C3P: Context-Aware Crowdsourced Cloud Privacy





Privacy Enhancing Technologies Symposium, 2014











20%

of files shared to the cloud contain protected data

20% of files shared to the cloud contain protected data

of sensitive files contain PII



Source: Elastica's Q2 2015 Shadow Data Report

20% of files shared to the cloud contain protected data

of sensitive files contain PII



Emergence of "Shadow IT"

Anti-Snooping Tools for the Cloud



What if Antivirus Software was Similar?









Obstacles



Obstacles



Manual effort and expertise for assessing data sensitivity

I dedicate the rest of my life for sorting out sensitive from non-sensitive files on my HD

8,8

What is needed?



Ensure serviceable protection instead of brute encryption.

What is needed?

Ensure serviceable protection instead of brute encryption.



flickr

write LAT_FX

Account for the metadata, sharing environment, and data content.

What is needed?

Ensure serviceable protection instead of brute encryption.

Account for the metadata, sharing environment, and data content.

I dedicate the rest of my life for sorting out sensitive from non-sensitiv files on my HD

flick

writeLATE

Automatically estimate the sensitivity of shared data.











Defining Data through Context





Defining Data through Context





Defining Data through Context







Privacy Preserving Crowdsourcing

















Sensitivity Estimation using Item Response Theory



Sensitivity Estimation using Item Response Theory


















Connecting the Dots ?



Client

14







Client

14







Client



Client















Evaluation



beta





• **Ex:** With which privacy level would you share a project presentation with a friend?

96

81



- **Ex:** With which privacy level would you share a project presentation with a friend?
- Standardized Infit Statistic:
 - (x-axis values should lie in [-2,2])





- **Ex:** With which privacy level would you share a project presentation with a friend?
- Standardized Infit Statistic:
 - (x-axis values should lie in [-2,2])



Artificial Artificial Intelligence



- **Ex:** With which privacy level would you share a project presentation with a friend?
- Standardized Infit Statistic:
 - (x-axis values should lie in [-2,2])



96





- **Ex:** With which privacy level would you share a project presentation with a friend?
- Standardized Infit Statistic:
 - (x-axis values should lie in [-2,2])







- **Ex:** With which privacy level would you share a project presentation with a friend?
- Standardized Infit Statistic:
 - (x-axis values should lie in [-2,2])





• Synthetic Dataset:





Zipf context distribution



av.: 1 Item/6 hours

• Synthetic Dataset:









av.: 1 Item/6 hours





Crowdsourcing cost: Hit rate (HR) from 0 to 90% in 10 days



Crowdsourcing cost: Hit rate (HR) from 0 to 90% in 10 days **Anonymity cost:** HR difference starts high but vanishes in 10 days.

17

Effect of Sharing Behavior on the Temporal Cost





Effect of Sharing Behavior on the Temporal Cost



Effect: Long tail distribution is of lower temporal cost.

Robustness Towards Malicious Users?

- Test:
 - Assign sensitivities to items and attitudes to people.
 - **Honest** users choose policies according to the model.
 - Malicious users choose policies at random.



Robustness Towards Malicious Users?

- Test:
 - Assign sensitivities to items and attitudes to people.
 - **Honest** users choose policies according to the model.
 - Malicious users choose policies at random.

PRobustness Towards Malicious Users?

- Test:
 - Assign sensitivities to items and attitudes to people.
 - **Honest** users choose policies according to the model.
 - Malicious users choose policies at random.

Robustness Towards Malicious Users?

- Test:
 - Assign sensitivities to items and attitudes to people.
 - **Honest** users choose policies according to the model.
 - Malicious users choose policies at random.



Robustness Towards Malicious Users?

- Test:
 - Assign sensitivities to items and attitudes to people.
 - **Honest** users choose policies according to the model.
 - Malicious users choose policies at random.





6

Tolerance: 25% malicious: ≈8% difference, 50% malicious: ≈17% difference

Future Work

- Recommendation of policies to users
- Batch sensitivity analysis



• Working with IRT alternatives.



ELO

One more thing...

Adam Tow | AllThingsD.com

PrivyShare - Desktop



PrivyShare Benefits

• Works with any cloud service



PrivyShare Benefits

- Works with any cloud service
- Provides "Sensitivity as a Service"


PrivyShare Benefits

- Works with any cloud service
- Provides "Sensitivity as a Service"
- Offers fine grained protection
 - Metadata cleaning
 - Face Blurring
 - Encryption
 - Encryption + Auxiliary Information (automatic summaries, blurred thumbnails)



PrivyShare - Browser



PrivySeal: Dealing with 3rd Party Cloud Apps



privyseal.epfl.ch

Questions



Images/Media Credits

- <u>Pixel77</u>
- Freepik