Privacy@LAAS

Marc-Olivier Killijian

(joint work with S. Gambs, C. Aguilar, J. Guiochet & students)

Why Privacy ?

- **Essential Right**
  - **Human Rights (art 12.)**
  - **I&L …**
  - **EU…**
- **Data-leaks**
  - **Phone makers (Apple, Google, MS)**
  - **Access Providers (Orange, Skyhook, etc.)**
  - **Service & Application Providers (YellowPages, GAFAs, etc.)**
  - **GeoSocial Nets (Facebook, GoWalla, Twitter, etc.)**
  - **Databases (Crawdad, Nokia, MS, etc.)**
  - **Datacenters get hacked - 1 billion records breached in 2014**

Why Privacy ?

- Essential Right
  - Human Rights (art 12.)
  - I&L …
  - EU…
- Data-leaks
  - Phone makers (Apple, Google, MS)
  - Access Providers (Orange, Skyhook, etc.)
  - Service & Application Providers (YellowPages, GAFAs, etc.)
  - GeoSocial Nets (Facebook, GoWalla, Twitter, etc.)
  - Databases (Crawdad, Nokia, MS, etc.)
  - Datacenters get hacked - 1 billion records breached in 2014

Why Privacy ?

- For Security !
- But for security we need to monitor individuals !
  - NSA, Snowden, CCTV, Black-boxes, Hadopi, etc.
- For « dumb » security maybe ….
- But next-generation security will be « end-to-end »
  - Most efficient security attacks involve social-eng.
  - Personal info. -> easier social engineering
  - e.g. the President Attack
- Identity theft is heavily used by mafias
- Breach -> Leaked personal info -> Bigger breach

LAAS Privacy Activities

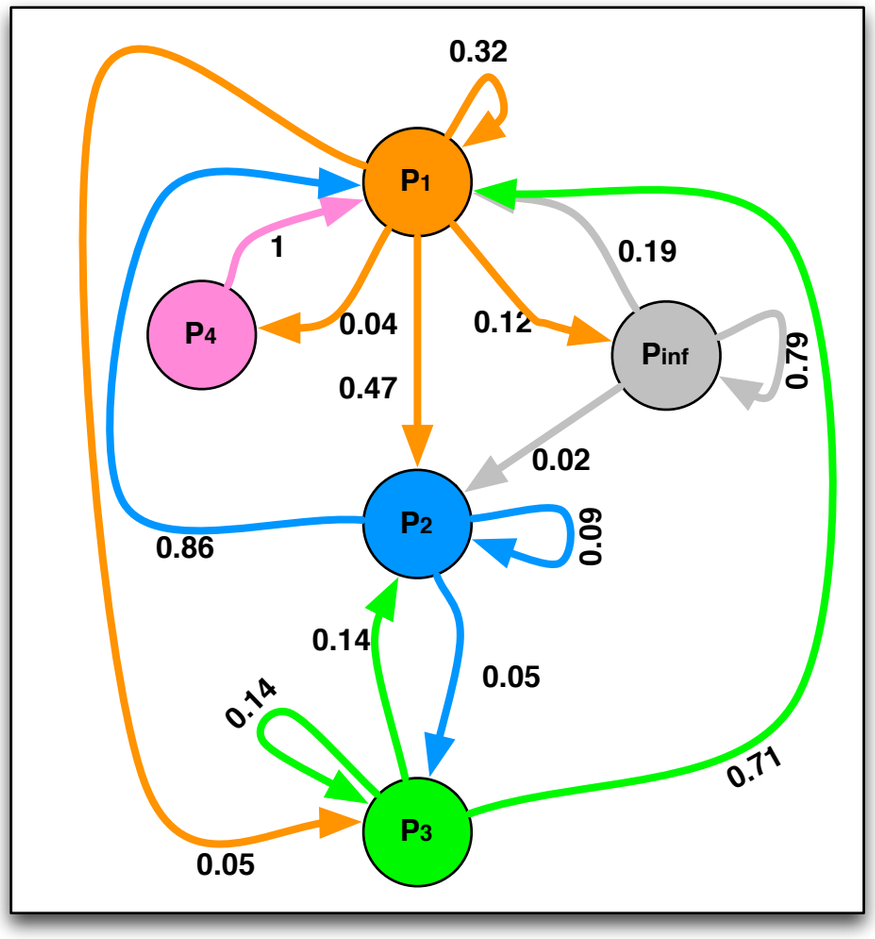
- Geoprivacy
  - Attacks (with S. Gambs-UQAM)
  - Protection mechanisms(with S. Gambs-UQAM)
  - Privacy Risk Analysis (with J. Guiochet-LAAS)
- Homomorphic cryptography & PETS (with C.Aguilar-IRIT)
  - NFLLib
  - XPIR

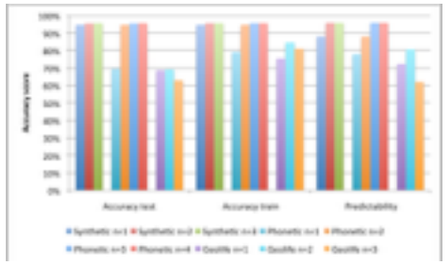
➡PSI, OT, ORAM, Sniffer, etc

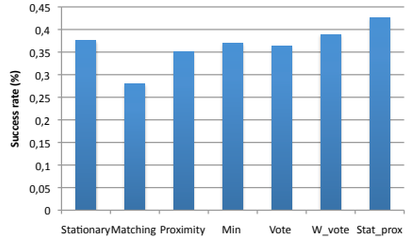
➡Genomic Privacy

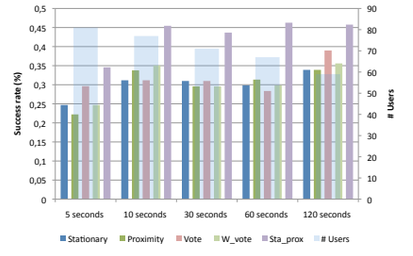
1 model to rule them all

- **Geoprivacy: A single Markov-based model for all type of attacks**
  - **POI identification, movement semantics**
  - **Next-place prediction ~90%**
  - **Re-identification (after sampling) ~45%**
  - **[JCSS'14]**



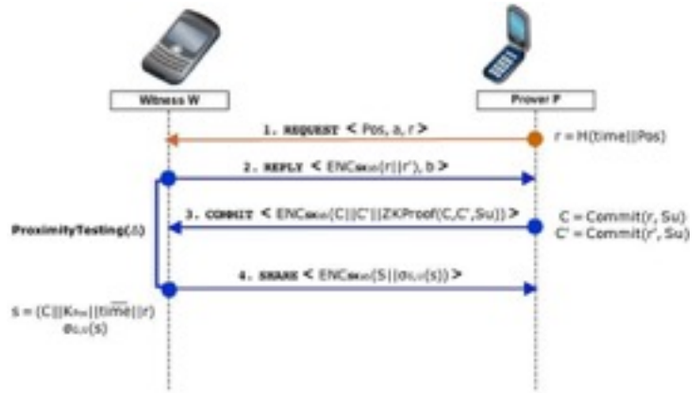






Protection of GeoPrivacy

- General Principles
  - Sovereignty and minimization: control dissemination of our data
  - Decentralization: favor local interactions
  - Privacy Risk Analysis [PMC'14]
- Mechanism-centered
  - Location-based services that don't need to send location
  - Verified positioning
- Data-centered
  - Sanitization: produce modified locations
- [SRDS'14, BalkanCryptSec'14]



Homomorphic Encryption 101

- (Not Homomorphic) Encryption
  - $D_k(C_k(m)) = m$
  - $D_k(C_k(m_1)+C_k(m_2)) = \text{noise}$
- Homomorphic Encryption
  - $D_k(C_k(m)) = m$
  - $D_k(C_k(m_1)\oplus C_k(m_2)) = m_1 + m_2$
  - $D_k(C_k(m_1)\odot C_k(m_2)) = m_1 \,.\, m_2$
- Somewhat Fully vs. Fully (SFHE/FHE)
  - bounded number of homomorphic ops vs. unbounded
- Cloud computing on private data, searchable encryption, etc.

NFLlib is a C++ library

- (efficient) Polynomial calculus
- Specialized, i.e. not-generic
  - targeting ideal lattices in $R_p=Z_p[x]/(x^n+1)$
  - fixed degree polynomials (n power of 2)
  - fixed size coefficients (>p) for modular operations
  - (but several instances can co-exist and interact)
- [CT-RSA'16] + github.NFLlib (6 cites 22 Stars 7 Forks)

Private Information Retrieval 101

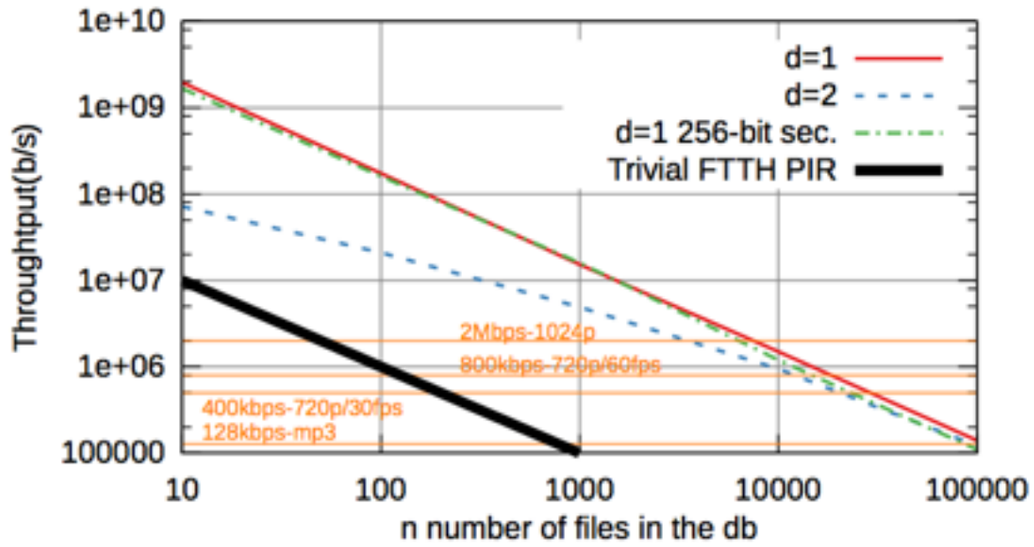
- Informally
  - A protocol allowing a user to retrieve an element from a database, without revealing which one

- {Private Information} Retrieval ?
  - No! The information retrieved is a priori public

- Private {Information Retrieval}
  - Element retrieved is unknown

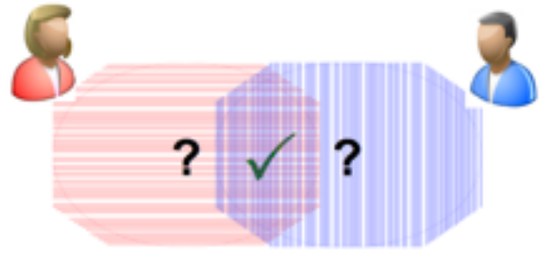
XPIR Tput on static data

- From 1Mbit/s to 25 Gbit/s
- Netflix-like : private streaming of HD movies

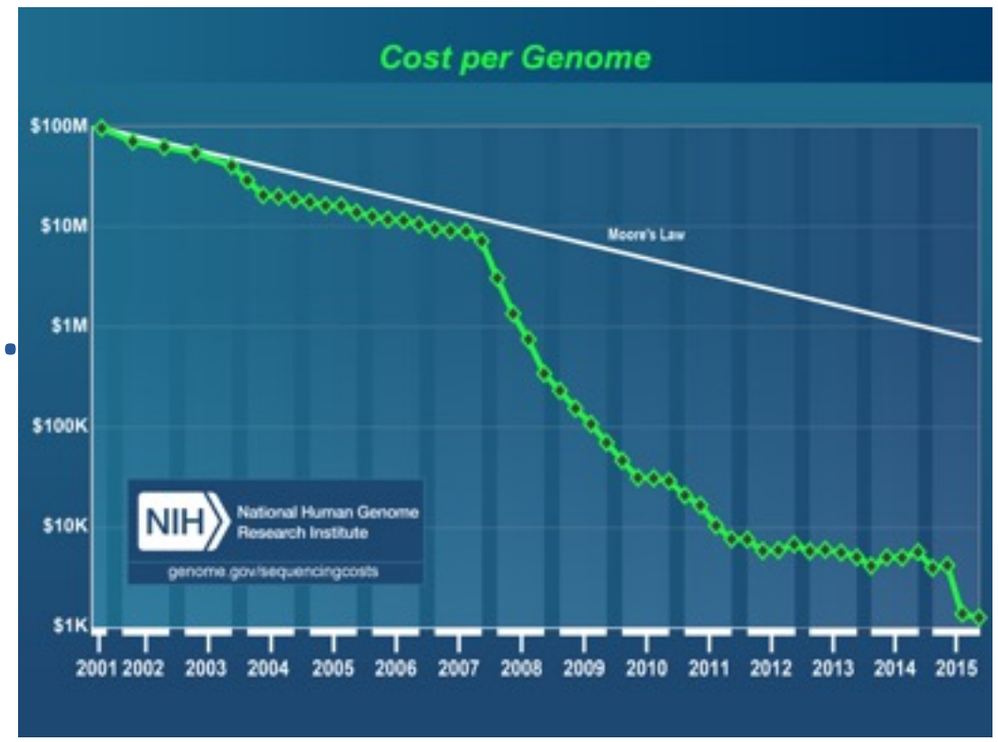


  - Static data : preprocessing occurs only once
  - Tput for 1 core (1 user) on a commodity CPU
    - 9000 HD movies (X100 trivial PIR)
- [PETS'16] + github.XPIR (11 cites 29 stars 7 forks)

Other PETs



- Bloom filters, Secure Scalar Products
- Private Set Intersection - (n=215)
  - Paillier-based [DeCristofaro-FC10]: 3 hours / 16 MB
  - OT-based [Pinkas-UsenixSec14]: 1458 ms / 8MB
  - NFLlib: 110ms / 128ko
- name your favorite PET, etc. ?
  - |PSI|, Private searching
  - ORAM, Anonymous direct downloads, ..
- Application to Genomic Privacy
  - iDash 2016 Challenge Finalist

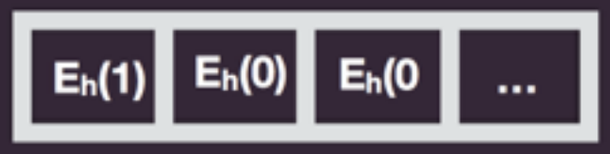


Genomic Example

1. Generate tag h for variant to be queried (e.g h_2)
2. Retrieve index (e.g. 0)

3. Homomorphically encrypt an array of 0's and a 1 in the respective index

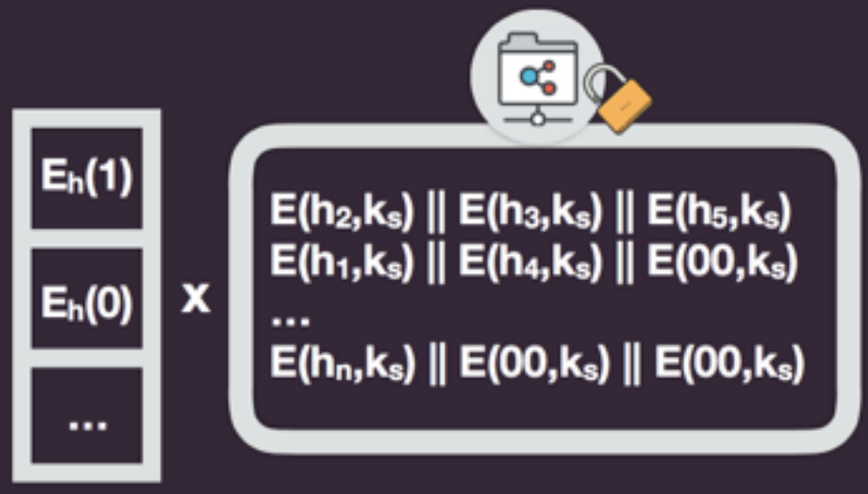
XPIR (Private Information Retrieval)



(optionally) 3b. $v=[E_h(E(h_2,k_s)), E_h(E(h_2,k_s)), \ldots]$

(optionally) 4.b Subtract v to the result and then multiply by random scalar

4.



5. Decrypt result and fetch variant

préGDR Sécurité Informatique

- Crée en 01/2016 par l'INS2I
- Dirigé par Gildas Avoine(IRISA)
- Structurer la communauté Sécurité française
  - crypto, privacy, données multimédia, réseaux et infra, systèmes logiciels, systèmes matériels, vérification
- Colloque "Sécurité informatique: mythes et réalité
  - 9-10/12/16 - (Preneel, Hubaux, Joux, etc.)
  - vidéos prochainement en ligne
- Rencontres Entreprises DOCtorants Sécurité (REDOCS)
  - début novembre - attention 2017
- École d'été Cyber in Bretagne du pré-GDR Sécurité et du PEC
  - 2017 - Cyber in * ?

préGDR Sécurité Informatique

- Crée en 01/2016 par l'INS2I
- Dirigé par Gildas Avoine(IRISA)
- Structurer la communauté Sécurité française
  - crypto, privacy, données multimédia, réseaux et infra, systèmes logiciels, systèmes matériels, vérification
- Colloque "Sécurité informatique: mythes et réalité
  - 9-10/12/16 - (Preneel, Hubaux, Joux, etc.)
  - vidéos prochainement en ligne
- Rencontres Entreprises DOCtorants Sécurité (REDOCS)
  - début novembre - attention 2017
- École d'été Cyber in Bretagne du pré-GDR Sécurité et du PEC
  - 2017 - Cyber in * ?

http://gdr-securite.irisa.fr -> Mailing-liste + Forum