Consensual and Privacy-Preserving Sharing of Multi-Subject and Interdependent Data

Alexandra-Mihaela Olteanu\textsuperscript{1,2}, Kévin Huguenin\textsuperscript{2}, Italo Dacosta\textsuperscript{1}, and Jean-Pierre Hubaux\textsuperscript{1}

\textsuperscript{1}EPFL, \textsuperscript{2}UNIL
Multi-Subject and interdependent data is abundant

• Individuals share increasing amounts of data online

• Other individuals are irrevocably affected by the shared data
  – The uploader is not the owner of the data
  – The data has consequences for other individuals due to interdependencies
(Privacy) problems

• For the users
  – Sensitive information leaked without users’ consent or even awareness
  – Prejudice and discrimination
  – Revenge pornography

• For the service providers
  – Lawsuits
Facebook pays costs in naked photo settlement

A 14-year-old girl who sued Facebook after a man allegedly posted a naked photo of her on the website has settled the landmark legal action out of court.

Facebook will pay the Northern Ireland teenager’s legal costs, under the terms of the settlement which a court heard was the first of its kind in the world.

The photo was allegedly posted on a so-called “shame” page on Facebook several times between 2014 and 2016.

Facebook warned it faces legal action from ‘revenge porn’ victims

Settlement with teenager over naked images of her posted online on networking site has ‘moved goalposts’, say lawyers

Facebook is facing a number of lawsuits from victims of “revenge porn”, a leading libel lawyer has warned, after a teenager reached a settlement with the social networking site over naked images of her that were posted online.
Facebook is working on a new way to rid the platform of 'revenge porn,' and giving it a first whirl in Australia (where, incidentally, Domino’s Pizza and McDonald’s have also experimented with unusual deliveries); specifically, they’ll be asking for nudes.

The Australia Broadcasting Corporation (ABC) reported this week that Facebook will be testing out a new preemptive method for blocking the unauthorized distribution of nude photos on the site, and (they hope) with the help of Australian Facebook users. According to ABC, Facebook has partnered with the country’s e-Safety Commissioner to develop the program, aimed at protecting not just extant victims of vindictive image-sharing, but also potential ones.
(Privacy) problems

• For the users
  – Sensitive information leaked *without* users’ *consent* or even *awareness*
  – *Prejudice* and *discrimination*
  – *Revenge pornography*

• For the service providers
  – Lawsuits

• Amazon MTurk survey (N=321)
  – *10.3%* of the participants claimed they were *victims* of *discrimination* or *prejudice* based on online content
    – *66.7%* of these reported as *cause content shared by others*
    – *4.1%* of the participants reported being *victims of revenge pornography*
Some good elements do exist

- Tagging can be automated
- Companies start to work around the notion of consent
You looked so much younger before your PhD, Alexandra — with Alexandra Oltanu.

Please note that this channel is reserved for people reporting potential violations of their privacy rights concerning their image on Facebook. If you're writing in about something else, please return to the Help Center:

www.facebook.com/help

If you need help because someone is threatening to share something you want to keep private, follow these steps outlined in this form:

www.facebook.com/help/561743407175049

Also note that while we do review all reports, you won't receive a confirmation email if we take action on your report.

What are you trying to report?
- Photo
- Video
- Other

What type of photo are you trying to report?
- Profile picture
- Other photo

Where do you live?
- In the US
- Outside the US

Please provide a link to the content you're trying to report so we can investigate. To get a link to the exact content you want to report:

1. Find the content (ex: photo, video, comment) you want to report
2. If this content is on someone's Timeline, click on the date/time it was posted (ex: 27 minutes, May 30 at 7:30pm)
3. Copy the URL from your browser's address bar:
Pornhub is using machine learning to tag its 5 million videos

The company is starting with facial recognition, but wants its AI to recognize:

By James Vincent | @vincent Oct 11, 2017, 2:03pm EDT

Content Removal Request?

Pornhub takes all content removal requests seriously. Should you be a victim of revenge porn, blackmailing or intimidation because of a video or photo of yourself on our sites that you did not authorize, please complete the form below and we will remove the content expeditiously.

For all other content removal requests related to copyright infringement, please contact copyright@pornhub.com or use the DMCA takedown request form on /information#dmca.

Email:

Name:
But ...
Goals

*Privately detect & inform* all affected individuals when content regarding them is submitted

Enable individuals to grant their *consent before* such content is made available
Challenges

• Identity claim

• *Privately & securely determine* the affected users by some piece of data

• *Privately & securely contact users* and provide them with enough context to make an *informed decision*

• Privately & securely collect and enforce consent decisions

⚠ *Data dependent*
System model & adversarial assumptions

• Key features
  – Effectiveness
  – User privacy through *anonymity* and *unlinkability*
    – wrt to any other users and to any service providers
  – Robustness against malicious user behavior
  – Usability and transparency

• Service providers: *honest-but-curious*
• Users of the system: *malicious*
System overview

• **Content Management Service** (CMS) – user-generated content sites (e.g., flickr, Pornhub, YouTube, openSNP, etc.)

• **Identity Management Service** (IMS) – e.g., Facebook, government agencies
  – Manages users’ identities
  – Offers services to identify the users associated with a particular content
  – In charge of users’ relationships (social and family)

• **User applications** (CMS and IMS) – components that users interact with to publish content and to review consent requests
ConsenShare: an example for photos

User registration protocol

The user provides a webcam photo/video to the IMS

1. Bob provides P, e-mail, ...
2. IMS detects faces in P
   If only one viable face is detected
   Compute face representation
   Store + e-mail + ...
3. IMS sends ok
Photo upload protocol

Alice - wants to upload $P$

Bob - appears in $P$

CMS

IMS

 encrypt $P_B$, context, $E Pk_{sid}$

$\sigma = \text{sig}_{\text{ims}}(m)$

$m = \text{hash}(P_B)$

If $P_B$ is a valid background

otherwise

$m = \text{invalid\_background'}$

publish $P_V$

$m = \text{ver}_{\text{ims}}(\sigma)$

verify $\text{hash}(P_B) = m$

store $\text{hash}(P_B)$, context

$\sigma = \text{sig}_{\text{ims}}(m)$
Alice wants to upload $P$

Bob appears in $P$

- CMS
- IMS

(4) $hash(P_B), E_{Pk_{IM}}(r_2)$

verify $hash(P_B)$ is stored
generate random $sid$
store $sid, E_{Pk_{IM}}(r_2)$ linked to $hash(P_B)$
status(sid) = Pending

(5) $sid, E_{Pk_{IM}}(r_2)$

$r_2 = D_{Sk_{IM}}(E_{Pk_{IM}}(r_2))$
lookup best match($r_2$)

(6) $sid, CMS$

Generate key pair ($P_{k_{sid}}, Sk_{sid}$)

(8) $sid, P_{k_{sid}}$

Encrypt using $P_{k_{sid}}$

(9) $sid, E_{P_{k_{sid}}}(\bullet)$

store $E_{P_{k_{sid}}}(\bullet)$ linked to $sid$

(10) $P_B, Alice$, context, $E_{P_{k_{sid}}}(\bullet)$

(11a) $sid, Sk_{sid}$

If accept

$x = D_{Sk_{sid}}(E_{P_{k_{sid}}}(\bullet))$

check $x$
verify $E_{P_{k_{sid}}}(r_2) = E_{P_{k_{IM}}}(P)$
status(sid) = Accepted
update $P_V$

(11b) $sid, 'deny'$

status(sid) = Denied

---

- spam filtering
- manual decision
- policy (e.g., "accept if not nude", "accept if I am the uploader")
- automatic decision (e.g., machine learning)
- content obfuscation
- specific audience
Privacy analysis

• Information observed by the CMS (flickr)
  – Background image
  – Encrypted faces and encrypted face vectors

• Information observed by the IMS (facebook)
  – Background image
  – Face vectors

• Information observed by a consenter (👨‍💻)
  – Context
  – Background image
  – His own face

---

**Cannot link encrypted faces to users’ identities**

**Cannot link users’ identities to photos**

**Guaranteed user anonymity, unlinkability and robustness against malicious users**
Potential for adoption

- CMS incentives (flickr, Pornhub)
  - Follow new trends and/or regulations, lawsuit avoidance, maintain good reputation, increase user base

- IMS incentives (Facebook)
  - B2B arrangements with different CMS, increase user base

- User incentives (People)
  - Awareness and control over their data, usability and transparency
Surveyed potential interest

Q27: Do you find ConsenShare useful?

Q29: Would you use ConsenShare with Machine Learning automation?

Q32: Would you use ConsenShare?

Q28: Which policies would you use with ConsenShare?
Implementation

• 128-bits security for the basic cryptographic operations (hash, sign/verify, encrypt/decrypt and generate keys)

• Python with PyNaCl* and OpenFace**

• W/o any CPU optimizations

* PyNaCl - Python binding to the Networking and Cryptography library. https://github.com/pyca/pynacl

Evaluation

• Dataset: random sample of 17k+ photos from Yahoo’s YFCC100m
  – Original photos uploaded by real users on Flickr

• Efficiency of ConsenShare
  – Per-photo CPU and bandwidth consumption (CMS, IMS, uploader, consenter)
  – Measured on Intel i7 CPU, 2.8 GHz, 8GB RAM
  – Reasonable bandwidth and CPU requirements
Conclusion

• ConsenShare: practical and privacy-preserving
  – Preserves the main features of existing CMS
  – Provides privacy guarantees (wrt to service providers and to other users)
  – Has negligible overhead

• Next steps
  – Include more human interaction aspects
  – Perform a usability study for a fully integrated prototype

• Extensible to other data types
  – Audio, video, genomic, co-location data