Analysis of Visual Privacy Leaks on Twitter

Makya Stell; Jasmine DeHart; Christan Grant, Ph.D.
The University of Oklahoma, School of Computer Science

Introduction

- Users expose private information on social media networks (SMNs) both knowingly and unknowingly
- People are using trending hashtags and phrases to upload visual privacy leaks to Twitter
  - Ex: #stayoffthesidewalk, #raciststwitter
- Private visual content (images and videos) exposes sensitive information that can be detrimental to one's finances, personal life, and reputation

Methods & Materials

- Deployed a survey (IRB #11349) with the goal of understanding the hashtags and keywords associated with visual privacy leaks on Twitter
- Used the 126 survey responses to create text files with synonymous words
- Initialized the Web Scraper (fig. 1) - used the text files that were created to collect images and data from Twitter
- Three individuals classified the data and placed it into one of three categories based on the average agreement

Results

- Collected 18,751 images from Twitter and placed them into three categories: *severe, moderate, and no risk*
  1. **Severe** - images that include government issued identification (i.e., social security cards, driver's license, etc.), items that can be used to identify a person and/or used for facial recognition (i.e. work or school identification cards), or items that contain insight to a person's location and/or place of residence
  2. **Moderate** - images that include items that can ***ONLY*** be used to identify a person and/or used for facial recognition
  3. **No Risk** - images that do not include any of the above items
- Refer to tables 1 and 2 for final results

Discussion

- From the survey, we found that:
  - ✓ Though each user's perception of privacy varied, their definition of privacy leaks still overlapped
  - ✓ 65% of participants reported that they seen privacy leaks on Twitter; however, we were unable to collect a corresponding amount of visual privacy leaks
- We found that users will not call the private information that they are posting only by its actual name, but they will possibly describe it by location or other current posting trends that are occurring on Twitter

Conclusion

- Users have solid personal notions of privacy, but they do not yet understand how visual privacy leaks can affect them
- As new technologies arise, application developers must implement mitigation techniques that allow users to explore the trade-offs between privacy and sharing
- Future work focuses on identifying trending hashtags and keywords that are closely associated with visual privacy leaks on Twitter

Twitter users are using **hashtags and keywords** when posting **visual** tweets (images and videos) that contain **privacy leaks** both knowingly and **unknowingly.**

For more information about the VIPER Project and affiliated works

SCAN ME

Makya Stell
makyastell@ou.edu

Jasmine DeHart
dehart.jasmine@ou.edu

Christan Grant, Ph.D.
cgrant@ou.edu

This research is supported in part by OK-LSAMP and the School of Computer Science

Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the supporters

Tables & Figures

Figure 1. Scraper Implementation

Figure 2. Severe Visual Privacy Leak

Figure 3. Moderate Visual Privacy Leak

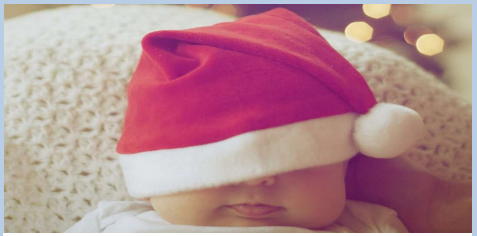

Figure 4. No Risk Example

Category	# of Images Collected
Severe	160
Moderate	327
No risk	18,264

Table 1. Risk Classification From Web Scraping via Twitter

Category		Keyword (Count)	
Severe	(160)	Baby	71
		Driver's License	12
		Financial Document	2
		Hospital	54
		Job	4
		Keys	1
		License Plate	4
		Medication	10
		Medical Records	6
Moderate	(327)	Baby	45
		College Letter	6
		Driver's License	24
		Hospital	123
		Job Promotion	7
		Medical Information	52
		Medication	43
		Work Identification	12
		Workplace	15

Table 2. Distribution of Content for Risk Categories. This table lists the frequency of the content within the Moderate and Severe categories

References

1. DeHart, J.; Stell, M.; Grant, C. *Social Media and the Scourge of Visual Privacy.* Information (MDPI). Special Issue: End of Privacy? 11(2), 57. 2020.
2. DeHart, J; Grant. C. *Visual Content Privacy Leaks on Social Media Networks.* The 39th IEEE Symposium on Security and Privacy (S&P). San Francisco, CA. 2018.