Use-Cases for Private Information Retrieval and Secure Multiparty Computation in Modern Network Architecture

Christian Grothoff

Berner Fachhochschule
Technik und Informatik

11.02.2020

Context

Design Choices for a Civil Network!

Internet Design Goals (David Clark, 1988)

1. **Internet communication must continue despite loss of networks or gateways.**
2. The Internet must support multiple types of communications service.
3. The Internet architecture must accommodate a variety of networks.
4. The Internet architecture must permit *distributed management* of its resources.
5. The Internet architecture must be cost effective.
6. The Internet architecture must permit host attachment with a low level of effort.
7. **The resources used in the internet architecture must be accountable.**

Design Choices for a Civil Network!

Internet Design Goals (David Clark, 1988)

1. **Internet communication must continue despite loss of networks or gateways.**
2. The Internet must support multiple types of communications service.
3. The Internet architecture must accommodate a variety of networks.
4. The Internet architecture must permit *distributed management* of its resources.
5. The Internet architecture must be cost effective.
6. The Internet architecture must permit host attachment with a low level of effort.
7. **The resources used in the internet architecture must be accountable.**

GNUnet Design Goals

1. GNUnet must be implemented as Free Software.
2. GNUnet must minimize the amount of personally identifiable information exposed.
3. The GNUnet must be fully distributed and resilient to external attacks and rogue participants.
4. GNUnet must be self-organizing and not depend on administrators or centralized infrastructure.
5. GNUnet must inform the user which other participants have to be trusted when establishing private communications.
6. GNUnet must be open and permit new peers to join.
7. GNUnet must support a diverse range of applications and devices.
8. GNUnet must use compartmentalization to protect sensitive information.
9. The GNUnet architecture must be resource efficient.
10. GNUnet must provide incentives for peers to contribute more resources than they consume.

Applications in GNUnet (under development)

- Anonymous and non-anonymous publishing
- IPv6–IPv4 protocol translation and tunnelling
- **GNU Name System**: censorship-resistant replacement for DNS
- Conversation: secure, decentralized voice communication
- **SecuShare**: social networking
- GNU Taler: privacy-friendly payments
- ...

Part I: Private Information Retrieval

Back to the Internet: DNS troubles

- DNS remains a source of traffic amplification for DDoS
- DNS censorship (i.e. by China) causes collateral damage in other countries
- DNS is part of the mass surveillance apparatus (MCB)
- DNS is abused for the offensive cyber war (QUANTUMDNS)

Band aid solutions[1] will **not** fix this.

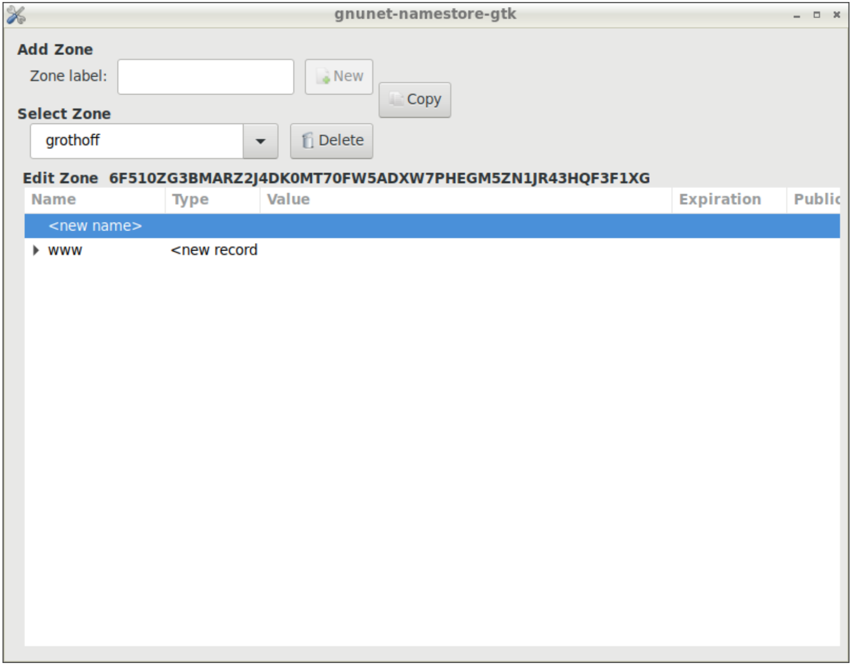
[1] DNS-over-TLS, DoH, DNSSEC, DPRIVE, ...

The GNU name system[2]

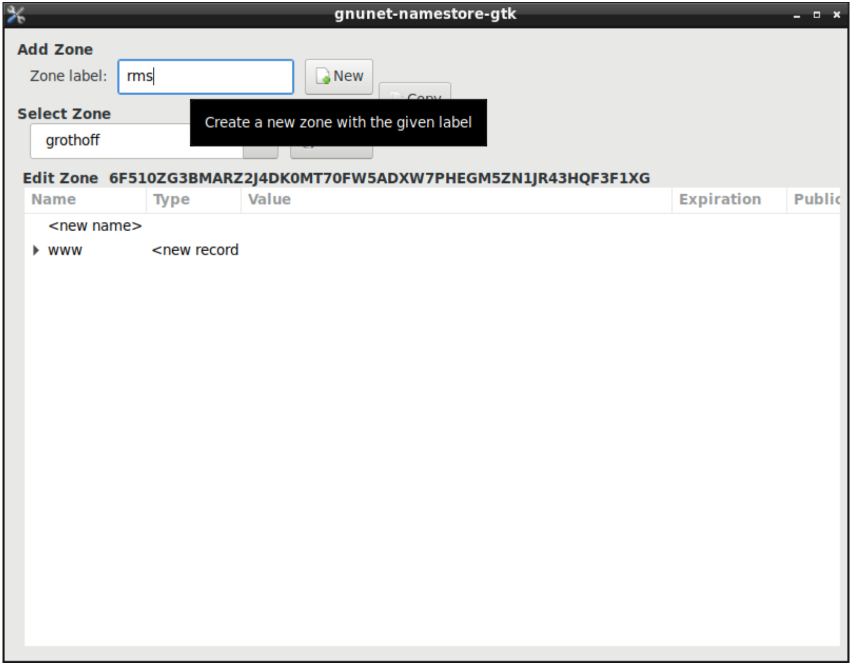
- Decentralized name system ⇒ Names are not global
- Supports globally unique (& secure) identification
- Achieves query and response privacy
- Provides public key infrastructure
- Interoperable with DNS

[2] Joint work with Martin Schanzenbach, Matthias Wachs and Patrick Gerber

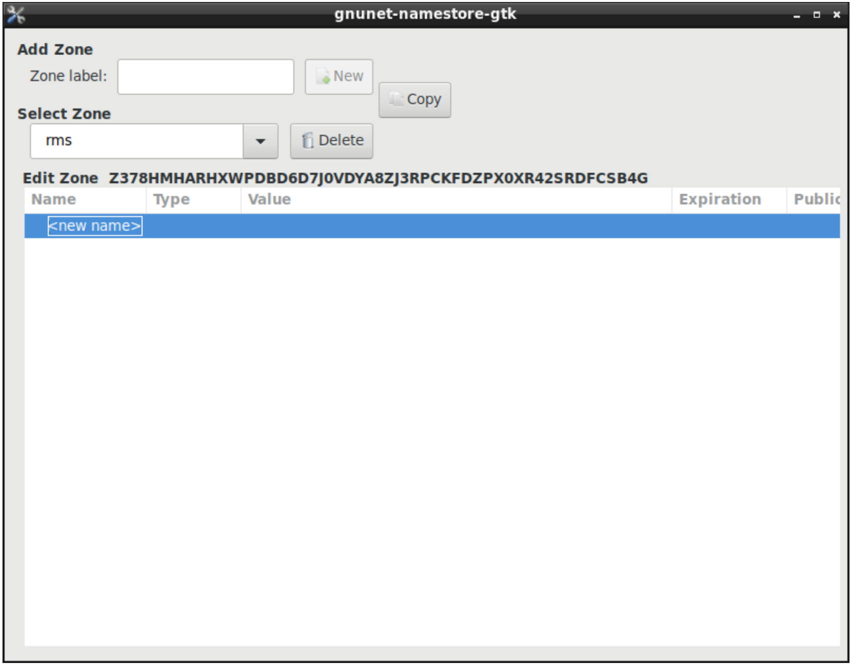
Zone management



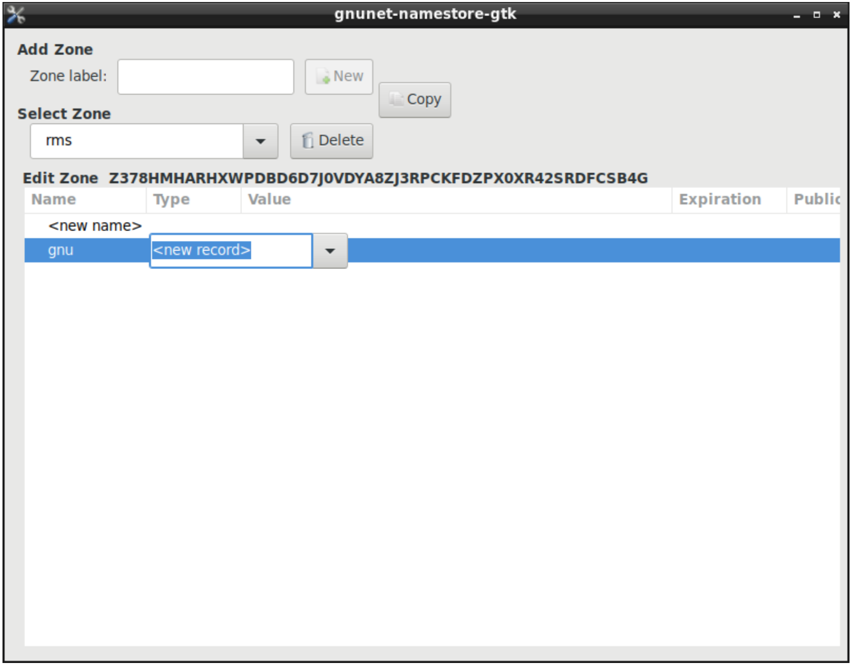
Zone management



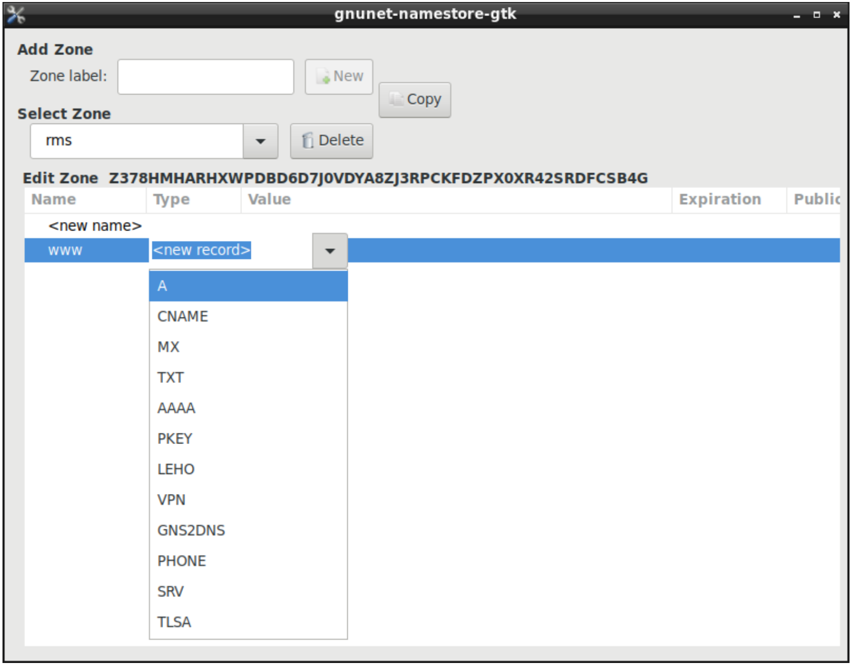
Zone management



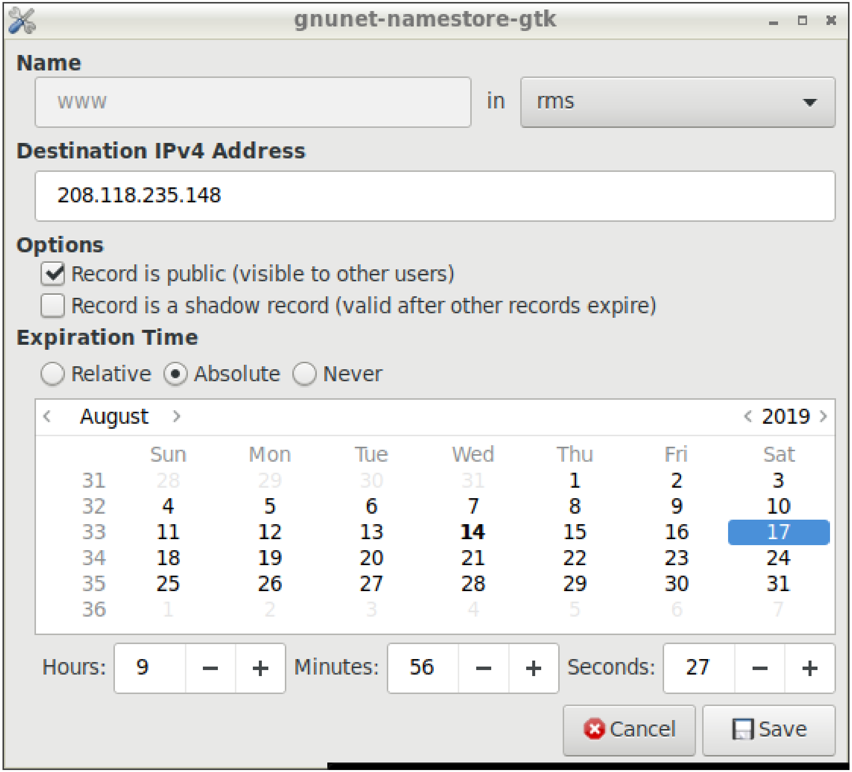
Zone management



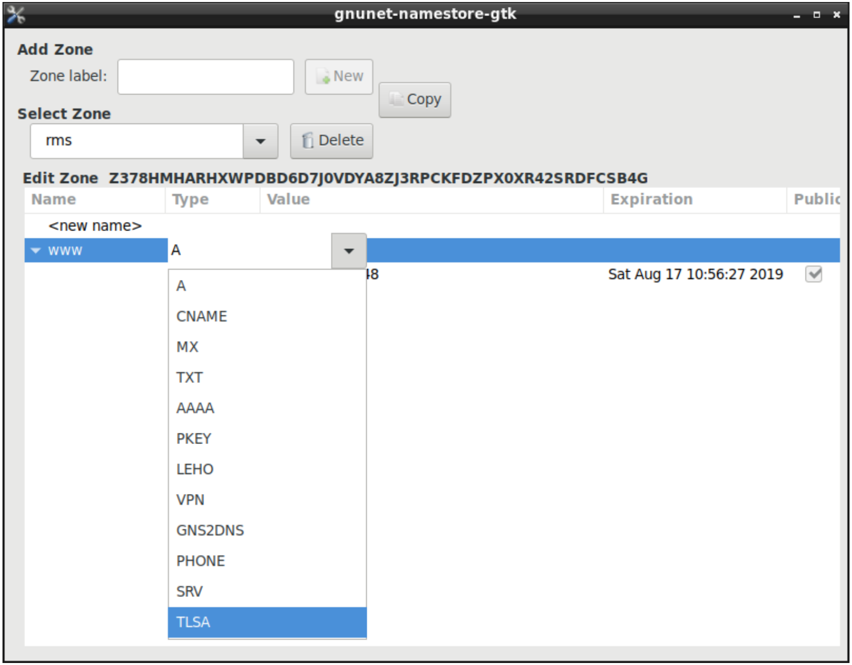
Zone management



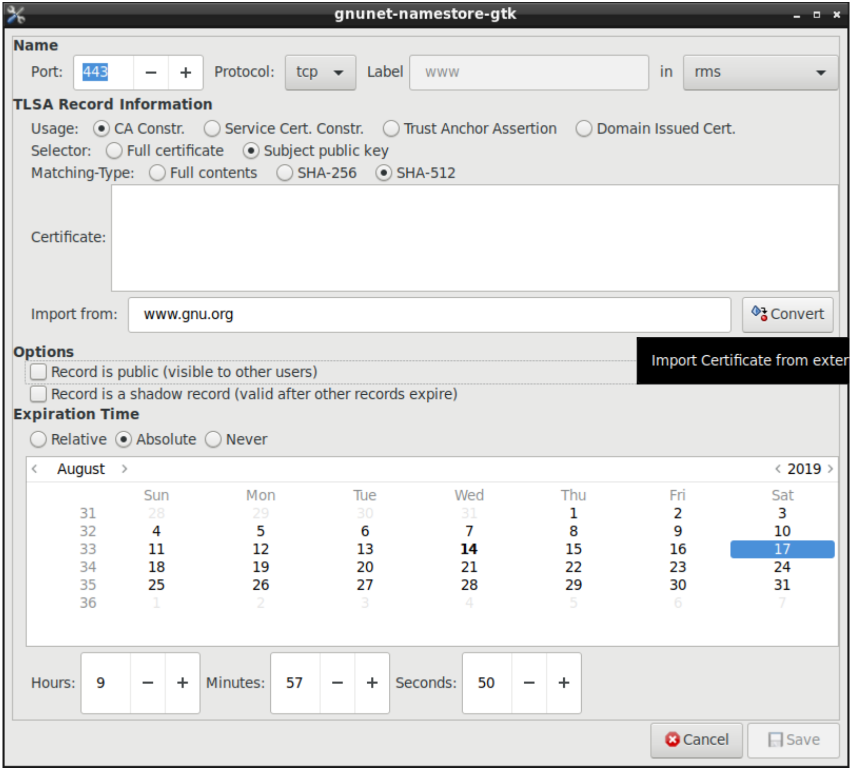
Zone management



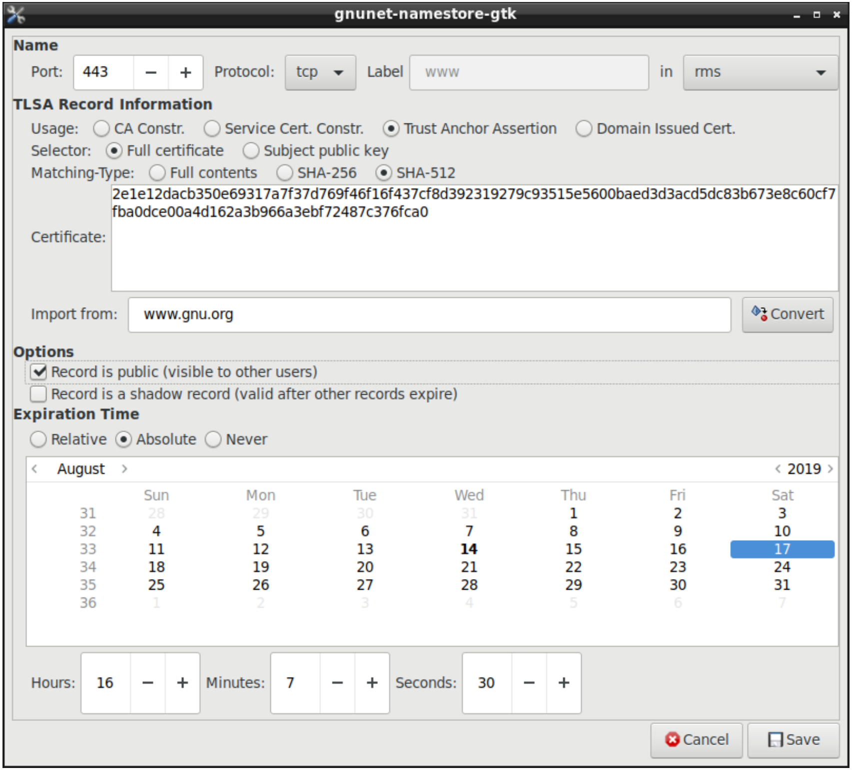
Zone management



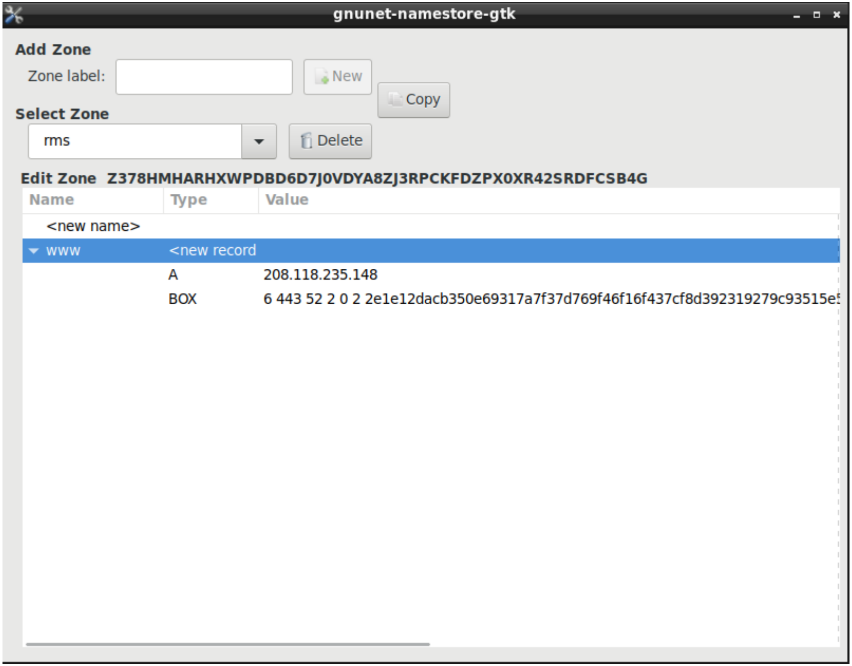
Zone management



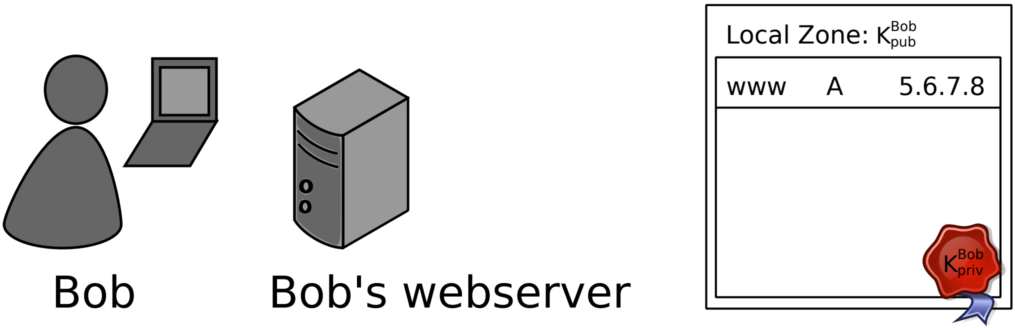
Zone management



Zone management



Name resolution in GNS

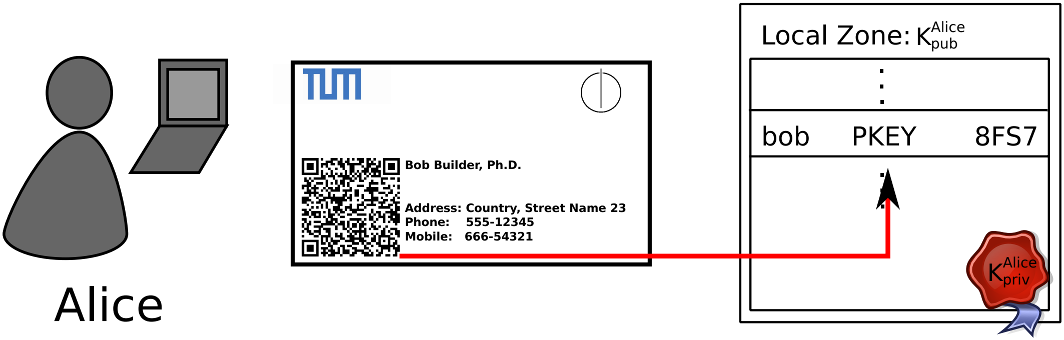


- Bob can now reach his Web server under **www.bob**

Secure Introduction

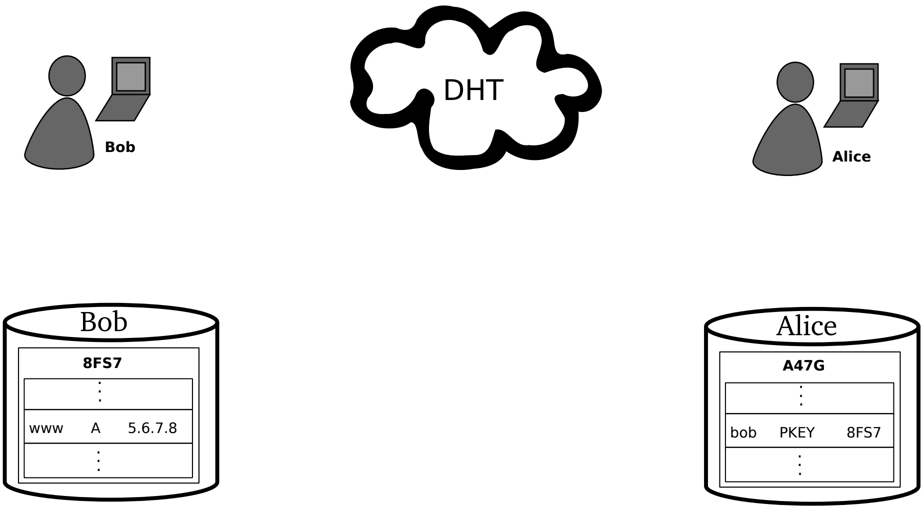
- Bob provides his public key to his **friends**, i.e. via QR code

Delegation

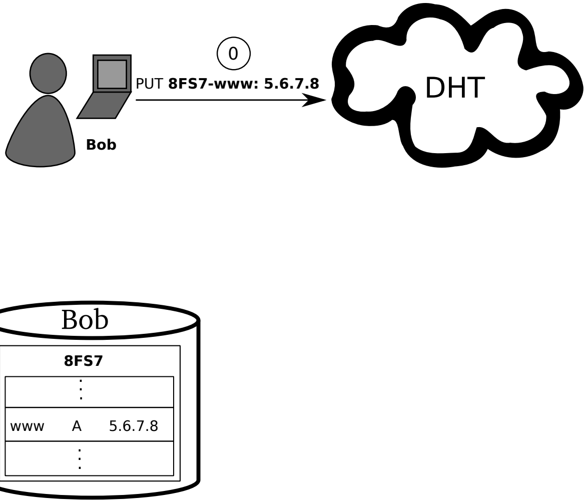


- Alice learns Bob's "public" key
- Alice creates a delegation to zone K_{pub}^{Bob} under the label **bob**
- Alice can then reach Bob's Web server under **www.bob.alice**

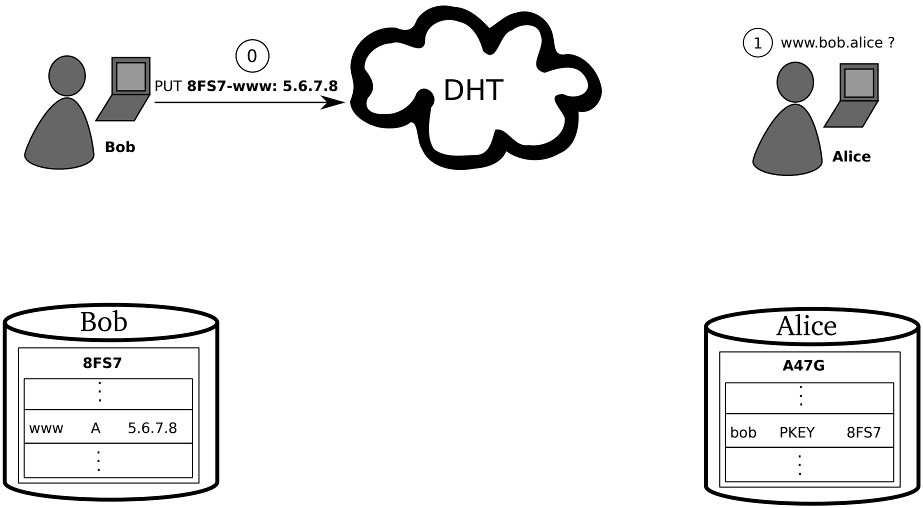
Name resolution



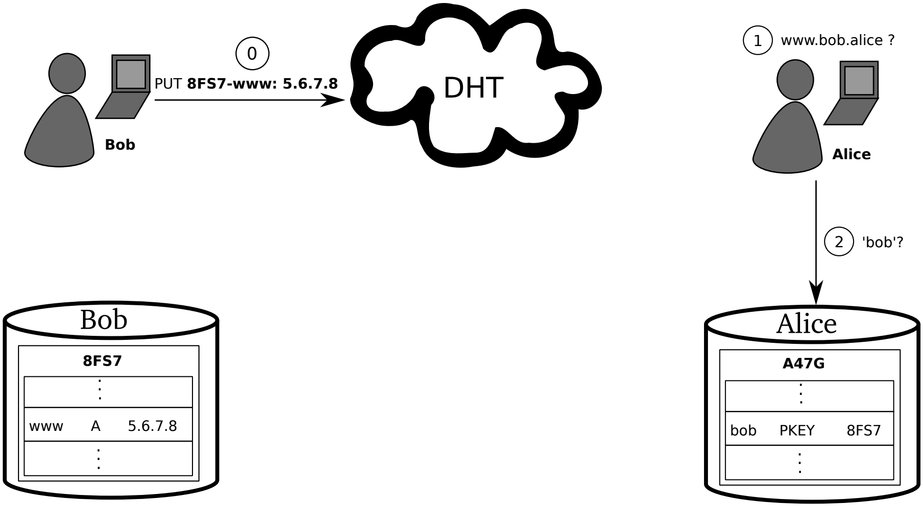
Name resolution



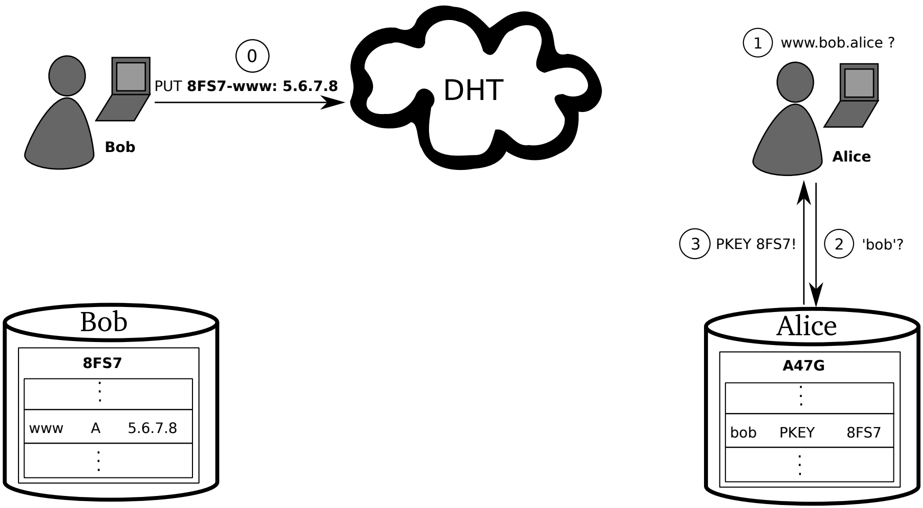
Name resolution



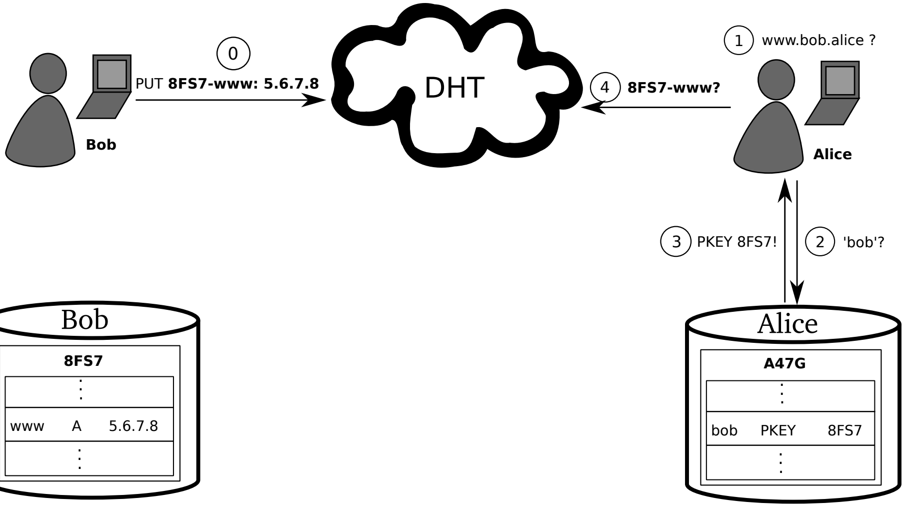
Name resolution



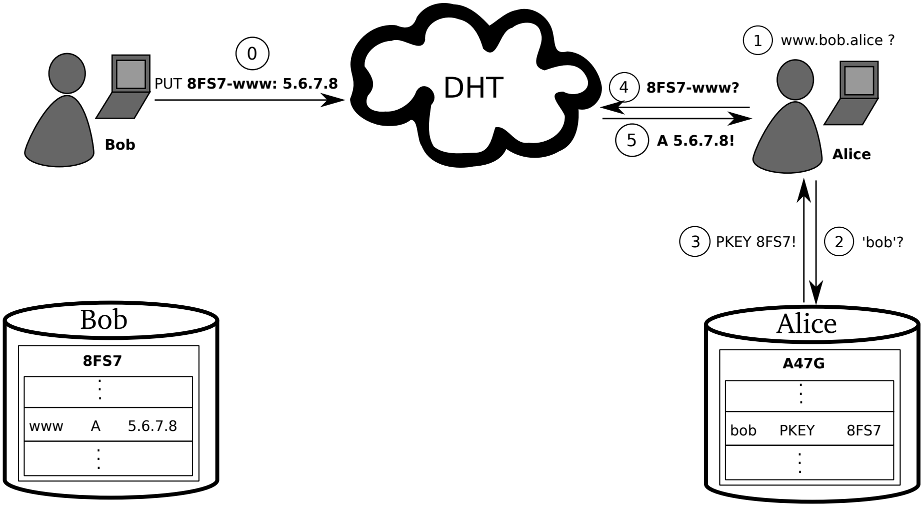
Name resolution



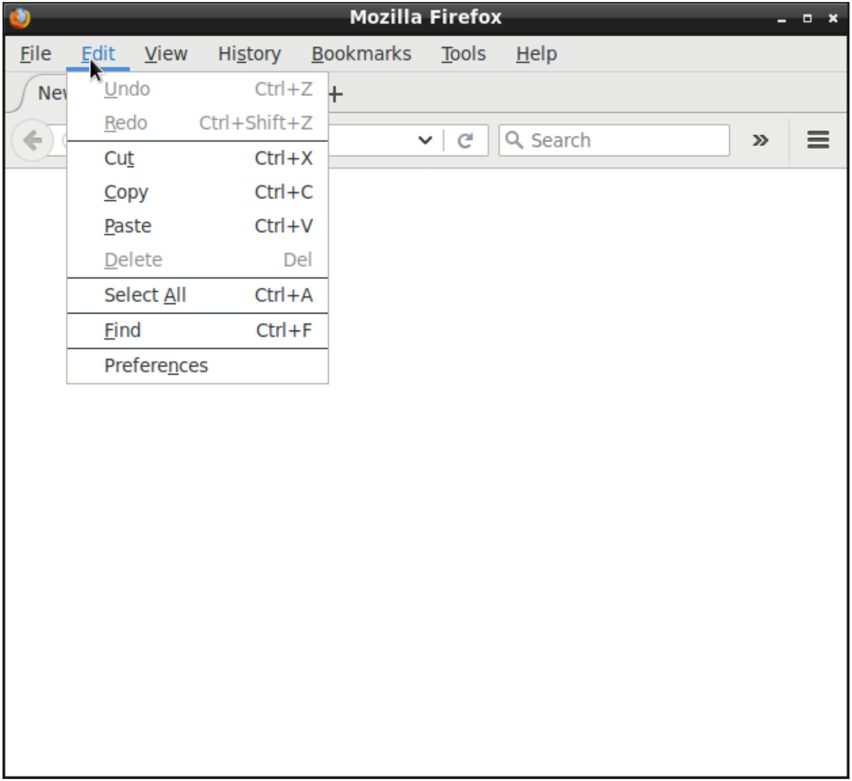
Name resolution



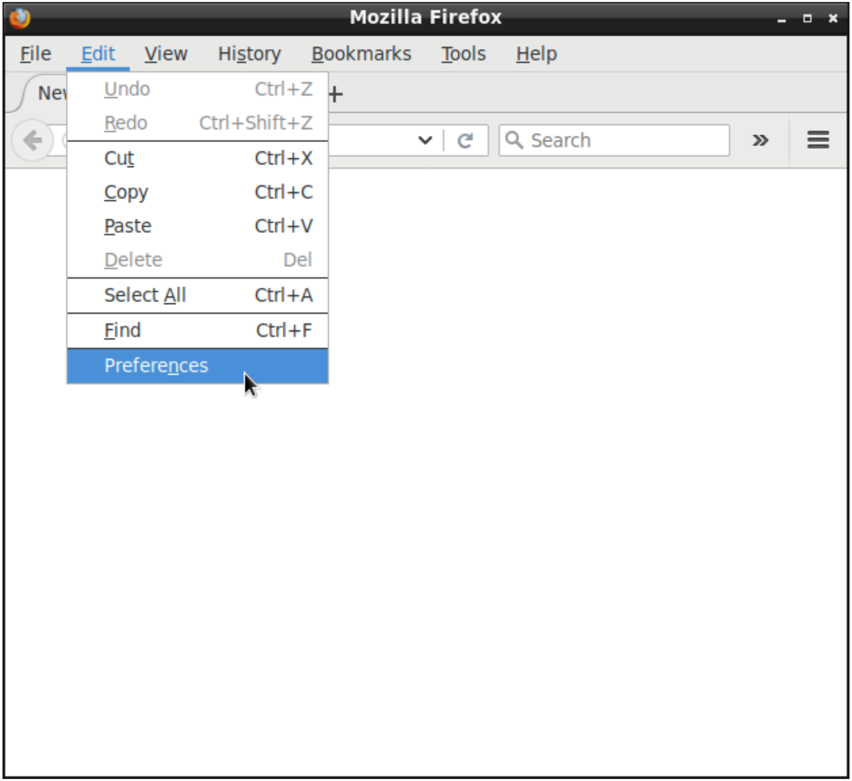
Name resolution



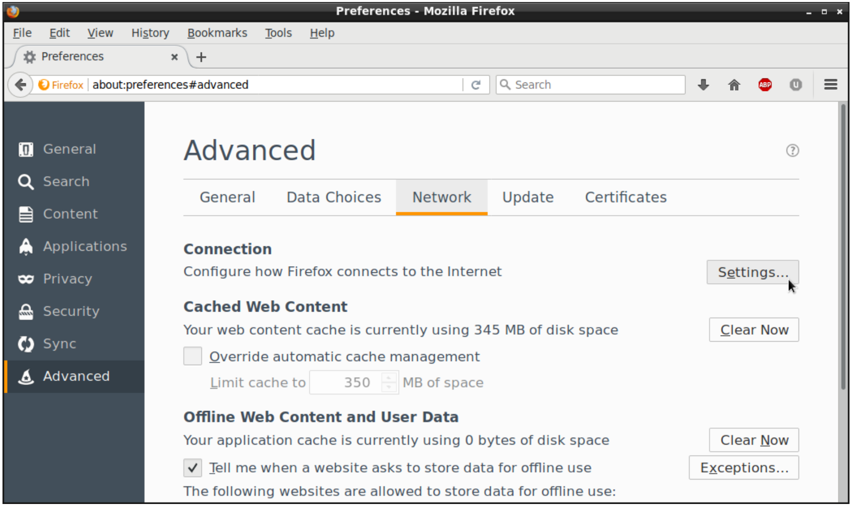
Browser Configuration



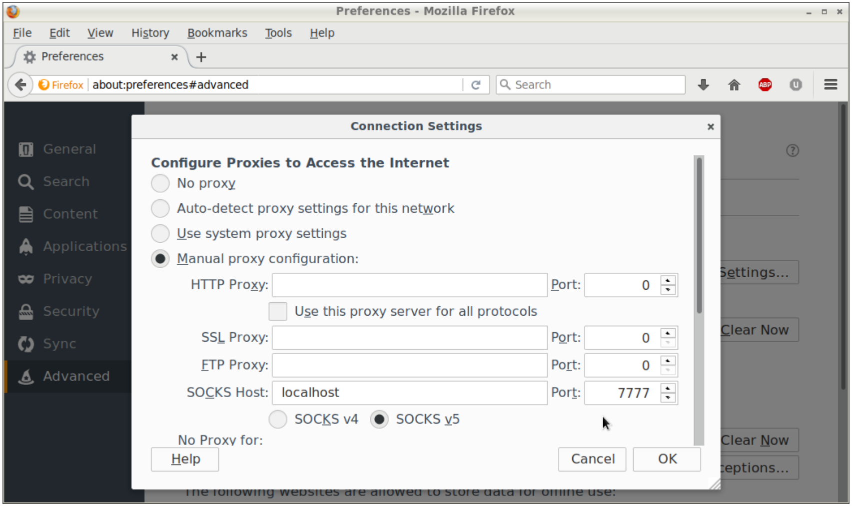
Browser Configuration



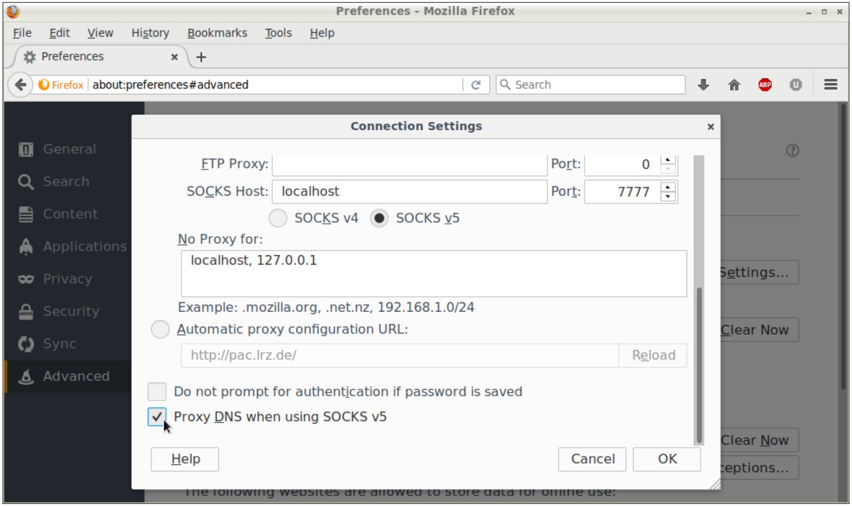
Browser Configuration



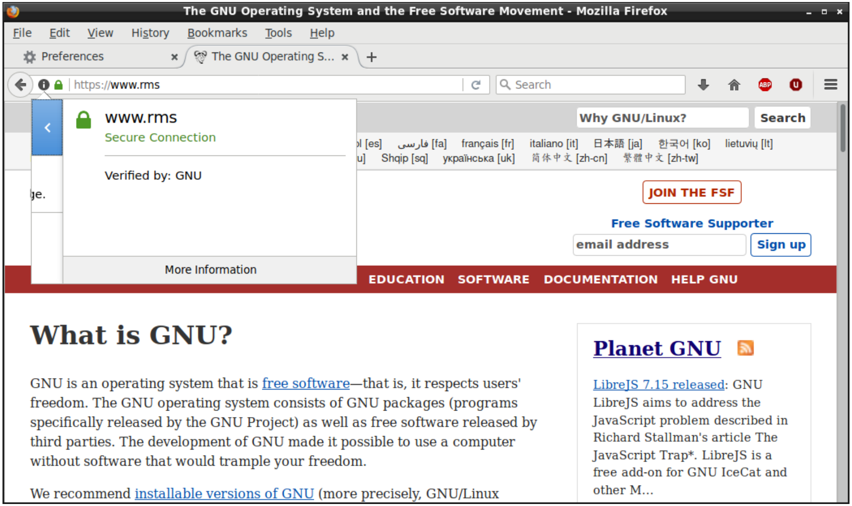
Browser Configuration



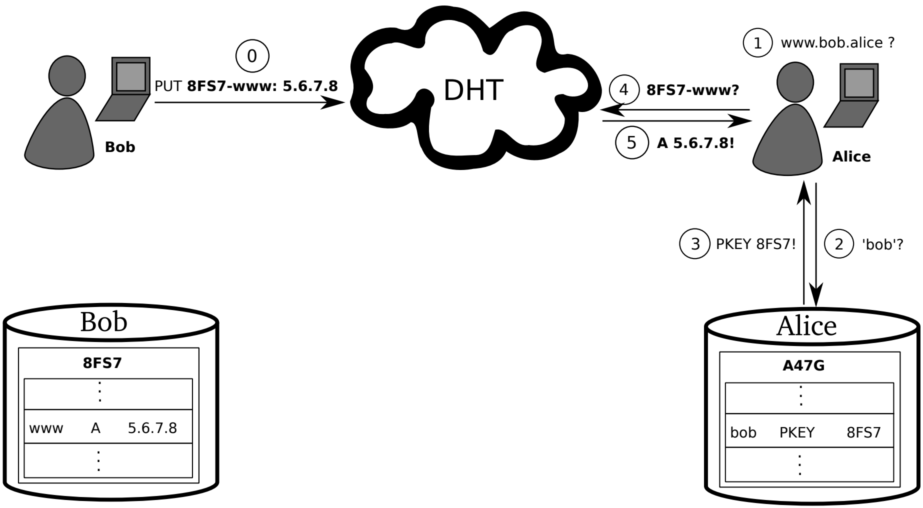
Browser Configuration



Browser Usage



Privacy issue: DHT



Terminology

- G generator in ECC curve, a point
- n size of ECC group, $n := |G|$, n prime
- x private ECC key of zone ($x \in \mathbb{Z}_n$)
- P public key of zone, a point $P := xG$
- l label for record in a zone ($l \in \mathbb{Z}_n$)
- $R_{P,l}$ set of records for label l in zone P
- $q_{P,l}$ query hash (hash code for DHT lookup)
- $B_{P,l}$ block with encrypted information for label l in zone P published in the DHT under $q_{P,l}$

Private Information Retrieval

Publishing records $R_{P,l}$ as $B_{P,l}$ under key $q_{P,l}$

$$
\begin{align}
h &:= H(l, P) \tag{1}\\
d &:= h \cdot x \mod n \tag{2}\\
B_{P,l} &:= S_d(E_{HKDF(l,P)}(R_{P,l})), dG \tag{3}\\
q_{P,l} &:= H(dG) \tag{4}
\end{align}
$$

Private Information Retrieval

Publishing records $R_{P,l}$ as $B_{P,l}$ under key $q_{P,l}$

$$
\begin{align}
h &:= H(l, P) \tag{1}\\
d &:= h \cdot x \mod n \tag{2}\\
B_{P,l} &:= S_d(E_{HKDF(l,P)}(R_{P,l})), dG \tag{3}\\
q_{P,l} &:= H(dG) \tag{4}
\end{align}
$$

Searching for records under label l in zone P

$$
\begin{align}
h &:= H(l, P) \tag{5}\\
q_{P,l} &:= H(hP) = H(hxG) = H(dG) \Rightarrow \texttt{obtain } B_{P,l} \tag{6}\\
R_{P,l} &= D_{HKDF(l,P)}(B_{P,l}) \tag{7}
\end{align}
$$

Part II: Secure Multiparty Computation

Core features of social networking applications

- Users create profiles and messages (“user-generated content”)
- Users connect to each other and/or subscribe to channels
- Communication happens over those connections

Why?

Core features of social networking applications

- Users create profiles and messages (“user-generated content”)
- Users connect to each other and/or subscribe to channels
- Communication happens over those connections

Why?

Management of information overload via collaborative filtering

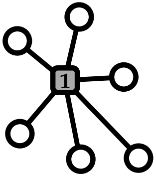
Application Domains

Business

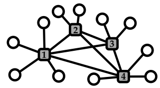
News

Friendship

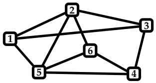
Architectures for Social Networks



Centralized



Federated



Decentralized

diaspora*

matrix

gnusocial

Bitmessage

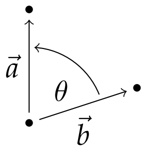
Secure Multiparty Computation

- Alice and Bob have private inputs a_i and b_i.
- Alice and Bob run a protocol to collaboratively compute $f(a_i, b_i)$.
- Only one of them learns the result
- Adversary model: honest but curious

Collaborative Filtering $\equiv$ Scalar product

Motivation

- Scalarproduct $\Rightarrow$ Cosinus-Similarity:



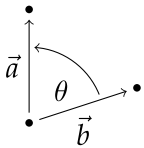
$$\vec{a} \cdot \vec{b} = ||\vec{a}|| \cdot ||\vec{b}|| \cos\theta \tag{8}$$

$$\Leftrightarrow \cos\theta = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \cdot ||\vec{b}||} \tag{9}$$

Collaborative Filtering ≡ Scalar product

Motivation

- Scalarproduct ⇒ Cosinus-Similarity:



$$\vec{a} \cdot \vec{b} = ||\vec{a}|| \cdot ||\vec{b}|| \cos\theta \qquad (8)$$

$$\Leftrightarrow \cos\theta = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \cdot ||\vec{b}||} \qquad (9)$$

Properties

- Private inputs remain protected (given limited number of interactions)
- Efficient in bandwidth and computation

The Protocol[3]

Alice's public key is $A = g^a$, her private key is a. Alices sends to Bob $(g_i, h_i) = (g^{r_i}, g^{r_i a + a_i})$ with random values r_i for $i \in M$.
Bob replies with:

$$\left(\prod_{i\in M} g_i^{b_i}, \prod_{i\in M} h_i^{b_i}\right) = \left(\prod_{i\in M} g_i^{b_i}, (\prod_{i\in M} g_i^{b_i})^a g^{\sum_{i\in M} a_i b_i}\right)$$

Alice can then compute:

$$\left(\prod_{i\in M} g_i^{b_i}\right)^{-a} \cdot \left(\prod_{i\in M} g_i^{b_i}\right)^{a} \cdot g^{\sum_{i\in M} a_i b_i} = g^{\sum_{i\in M} a_i b_i}.$$

If $\sum_{i\in M} a_i b_i$ is sufficiently small, Alice can efficiently compute the scalar product by solving the DLP.

[3] Joint work with Tanja Lange

Preliminary experimental results

Länge	ECC-2^{20}	ECC-2^{28}
25	2 s	29 s
50	2 s	29 s
100	2 s	29 s
200	3 s	30 s

The pre-computation for ECC-2^{28} is $\times 16$ more expensive than for ECC-2^{20}, as the table grows with $\sqrt{n}$.

Conclusion

- GNU name system is a PKI using private information retrieval
- SMC can be used to efficiently perform collaborative filtering
- Cryptography can help us build better privacy-preserving decentralized networks!

This was **only** a short introduction.
GNUnet includes other cool cryptographic protocols.

Questions?

Literature:

- Matthias Wachs, Martin Schanzenbach and Christian Grothoff. *A Censorship-Resistant, Privacy-Enhancing and Fully Decentralized Name System*. **13th International Conference on Cryptology and Network Security**, 2014.
- Geert Lovink and Miriam Rasch. *Unlike Us Reader: Social Media Monopolies and their Alternatives*. Institute of Network Cultures, 2013.
- John Naughton. *Death by drone strike, dished out by algorithm*, **The Guardian**, 21.2.2016.
- Lee Fang. *The CIA is investing in firms that mine your Tweets and Instagram photos*. **The Intercept**, 14.4.2016.

More Information on the Web:

- `https://gnunet.org/`
- `https://taler.net/`
- `https://grothoff.org/christian/`