GALLOGLY COLLEGE OF ENGINEERING
SCHOOL OF COMPUTER SCIENCE
The UNIVERSITY of OKLAHOMA

VIPER

Visual Inspection of Personal Exposed Records

Jasmine DeHart

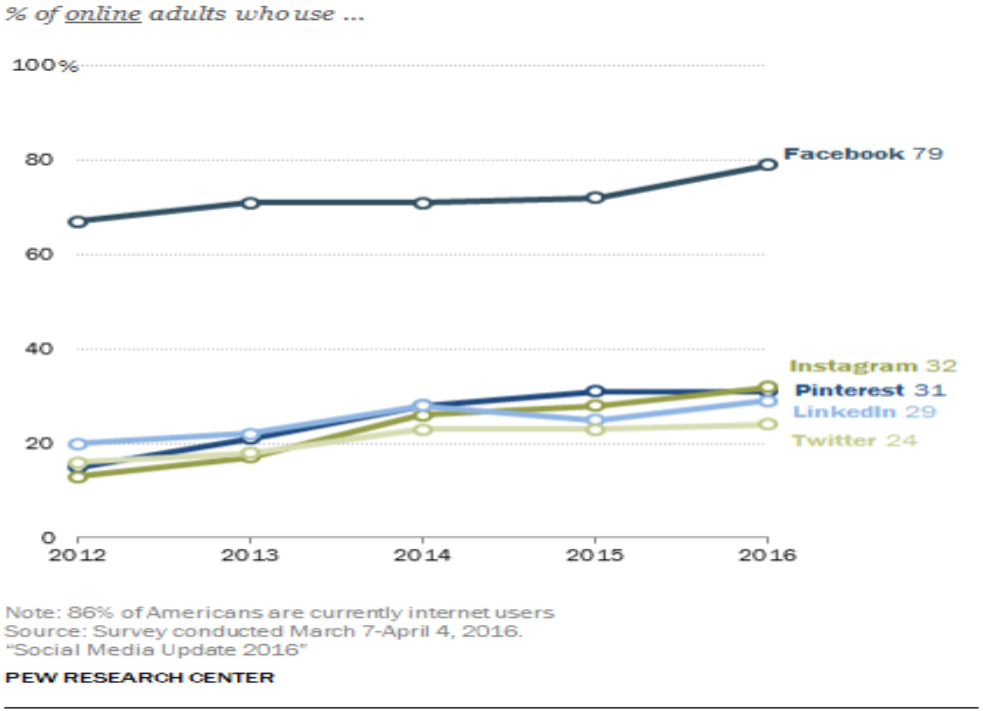
Advisor: Christan Grant

OU Data Analytics Lab

OUTLINE

- Background and Overview
- Openings for Attacks
- What is VIPER?
  - Mitigation Techniques
  - Redaction Spectrum
- Related Works

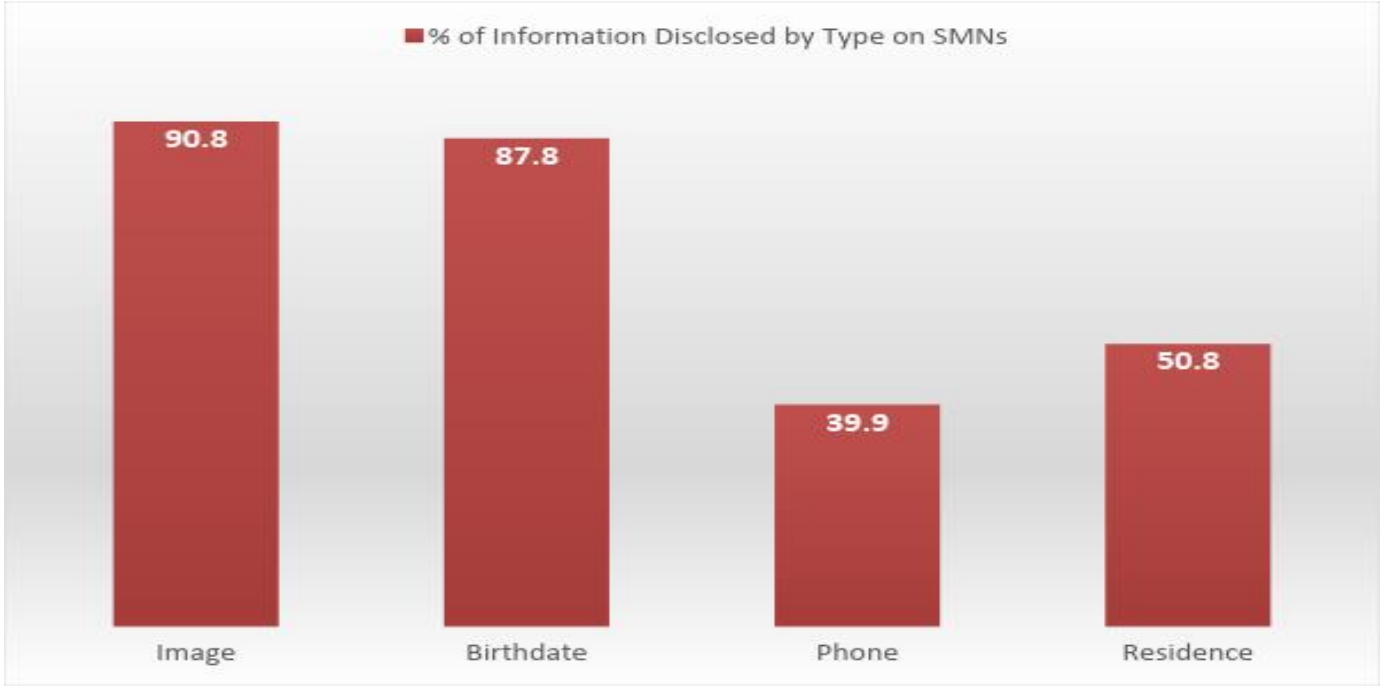
Overview



Greenwood, S., Perrin, A., & Duggan, M. (2016). Social media update 2016. Pew Research Center, 11, 83.

Visual Privacy Leak Example
Hawaii Emergency Agency
source: Twitter

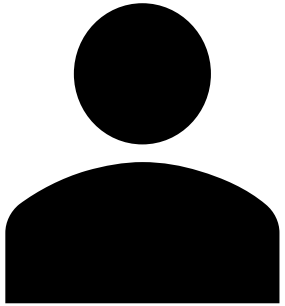
Types of User Disclosed Information

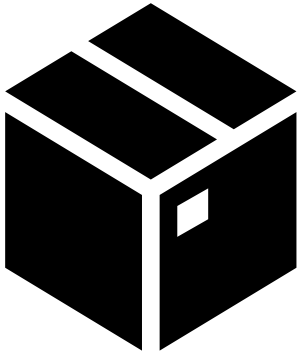


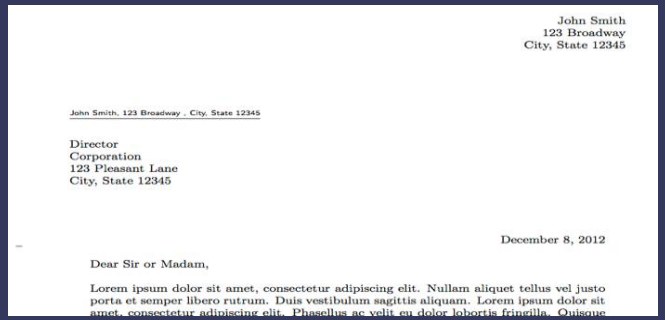
- Relationship Status
- Dating Preference
- Political Views

Gross, R., & Acquisti, A. (2005, November). Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM workshop on Privacy in the electronic society (pp. 71-80). ACM.

Openings for Attacks







John Smith
123 Broadway
City, State 12345

John Smith, 123 Broadway , City, State 12345

Director
Corporation
123 Pleasant Lane
City, State 12345

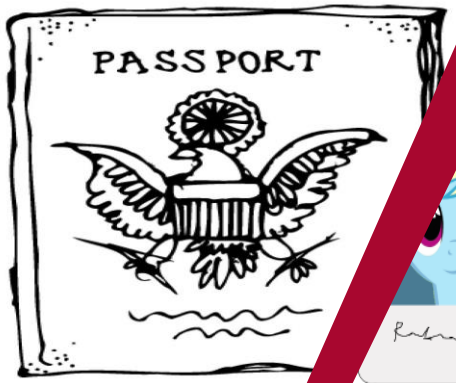
December 8, 2012

Dear Sir or Madam,

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nullam aliquet tellus vel justo porta et semper libero rutrum. Duis vestibulum sagittis aliquam. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Phasellus ac velit eu dolor lobortis fringilla. Quisque

Location

- An attacker can use this opening to find out where you live and/or your current location.
- Dangers:
  - burglary
  - stalking
  - kidnapping



Identity

An attacker can use this opening to exploit your identity.

Dangers:

- identity theft
- financial threat
- burglary

Asset

- An attacker can use this opening to exploit your possessions and valuables.
- Dangers:
  - financial threat
  - burglary
  - digital kidnapping/explicit websites

WHAT IS VIPER?

An object detection model used to identify privacy leaks from visual content (images and videos) on SMNs.

VIPER System

DEPLOY SURVEYS

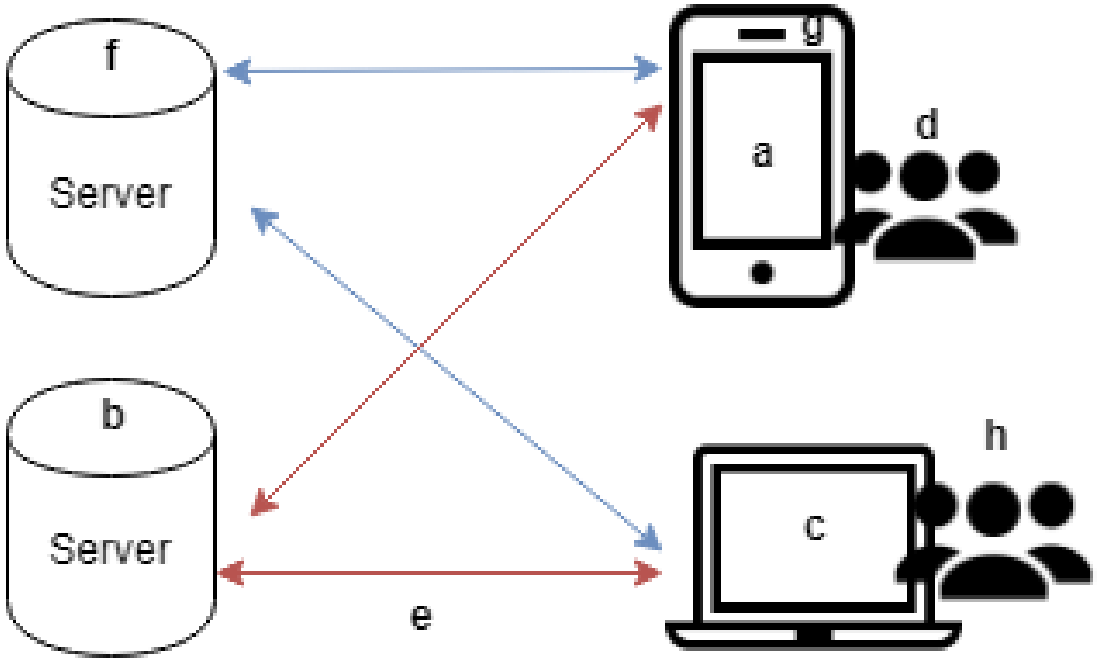
BUILD AN IMAGE DATASET

CREATE OBJECT DETECTION MODEL

ESTABLISH PRIVACY SCORE

IMPLEMENT MITIGATION TECHNIQUES/APPLICATION

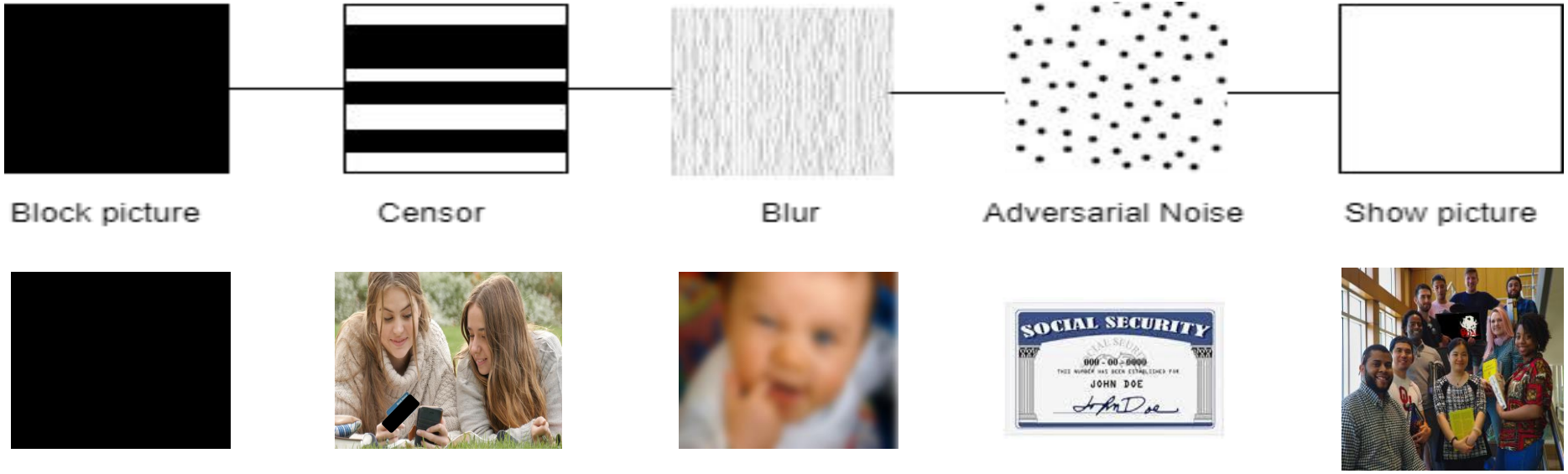
Mitigation Techniques

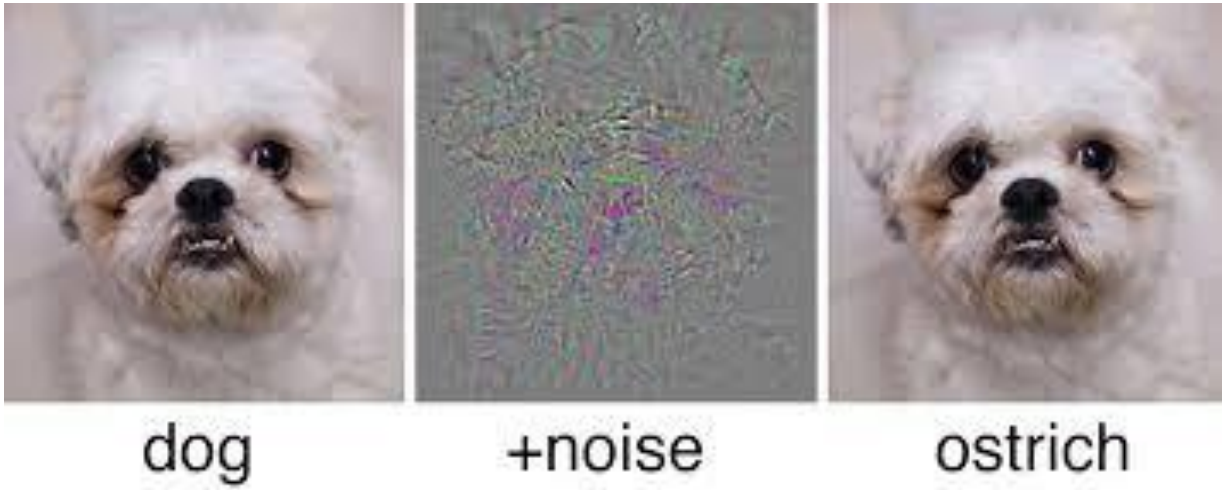


a - *Client app*
b - *Privacy Patrol*
c - *Chaperone bot*
d - *Category tagger*
e - *Privacy Scorer*
f - *Server app*
g - *Interceptor*
h - *Redactor*

DeHart, J., & Grant, C. (2018). Visual Content Privacy Leaks on Social Media Networks. *arXiv preprint arXiv:1806.08471*.

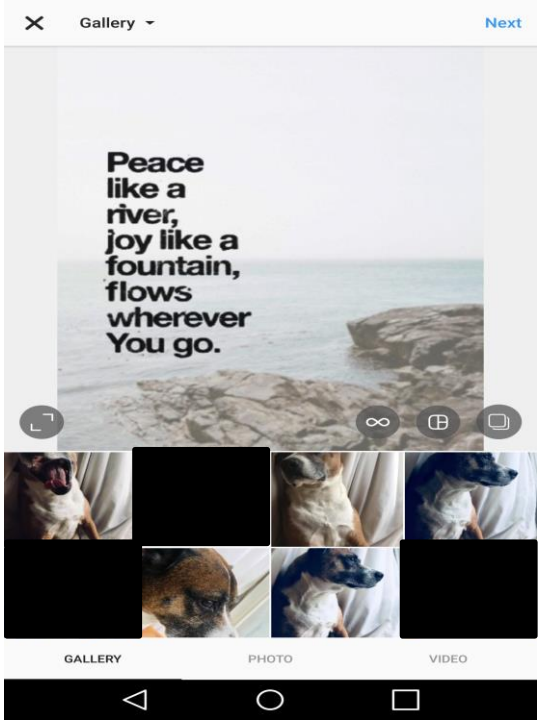
Redaction Spectrum

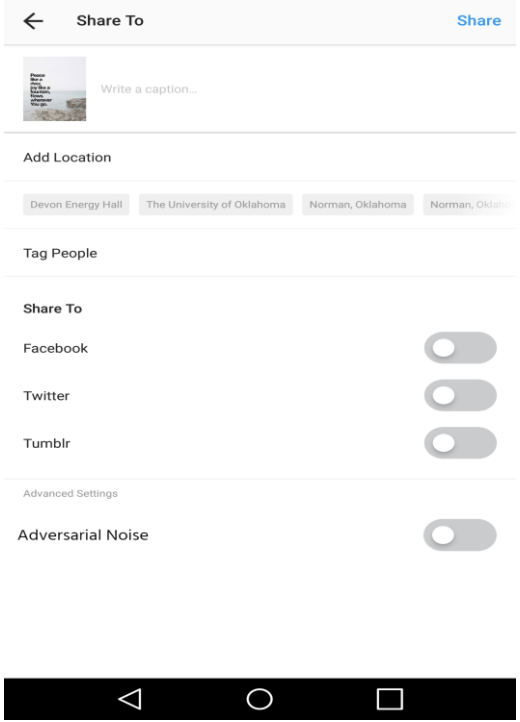




Adversarial Noise Example

Application Design





Central Research Questions

- How pervasive are privacy leaks on SMNs?
- What reasonable techniques can be built to decrease the amount of privacy leaks?

Related Works

Analyzing images' privacy for the modern web (Scquinnari, 2014)

Flickr images from the PiCalert dataset and Visual Sentiment Ontology repository

Analyze based on features and metadata

Used linear support vector machines (SVMs) classifiers

Squicciarini, A. C., Caragea, C., & Balakavi, R. (2014, September). Analyzing images' privacy for the modern web. In Proceedings of the 25th ACM conference on Hypertext and social media (pp. 136-147). ACM.

Information revelation and privacy in online social networks (Gross, 2005)

- Used Facebook
- Studied Carnegie Mellon University (CMU) students
- Found patterns of information revelation and privacy implications
- Evaluated the amount of information disclosed and studied usage of the site's privacy settings

Gross, R., & Acquisti, A. (2005, November). Information revelation and privacy in online social networks. In Proceedings of the 2005 ACM workshop on Privacy in the electronic society (pp. 71-80). ACM.

Photo Privacy Conflicts in Social Media: A Large-scale Empirical Study (Such, 2017)

MULTIPARTY PRIVACY CONFLICTS (MPCS) IN SOCIAL MEDIA

CRITICAL INCIDENT TECHNIQUE FOR QUALITATIVE STUDY

EMPIRICAL BASIS FOR THE PREVALENCE, CONTEXT AND SEVERITY OF PHOTO

Such, J. M., Porter, J., Preibusch, S., & Joinson, A. (2017, May). Photo privacy conflicts in social media: A large-scale empirical study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3821-3832). ACM.

Enhancing Lifelogging Privacy by Detecting Screens (Korayem, 2016)

LOW-COST, LIGHTWEIGHT WEARABLE CAMERAS

USED COMPUTER VISION TO DETECT COMPUTER SCREENS

MANAGING PRIVACY OF WEARABLE CAMERAS

Korayem, M., Templeman, R., Chen, D., Crandall, D., & Kapadia, A. (2016, May). Enhancing lifelogging privacy by detecting screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4309-4314). ACM.

Questions?

Thank you!

Image Citations:
[1] https://www.videoblocks.com/video/young-teenage-friends-online-retail-shopping-purchasing-item-with-credit-card-h6qvipcgiodtlfj6
[2] https://mom.girlstalkinsmack.com/family/how-to-know-your-baby-is-teething.aspx
[3] https://www.kissclipart.com/snake-vector-png-clipart-snakes-clip-art-ak5tsa/
[4] https://medium.com/@samim/adversarial-machines-998d8362e996
[6]https://www.google.com/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwih_di6zMbeAhXmT98KHZa_DzcQjhx6BAgBEAM&url=https%3A%2F%2Fpropertymash.com%2Fnews%2Fapartment-prices-remain-flat-in-december%2F&psig=AOvVaw3_PU_L3Q6U2DPkyyGde9Ui&ust=1541828354457178

Evaluation

- Taxonomy - privacy leak categories
  - Survey from users
  - Application feedback
- VIPER System
  - Model performs competitively with other related object detection techniques
  - Decrease in privacy score from users
  - Scalability of each mitigation technique

Explored Options

- Models: YOLO, Image AI, ml5
- Datasets: COCO, ImageNet