University of Glasgow

School of Computing Science

Human-Centered Security essay:

Obfuscating sensitive content: Some systems obfuscate sensitive content to protect privacy. This can be done by, for example, blurring faces of strangers that appear in photos

Uddeshya Sinha – 2397282S

Word Count: 1636

Dr. Mohamed Khamis

Office 204, Sir Alwyn Williams Building

School of Computing Science

University of Glasgow

I Introduction

The advent of social media has led to raising privacy concerns among internet users. However, this is not new. It was first considered to be a major issue in 1998, the year which saw the rise of privacy and data protection laws all around the world. Fast-forward to 2019, privacy is still one of the biggest concerns even in the presence of rigorous GDPR in the EU, for example. Security and privacy issues can be observed among a wide range of spectrum ranging from social media to the internet of things. It’s no longer a revelation that humans are the most critical asset that needs to be protected from malicious actors. In light of this, privacy researchers have been introducing new groundbreaking techniques for protecting the privacy of individuals while preserving good user experience. This essay summarizes the use of privacy protection solutions for obfuscating sensitive content as proposed by researchers and identifies the usability, privacy and security issues in these solutions by suggesting improvements.

The rest of this essay is organized as follows. Section II is divided into 4 sub-sections presenting the literature review and critical analysis of 4 research papers. Finally, Section III gives a conclusion and proposes some future research directions.

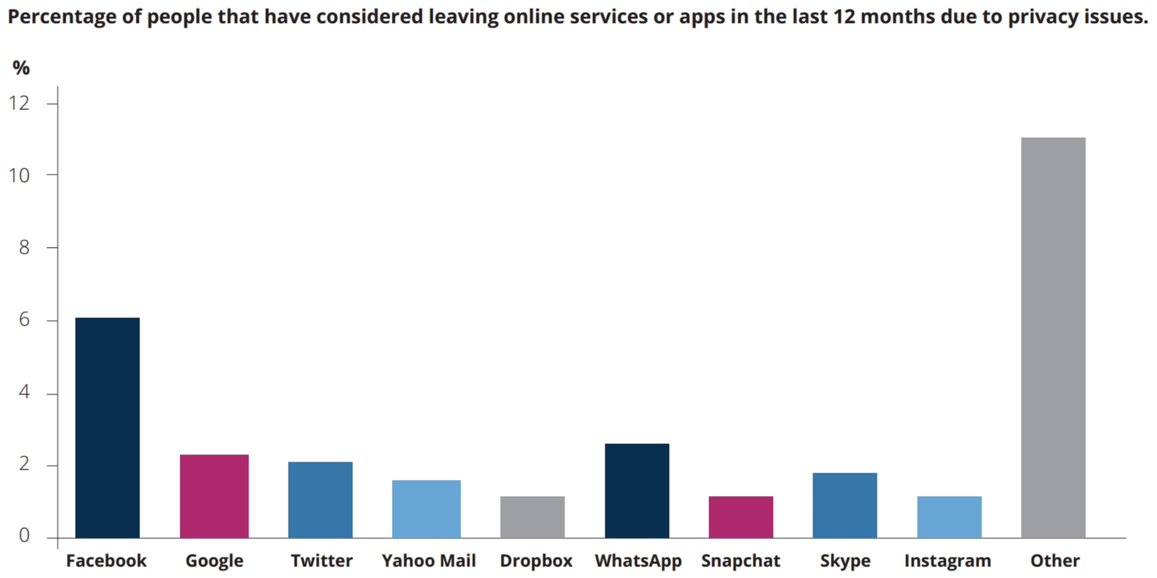

Image showing the percentage of people considered to leave online services/apps due to privacy issues. [1]

II Literature Review and Critical Analysis

a) Paper 1: Enhancing Lifelogging Privacy by Detecting Screens [5]

Summary: In this paper, the authors argue that computer screens are the biggest privacy concerns for lifeloggers since they spend most of their time in front of their computer screens and this could result in the leakage of sensitive and private information such as entered passwords, emails, financial records, pictures of bystanders etc. They also raise an important problem with modern apps such as Google Photos and Lens that upload the captured feed to the cloud automatically without the user's consent which might contain sensitive data. [2] For countering this, the authors have employed a solution using Computer Vision to detect computer monitors and Convolutional Neural Network to detect sensitive contents in these identified computer monitors.

Issues:

Usability: Although the proposed solution is highly accurate in identifying computer monitors in lifelogs they failed to demonstrate the implication of this technique as an end product. There is no next step on what users should do after they identify feeds with sensitive content such as obfuscation or image filtering or restricting access to certain users.

Security: The authors have not considered that Computer Vision can also be applied to reverse engineer and extract the original content of images. The research study is only limited to computer monitors but statistics suggests that people spend more time on their smartphones than computers or laptops. [3]

Privacy: The scope of this research doesn't incorporate smartphone or TV screens which could also leak sensitive information. Additionally, lots of bystanders are captured by lifeloggers and this paper does not suggest any means of protecting their privacy. While classifying sensitive contents among apps such as Facebook and Gmail, they failed to achieve significant accuracy which could lead to serious trust issues among users towards the reliability of the classifier. The experiment is conducted on a very small dataset and it might fail when dealing with exponential data in the real world. The authors also found a huge number of false positives such as physical windows, photos, and mirrors further questioning the reliability of the privacy-protecting technique.

Suggested Improvements:

1) User experience should be enhanced by suggesting and evaluating techniques such as image obfuscation, image filtering, and access control mechanisms.

2) The study should be extended to the detection of smartphone and TV screens.

3) The classifier should be applied to a bigger dataset so that the classification of data among sensitive apps such as Facebook and Gmail could be achieved with higher accuracy.

4) The CNN should be trained on different datasets to identify objects and bystanders in order to build a more robust tool for the protection of privacy.

5) More research should be done to protect the proposed technique from advanced reverse engineering methods based on Computer Vision.

b) Paper 2: Blur vs. Block: Investigating the Effectiveness of Privacy-Enhancing Obfuscation for Images. [6]

Summary: In this paper, the authors discuss how Computer Vision and Deep Learning could reveal sensitive information from images such as object extraction and learn relationships between persons with high precision. This attacks the privacy of individuals and hence they present two approaches to protect privacy – 1) Controlling access to an image
2) Controlling image content
The authors found that the first approach is not very effective and therefore they focused on the second approach for the research. For people who do not wish to have their pictures uploaded on social media, the proposed solution employs two image obfuscation techniques – Blocking and Blurring. The paper analyzes the effectiveness of the above-mentioned image obfuscation techniques in preserving privacy. Moreover, the authors also present viewer perception tests which are image satisfaction, information sufficiency, image enjoyment, social presence, and obfuscation likeability. 53 participants were hired from Amazon Mechanical Turk for this experiment. The authors found that Blocking is a better obfuscation technique than blurring for protection against human and machine identification. However, blocking performed very poorly in all of the viewer's perception tests when compared to blurring and was disliked by users.

Issues:

Usability: The paper declares blocking as the more superior technique for image obfuscation but most users disliked blocking. This is a typical example of security – usability tradeoff. This was because of numerous reasons – 1) Users were not satisfied with Blocking because it reduces the aesthetic and integrity of photos. 2) Blocking reduces the information in a picture because it can also cover adjacent people. 3) Blocking lowers the level of human contact in pictures which may lead to the reduction of users participation and motivation to continue using the medium. 4) Participants of the user study were unfamiliar with people in the pictures which does not present a very likely scenario. 5) People in the pictures were from different races and hence participants have a better chance of identifying someone from their own race.

Security: Blurring is very popular but it is reversible and can lead to reidentification by neural networks. Even though security provided by blocking is a lot better than blurring, users are reluctant to use it and may circle back to a less secure medium such as blurring.

Privacy: Authors have only considered obfuscation of people, not objects or visual textual private information. This could have serious implications since they can also reveal sensitive data and compromise privacy. They used lower intensity blurring which may lead to higher identification success for adversaries.

Suggested Improvements:

1) Obfuscation of objects and visual textual information should be implemented.

2) High-Intensity blurring should be used for image obfuscation.

3) People in the pictures should be a mixture of familiar and unfamiliar people.

4) Square blurring should be used instead of blurring along the body shape as it can still give away information from the person's body shape.

5) Alternative blurring techniques should be advised which provide both good security and usability.

c) Paper 3: Effectiveness and Users' Experience of Obfuscation as a Privacy-Enhancing Technology for Sharing Photos. [7]

Summary: In this paper, the authors have investigated eight image obfuscation techniques and their effectiveness both in terms of security and usability for image privacy. The obfuscation techniques were applied to both bodies and face separately for comparison. They have also debated the obfuscation timings such as at the time of upload, at time of capture etc. It was found that despite popular belief pixelating and blurring are the most ineffective obfuscation techniques. Avatar and Inpainting are the most successful technique both in terms of security and user satisfaction. Lastly, they argue that it is most suitable to obfuscate the images before uploading to social media so that no one can access the raw image data.

Issues:

Usability: Although it was found that inpainting and avatar are the most secure and usable techniques, a huge amount of users rated insecure techniques such as pixelating and blurring very highly in viewers perception tests. Blurring is still very popular that even Google Street view and Youtube uses it.

Security: Other information than face or body in an image can reveal contextual cues leaking sensitive information. Blurring is very popular techniques among users and it was found that it could be reversed by AI and machine learning techniques.

Privacy: Avatar and Inpainting achieved the intended success however they are still very unpopular among users when compared to inefficient techniques such as blurring and pixelating. Face obfuscation can very easily reveal sensitive information when compared to body obfuscation. Even when using highly effective obfuscation techniques an image can still reveal information about the activity and the environment that a person is engaged in.

Suggested Improvements:

1) People should be made aware of highly effective obfuscation techniques and the implications of using less effective ones.

2) Contextual cues which might possibly reveal sensitive information should be hidden along with face and body.

3) Advancement of AI and machine learning techniques as countermeasures against obfuscation should be considered thoroughly and the proposed techniques for obfuscation should not be reversible.

4) Other than concealing the body and face, techniques to change the activity and the environment in the image. For example, changing whiskey to coke or a bar to a home party should be investigated.

Conclusion

Image obfuscation is an important step in privacy protection. However, the researchers are lacking in building strong threat models against sophisticated adversaries such as AI and machine learning which could very easily reverse obfuscation techniques. [4] Since, humans are the most important asset, security and usability need to work together. Image obfuscation techniques can prove to be very effective in privacy protection and should be applied to both people and objects.

References (IEEE style):

[1][Online]. Available: https://blog.digi.me/2014/12/10/research-insights-one-in-five-has-left-social-media-services/percentage-of-people-considering-leaving-online-services-due-to-privacy/. [Accessed: 24- Feb- 2019].

[2]S. Hill, "Google Photos: Should you be worried about privacy?", Android Authority, 2015. [Online]. Available: https://www.androidauthority.com/google-photos-worried-privacy-616339/. [Accessed: 24- Feb- 2019].

[3]"Smartphone internet usage surpasses desktop and laptop for first time in the UK | Mobile Marketing Magazine", Mobilemarketingmagazine.com. [Online]. Available: https://mobilemarketingmagazine.com/smartphone-internet-usage-surpasses-desktop-laptop-uk-emarketer. [Accessed: 24- Feb- 2019].

[4]R. McPherson, A. Houmansadr and V. Shmatikov, "Defeating Image Obfuscation with Deep Learning", Published in ArXiv 2016, vol. 2016, no. 3, pp. 212-225, 2016.

[5]M. Korayem, R. Templeman, D. Chen, D. Crandall and A. Kapadia, "Enhancing Lifelogging Privacy by Detecting Screens", Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, vol. 2016, pp. 4309-4314, 2016. Available: 10.1145/2858036.2858417.

[6]Y. Li, N. Vishwamitra, B. Knijnenburg, H. Hu and K. Caine, "Blur vs. Block: Investigating the Effectiveness of Privacy-Enhancing Obfuscation for Images", 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017. Available: 10.1109/cvprw.2017.176 [Accessed 24 February 2019].

[7]Y. Li, N. Vishwamitra, B. Knijnenburg, H. Hu and K. Caine, "Effectiveness and Users' Experience of Obfuscation as a Privacy-Enhancing Technology for Sharing Photos", Proceedings of the ACM on Human-Computer Interaction, vol. 1, no., pp. 1-24, 2017. Available: 10.1145/3134702 [Accessed 24 February 2019].