
Q2	

Anonymization B

Q1	Flu
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Cancer
Q2	Flu

50% cancer $\Rightarrow$ quasi-identifier group is “diverse”

This leaks a ton of information

l-diversity does not consider distribution of semantic values!

ℓ-Diversity: problems

Q1: 423, >50**
Q2: 423, <60**

Original dataset

…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Cancer
…	Flu
…	Flu

99% have cancer

Anonymization A

Q1	Flu
Q1	Flu
Q1	Cancer
Q1	Flu
Q1	Cancer
Q1	Cancer
Q2	Cancer

Anonymization B

Q1	Flu
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q1	Cancer
Q2	Cancer
Q2	Flu

99% cancer ⇒ quasi-identifier group is not "diverse"
…yet anonymized database does not leak anything

50% cancer ⇒ quasi-identifier group is "diverse"
This leaks a ton of information

l-diversity does not consider distribution of semantic values!

t-Closeness

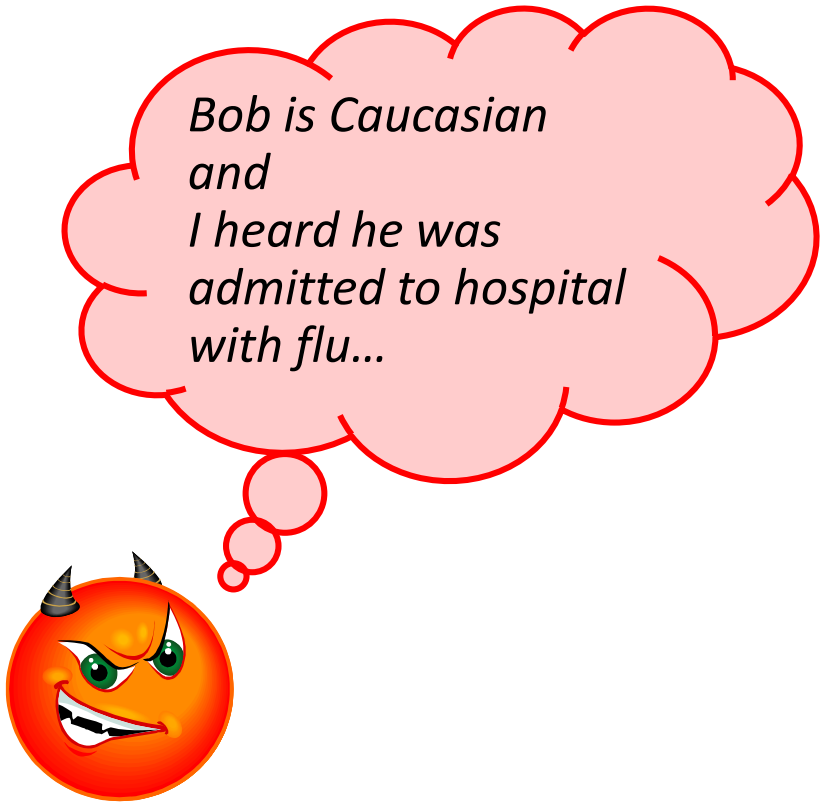
- An equivalence class has *t-closeness* if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole table is no more than a *threshold* t.
- A table has *t-closeness* if all equivalence classes have t-*closeness.*

Caucasian	787XX	HIV+	Flu
Asian	787XX	HIV-	Flu
Asian	787XX	HIV+	Herpes
Caucasian	787XX	HIV-	Acne
Caucasian	787XX	HIV-	Herpes
Caucasian	787XX	HIV-	Acne

This is k-anonymous, l-diverse and t-close...

...so secure, right?

What Does the Attacker Know?



Caucasian	787XX	HIV+	Flu
Asian	787XX	HIV-	Flu
Asian	787XX	HIV+	Herpes
Caucasian	787XX	HIV-	Acne
Caucasian	787XX	HIV-	Herpes
Caucasian	787XX	HIV-	Acne

Takeaways

Anonymizing a dataset via generalization and suppression is extremely hard

The k-anonymity idea focuses on transforming the dataset not its semantics

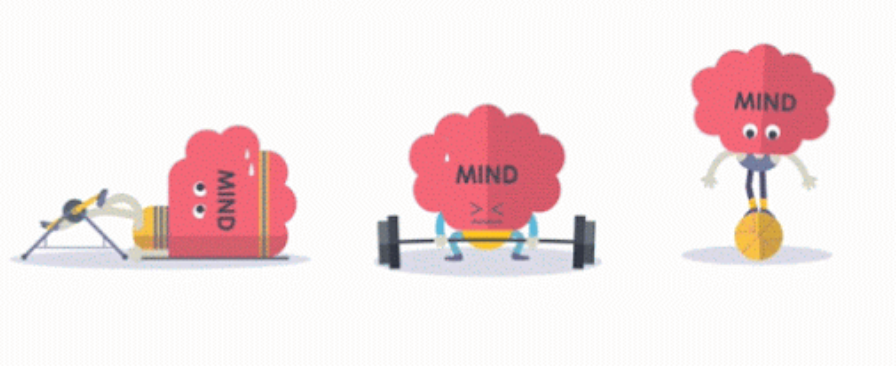
Achieving k-anonymity, l-diversity, t-closeness is hard, and still does not guarantee privacy

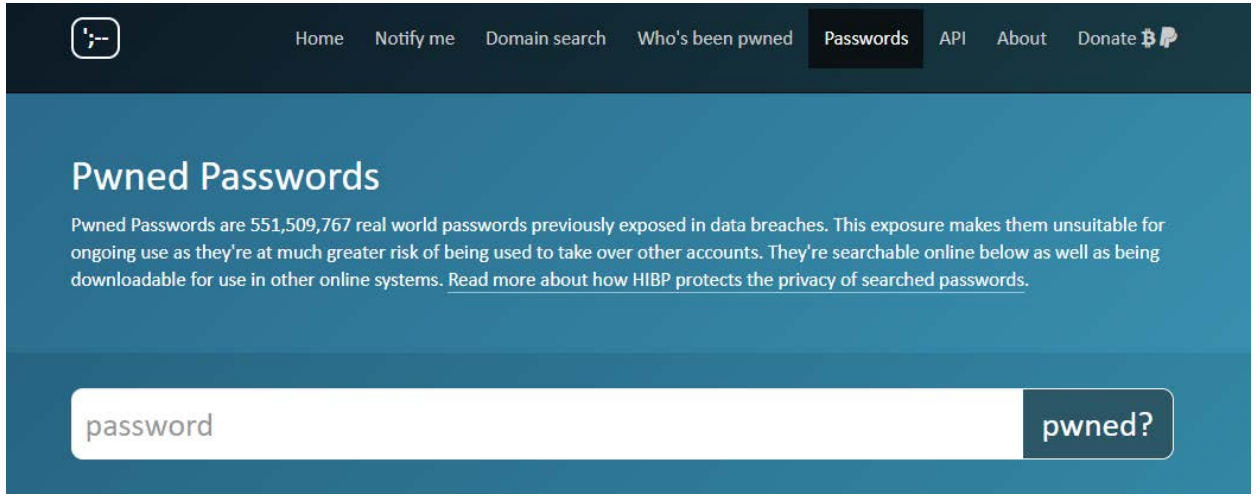
The adversary's background **can be anything**

BEING ABLE TO FULLY ANONYMIZE A HIGH-DIMENSIONAL DATABASE IS AS LIKELY AS BEING ABLE TO FIND A UNICORN IN THE GALAXY

IF WE CANNOT PUBLISH THE DATA, CAN WE DO SOMETHING WITH IT?

Let’s exercise your privacy brain

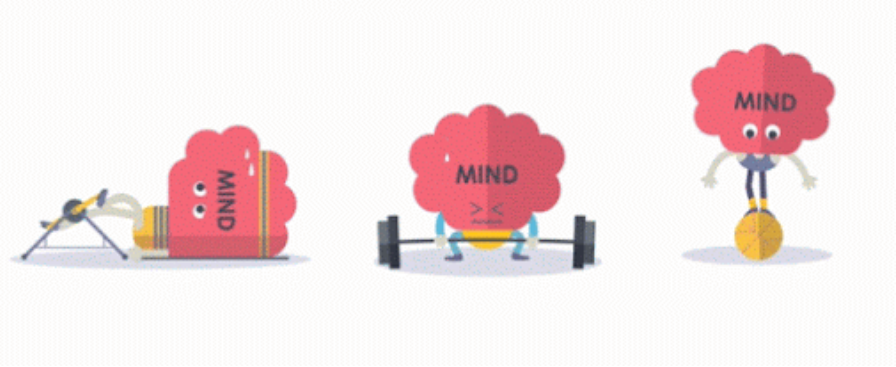


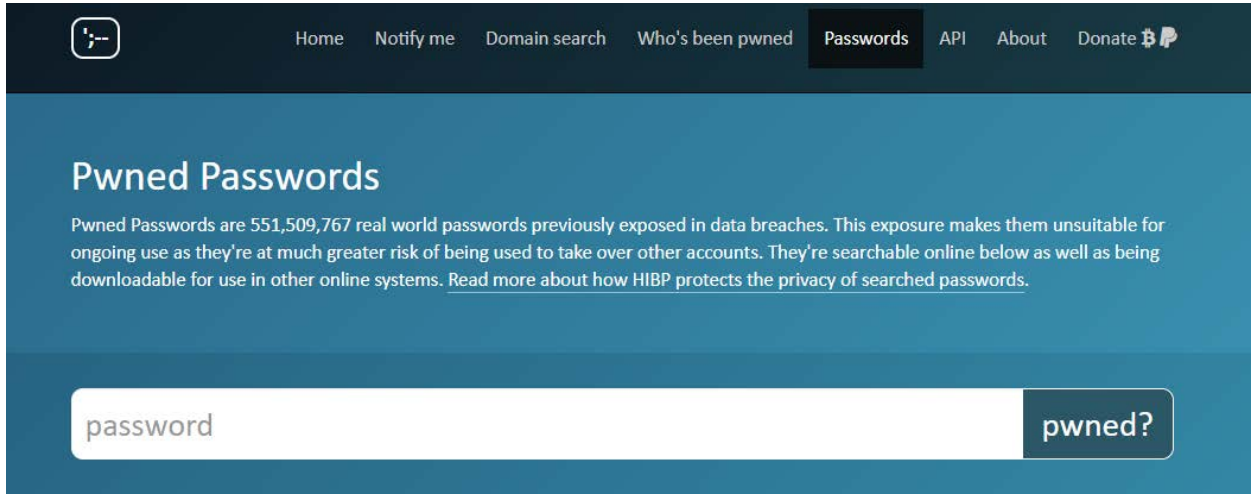


Would you send your password in the clear?

https://haveibeenpwned.com/

Let’s exercise your privacy brain



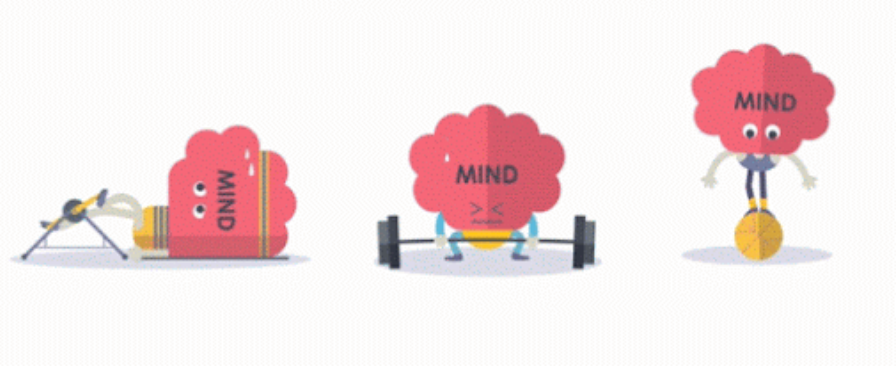


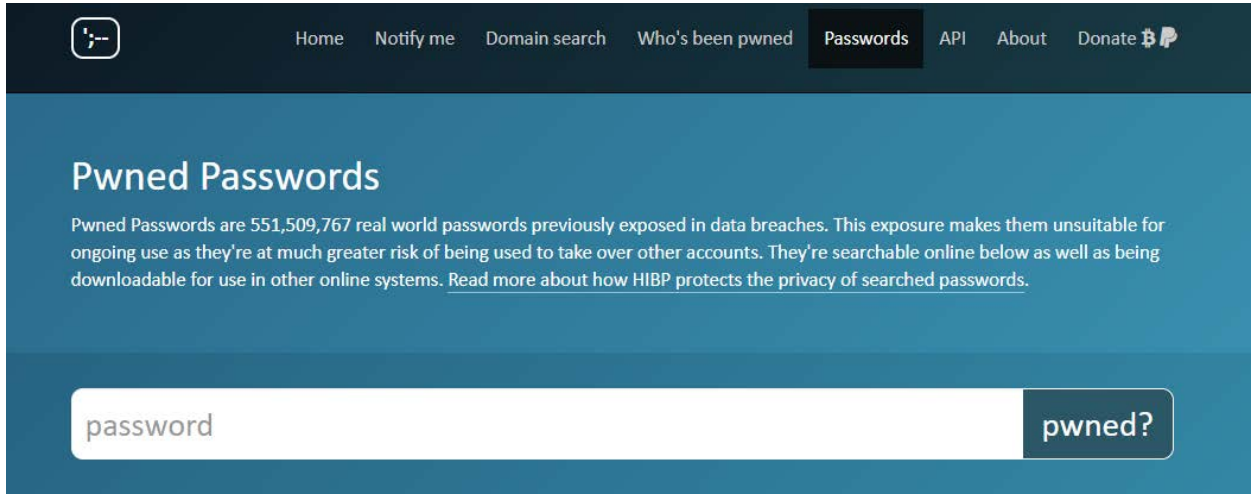
Would you send your password in the clear?

Would you send a hash?

https://haveibeenpwned.com/

Let’s exercise your privacy brain





Would you send your password in the clear?

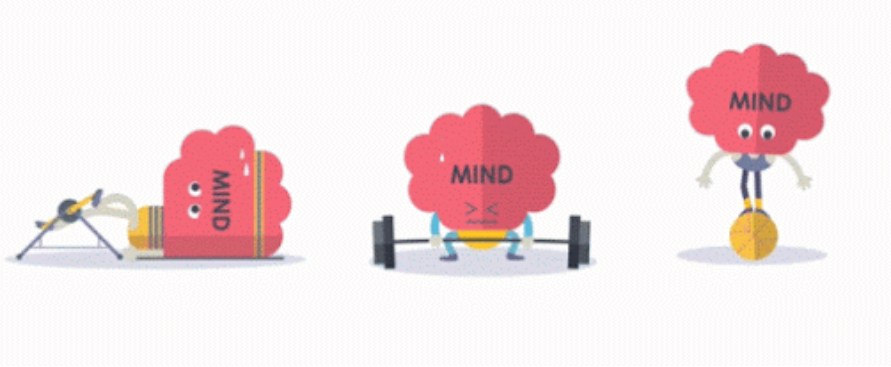
Would you send a hash?

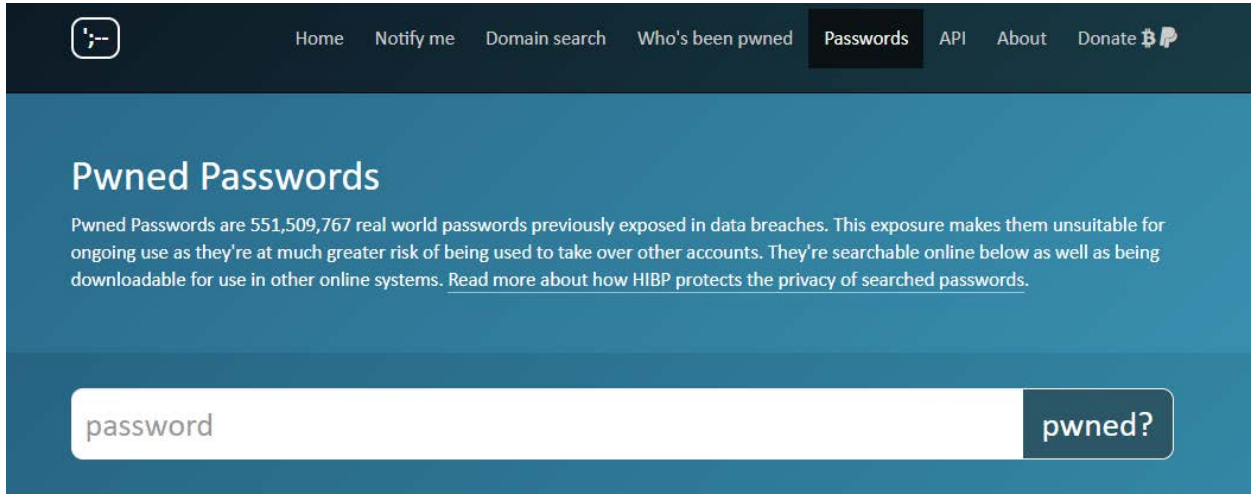
What they do: send the first 5 bytes of the hash of the password and receive a list of 475 suffixes to check offline

1. (21BD1) **0018A45C4D1DEF81644B54AB7F969B88D65:1** (password "lauragpe")
2. (21BD1) **00D4F6E8FA6EECAD2A3AA415EEC418D38EC:2** (password "alexguo029")
3. (21BD1) **011053FD0102E94D6AE2F8B83D76FAF94F6:1** (password "BDnd9102")
4. (21BD1) **012A7CA357541F0AC487871FEEC1891C49C:2** (password "melobie")
5. (21BD1) **0136E006E24E7D152139815FB0FC6A50B15:2** (password "quvekyny")
6. …

https://haveibeenpwned.com/
https://blog.cloudflare.com/validating-leaked-passwords-with-k-anonymity/

Let’s exercise your privacy brain





Would you send your password in the clear?

Would you send a hash?

What they do: send the first 5 bytes of the hash of the password and receive a list of 475 suffixes to check offline

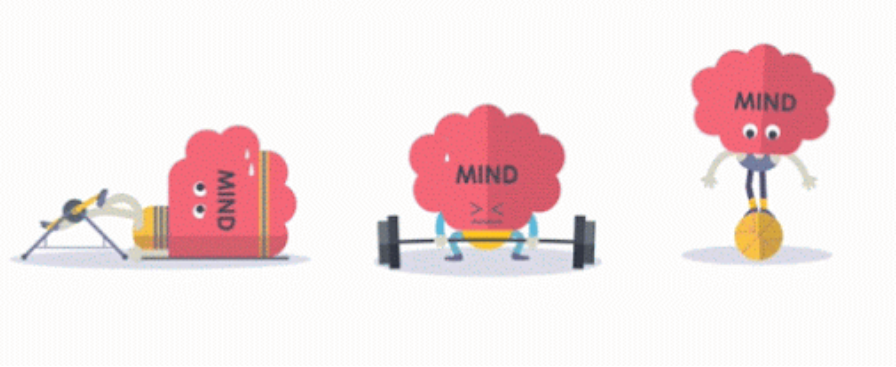
Send

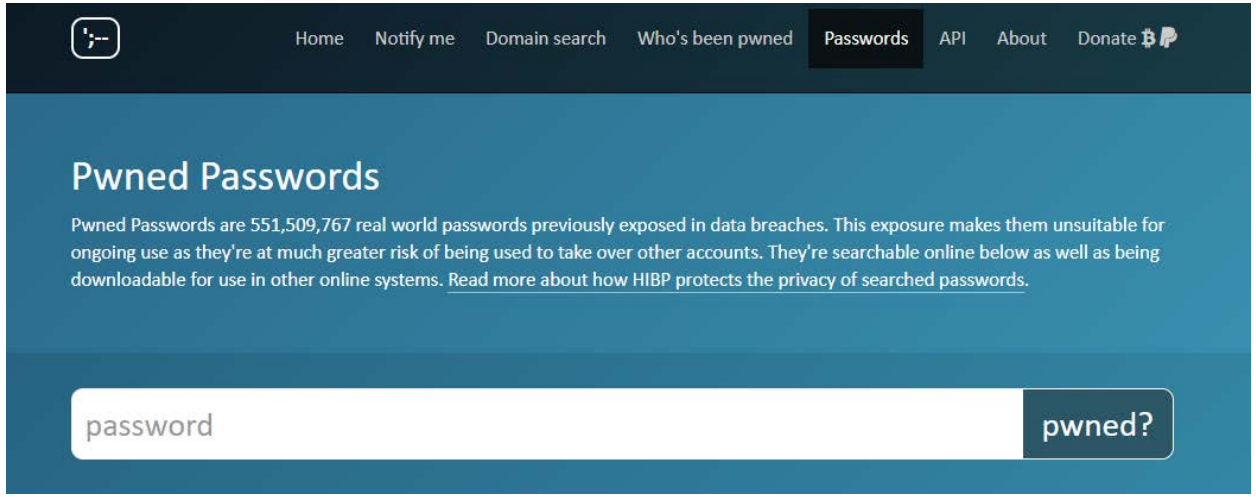
Receive

1. (21BD1 0018A45C4D1DEF81644B54AB7F969B88D65:1 (password "lauragpe")
2. (21BD1 00D4F6E8FA6EECAD2A3AA415EEC418D38EC:2 (password "alexguo029")
3. (21BD1 011053FD0102E94D6AE2F8B83D76FAF94F6:1 (password "BDnd9102")
4. (21BD1 012A7CA357541F0AC487871FEEC1891C49C:2 (password "melobie")
5. (21BD1 0136E006E24E7D15213981 5FB0FC6A50B15:2 (password "quvekyny")
6. …

https://haveibeenpwned.com/
https://blog.cloudflare.com/validating-leaked-passwords-with-k-anonymity/

Let’s exercise your privacy brain





From the point of view of the server (that receives the 5-bytes suffix

What is the privacy of the password?

Would you send your password in the clear?

Would you send a hash?

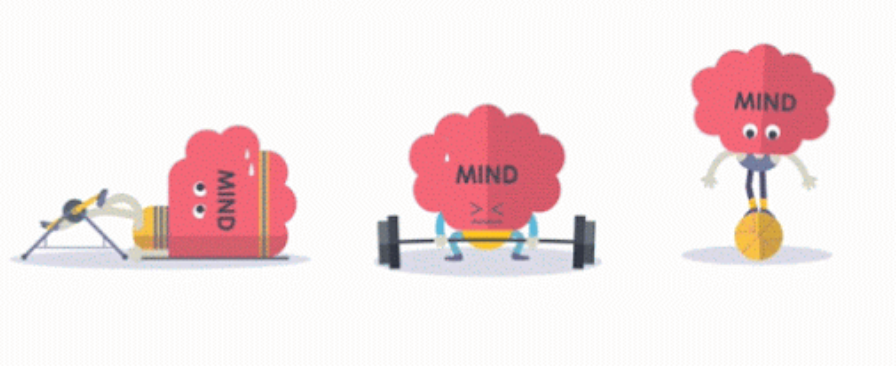
What they do: send the first 5 bytes of the hash of the password and receive a list of 475 suffixes to check offline

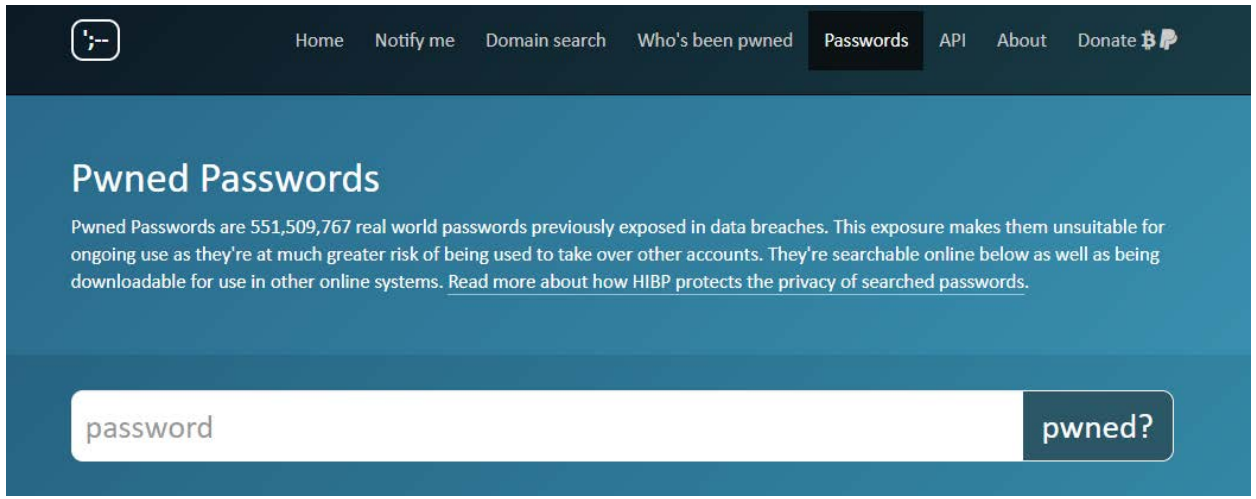
Send Receive

1. (21BD1 0018A45C4D1DEF81644B54AB7F969B88D65:1 (password "lauragpe")
2. (21BD1 00D4F6E8FA6EECAD2A3AA415EEC418D38EC:2 (password "alexguo029")
3. (21BD1 011053FD0102E94D6AE2F8B83D76FAF94F6:1 (password "BDnd9102")
4. (21BD1 012A7CA357541F0AC487871FEEC1891C49C:2 (password "melobie")
5. (21BD1 0136E006E24E7D15213981 5FB0FC6A50B15:2 (password "quvekyny")
6. …

https://haveibeenpwned.com/
https://blog.cloudflare.com/validating-leaked-passwords-with-k-anonymity/

Let’s exercise your privacy brain





Would you send your password in the clear?

Would you send a hash?

What they do: send the first 5 bytes of the hash of the password and receive a list of 475 suffixes to check offline

From the point of view of the server (that receives the 5-bytes suffix

What is the privacy of the password?
The password is 475-anonymous!!

Send **Receive**

1. (21BD1) 0018A45C4D1DEF81644B54AB7F969B88D65:1 (password "lauragpe")
2. (21BD1) 00D4F6E8FA6EECAD2A3AA415EEC418D38EC:2 (password "alexguo029")
3. (21BD1) 011053FD0102E94D6AE2F8B83D76FAF94F6:1 (password "BDnd9102")
4. (21BD1) 012A7CA357541F0AC487871FEEC1891C49C:2 (password "melobie")
5. (21BD1) 0136E006E24E7D15213981 5FB0FC6A50B15:2 (password "quvekyny")
6. …

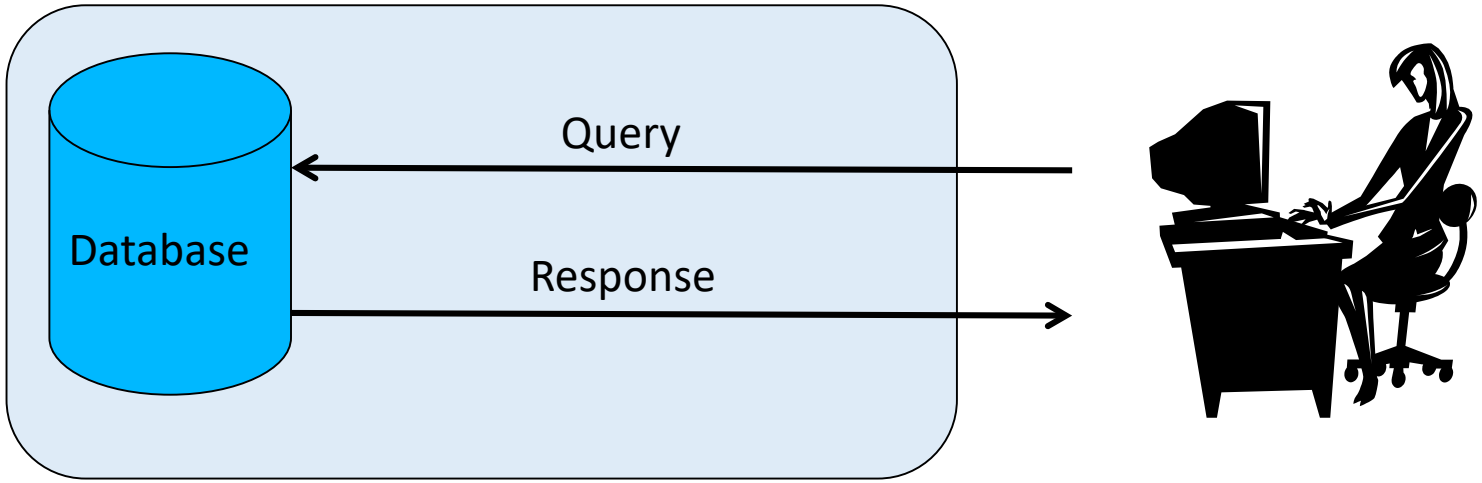
https://haveibeenpwned.com/
https://blog.cloudflare.com/validating-leaked-passwords-with-k-anonymity/

The interactive scenario

Many times we do not want the data, we want statistics!

Redefined Goal for the interactive case:

Produce an **answer** that **preserves the utility** of the **statistics** **without leaking information** about individuals.

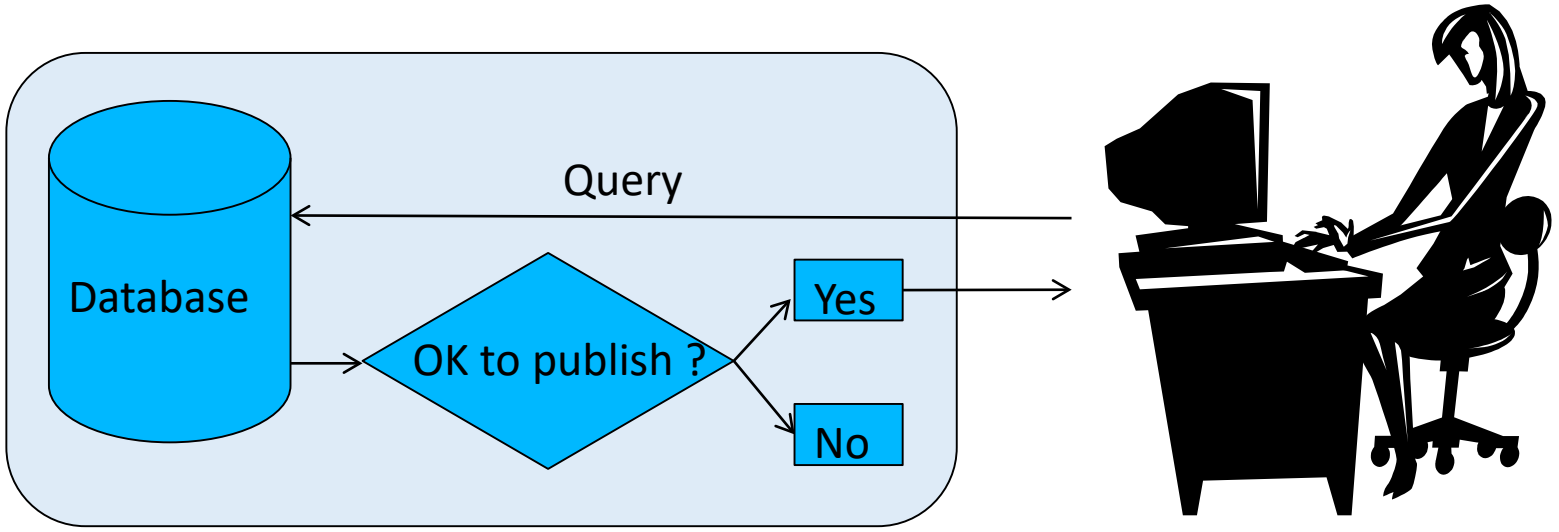


Query = **What is the average salary of women professors at IC@EPFL with Spanish nationality?**

Is there a privacy problem?

The interactive scenario

Let’s audit the queries, if the query will leak, deny!
Either answer truthfully or state that there will be no answer

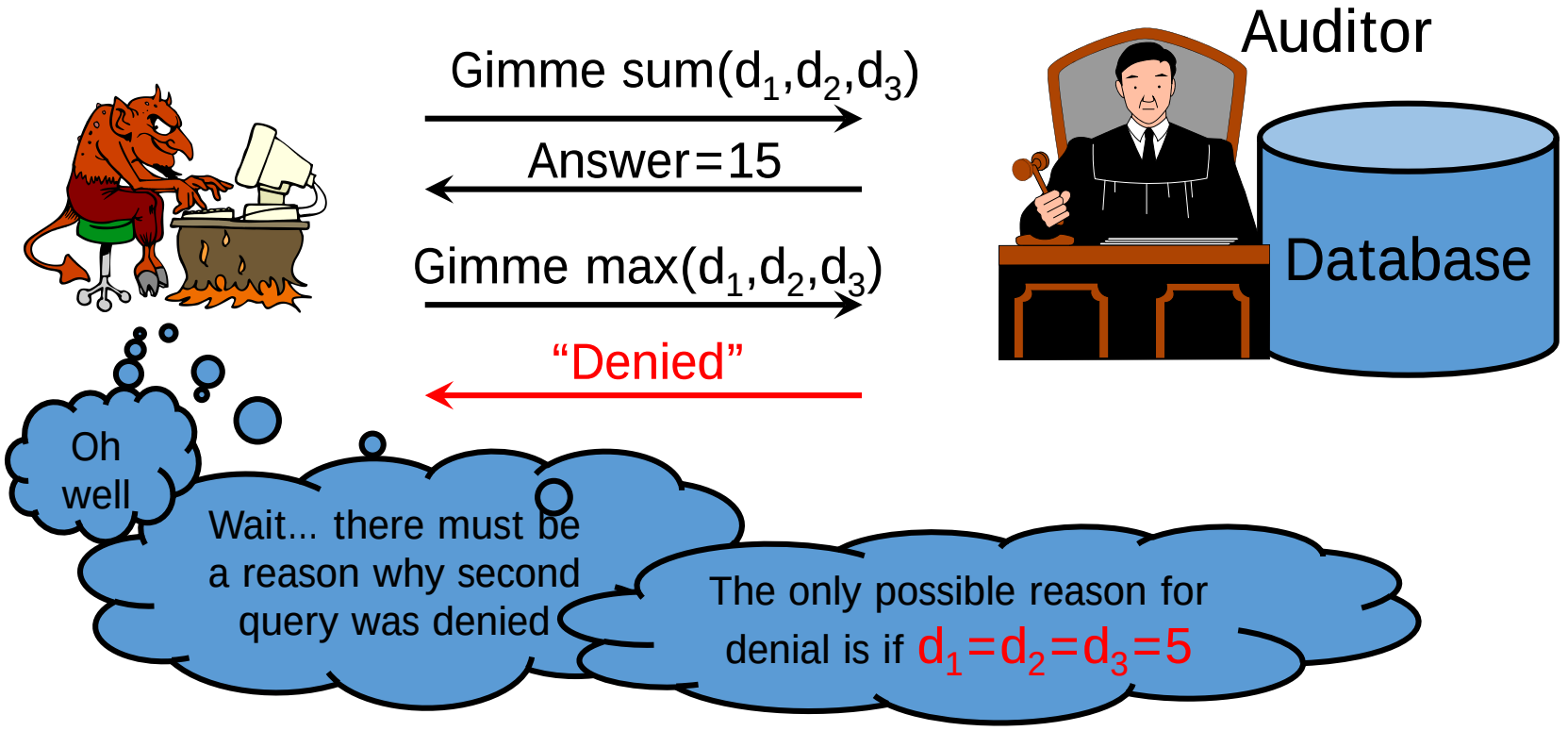


Database assumed to contain *numeric* values.

! Not answering already reveals some information !

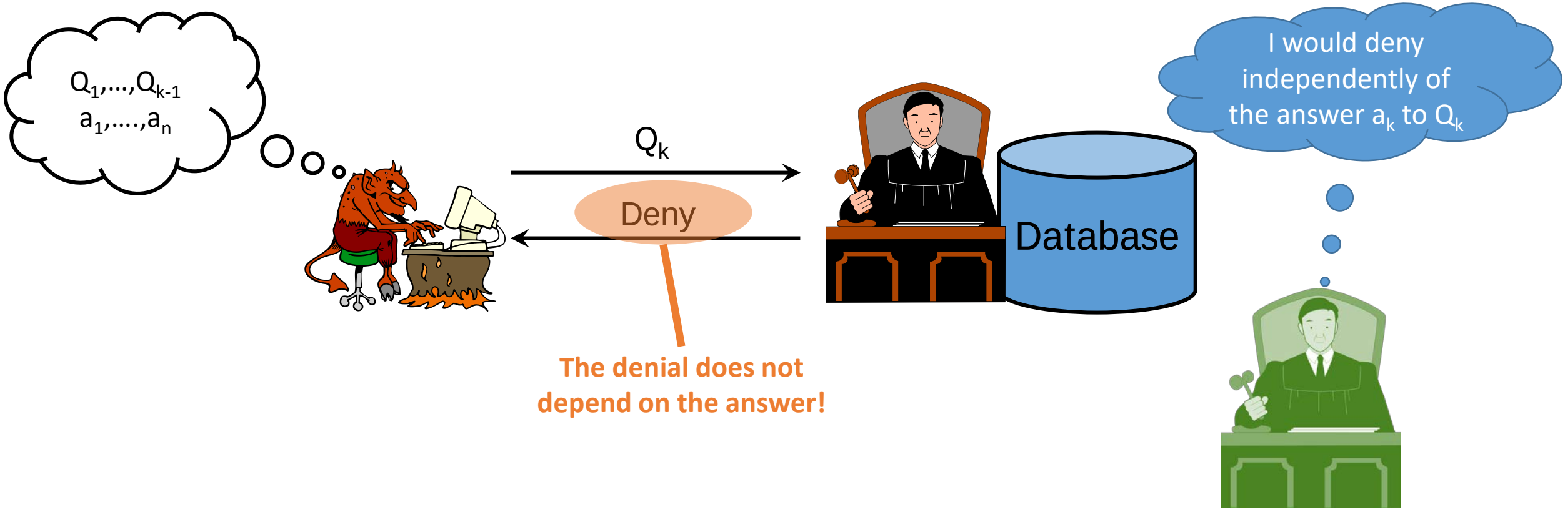
When denying fails: learning exact values

- Variables d_i are real, privacy breached if adversary learns some d_i



Can I make sure that the next query does not leak? Simulatable auditing

One cannot learn anything from the denial **if the decision to deny or give an answer is independent of the actual data set and the real answer**.



https://theory.stanford.edu/~nmishra/SimulatableAuditing.pdf

Auditing has problems

- Privacy definition? Privacy of Values? Groups? Exact?
- Algorithmic limitations
  - Secure deniability implies using algorithms computationally prohibitive
  - Feasible focus mostly on simple queries
- Collusion? Either high cost or no security
- Utility?
  - Percentage of denials may not be the best measure

What else can we do? Modifying inputs

- **Subsampling**
  - A subset of the rows is chosen at random and released and **statistics are computed on the subsample**
  - Uneven privacy for users, being in a subsample may have unfortunate consequences
    - ➢ Not being may too!
- **Input perturbation**
  - **Data or queries are modified before** a response is generated
  - How can we quantify the leakage?
  - How to balance for utility?

What else can we do? Modifying outputs

- **Adding random noise to the output**
  - **Naively**, this approach will fail
    - E.g., if the same query is asked repeatedly, then the responses can be averaged, and the true answer will eventually emerge.
  - This cannot be fixed by recording each query and providing the same response each time a query is re-issued.
    - **Syntactically different queries may be semantically equivalent**, and, if the query language is sufficiently rich, then the equivalence problem itself is undecidable.
- **Randomized response**
  - Respondents to a query **flip a coin and, based on the outcome, they either honestly respond or respond randomly**
  - Privacy comes from the uncertainty of how to interpret a reported individual value.
  - Yet, data can be useful because **randomness can be averaged out**
  - **Not usable for every case, or combined with other techniques**

Differential privacy

Remember the Goal for the interactive case:

Produce an **answer** that preserves the utility of the statistics without leaking information about individuals.

To have any utility we must allow the leakage of some information, but we can set a bound on the extent of leakage!

Differential Privacy:

Output is similar whether any single individual's record is included in the database or not.

Guarantees minimal similarity

Differential Privacy

- **Basic philosophy:** instead of the real answer to a query, output a random answer, such that by a small change in the database (someone joins or leaves), the distribution of the answer does not change much.

- **A new privacy goal:** minimize the increased risk incurred by an individual when joining (or leaving) a given database.

- **Motivation:** A privacy guarantee that limits risk incurred by joining, therefore encourages participation in the dataset, increasing social utility.

Important!!!!

Differential Privacy is a privacy notion **NOT** a mechanism

You use mechanisms to achieve differential privacy

Differential Privacy - Informal Definition

Output is similar whether any single individual's record is included in the database or not.

C's inclusion of her record in the computation does not make her *significantly* worse off.

If there is already some risk of revealing a secret of C by combining auxiliary information and something learned from DB, then **that risk is still there** but not *significantly* increased by C's participation in the database.

ϵ-Differential Privacy – Formal Definition

- $\mathcal{D}$: The set of input databases
- R: Output space of the query
- F: Query function

  $F: \mathcal{D} \rightarrow R$
- d: Distance function on the set of databases
- *Neighboring databases*: Pairs of databases ($\mathcal{D}$, $\mathcal{D}_{-r}$) differing only in one row *r* (e.g., individual)

  $d(\mathcal{D}\text{-}\mathcal{D}_{-r}) = 1$

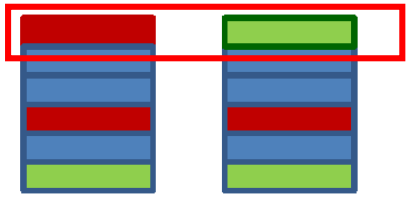
ϵ-Differential Privacy – Formal Definition

- Principle
  - The **removal/addition** of a **single record** in the **database should/does *not* substantially affect the values** of the computed function/statistics.
- Formalization
  - Let A be the **randomized function** (namely a **mechanism**) to be computed on a set of records.
    - A is the actual function to be computed f + **noise**.
  - Let S be a subset of the possible values taken by A.
  - A provides ϵ-differential privacy if for all r, S:

$$\mathrm{P}[A(D) \in S] \leq e^{\epsilon} \times \mathrm{P}[A(D_{-r}) \in S]$$

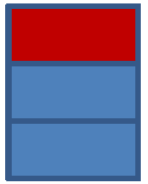
Differential Privacy - Intuition

For every pair of inputs that differ in one value



D_1 D_2

For every output …

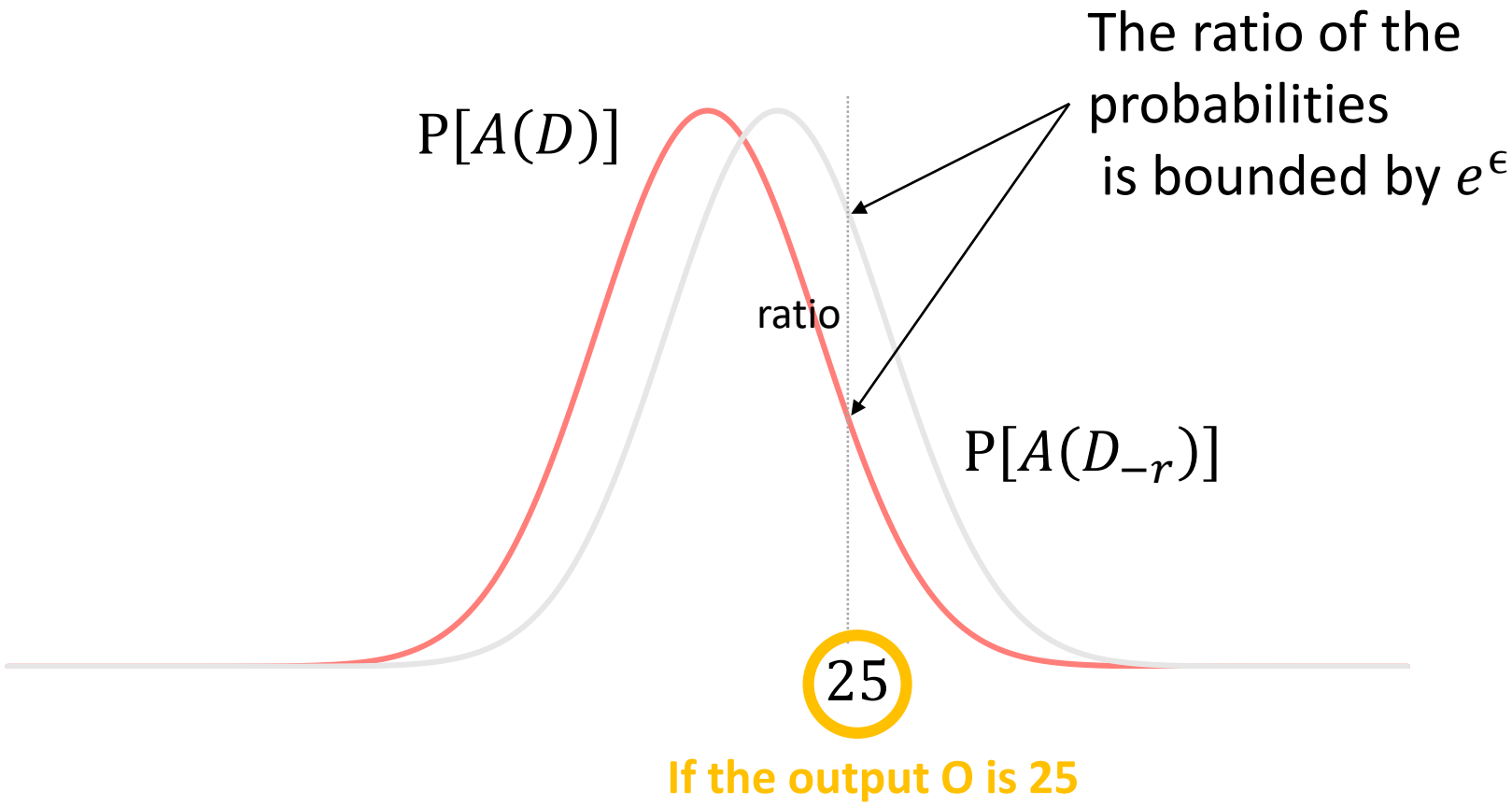


O

Adversary should not be able to distinguish between any D_1 and D_2 based on any O

$$\log\left(\frac{\Pr[A(D_1) = O]}{\Pr[A(D_2) = O]}\right) < \varepsilon \qquad (\varepsilon>0)$$

ϵ-Differential Privacy



How to achieve ϵ-Differential Privacy (simple case)

How to achieve ϵ-differential privacy (simple case)?

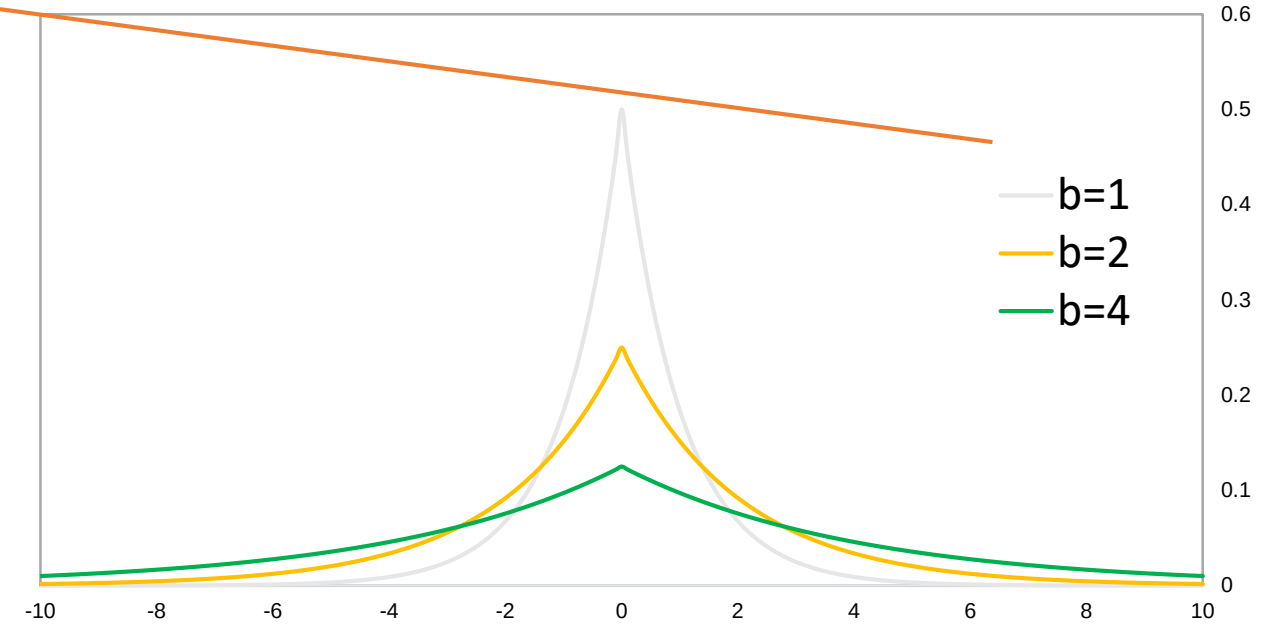
Assume f is a scalar function, i.e., $f: \mathcal{D} \rightarrow \mathbb{R}$ (e.g., "number of records with cancer"?)

Return $\boldsymbol{A(D) = f(D) + \mathbf{Lap}\left(\frac{\Delta f}{\epsilon}\right)}$

$\mathrm{Lap}\left(\frac{\Delta f}{\epsilon}\right)$ is **noise** drawn from a **Laplacian** distribution of parameter $\frac{\Delta f}{\epsilon}$

Δf is the **sensitivity** of function f:

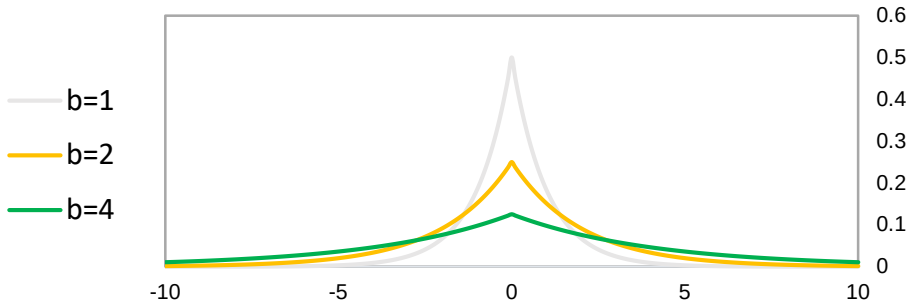
$$\Delta f = \max_r |f(D) - f(D_{-r})|$$



Why Laplacian Distribution?

- The Laplacian distribution is: $\mathrm{Lap}\left(\frac{\Delta f}{\epsilon}\right) = \frac{\epsilon}{2\Delta f} \exp(-\frac{x\epsilon}{\Delta f})$.
- The distortion of the result depends on both the sensitivity and privacy guarantee:
  - The higher the sensitivity, the higher the distortion
  - The higher the privacy guarantee (the lower ϵ), the higher the distortion

- This distribution has highest density at 0 (good for accuracy).
- This distribution is symmetric about 0 and has a heavy tail.



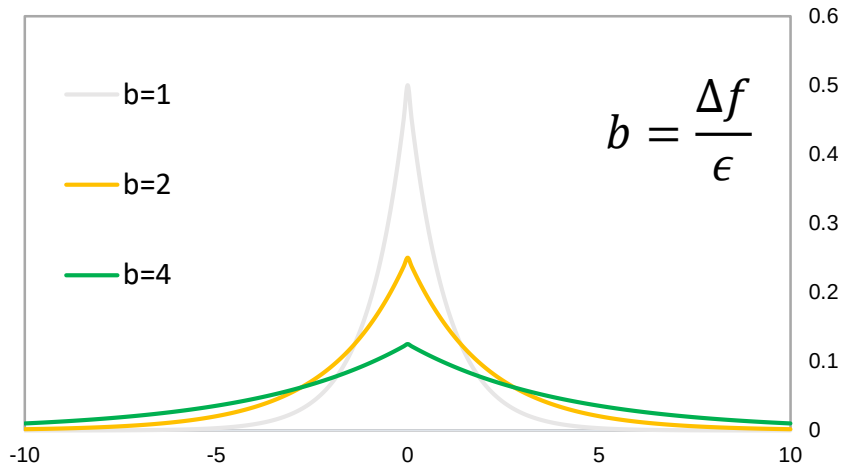
How to choose the parameters ?

Selecting ε

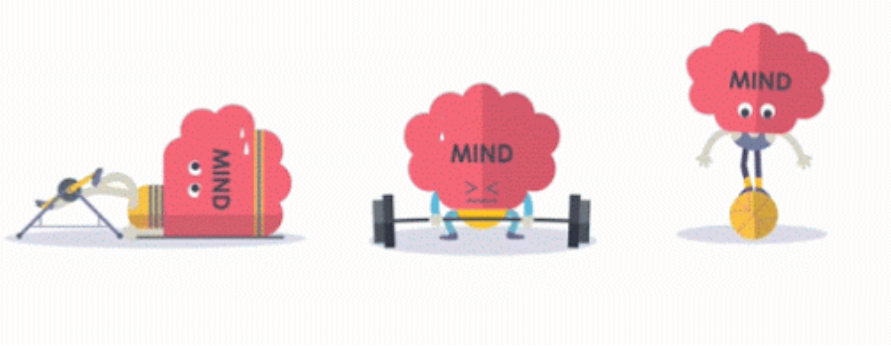
- The parameter ε is public (remember: no security by obscurity
- Selection of ε by Cynthia Dwork:
  - *"We tend to think of ε as 0.01, 0.1, or in some cases, ln2 or ln3"*
  - Smaller ε means better privacy
  - But, what about the utility ?

It depends on the sensitivity!

$$\Delta f = \max_{r} |f(D) - f(D_{-r})|$$



What is the sensitivity of… ?



For any two neighboring databases (D, D_{-r}):

$$\Delta f = \max_{D, D_{\pm r}} || F(D) - F(D_{-r}) ||$$

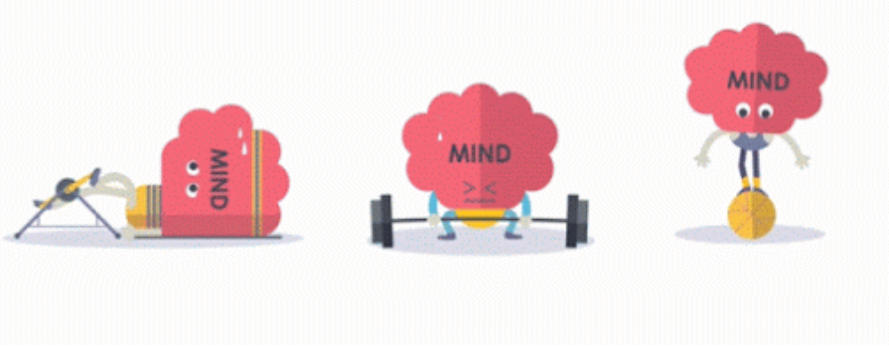
Sensitivity of counting queries:

- The number of elements in the database with a given property *P.*

Sensitivity of histogram queries:

- Suppose each entry in d takes values in $\{c_1, c_2, \ldots, c_n\}$.
- *Histogram*(d) = $\{m_1, m_2, \ldots, m_n\}$, m_i = (#entries in d with value c_i)

What is the sensitivity of… ?



For any two neighboring databases (D, D_{-r}):

$$\Delta f = \max_{D, D_{\pm r}} || F(D) - F(D_{-r}) ||$$

Sensitivity of counting queries:

- The number of elements in the database with a given property *P.*
- By adding or deleting one element of the database, *F* can change by at most 1.
- **Δf(counting) = 1**

Sensitivity of histogram queries:

- Suppose each entry in d takes values in $\{c_1, c_2, \ldots, c_n\}$.
- *Histogram*(d) = $\{m_1, m_2, \ldots, m_n\}$, m_i = (#entries in d with value c_i)
- **If adding/removing one entry: Δf(histogram) = 2**
- **If elements move from one entry to another: Δf(histogram) = 2**

Composability of Differential Privacy

Theorem: If algorithms $F_1, F_2, \ldots, F_k$ use independent randomness and each F_i satisfies ε_i-differential privacy, respectively. Then outputting all the answers together satisfies differential privacy with

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_k$$

Does privacy increase or decrease?

How to ensure differential privacy ?

- **Input perturbation**
  - Add noise directly to the database (≠ perturbed dataset can be published)
    + independent of the algorithm & easy to reproduce
    - determining the amount of required noise is difficult

- **Output perturbation**
  - Add noise to the function (statistic) output
    + easier to control privacy & better guarantees than input perturbation
    - results cannot be reproduced

- **Algorithm Perturbation**
  - Inherently add noise to the algo
    + algorithm can be optimized with the noise addition
    - difficult to generalize & depends on the inputs

More on these algorithms and variants in CS-523

Why is DP possible (while anonymization was impossible):

The final result depends on multiple personal records

However it does not depend much on any particular one (sensitivity)

Therefore adding a little bit of noise to the result, suffices to hide any record contribution

For full anonymization.... one would need to add a lot of noise to all the entries

But… the architecture is different: **one Trusted-Third-Party holds the data**!

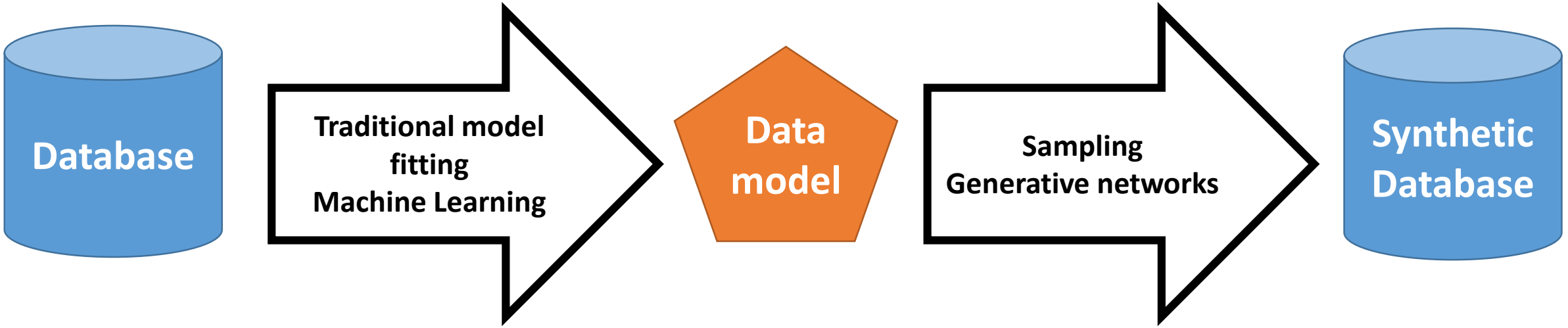
Also… after some uses utility drops

best use: one-time, **data collection**!!

Google RAPPOR ← Collect data from phones
Apple ← Collect data from phones
Federated learning ← Share models
Smart energy ← Collect measurements

Synthetic databases: A new hope

Synthetic databases: A new hope



Problems:

How to know which features to model? Features determine utility!

How to measure privacy? Some processes provide DP, but for what attribute?

Takeaways

Anonymizing is difficult, but privacy-preserving statistical querying is possible

Differential privacy: a notion to reason about privacy

- Key idea: given an output not possible to learn about one individual's participation
- Algorithms available for protecting different types of queries

Synthetic data can be an option for improving data sharing

- Very early days…

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

SPRING
SECURITY AND PRIVACY ENGINEERING LABORATORY

Information Security and Privacy (COM-402)

Part 6: Privacy enhancing technologies

Metadata protection

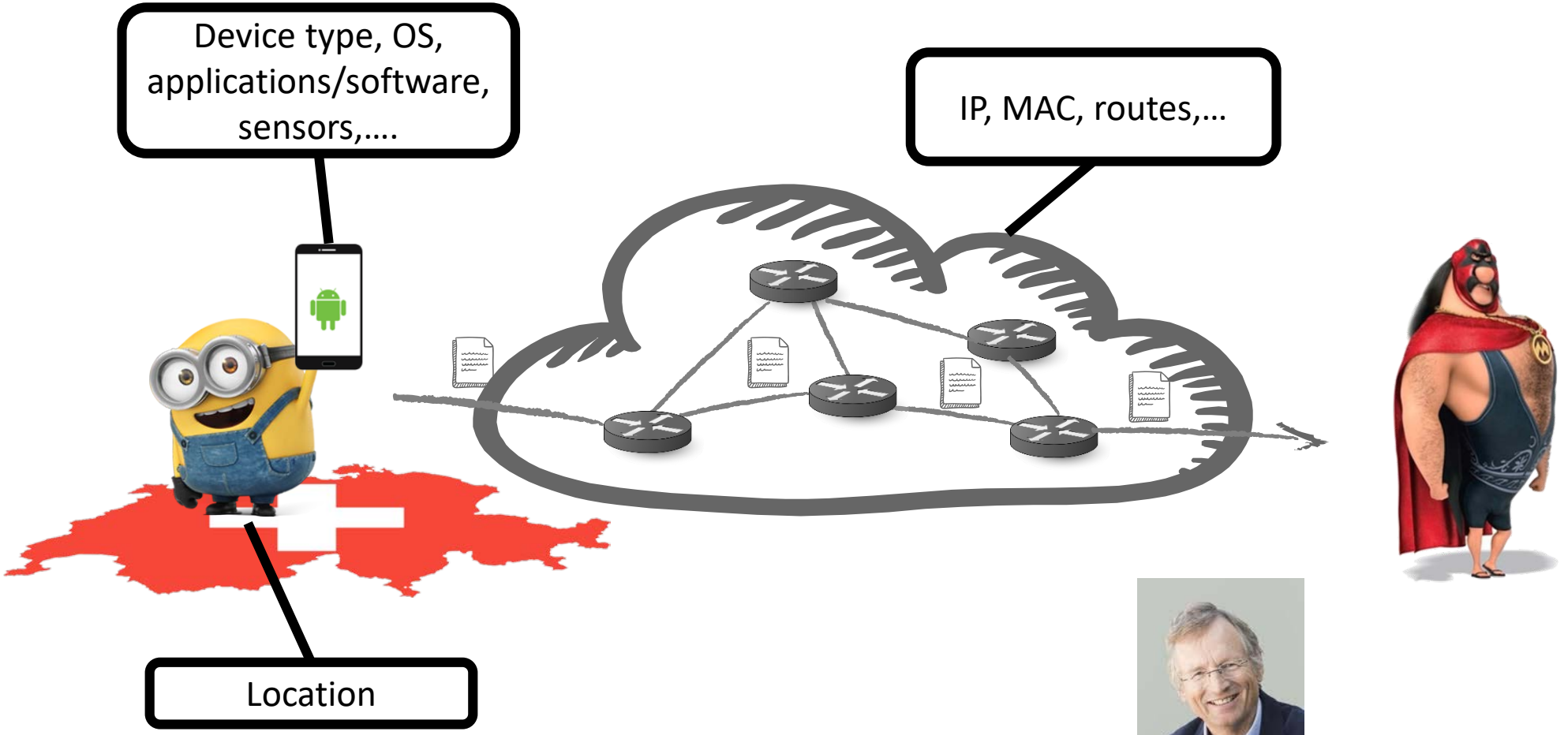
Carmela Troncoso

SPRING Lab

carmela.troncoso@epfl.ch

Some slides/ideas adapted from: JP Hubaux, Bryan Ford, Vitaly Shmatikov

Beyond data publication… privacy when data is in transit



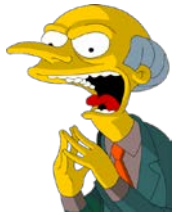
Computing privately on data

The adversary is anyone and VERY powerful

Your Parents

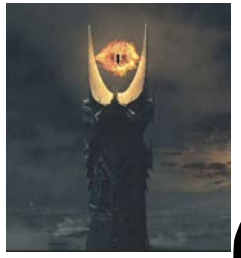
Your Children

Intelligence agencies



The Boss

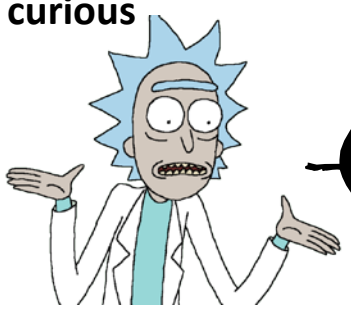
Your Roomates



Anybody curious

ISPs

SysAdmins

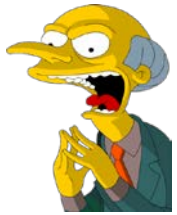


The adversary is anyone and VERY powerful

Intelligence agencies

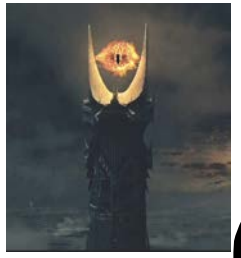
Your Parents

Your Children



The Boss

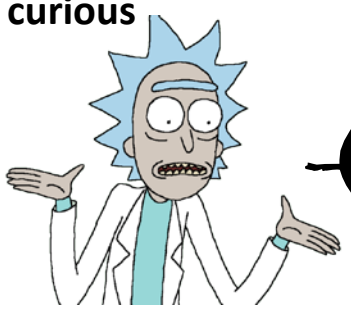
Your Roomates



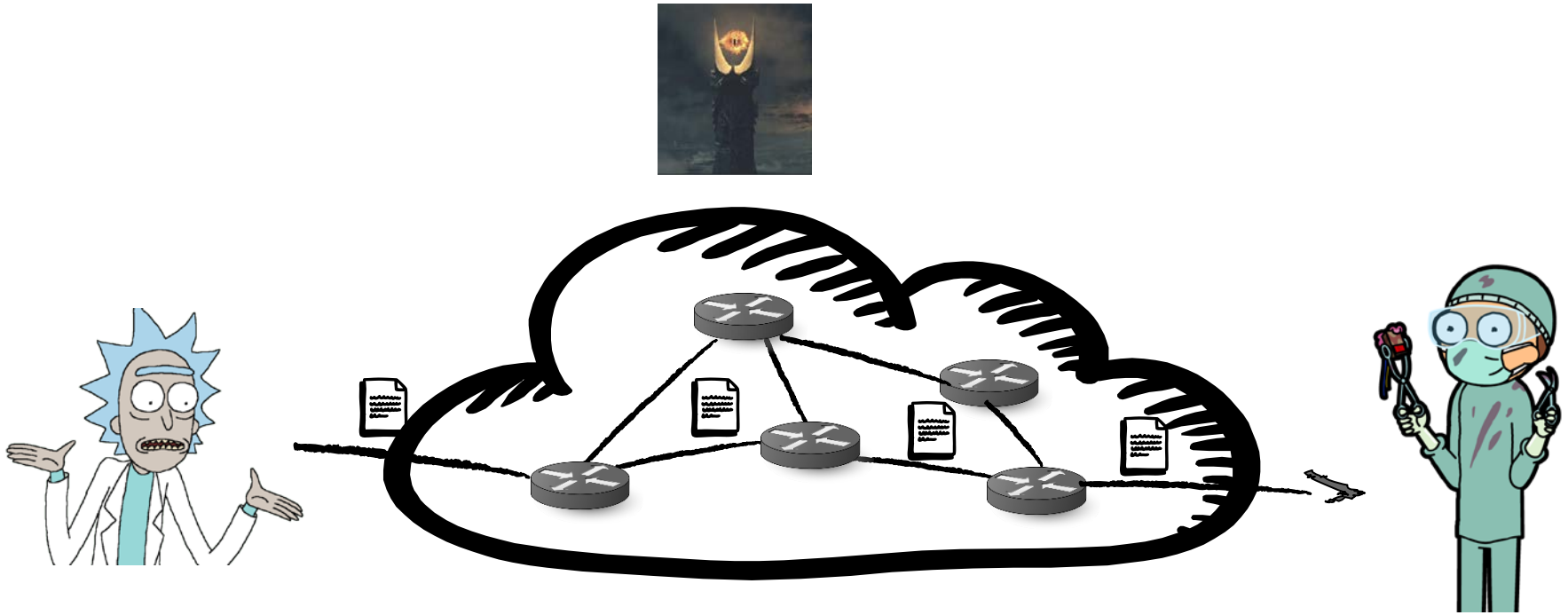
Anybody curious

ISPs

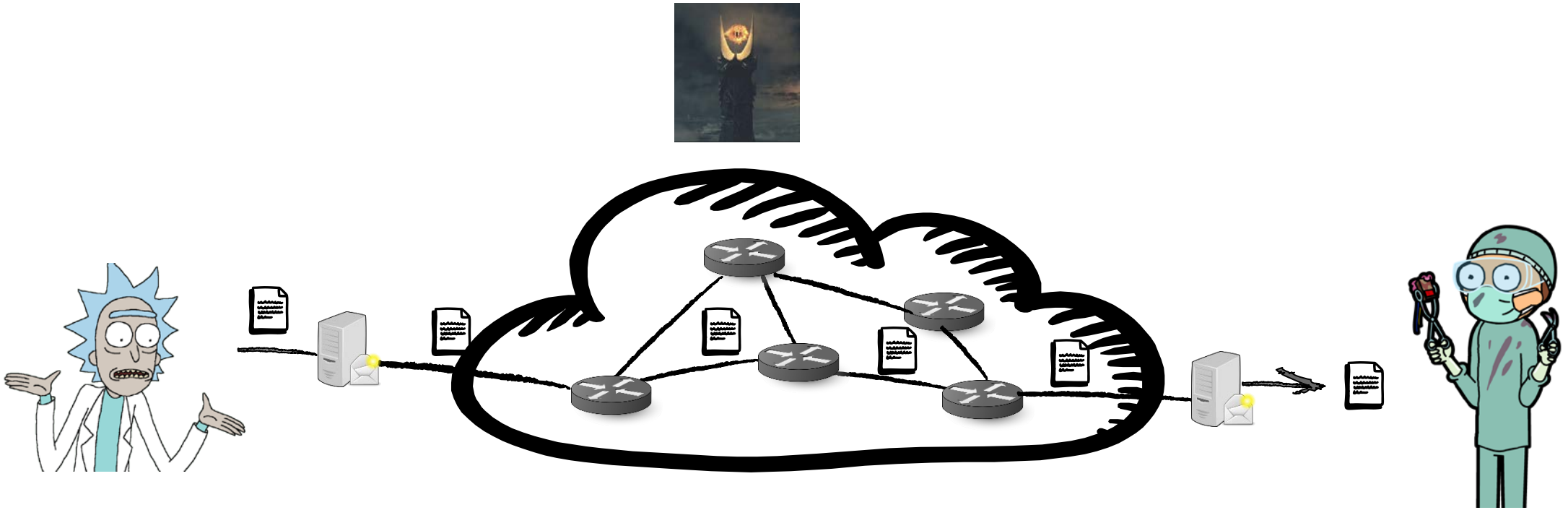
SysAdmins



End to End Encryption



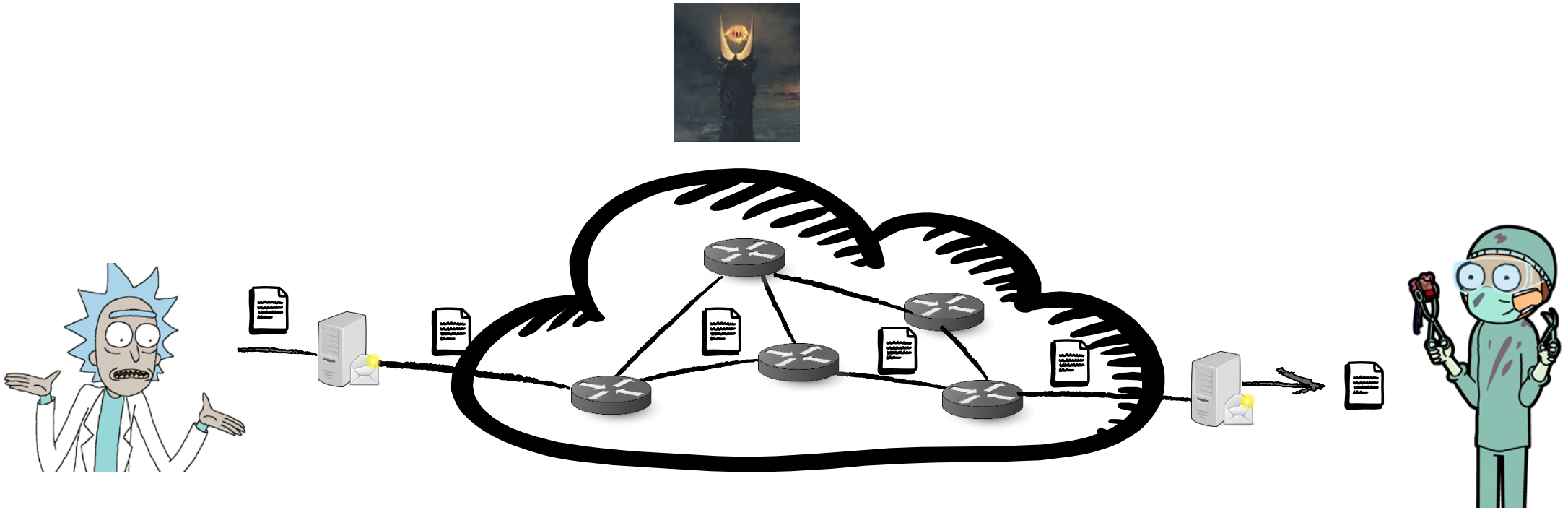
End to End Encryption



Cryptography → Confidentiality!
(and integrity and authenticity)

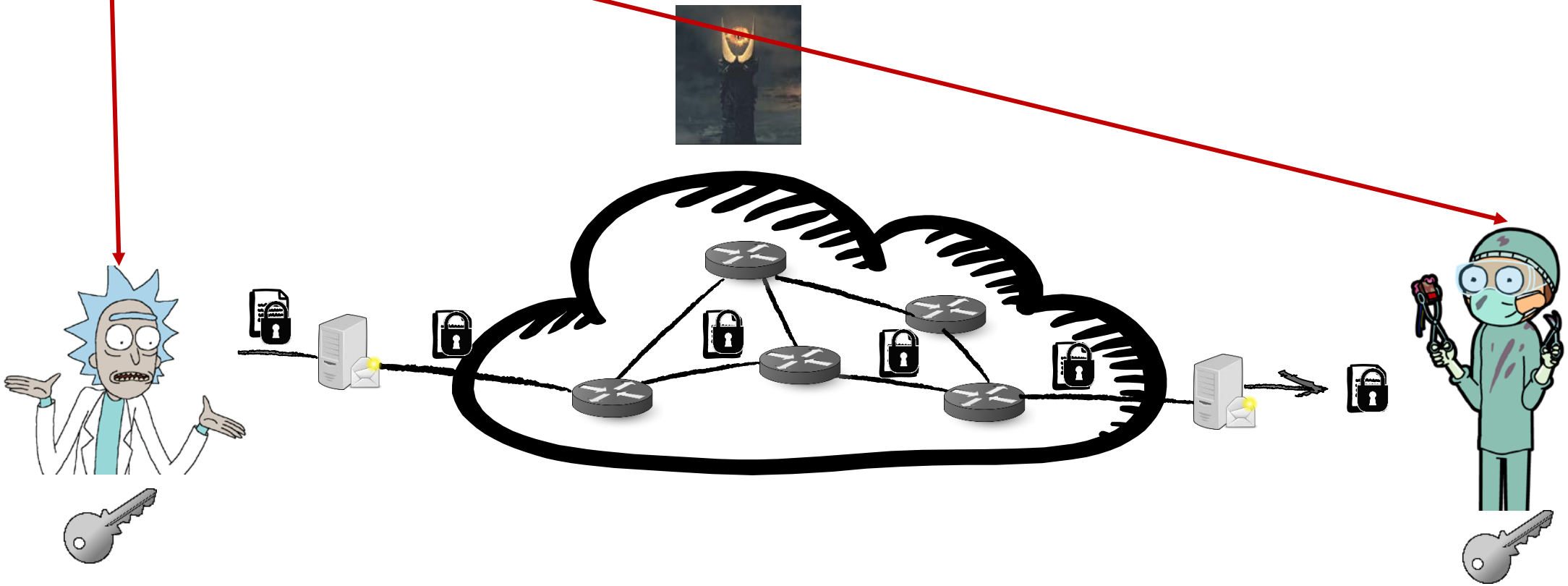
End to End Encryption

What is an End?



Cryptography → Confidentiality!
(and integrity and authenticity)

End to End Encryption



End to End Encryption

Perfect forward secrecy

Cryptography → Confidentiality!

BUT WHAT IF SOMEONE FORCES YOU TO DISCLOSE THE KEY?



1) Start with keys that allow Alice to authenticate Bob.
   - Public key encryption

2) Alice and Bob create fresh public keys and exchange them

3) They establish fresh shared keys, and talk secretly
   - Diffie Hellman

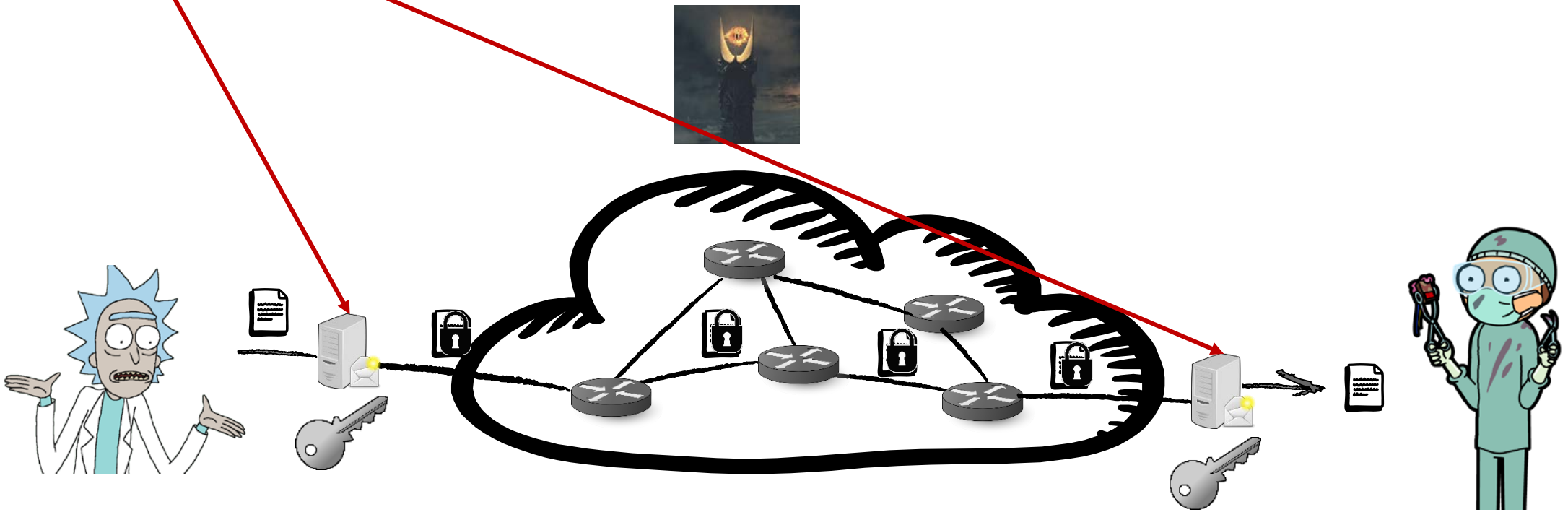
4) Once done, they delete the shared keys.

ONE-TIME USE KEYS:
EPHEMERAL KEYS

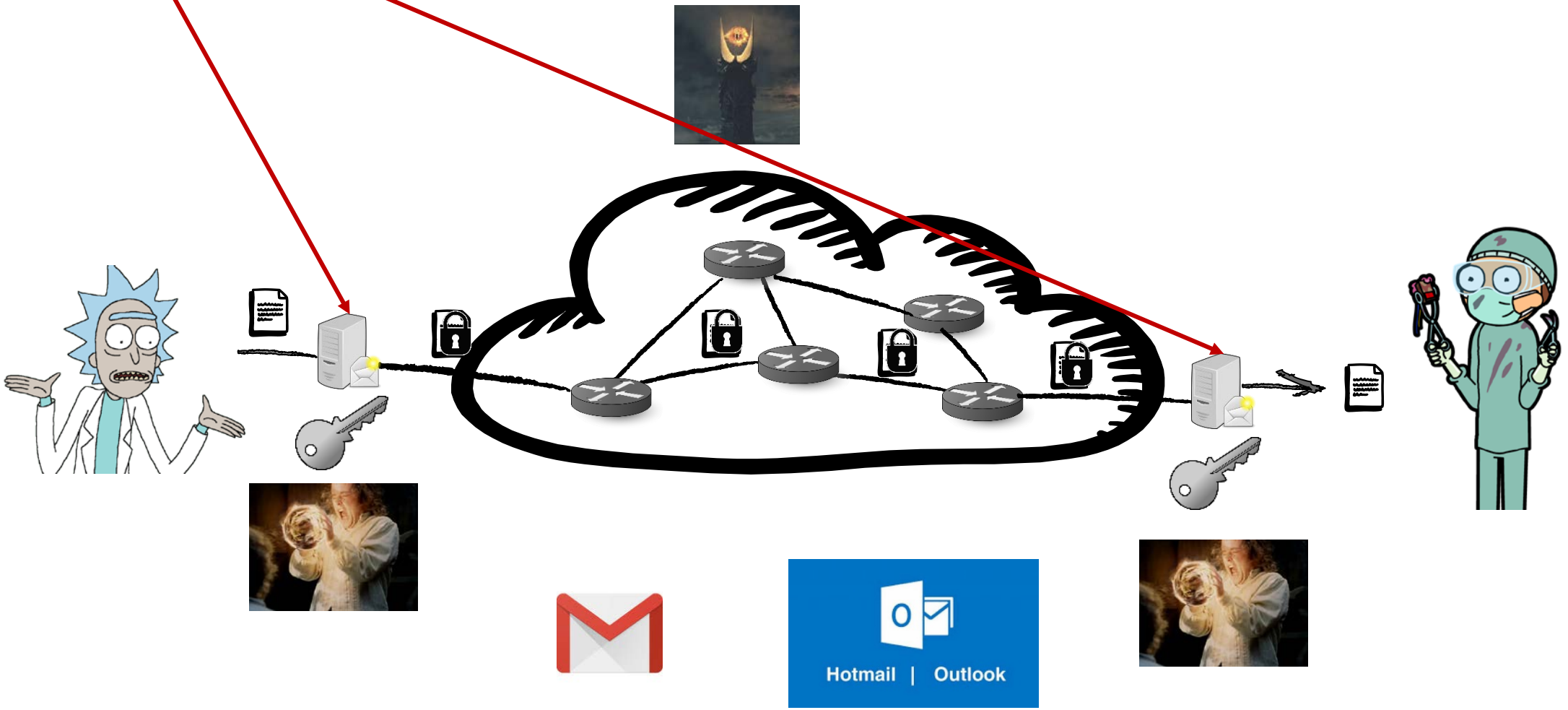
AFTER A CONVERSATION IS OVER
NO-ONE CAN DECRYPT WHAT WAS SAID!!!

PLAUSIBLE DENIABILITY!!

End to End Encryption



End to End Encryption



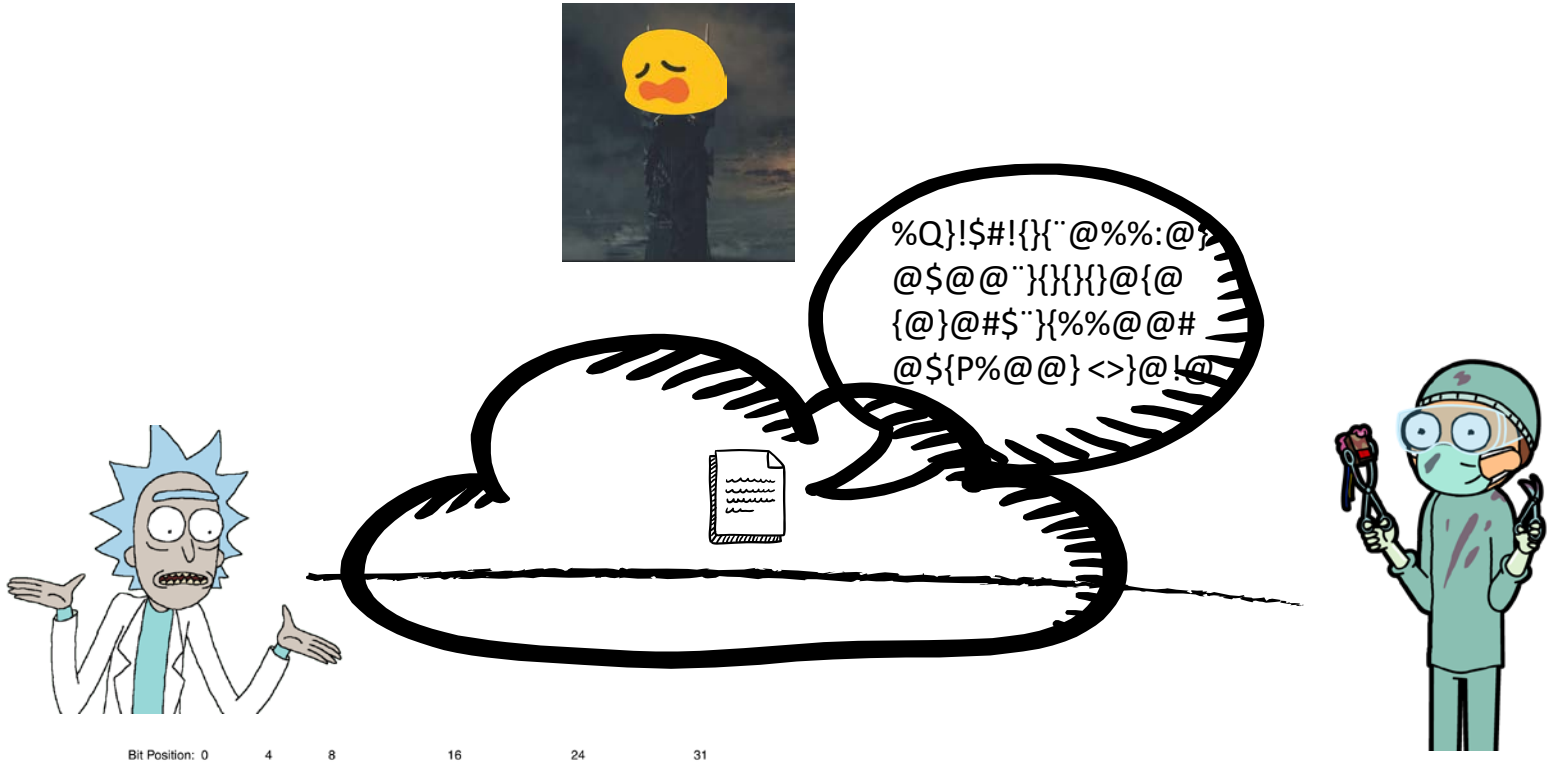
So we can encrypt! What is the problem?

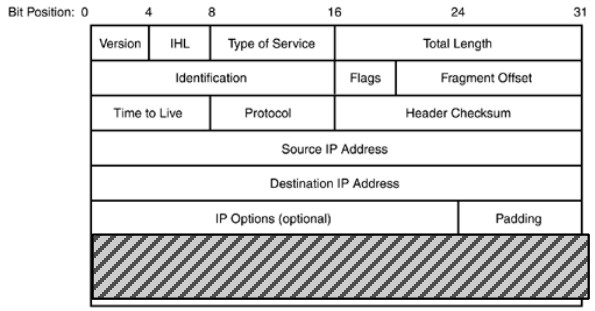
So we can encrypt! What is the problem?

Bit Position: 0	4	8	16	24	31
Version	IHL	Type of Service	Total Length		
Identification			Flags	Fragment Offset	
Time to Live		Protocol	Header Checksum		
Source IP Address					
Destination IP Address					
IP Options (optional)					Padding
Data					
More Data...?					

IPv4 Header (RFC 791, 1981)

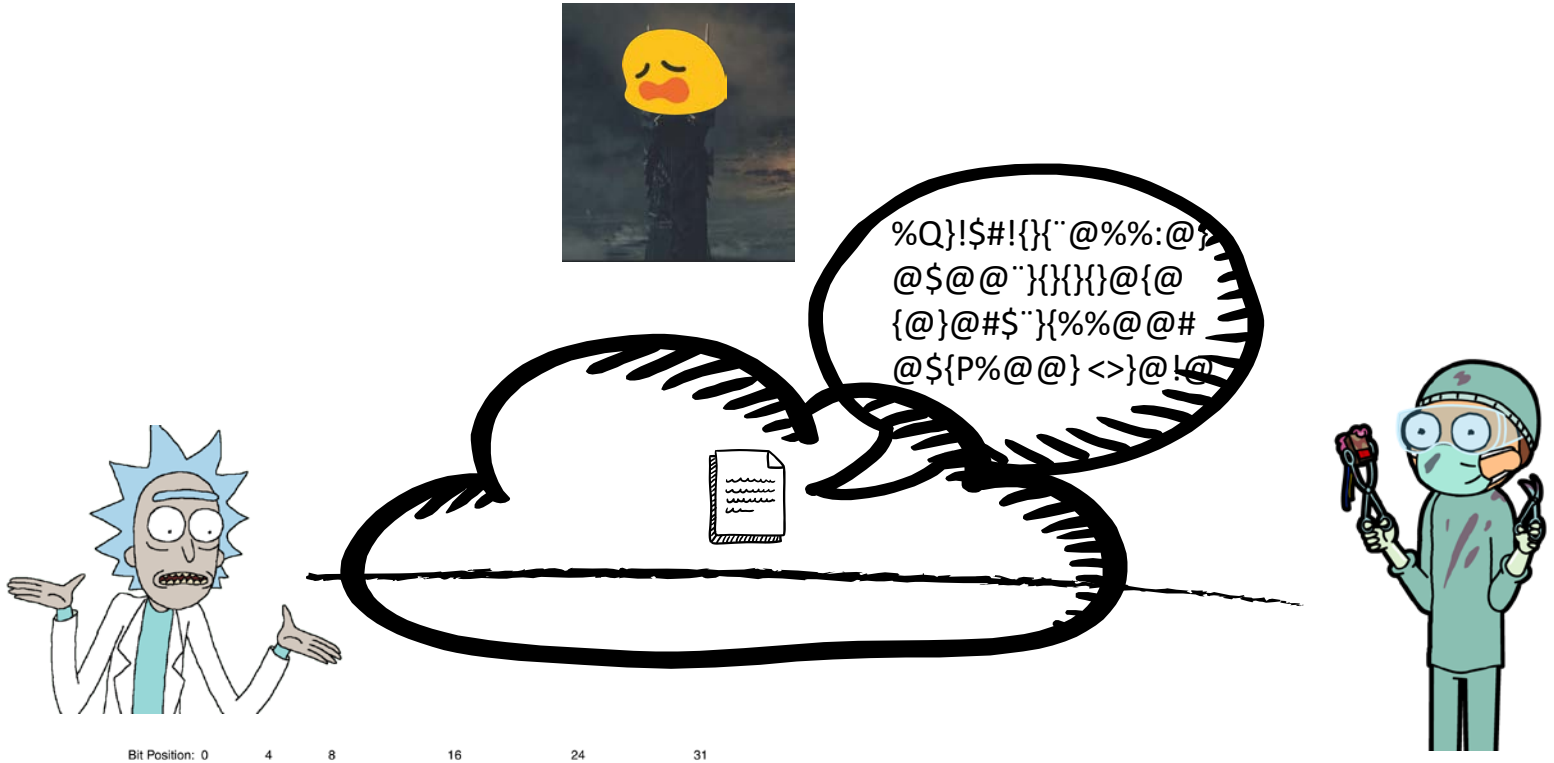
So we can encrypt! What is the problem?

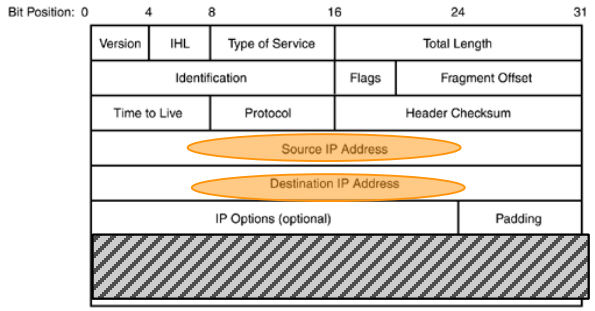




IPv4 Header (RFC 791, 1981)

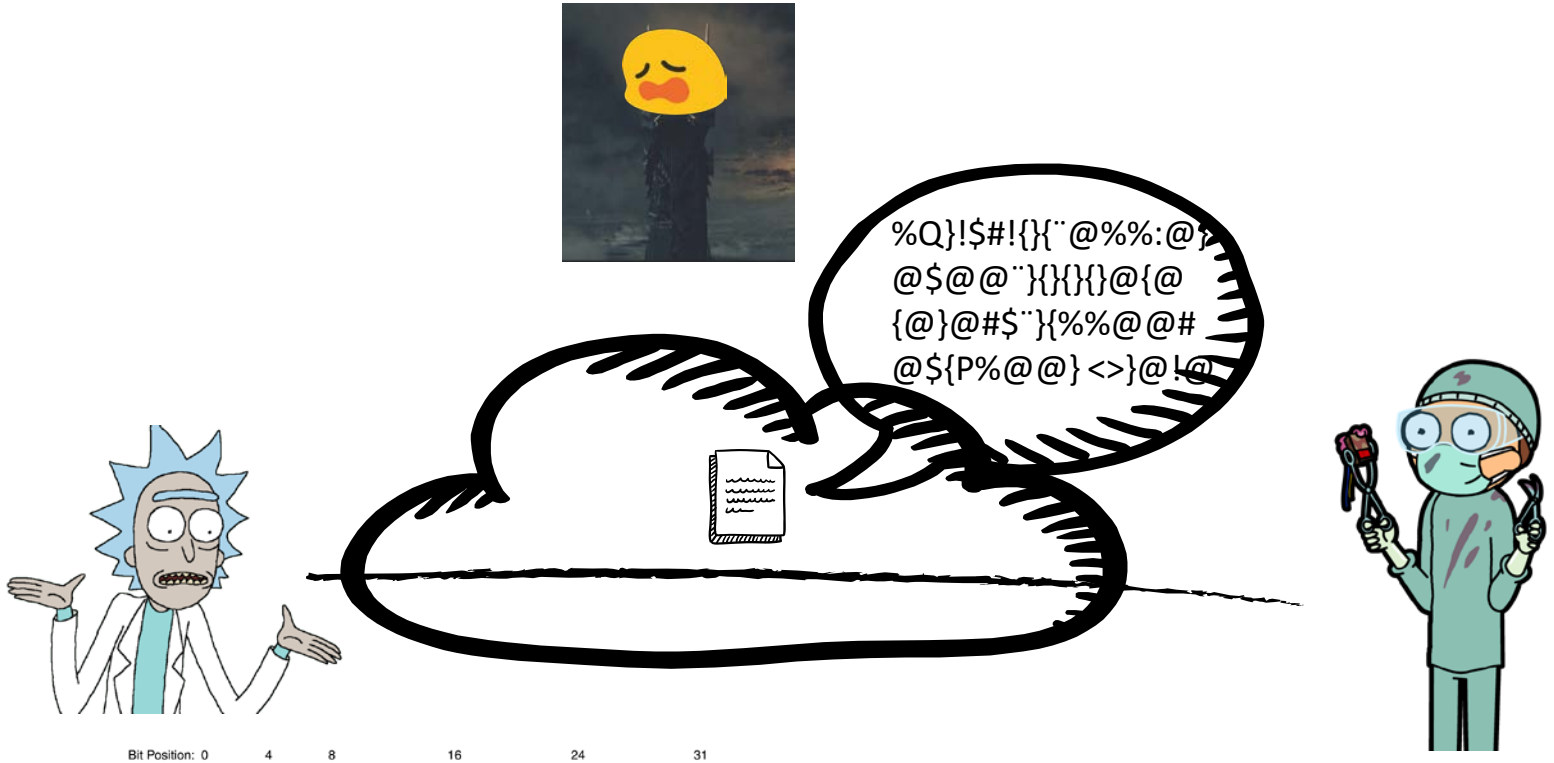
So we can encrypt! What is the problem?

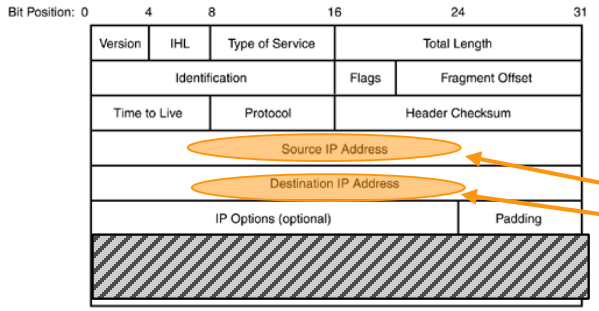




IPv4 Header (RFC 791, 1981)

So we can encrypt! What is the problem?

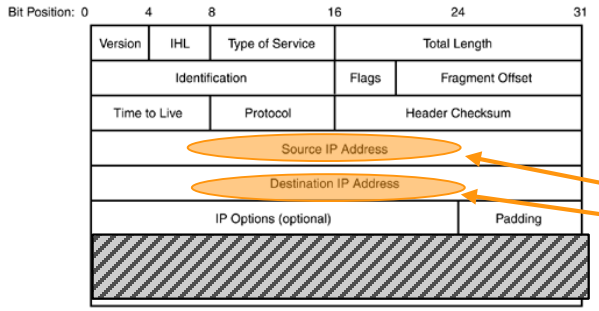




IPv4 Header (RFC 791, 1981)

Same for Ethernet, TCP, SMTP, IRC, HTTP, ...

The problem is Traffic Analysis



IPv4 Header (RFC 791, 1981)

Same for Ethernet, TCP, SMTP, IRC, HTTP, ...

Traffic WHAT?

Wikipedia: traffic analysis is the process of intercepting and examining messages in order to deduce information from patterns in communication

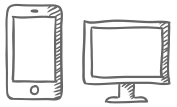
Making use of "just" traffic data of a communication (aka metadata) to extract information (as opposed to analyzing content or perform cryptanalysis)


Identities of communicating parties


Timing, frequency, duration


Location


Volume


Device

MILITARY ROOTS

M. Herman: "These non-textual techniques can establish **targets' locations**, order-of-battle and **movement**. Even when messages are not being deciphered, traffic analysis of the target's Command, Control, Communications and intelligence system and its patterns of behavior provides indications of his **intentions** and **states of mind**"

WWI: British troops finding German boats.

WWII: assessing size of German Air Force, fingerprinting of transmitters or operators (localization of troops).

NOWADAYS

Diffie&Landau: "Traffic analysis, not cryptanalysis, is the backbone of communications intelligence"

Stewart Baker (NSA): "metadata **absolutely tells you everything about somebody's life**. If you have enough metadata, you don't really need content."

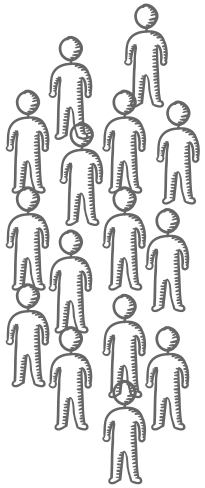
Tempora, MUSCULAR → XkeyScore, PRISM

Herman, Michael. Intelligence power in peace and war. Cambridge University Press, 1996.
Diffie, Whitfield, and Susan Landau. Privacy on the line: The politics of wiretapping and encryption. MIT press, 2010.
http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded

Anonymous communications – Abstract model

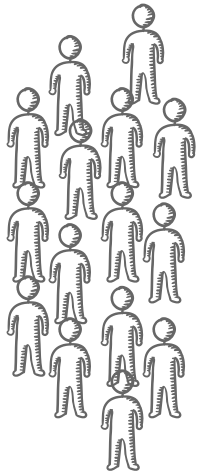
IDs
Timing
Volume
Length
…

SENDERS



ANONYMOUS COMMUNICATION SYSTEM

RECEIVERS



Bitwise unlinkability

Use cryptography to make inputs and outputs to the anonymous communication systems appearance (bits) different

(re)packetizing + (re)schedule

Destroy patterns (traffic analysis resistance)

Anonymous communications – Abstract model

IDs
Timing
Volume
Length
…

Bitwise unlinkability

Use cryptography to make inputs and outputs to the anonymous communication systems appearance (bits) different

(re)packetizing + (re)schedule

Destroy patterns (traffic analysis resistance)

ONE-PROXY PROBLEMS

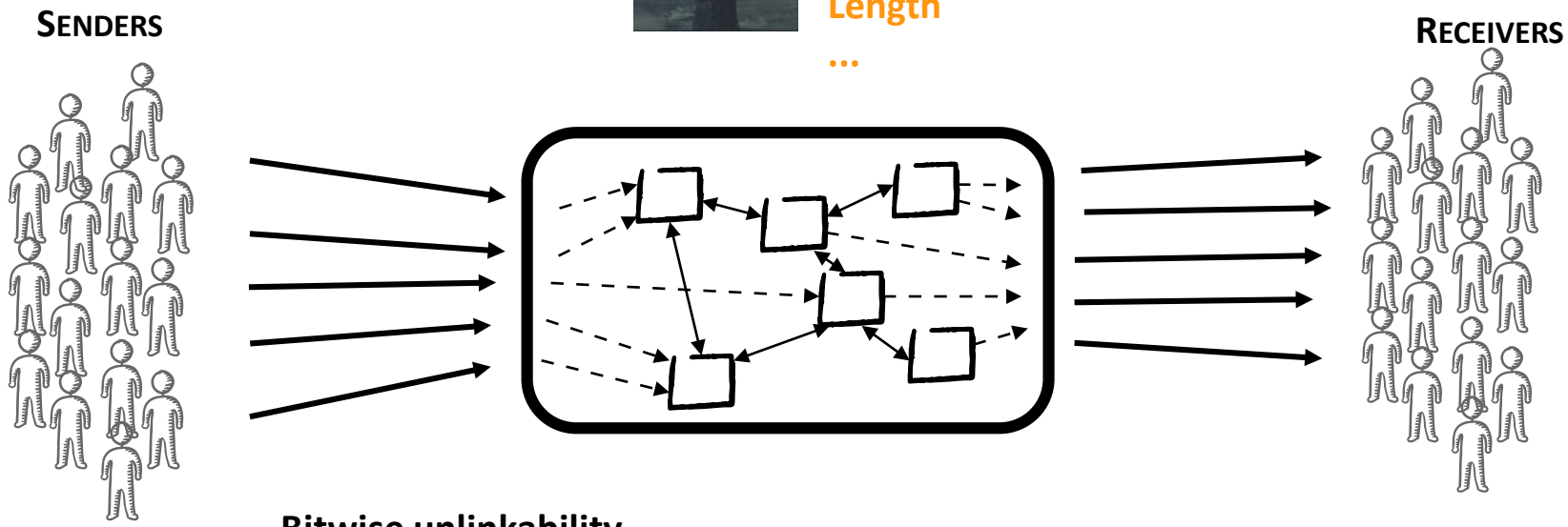
LOW THROUGHPUT

CORRUPT PROXY OR PROXY HACKED / COERCED

REAL CASE: PENET.FI VS THE CHURCH OF SCIENTOLOGY (1996)

Anonymous communications – Abstract model

IDs
Timing
Volume
Length
…



Bitwise unlinkability

Use cryptography to make inputs and outputs to the anonymous communication systems appearance (bits) different

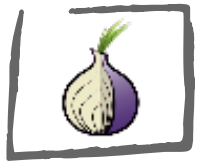
(re)packetizing + (re)schedule + (re)routing

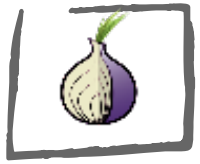
Destroy patterns (traffic analysis resistance)

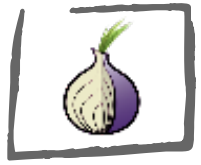
Load balancing

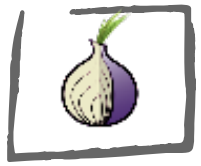
Distribute trust

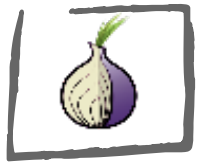
The Tor network – Onion routing

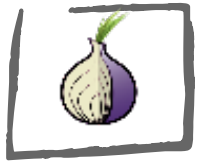




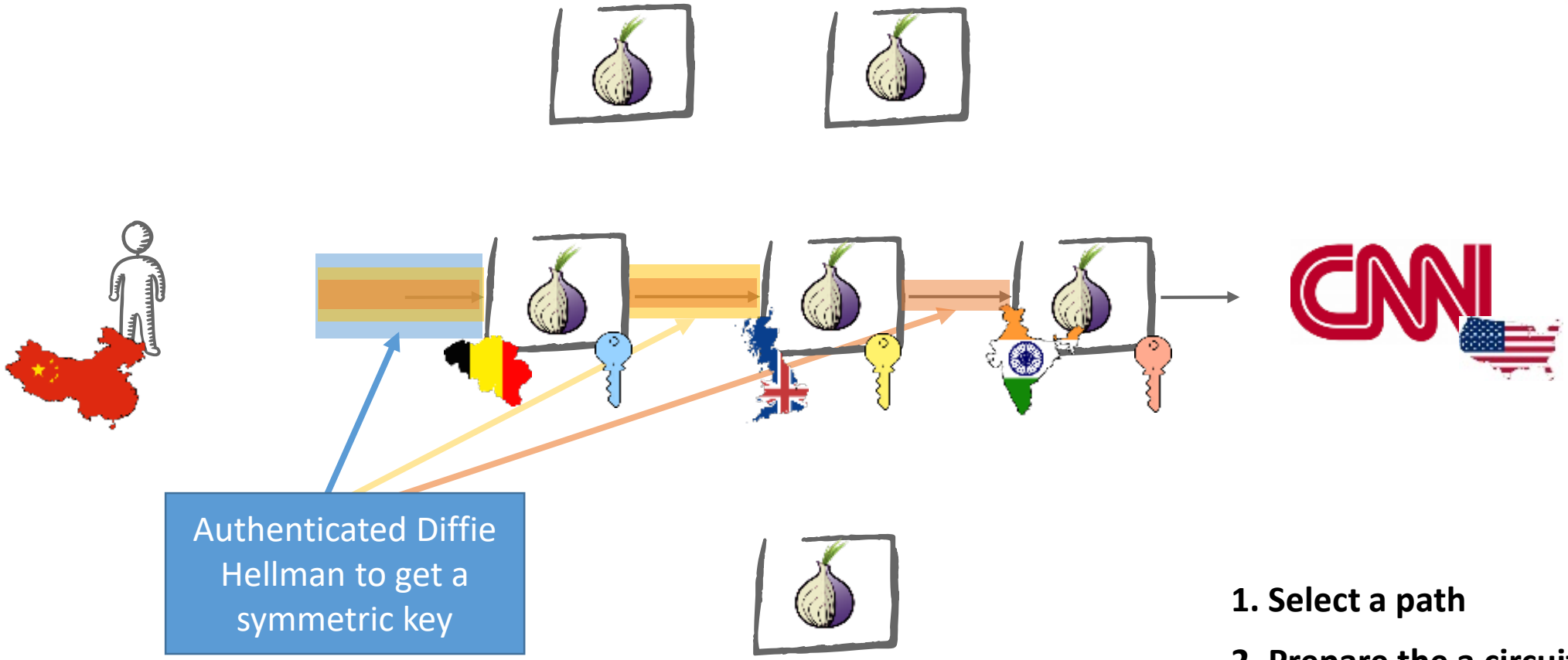








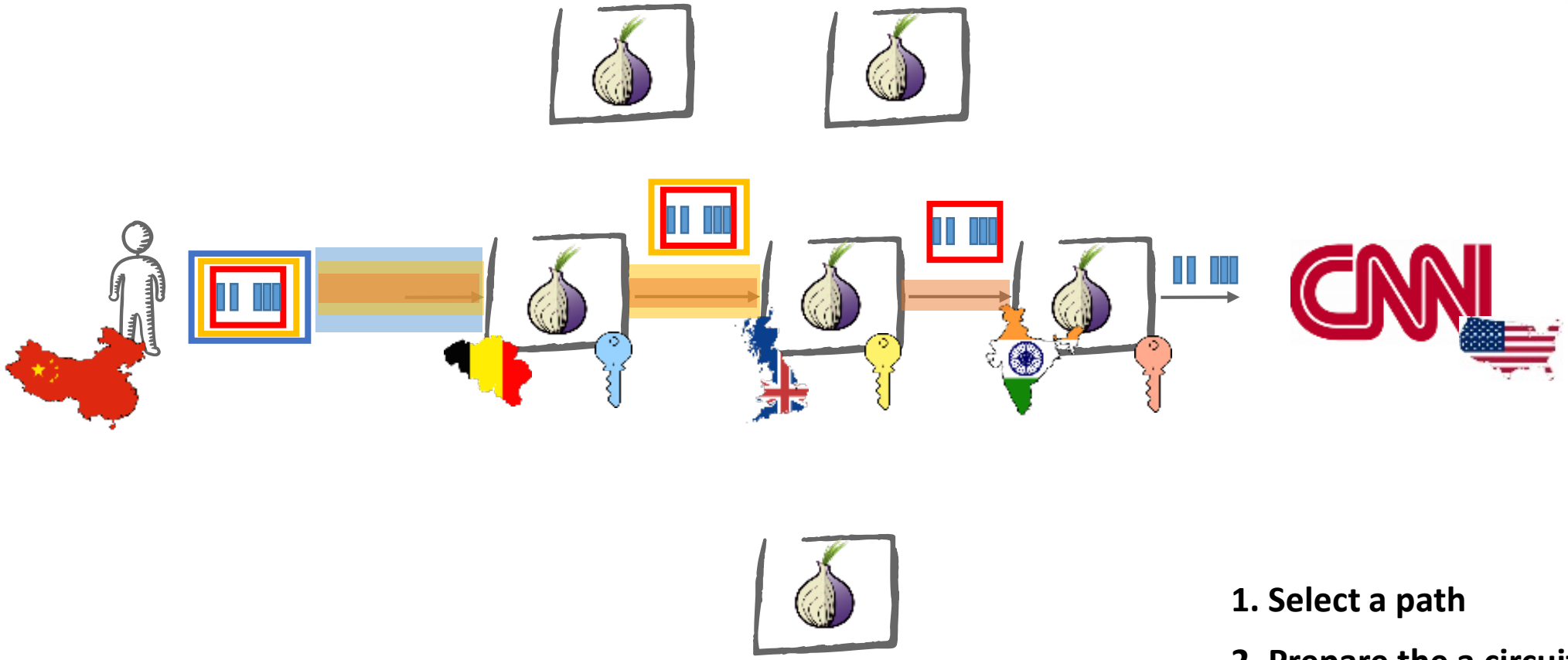
The Tor network – Onion routing



1. Select a path

2. Prepare the a circuit

The Tor network – Onion routing



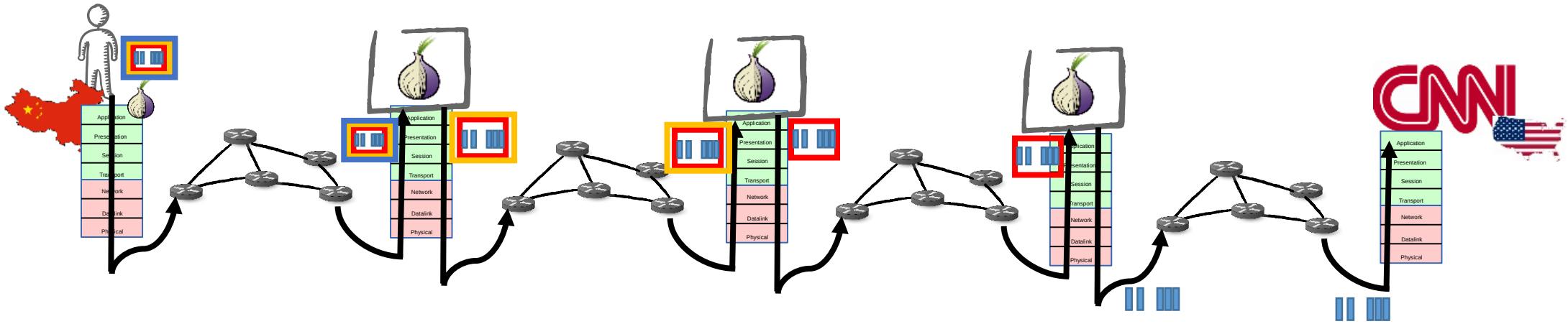
1. **Select a path**
2. **Prepare the a circuit**
3. **Send stream**

Anonymous communication networks are overlay networks

Nodes in anonymous communication networks (e.g., onion routers in Tor) are **not** internet routers. They work at the application layer!

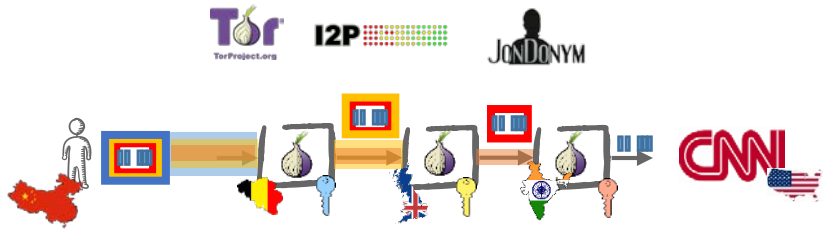
(overlay network = a computer network that is built on top of another network)

A more realistic view of how Tor traffic travels would be this



Anonymous communications out there

Low Latency

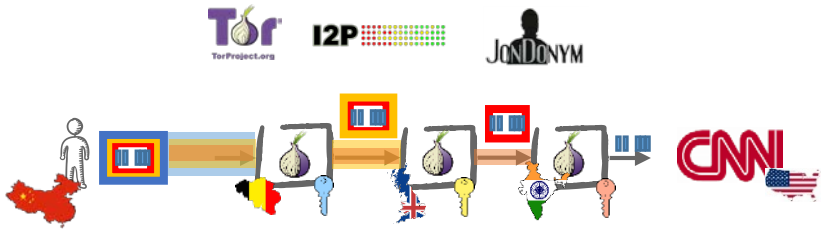


Web browsing, Instant Messaging, streaming

High Latency

Anonymous communications out there

LOW LATENCY

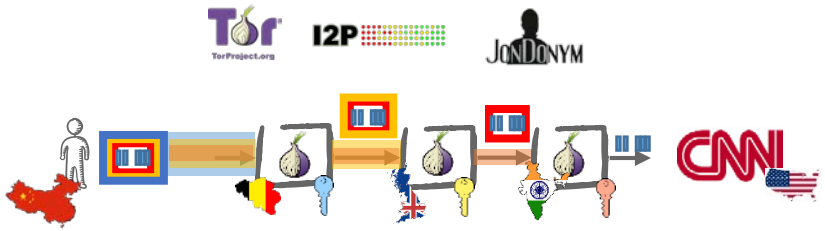


Web browsing, Instant Messaging, streaming

HIGH LATENCY

Anonymous communications out there

LOW LATENCY

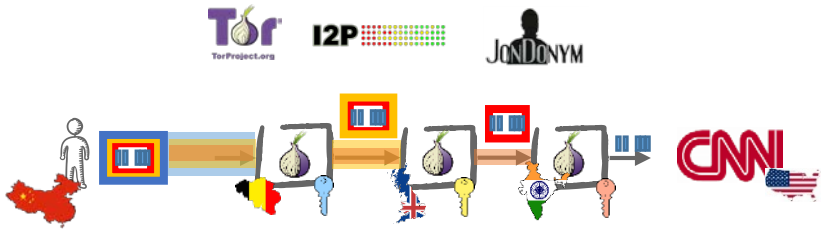


Web browsing, Instant Messaging, streaming

HIGH LATENCY

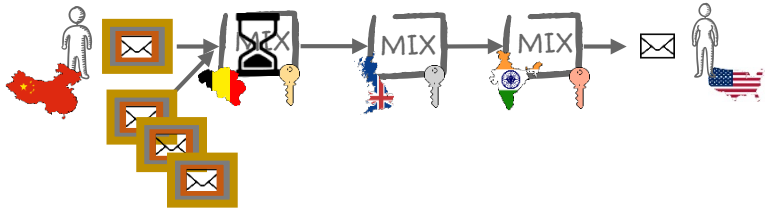
Anonymous communications out there

Low Latency



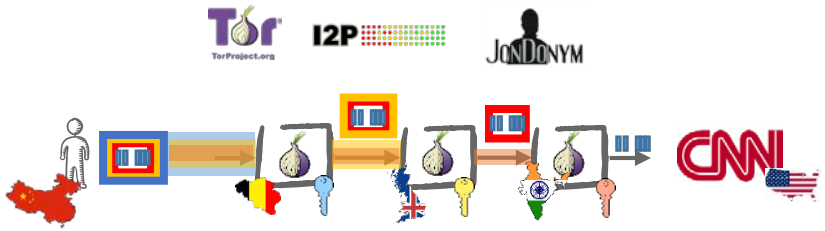
Web browsing, Instant Messaging, streaming

High Latency



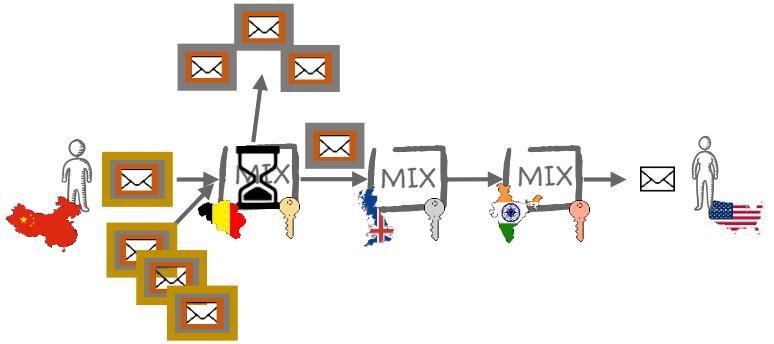
Anonymous communications out there

LOW LATENCY



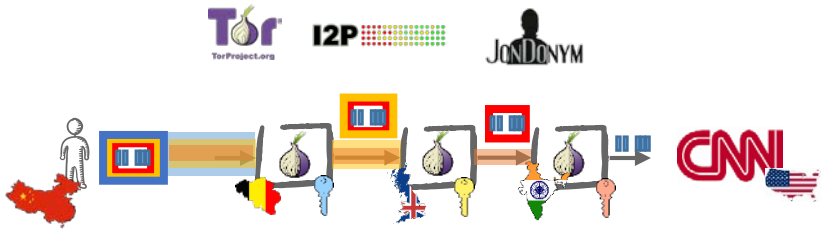
Web browsing, Instant Messaging, streaming

HIGH LATENCY



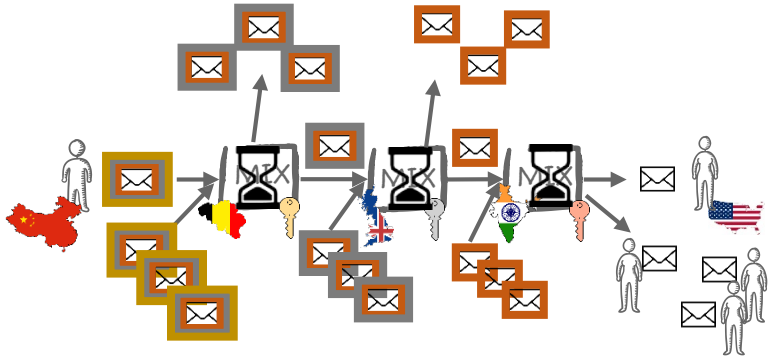
Anonymous communications out there

LOW LATENCY



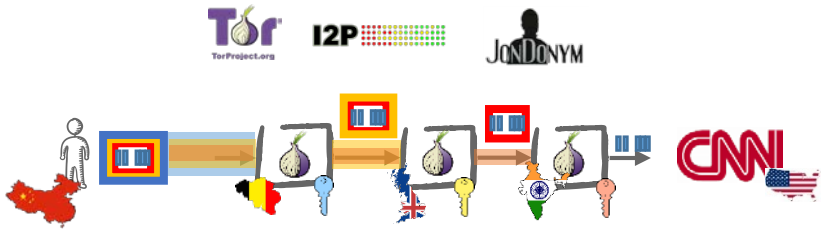
Web browsing, Instant Messaging, streaming

HIGH LATENCY



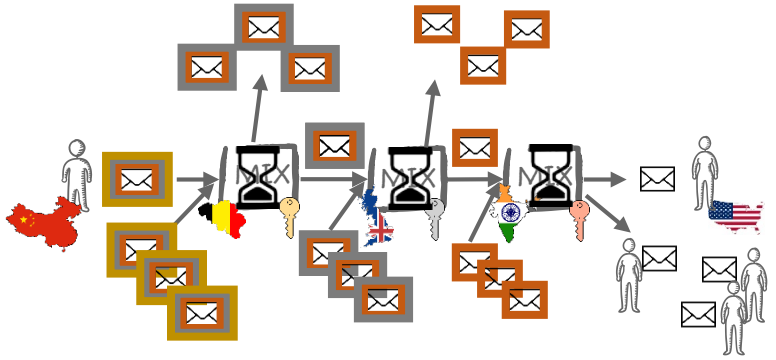
Anonymous communications out there

LOW LATENCY



Web browsing, Instant Messaging, streaming

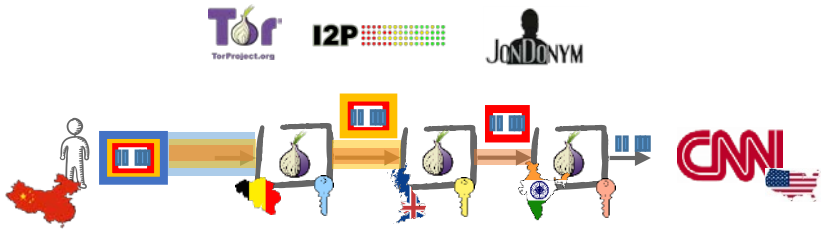
HIGH LATENCY



Email, Voting

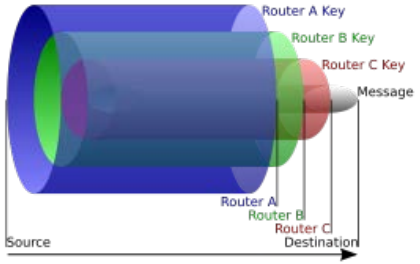
Anonymous communications out there

Low Latency

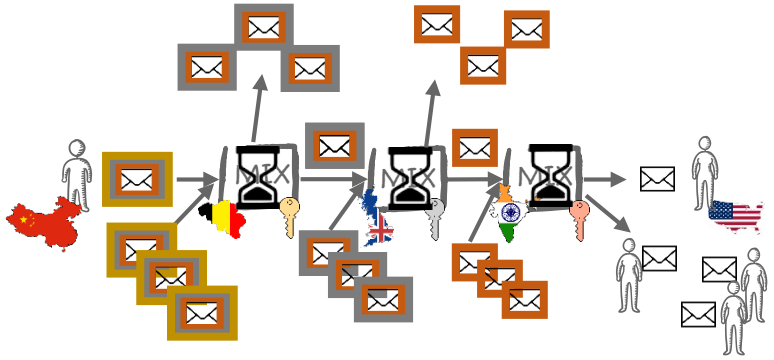


Web browsing, Instant Messaging, streaming

STREAM-based:  **fixed for the stream**



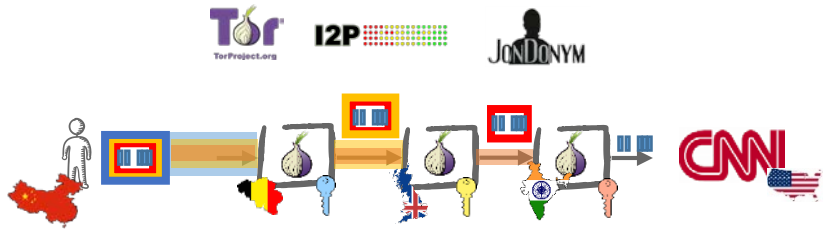
High Latency



Email, Voting

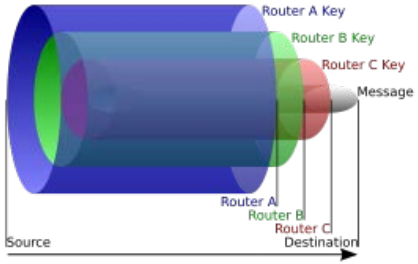
Anonymous communications out there

LOW LATENCY

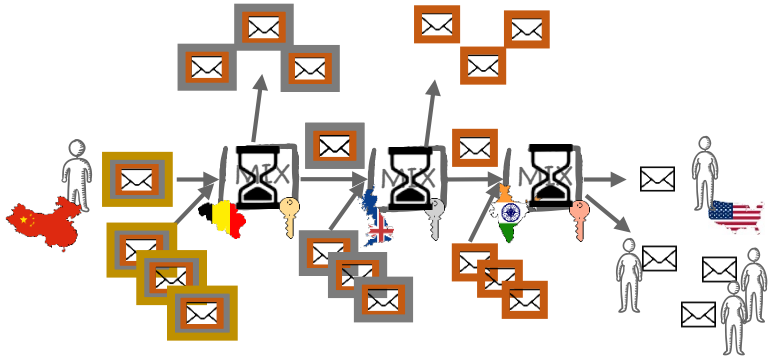


Web browsing, Instant Messaging, streaming

STREAM-based:  **fixed for the stream**



HIGH LATENCY



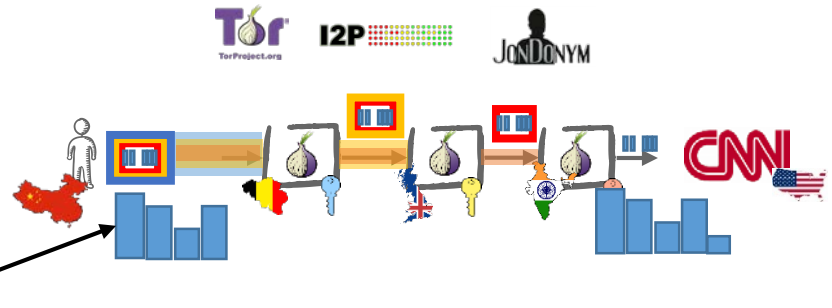
Email, Voting

MSG-based:  **vary every message**

One route per message + delays (slower!)

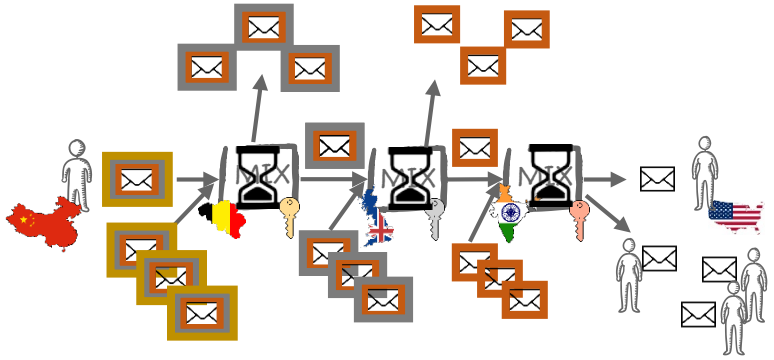
Anonymous communications out there

Low Latency



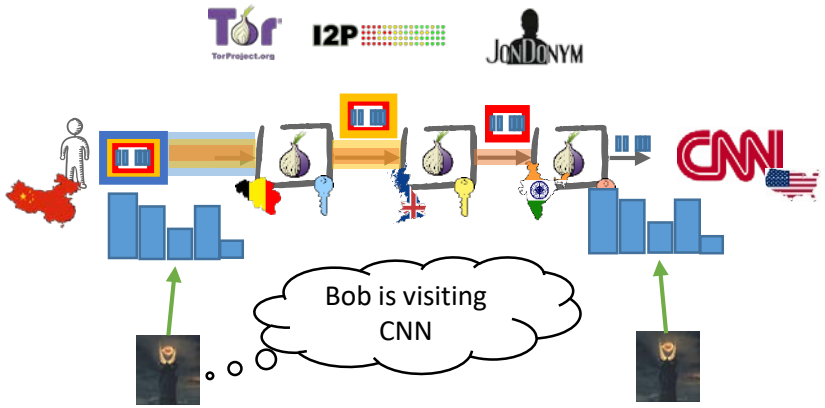
Volume of cells:
(e.g. how many cells per second)

High Latency

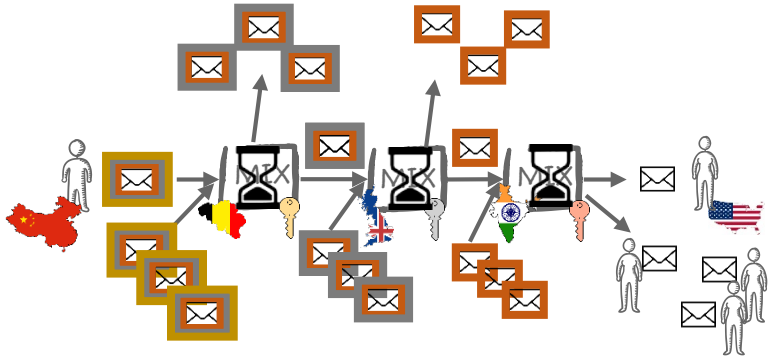


Anonymous communications out there

LOW LATENCY

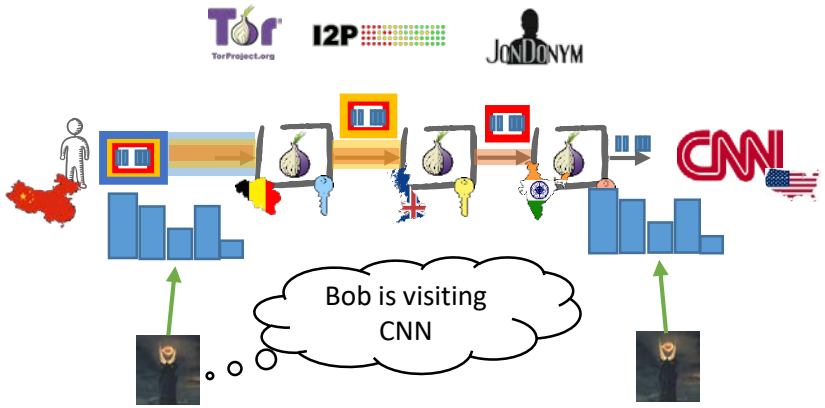


HIGH LATENCY



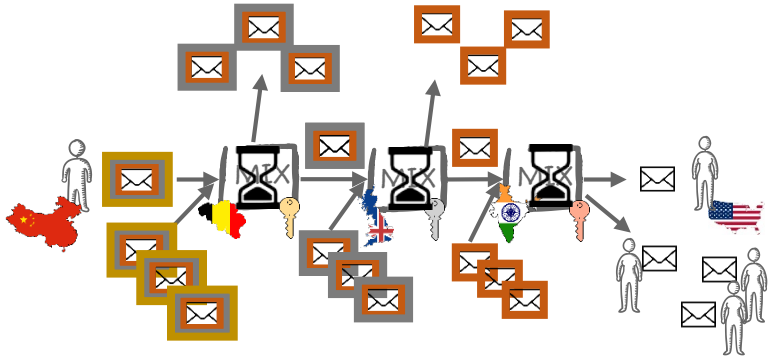
Anonymous communications out there

Low Latency



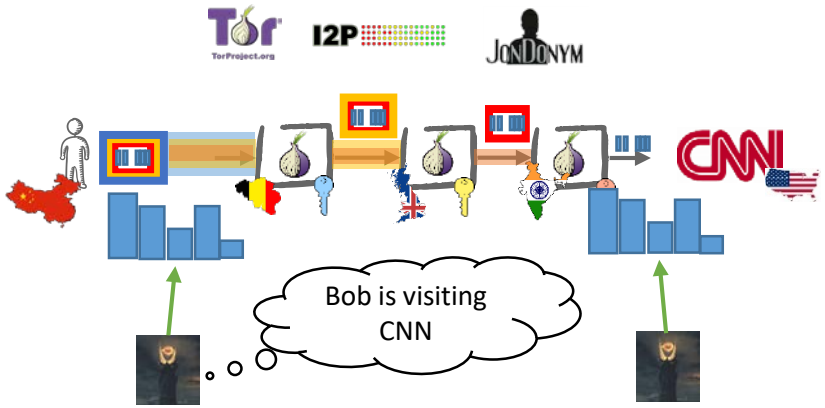
Cannot resist **Global** Adversary
(Tor assumes that the adversary cannot see both edges)

High Latency



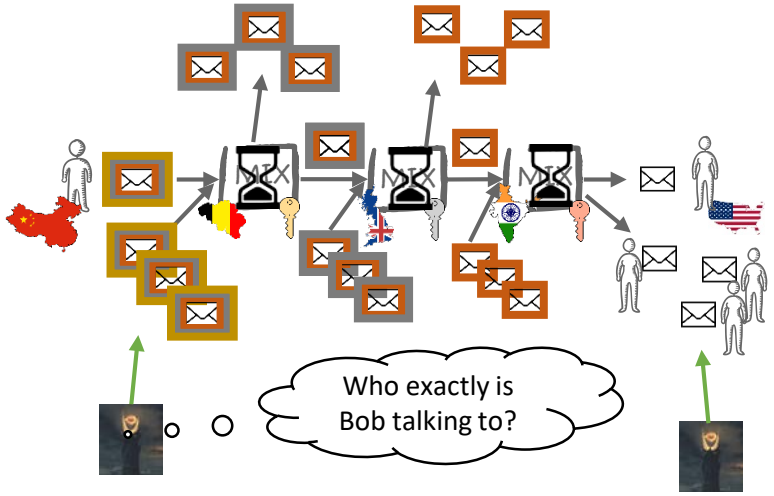
Anonymous communications out there

Low Latency



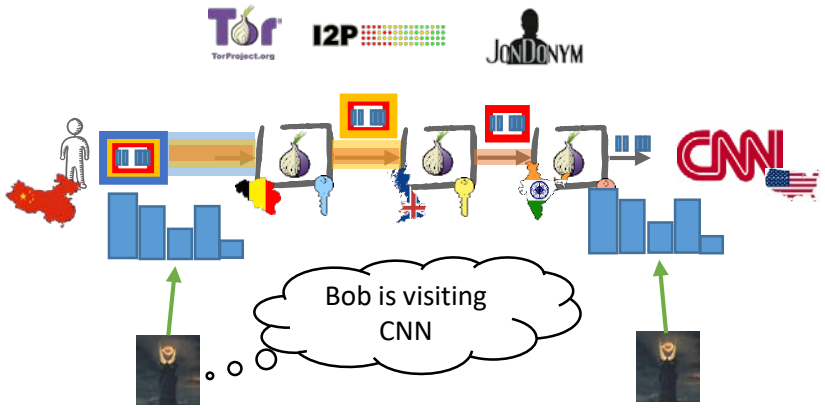
Cannot resist **Global** Adversary
(Tor assumes that the adversary cannot see both edges)

High Latency



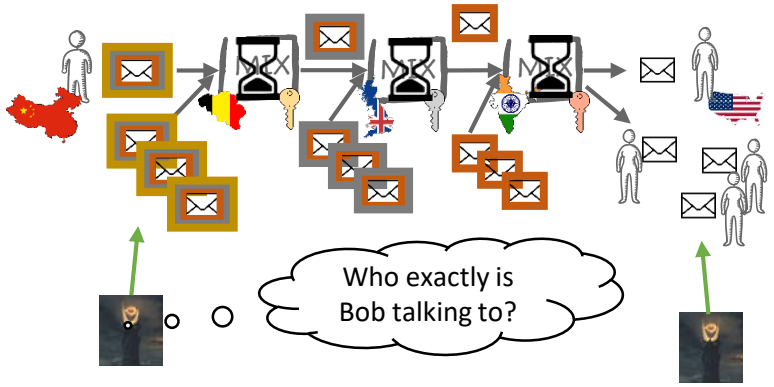
Anonymous communications out there

Low Latency



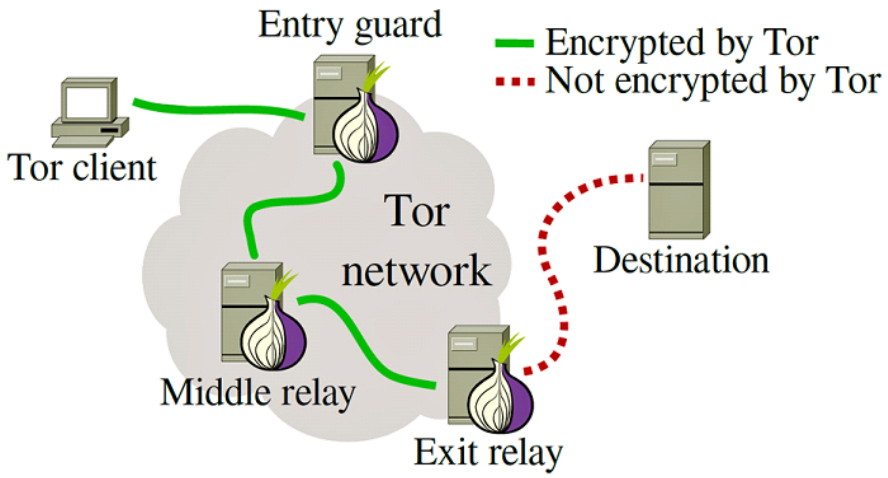
Cannot resist **Global** Adversary
(Tor assumes that the adversary cannot see both edges)

High Latency



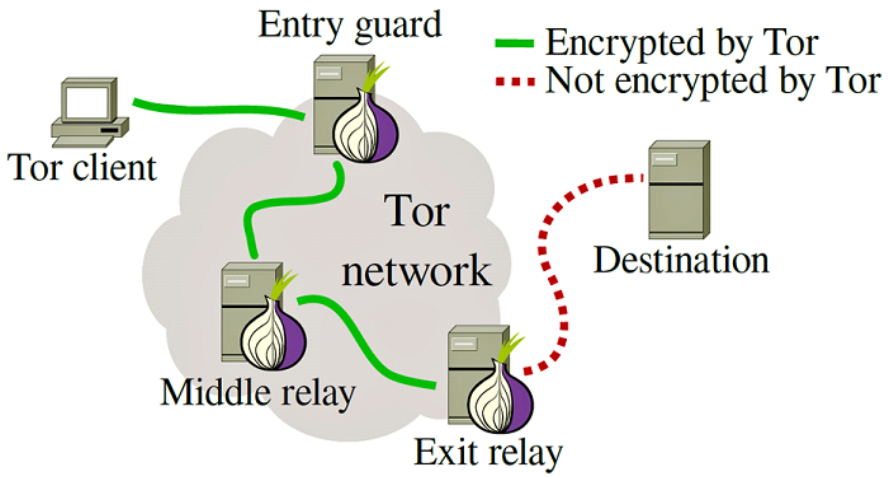
Global Adversary resistance
at the cost of latency
(and long term patterns revealed)

Anonymous communications vs. VPN



https://www.privacyend.com/difference-between-tor-and-vpn/

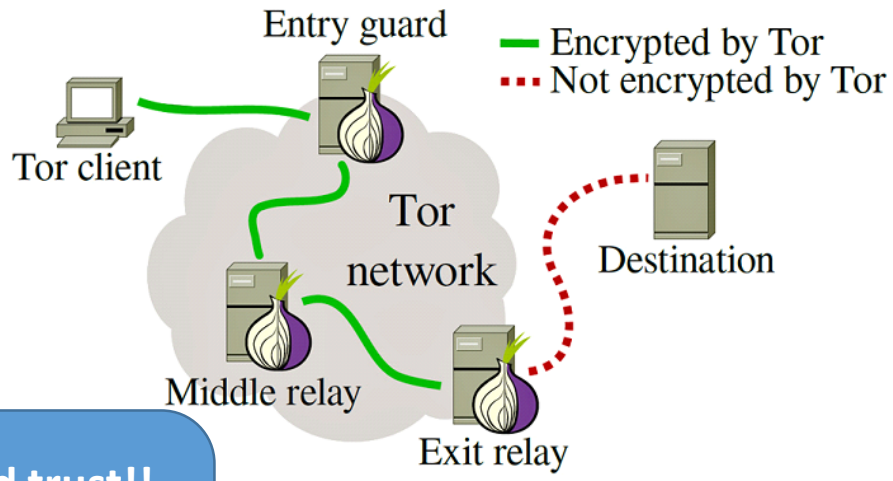
Anonymous communications vs. VPN



Different trust models!! Who is the adversary?

https://www.privacyend.com/difference-between-tor-and-vpn/

Anonymous communications vs. VPN



Decentralized trust!! Provides privacy as long as the adversary cannot see both edges

Centralized trust. No anonymity vs the VPN, or anyone seeing the VPN

https://www.privacyend.com/difference-between-tor-and-vpn/

Takeaways

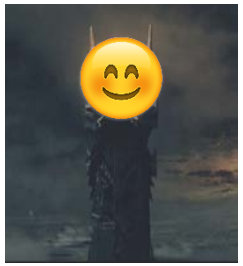
When thinking about end-to-end encryption it is important to think:

- Who are the ends of the communication vs. who is the adversary
- Forward secrecy, what happens if one key is compromised

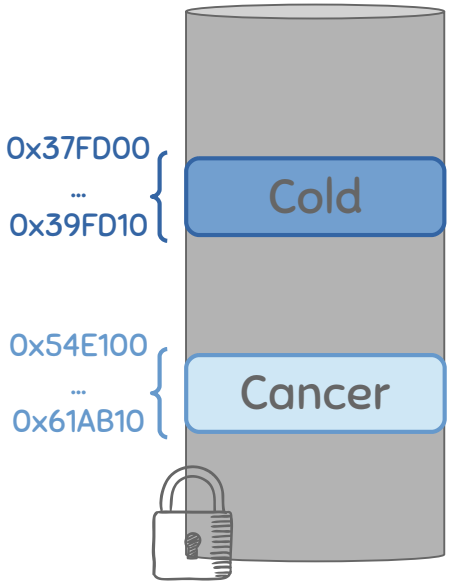
Encryption is great, but for privacy **protecting traffic metadata is as important**

Low-latency communications – fast, but **only protect from partial adversaries**
High-latency communications – slow, but **protect from global adversaries**

Other metadata is also sensitive!!



Implicit data is as important as explicit data!



The hardware, software, firmware of a device is a unique identifier that encodes information
Example: medical app installed.

The address where data is stored may reveal information about the content.
Example: medical database with patients with mild and severe diseases in different locations

The address where an action happens may reveal information about the action / user.
Example: sending a message from an Oncologist clinic reveals information about the sender

Tracking anonymous users

The cookie zoo

Normal cookies

- persistent identifiers

Third party cookies

- webs establish cookies for other webs
- webs have identifiers for other webs

Evercookies!

Tracking anonymous users without cookies

Would tracking be solved??

Fingerprinting Web Browsers

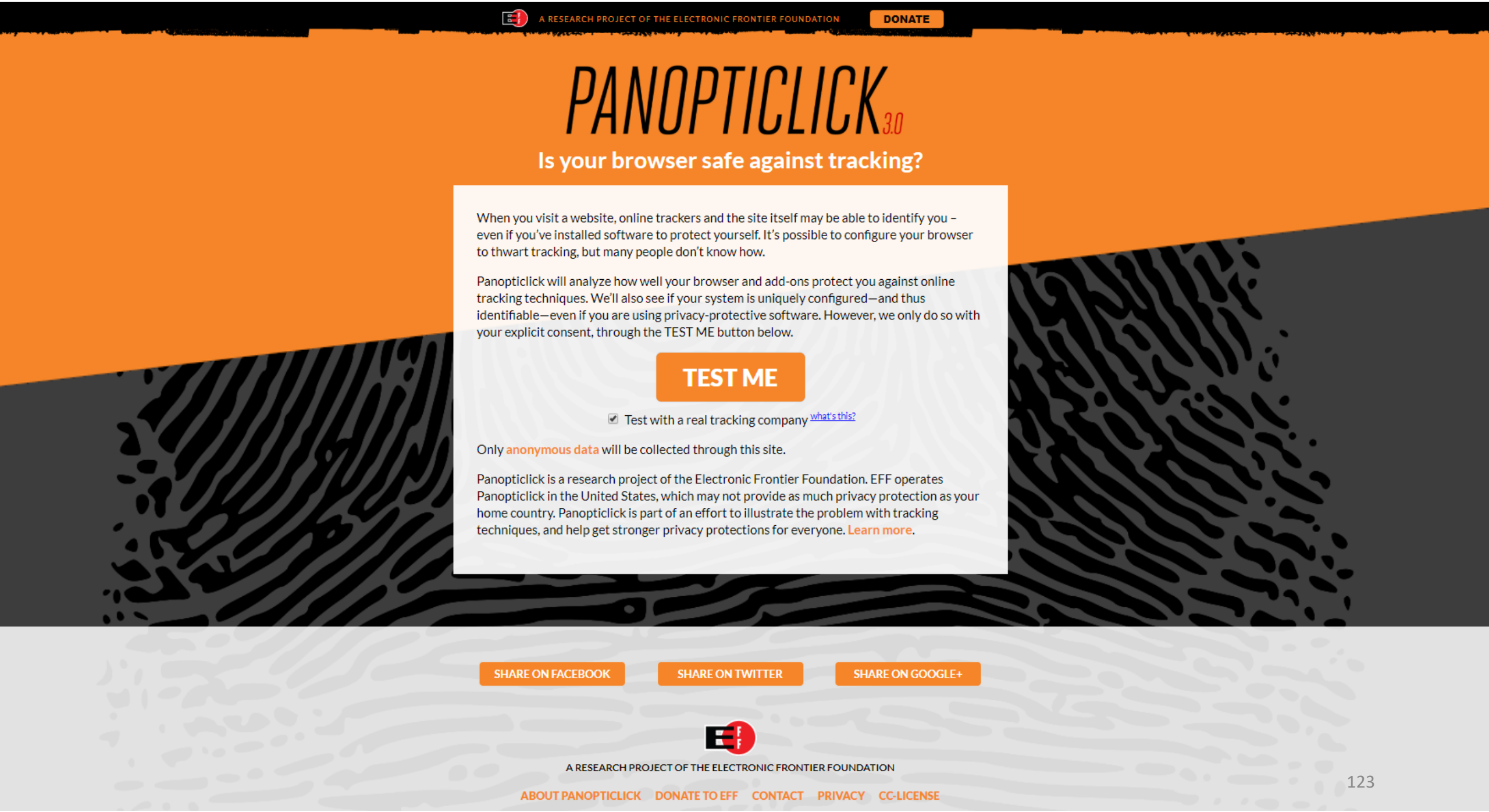
- User agent
- HTTP ACCEPT headers
- Browser plug-ins
- MIME support
- Clock skew

- Installed fonts
- Cookies enabled?
- Browser add-ons
- Screen resolution

Mobile phones are also fingerprintable!

Know how mobile device fingerprinting or digital fingerprinting can assist businesses to enhance advertising and diminish fraud risk

http://www.cybage.com/sites/cybage.com/files/whitepaper/2016-03/Mobile%20Device%20Fingerprinting%20and%20Unique%20User%20Identification.pdf


A RESEARCH PROJECT OF THE ELECTRONIC FRONTIER FOUNDATION
DONATE
PANOPTICLICK 3.0
Is your browser safe against tracking?
When you visit a website, online trackers and the site itself may be able to identify you – even if you've installed software to protect yourself. It's possible to configure your browser to thwart tracking, but many people don't know how.
Panopticlick will analyze how well your browser and add-ons protect you against online tracking techniques. We'll also see if your system is uniquely configured—and thus identifiable—even if you are using privacy-protective software. However, we only do so with your explicit consent, through the TEST ME button below.
TEST ME
Test with a real tracking company what's this?
Only anonymous data will be collected through this site.
Panopticlick is a research project of the Electronic Frontier Foundation. EFF operates Panopticlick in the United States, which may not provide as much privacy protection as your home country. Panopticlick is part of an effort to illustrate the problem with tracking techniques, and help get stronger privacy protections for everyone. Learn more.
SHARE ON FACEBOOK
SHARE ON TWITTER
SHARE ON GOOGLE+
A RESEARCH PROJECT OF THE ELECTRONIC FRONTIER FOUNDATION
ABOUT PANOPTICLICK
DONATE TO EFF
CONTACT
PRIVACY
CC-LICENSE

Location privacy: Points of interest (POIs)

specific location that someone may find useful or interesting

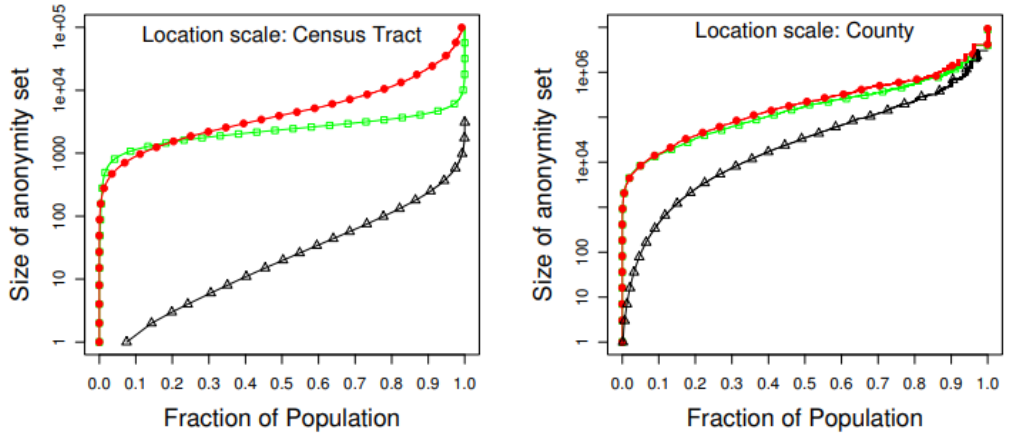
Why are POIs important?

- **Movements are unique** [De Montjoye et al 2013] [De Montjoye et al 2015]

  4 spatio-temporal points are enough to uniquely identify 95% of people in a mobile phone database of 1.5M people and to identify 90% of people in a credit card database of 1M people

- **Home and Work: unique identifier** [Golle & Partridge 2009]

  individual's anonymity set in the U.S. working population is 1, 21 and 34,980, for locations known at the granularity of a census block, census track and county respectively



Location privacy: Points of interest (POIs)

specific location that someone may find useful or interesting

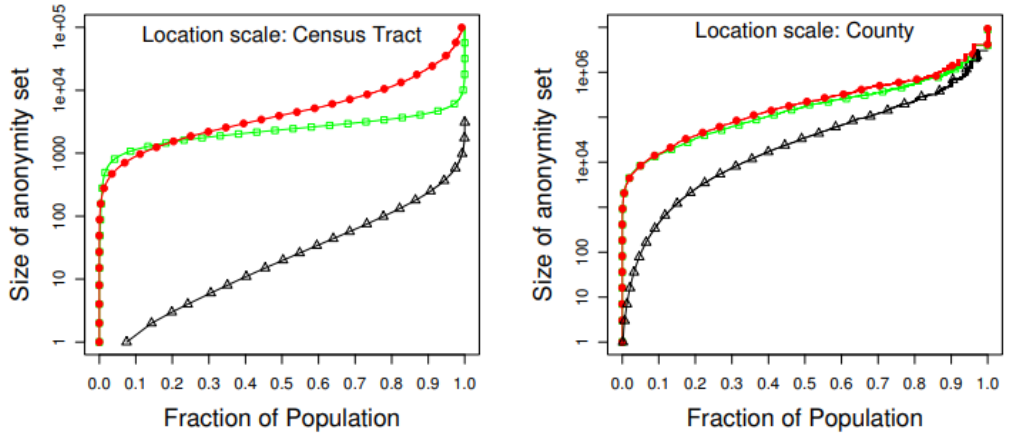
Why are POIs important?

- **Movements are unique** [De Montjoye et al 2013] [De Montjoye et al 2015]

  4 spatio-temporal points are enough to uniquely identify 95% of people in a mobile phone database of 1.5M people and to identify 90% of people in a credit card database of 1M people

- **Home and Work: unique identifier** [Golle & Partridge 2009]

  individual's anonymity set in the U.S. working population is 1, 21 and 34,980, for locations known at the granularity of a census block, census track and county respectively



Location privacy: Points of interest (POIs)

specific location that someone may find useful or interesting

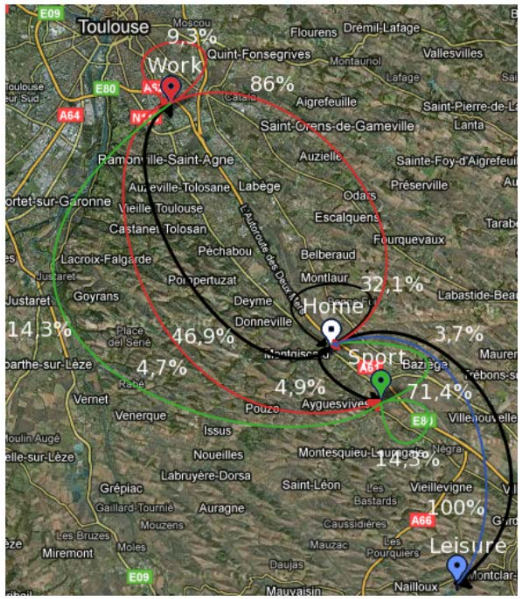
Why are POIs important?

- **N-top locations: unique identifiers** [Zhang & Bolot 2011]

  [call records] *“top 2" locations likely correspond to home and work locations, the "top 3" to home, work, and shopping/school/commute path locations*

- **Where a user will move next** [Gambs et al 2012]

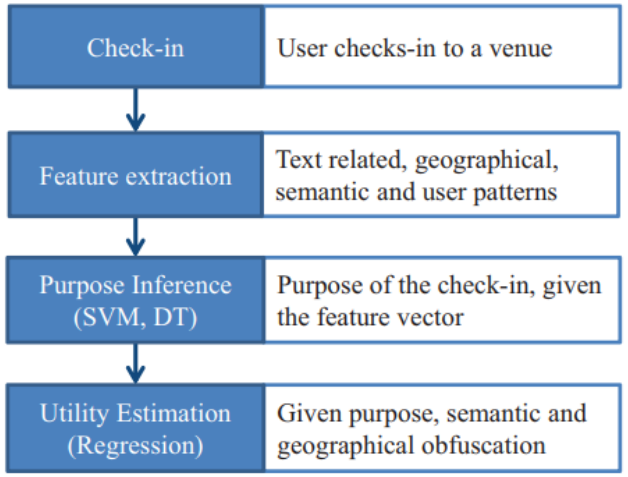
  Accuracy for the prediction of the next location in the range of 70% to 95%



Hidden Markov Model movement patterns

Location privacy: Points of interest (POIs)

specific location that someone may find useful or interesting



Why are POIs important?

- **Learning about users' motivation** [Bilogrevic et al 2015]

  43 % correct classification (22% baseline predict most likely "Inform I am here")
  Interesting utility impact study [complementary to this presentation]

- **Learning Demographics and other Patterns**

  [Pang and Zhang 2017] [Felbo et al 2017][Cho et al 2010] [Liao et al 2005] [Liao et al 2007]

  Machine-learning based frameworks

Location privacy: Defenses

- **Perturbation**: report a perturbed noise for the location
  - How to add the noise? (remember that the adversary knows!)

- **Generalization**: report a larger region instead of a point, i.e., reduce precision

- **Hiding**: do not report every single location
  - How to hide? What about recurrent patterns?

- **Add dummy locations**: hide the real trips among dummies
  - How to create the dummies in an undistinguishable manner?

Takeaways

Metadata is as important as content

Anonymous communications protect traffic data

- Low latency (Tor) vs. High latency (Mixes)
- Tradeoff performance for security (stronger adversary)

Tracking anonymous users is easy

- Devices reveal / collect too much information

Location is very revealing

- Hard to defend, correlations enable inferences

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

SPRING
SECURITY AND PRIVACY ENGINEERING LABORATORY

Information Security and Privacy (COM-402)

Part 7: Privacy enhancing technologies

Steganography

Carmela Troncoso

SPRING Lab

carmela.troncoso@epfl.ch

Some slides/ideas adapted from: JP Hubaux, Bryan Ford, Vitaly Shmatikov

What about Unobservability? Steganography

Goal: concealing a file, message, image, or video within another file, message, image, or video.

WWI example -Sent by a German spy during WWI:

“Apparently neutral’s protest is thouroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.”

What about Unobservability? Steganography

Goal: concealing a file, message, image, or video within another file, message, image, or video.

WWI example -Sent by a German spy during WWI:

“Apparently neutral’s protest is thouroughly discounted and ignored. Isman hard hit. Blockade issue affects pretext for embargo on byproducts, ejecting suets and vegetable oils.”

Pershing sails from NY June 1!

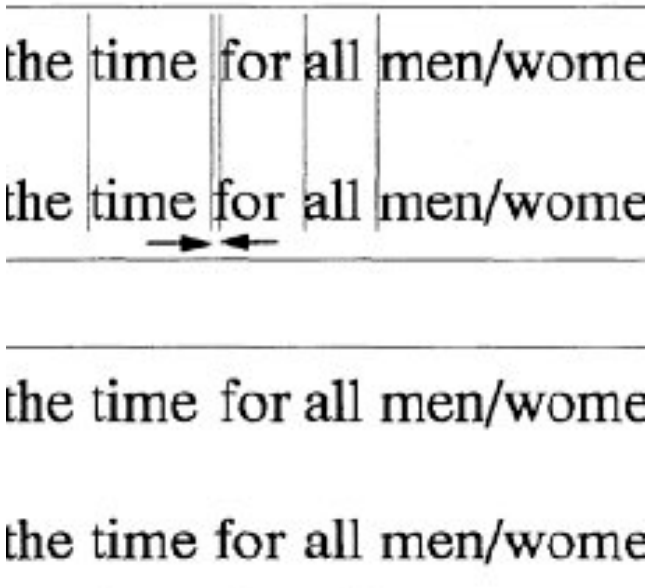
Steganography: example methods

Hide in LSB of an image

Alter the LSB of an image to encode words / other image

Hide in text

Modify spaces to encode words



Steganography: security

The goal is to hide the message

Unobservability - Pr[real action |observation]

Pr[observation | real action] = Pr [observation]

Analysis:

- Check whether there are anomalies in the text, image,…
- Requires to know how the original looks like!
  - Except for local correlations (e.g., color pixels in a photo)

steganography must take them into account!

But do people use this…?

1) Yes, this is important from a message secrecy point of view, and it is used by both governments and criminals

2) This is key for censorship resistance! The key to not be censored is to not be seen!
   - Steganographic properties and lessons are very much used in privacy technologies.

Takeaways PETs

Cryptography → Confidentiality!

Traditional: computer security context

Privacy goes BEYOND than traditional confidentiality.

What makes Privacy Enhancing Technologies (PETs) different:

- Threat model: **weak** actors, **powerful** adversaries.

- Susceptibility to **compulsion**.

- Cannot assume the existence of **Trusted Third Parties (TTP)**

- You should also worry about **Cost, Collusion, Corruption, Carelessness**.