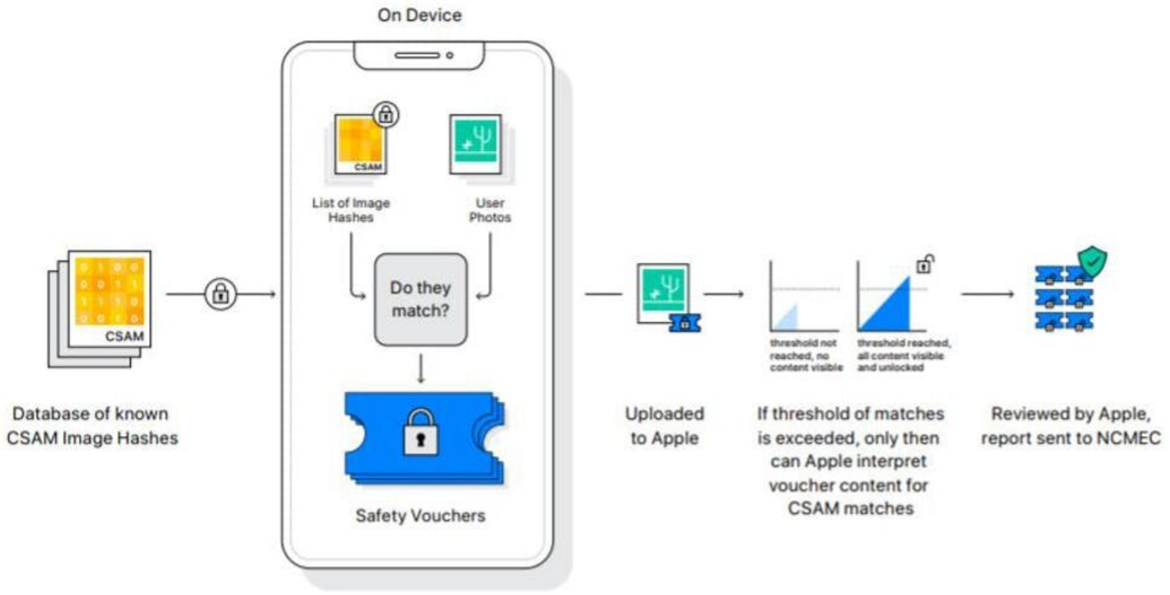
TechXplore

Apple can scan your photos for child abuse and still protect your privacy

August 11 2021, by Mayank Varia


On Device
CSAM
List of Image Hashes
User Photos
Do they match?
Safety Vouchers
CSAM
Database of known CSAM Image Hashes
Uploaded to Apple
threshold not reached, no content visible
threshold reached, all content visible and unlocked
If threshold of matches is exceeded, only then can Apple interpret voucher content for CSAM matches
Reviewed by Apple, report sent to NCMEC

Apple's new system for comparing your photos with a database of known images of child abuse works on your device rather than on a server. Credit: Apple

The proliferation of child sexual abuse material on the internet is harrowing and sobering. Technology companies send tens of millions of reports per year of these images to the nonprofit National Center for Missing and Exploited Children.

The way companies that provide cloud storage for your images usually detect child abuse material leaves you vulnerable to privacy violations by the companies—and hackers who break into their computers. On Aug. 5, 2021, Apple announced a new way to detect this material that promises to better protect your privacy.

As a computer scientist who studies cryptography, I can explain how Apple's system works, why it's an improvement, and why Apple needs to do more.

Who holds the key?

Digital files can be protected in a sort of virtual lockbox via encryption, which garbles a file so that it can be revealed, or decrypted, only by someone holding a secret key. Encryption is one of the best tools for protecting personal information as it traverses the internet.

Can a cloud service provider detect child abuse material if the photos are garbled using encryption? It depends on who holds the secret key.

Many cloud providers, including Apple, keep a copy of the secret key so they can assist you in data recovery if you forget your password. With the key, the provider can also match photos stored on the cloud against known child abuse images held by the National Center for Missing and Exploited Children.

But this convenience comes at a big cost. A cloud provider that stores secret keys might abuse its access to your data or fall prey to a data breach.

A better approach to online safety is end-to-end encryption, in which the secret key is stored only on your own computer, phone or tablet. In this case, the provider cannot decrypt your photos. Apple's answer to checking for child abuse material that's protected by end-to-end encryption is a new procedure in which the cloud service provider, meaning Apple, and your device perform the image matching together.

Spotting evidence without looking at it

Though that might sound like magic, with modern cryptography it's actually possible to work with data that you cannot see. I have contributed to projects that use cryptography to measure the gender wage gap without learning anyone's salary, and to detect repeat offenders of sexual assault without reading any victim's report. And there are many more examples of companies and governments using cryptographically protected computing to provide services while safeguarding the underlying data.

Apple's proposed image matching on iCloud Photos uses cryptographically protected computing to scan photos without seeing them. It's based on a tool called private set intersection that has been studied by cryptographers since the 1980s. This tool allows two people to discover files that they have in common while hiding the rest.

Here's how the image matching works. Apple distributes to everyone's iPhone, iPad and Mac a database containing indecipherable encodings of known child abuse images. For each photo that you upload to iCloud, your device applies a digital fingerprint, called NeuralHash. The fingerprinting works even if someone makes small changes in a photo. Your device then creates a voucher for your photo that your device can't understand, but that tells the server whether the uploaded photo matches child abuse material in the database.

If enough vouchers from a device indicate matches to known child abuse images, the server learns the secret keys to decrypt all of the matching photos—but not the keys for other photos. Otherwise, the server cannot view any of your photos.

Having this matching procedure take place on your device can be better for your privacy than the previous methods, in which the matching takes place on a server—if it's deployed properly. But that's a big caveat.

Figuring out what could go wrong

There's a line in the movie "Apollo 13" in which Gene Kranz, played by Ed Harris, proclaims, "I don't care what anything was designed to do. I care about what it can do!" Apple's phone scanning technology is designed to protect privacy. Computer security and tech policy experts are trained to discover ways that a technology can be used, misused and abused, regardless of its creator's intent. However, Apple's announcement lacks information to analyze essential components, so it is not possible to evaluate the safety of its new system.

Security researchers need to see Apple's code to validate that the device-assisted matching software is faithful to the design and doesn't introduce errors. Researchers also must test whether it's possible to fool Apple's NeuralHash algorithm into changing fingerprints by making imperceptible changes to a photo.

It's also important for Apple to develop an auditing policy to hold the company accountable for matching only child abuse images. The threat of mission creep was a risk even with server-based matching. The good news is that matching devices offers new opportunities to audit Apple's actions because the encoded database binds Apple to a specific image set. Apple should allow everyone to check that they've received the same encoded database and third-party auditors to validate the images contained in this set. These public accountability goals can be achieved using cryptography.

Apple's proposed image-matching technology has the potential to improve digital privacy and child safety, especially if Apple follows this move by giving iCloud end-to-end encryption. But no technology on its own can fully answer complex social problems. All options for how to use encryption and image scanning have delicate, nuanced effects on society.

These delicate questions require time and space to reason through potential consequences of even well-intentioned actions before deploying them, through dialog with affected groups and researchers with a wide variety of backgrounds. I urge Apple to join this dialog so that the research community can collectively improve the safety and accountability of this new technology.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation

Citation: Apple can scan your photos for child abuse and still protect your privacy (2021, August 11) retrieved 17 April 2024 from https://techxplore.com/news/2021-08-apple-scan-photos-child-abuse.html