Privacy Preserving Technologies

Associate Professor Mihai Chiroiu

SYSTEMS LABORATORY

Computer Science
& Engineering
Department

Objectives

- GDPR/IAPP
- Mix nets / Onion routing / TOR
- Private Information Retrieval (PIR)
- ORAM
- Bitcoin (next time)

Why?

- “There was of course no way of knowing whether you were being watched at any given moment...You had to live – did live, from habit that became instinct – in the assumption that every sound you made was overheard and, except in darkness, every movement scrutinized.“

George Orwell, 1984 (1948)

Why?

- UKUSA
  - The Five Eyes: USA, Australia, Canada, New Zealand, and the United Kingdom
  - In 1975 the United States revealed the existence of the National Security Agency (NSA).
- ECHELON
  - mass surveillance and industrial espionage
- TEMPEST (Telecommunications Electronics Material Protected from Emanating Spurious Transmissions)
  - Secret NSA standard for low-emissions computers

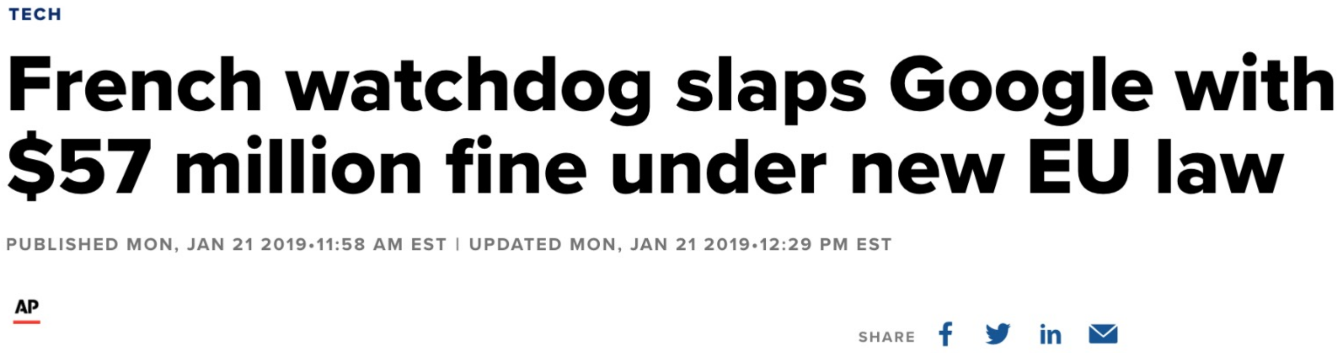
Why?

- “There is no privacy in the digital age.”
- “No one cares about privacy anymore.”
- “If you haven’t done anything wrong you have nothing to hide.”

GDPR

- “The General Data Protection Regulation (GDPR) is the toughest privacy and security law in the world.”
- The right to be forgotten
- Data minimisation

GDPR



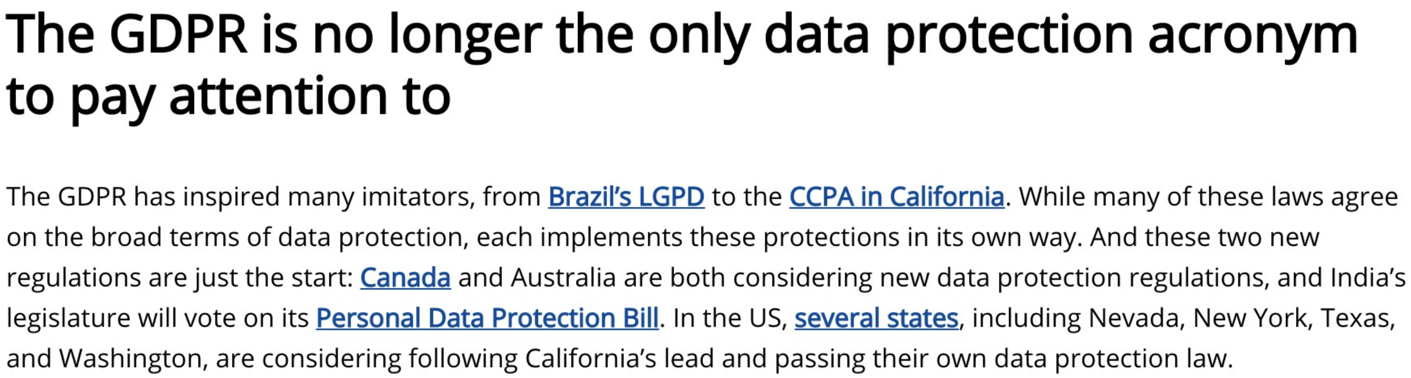
TECH

French watchdog slaps Google with $57 million fine under new EU law

PUBLISHED MON, JAN 21 2019•11:58 AM EST | UPDATED MON, JAN 21 2019•12:29 PM EST

AP

SHARE



The GDPR is no longer the only data protection acronym to pay attention to

The GDPR has inspired many imitators, from Brazil's LGPD to the CCPA in California. While many of these laws agree on the broad terms of data protection, each implements these protections in its own way. And these two new regulations are just the start: Canada and Australia are both considering new data protection regulations, and India's legislature will vote on its Personal Data Protection Bill. In the US, several states, including Nevada, New York, Texas, and Washington, are considering following California's lead and passing their own data protection law.

IAPP

- International Association of Privacy Professionals
- https://iapp.org/certify/programs/

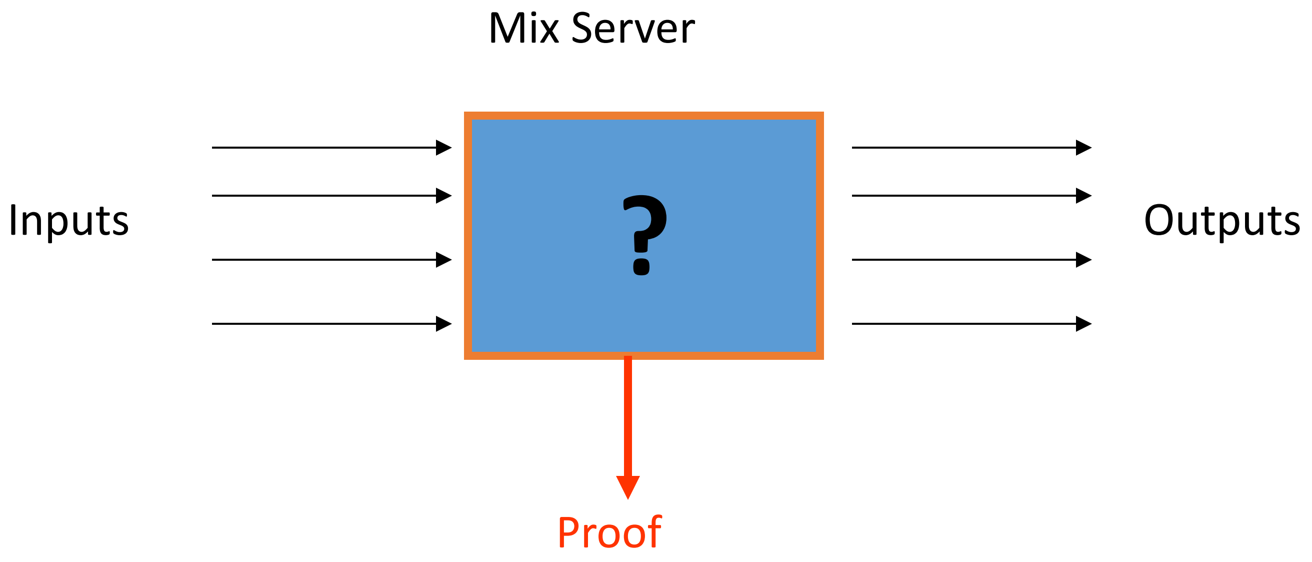
Definitions

- *Anonymity:* The state of being not identifiable
- *Unobservability* ensures that a user may use a resource or service without others being able to observe that the resource or service is being used
- *Unlinkability* of sender and recipient (relationship anonymity) means that it is untraceable who is communicating with whom
- Pseudonymity is the use of pseudonyms as IDs and allows to provide both privacy protection and accountability

Privacy-Enhancing Technologies

- 1. PETs for minimizing/ avoiding personal data (providing Anonymity, Pseudonymity, Unobservability, Unlinkability)
  - At communication level (Mix nets, Onion Routing, TOR, PGP, Crowds)
  - At application level (Anonymous Ecash, Private Information Retrieval, Anonymous Credentials)
- 2. PETs for the safeguarding of lawful processing
  - P3P
  - Privacy policy languages
  - Transparency Enhancing Tools (TETs)
- Combination of 1 & 2
  - Privacy-enhanced Identity Management

Mix Server



A mix server is a cryptographic implementation of a hat.

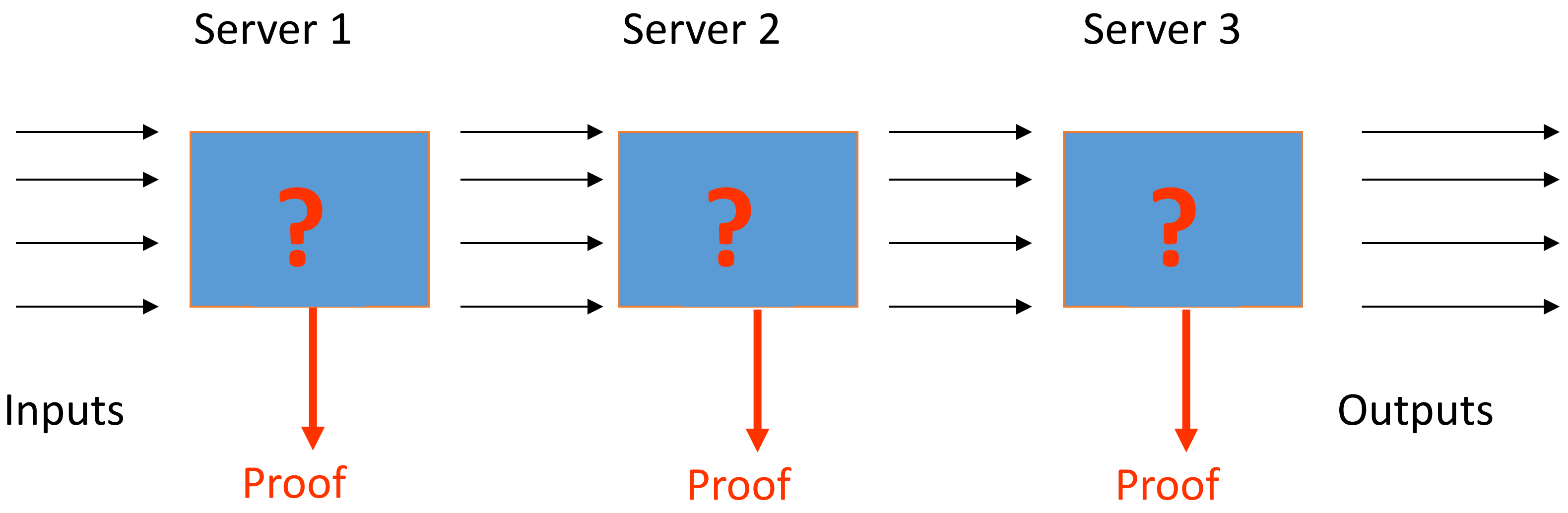
Mixnets properties

- The main privacy property desired of such a mixnet is that the permutation matching inputs to outputs should be known only to the mixnet, and no one else.

- Privacy in this mixnet construction derives from the fact that the ciphertext pair (*C,C'*) is indistinguishable from a pair (*C,R*) for a random ciphertext *R* to any adversary without knowledge of the private key.

- Inputs are ciphertext
- Outputs are a re-encryption of the inputs.

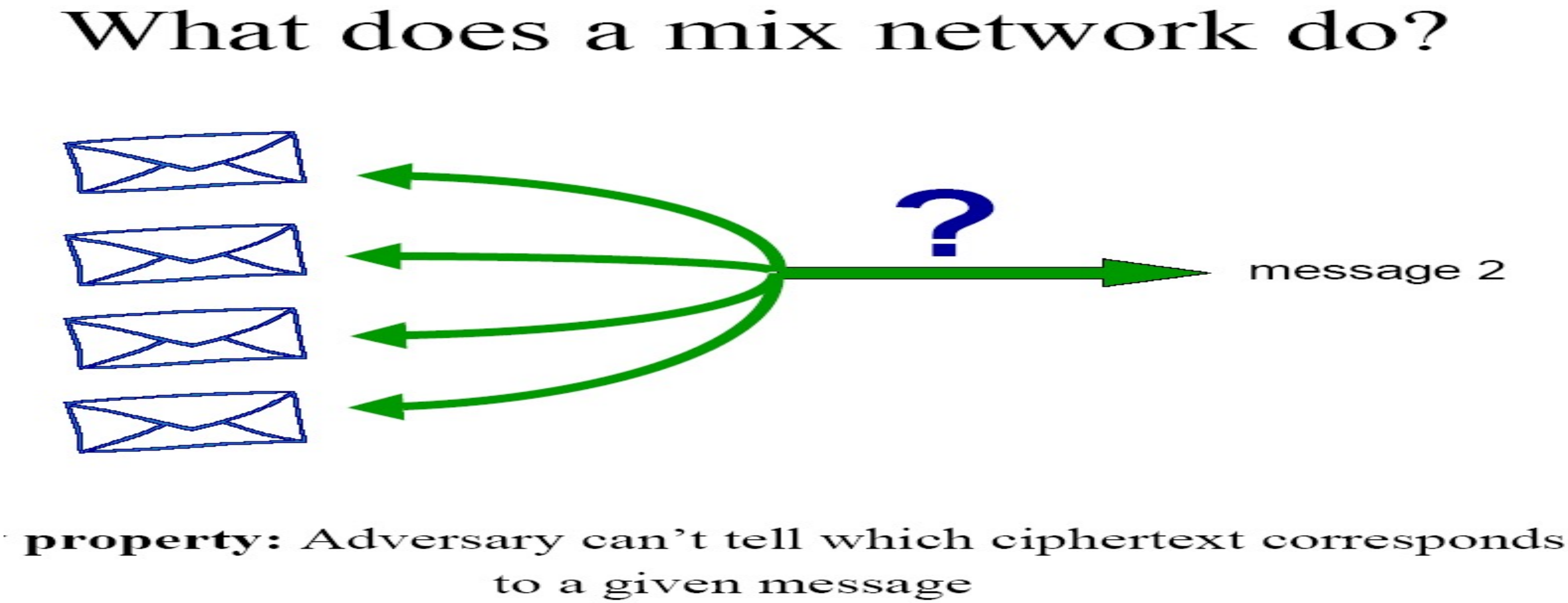
Mix Network



1. Servers sequentially mix the inputs
2. Verify the proofs of correct mixing:
   - OK: accept the output
   - Otherwise: remove cheaters and mix again

If a single mix server is honest, global permutation is secret.

What does a mix network do?



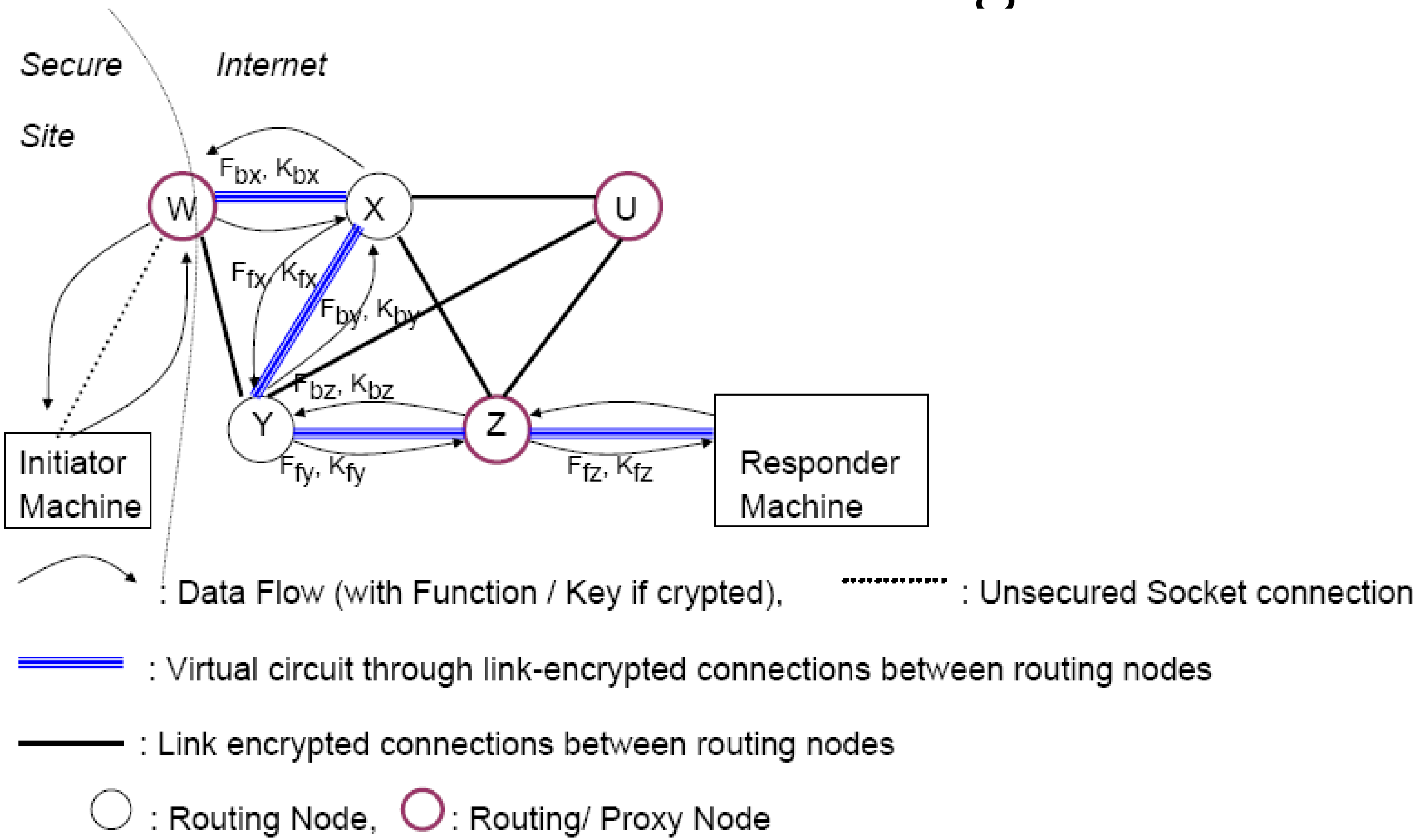
What does a mix network do?

property: Adversary can’t tell which ciphertext corresponds to a given message

Onion routing

- Initiator’s proxy sends the onion along that route to establish a virtual circuit between himself and the responder’s proxy.

Virtual Circuit with Onion Routing


Secure
Site
Internet
W
X
U
Y
Z
F_{bx}, K_{bx}
F_{fx}, K_{fx}
F_{by}, K_{by}
F_{bz}, K_{bz}
F_{fy}, K_{fy}
F_{fz}, K_{fz}
Initiator Machine
Responder Machine
: Data Flow (with Function / Key if crypted),
: Unsecured Socket connection
: Virtual circuit through link-encrypted connections between routing nodes
: Link encrypted connections between routing nodes
○ : Routing Node,
○: Routing/ Proxy Node

Onion routing

- Hiding of routing information in connection oriented communication relations
- Nested public key encryption for building up virtual circuit
- First/Last-Hop Attacks by
    - Timing correlations
    - Message length (No. of cells sent over circuit)

TOR

- No mixing is employed
- Can run any SOCKS application on top of Tor
- A request is routed to/from a series of a circuit of three routers (latency problems)

Mixnets vs Onion Routing

- “Tor plays with space alone (the bytes that you send across the network go through various other servers thus they aren’t where they are supposed to be) and mixnets play with both space and time (adds delays and shuffles the request through various servers as well).”

Pretty Good Privacy

- Is a trust model
- Web-of-Trust (PGP)
  - Peer-to-peer model
  - Individuals digitally sign each other keys
  - You would implicitly trust keys signed by some of your friends
- Time-consuming and general public doesn’t understand it
- https://keys.openpgp.org/

Off the record protocol [11]

- Perfect forward secrecy (forgetting keys)
  - A → B : g^{x1}
  - B → A : g^{y1}
  - A → B : g^{x2} , $E(M_1, k_{11})$
  - B → A : g^{y2} , $E(M_2, k_{21})$
  - A → B : g^{x3} , $E(M_3, k_{22})$
  - $k_{ij} = H(g^{xiyj})$
- Authentication
  - A → B : $Sign(g^{x1}, priv_A), pub_A$
  - B → A : $Sign(g^{y1}, priv_B), pub$
  - $g^{xi+1}, E(M_r, k_{ij}), MAC(\{g^{xi+1}, E(M_k, k_{ij})\}, H(k_{ij}))$

© Mihai Chiroiu

SYSTEMS LABORATORY

Computer Science & Engineering Department

Private Information Retrieval (PIR)

- Privacy for the item of interest:
  - Allows a user to retrieve an item from a database/news server without revealing which item he is interested in
- Application example: patent database
- Simple (but expensive) solution:
  - Download all database entries and make local selection

Oblivious RAM

- The access pattern is independent of the data
  - Probability distribution!

- Any sequence of locations i1, i2, …
  - induces a distribution on sequences of requests
  - q1, q2…

- Security: for any two sequence of locations
  - i1, i2, …
  - i’1, i’2, …
- induced distributions of requests should be indistinguishable

References

[1] https://csjourney.com/onion-routers-mix-networks-differences-explained/?fbclid=IwAR3DPEKLm1-Zee--9VbpWfdTH0qn-RRMG5e8K2CTjrPCBzBYgJI5PsTC6FQ

[2] https://gdpr.eu/

[3] https://iapp.org/

[4] https://svn-archive.torproject.org/svn/projects/design-paper/tor-design.pdf

[5] https://www.cs.kau.se/cs/education/courses/dvad03/p2/

References

[6] https://en.wikipedia.org/wiki/UKUSA_Agreement

[7] https://en.wikipedia.org/wiki/Tempest_(codename)

[8] https://petsymposium.org/2011/papers/hotpets11-final10Syverson.pdf

[9] https://link.springer.com/content/pdf/10.1007/3-540-47721-7_31.pdf

[10] https://www.activism.net/cypherpunk/manifesto.html

References

[11] https://otr.cypherpunks.ca/otr-wpes.pdf