To FLoC or Not?

Marc Langheinrich, *Università della Svizzera italiana, 6900, Lugano, Switzerland*

A large online advertisement company recently seems to have finally squared the circle: personalized advertisements without tracking cookies—code name "FLoC"! Is this too good to be true, and can (should?) this serve as a model for IoT privacy?

After years in the making, Google, the company behind the most popular browser software and the undisputed leader in online advertisement, finally, announced in March this year that its Chrome browser would soon stop accepting "third-party cookies." Third-party cookies are placed by ad networks to track consumer browsing behavior across multiple websites. By blocking these, Chrome will prevent these networks from building up comprehensive profiles of individual users, and subsequently sell targeted advertisements based on these profiles. Given that Google itself is running one of the biggest ad networks in the world, this move seems counterintuitive at first. Yet, Google thinks it has found a way to continue selling highly efficient (i.e., targeted) ads without the need for such tracking profiles. The solution is called "Federated Learning of Cohorts"—FLoC.

FLoC works like this: instead of accepting and subsequently replaying a tracking cookie that would allow an ad network to track one's "moves" across different websites (and thus using this to build up a personalized profile of a user), a FLoC-enabled browser instead *locally* learns a user's browsing behavior, and then classifies this behavior into one of a reasonably large set of user categories ("cohorts"). This learning is done in a distributed ("federated") fashion, without the need for a central server that might hold a user's personal data. Sites can, in turn, declare their interest in knowing this classification of their visitors, which the browser will then send along with each request (instead of the usual set of tracking cookies).

TRACK RECORD?

Federated learning, the technology underlying the FLoC approach, is a fascinating way to enable collaborative learning among multiple parties without disclosing one's individual data. Google is already using it in several of its products, most notably its Android keyboard app Gboard (to learn personalized word predictions). By using FLoC, Google is indeed ending third-party tracking cookies. Does that mean that privacy concerns over personalized advertisements are over?

Many consumer rights groups and civil liberties organizations worldwide have taken up the fight against FLoC technology. Several browser vendors, even those based on Chromium, have vowed to not support FLoC in their browsers. They all point out that FLoC fails to solve the fundamental issue in cookie-based tracking: behavioral classification. Yes, a FLoC-enabled browser *does* prevent third-party tracking—only to replace it with "first-party" tracking, i.e., the browser itself will continuously observe and classify one's surfing behavior and then notify every interested website of its findings.

While Google claims that FLoC will never generate cohorts with less than a few thousand users, this still allows for millions of user categories (current estimates put the number of Chrome users at over 2.5 billion), hinting at the level of detail that each category would convey about its members.

MODEL FOR THE IOT?

Federated learning promises to offer all the benefits of machine learning without the drawbacks of centralized data collections. With ever more applications in need of the "smarts" that only huge datasets can offer, being able to distribute the model learning across millions of clients might not only save bandwidth and speed up model learning but should also significantly improve personal privacy.

All this assumes a certain concept of privacy that is tightly bound to "personal data": only if one uses data that can identify an individual (even if only potentially), a privacy risk arises. This has always been the legal view of privacy (or "data protection"), and most privacy laws, such as the EU's GDPR or California's CCPA and CPRA, only apply to

1536-1268 © 2021 IEEE
Digital Object Identifier 10.1109/MPRV.2021.3076812
Date of current version 25 May 2021.

data that can somehow be attributed to an individual. Once data have been (sufficiently) anonymized, no privacy law applies.

Privacy scholars have started to question if data can ever truly be anonymous (while still offering useful information), motivated by the famous deanonymization work by Latanya Sweeny who in the mid-1990s identified then-governor of Massachusetts Bill Weld in an anonymized medical database by cross-referencing it with publicly available voter records. Since then, several "anonymized" datasets have been deanonymized, such as the AOL search queries dataset of 2006, the Netflix Prize dataset of 2007, or the NYC Taxi dataset of 2010–2013.

Also, today's ad networks have already developed user classification techniques that do not require any sort of cookie-based tracking but instead the "browser fingerprint"—a short description of the browser version and OS, its plugins, and preferences (e.g., languages), that browsers typically send along with their requests in order to allow servers to select the right version of a web page—to instantly classify website visitors. The entire process is devoid of any personal data, yet allows highly precise user categorization, including the prediction of, e.g., the user's age, gender, ethnicity, and education level.

In the context of an IoT environment, FLoC might run on the user's phone or smart watch, enabling these devices to announce the user's recent behavior (in the form of a FLoC cohort ID) to a smart environment. With the help of such a FLoC ID, the environment may adjust information provisioning on ambient displays, or inform its operators of typical behavioral patterns among smart city users. This clearly would be better than the active tracking of users with the help of, say, cameras or short-range wireless sensing. However, *how much* better?

WHAT'S IN A WORD?

Privacy concerns are often motivated by concrete risks that arise from a lack of privacy protection. A large database containing user data might become the target of a hacker break-in and thus make affected users vulnerable to identity theft, phising, and other cybercrimes. Another risk may be what David Lyon calls "social sorting," the act of social stratification by assigning both risk and worth to an individual. While an approach like FLoC will surely minimize the risk of data disclosure (as no personal data leave the user's device), it does little (by design) to address social sorting concerns. Whether with the help of cookies or fully decentralized: the classification of website visitors is the end-goal of both approaches. Would the technical details really make a difference to the actual experience of a website visitor being served an ad matching their recent browsing behavior in the past days and weeks?

FLoC is hardly the first attempt to "solve" the privacy problem, nor will it be the last. Limiting the amount of personal data that leave a user's devices is clearly a good thing. Yet keeping the overall process of user classification in place does not fundamentally change the problem of telling every site (or ad network) you visit just what kind of person you are. Before we can adopt this for the IoT, we will need to fundamentally revisit the concept of personalized advertisement so that it cannot be used to exploit, discriminate, or harm people.

Federated learning seems like a much-needed way to enable truly smart edge computing. Cohorts we might do without.

IN THIS ISSUE

The title of this Special Issue is "Pervasive Video and Audio," and Guest Editors Fahim Kawsar, Romit Roy Choudhury, and Ganesh Ananthanarayanan present four articles that offer insights into key challenges of putting cameras and microphones everywhere. You can find more details in their Guest Editors' Introduction later in this issue.

We also have four departments in this issue: Conferences, Education & Training, Smart Home, and Wearable Computing.

The Conferences department features a report by Jonathan Liebers and Roman Heger on the 19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020). Originally slated to take place in Essen, Germany, it was converted to a virtual event using a map-based meeting tool that tried to mimic the originally envisioned venue.

The Education and Training department again offers an insight into learning in times of COVID-19: authors Orit Shaer and Diana Tosca share their experience teaching tangible interaction remotely during COVID-19. It is another impressive example on how something so inherently physical (i.e., prototyping) can still be taught effectively even in an online context!

In our Smart Home department, author Mike Hazas investigates whether autonomous household robots (e.g., vacuums and mowers) help to save energy. While weekly figures do indeed show savings over the "dumb" (i.e., manual) alternatives, Hazas points out that the lifetime of those smart counterparts might not only be lower to begin with, but might

also be cut short prematurely, courtesy of a manufacturer ending their online support of a model.

Finally, our Wearable Computing department features a report by authors Melanie Baldinger, Axel Heinrich, Tim Adams, Eimo Martens, Michael Dommasch, Alexander Muller, Adonia Siegmann, and Georg Schmidt on using in-ear wearables to monitor vital signs of COVID-19 patients at home.

This issue also contains two feature articles. In "Beyond the Star of Stars: An Introduction to Multi-Hop and Mesh for LoRa and LoRaWAN," authors Roger PueyoCentelles, Felix Freitag, Roc Meseguer, and Leandro Navarro review the state of the art of the design and implementation of multihop and mesh solutions for LoRa and LoRaWAN, and identify the challenges and opportunities this technology can offer to IoT deployments.

In the very timely "Smart Multi-Modal Telehealth IoT System for COVID-19 Patients," authors Lloyd Emokpae, Wassila Lalouani, Mohamed Younis, and Roland Emokpae describe the development of a wearable shirt (vest) for diagnosing COVID-19 related symptoms and similar contagious diseases.

TEAM UPDATES

In this issue, we say goodbye to our Editorial Board member Yoshihiro Kawahara. Yoshihiro has been on the Editorial Board for over six years and has been instrumental in several magazine activities. I hope that we will be able to draw on his continued support when it comes to promoting our publication. Thank you!

I am excited that two of our Editorial Board Members recently accepted my invitation to become Associate Editor-in-Chief (AEIC): Dr. Claudio Pinhanez and Dr. Oliver Amft! Claudio and Oliver augment our AEIC team after long-term board member Nicholas Lane stepped down in February. I very much appreciate Nic's continued support of the magazine over the years and wish him all the best at his new place of work, Cambridge University!

Claudio is with IBM Research, Brazil, where he works on next-generation conversational systems. He holds a PhD degree from the MIT Media Laboratory. You can contact Claudio at claudio@pinhanez.com. Oliver holds the Chair for Digital Health at the Friedrich Alexander University Erlangen-Nurnberg (FAU), Germany. He received the PhD degree from ETH Zürich. You can contact Oliver at oliver.amft@fau.de.

The current AEICs and their areas of expertise are thus as follows:

- Oliver Amft: Pervasive Health and Wearable Computing
- Stephen Intille: Pervasive Health (Departments Editor)
- Fahim Kawsar: Smart Homes, Enterprises, and Cities (Community Efforts)
- Mirco Musolesi: Mobile and Sensing Systems
- Florian Michahelles: Internet of Things
- Claudio Pinhanez: AI-based Pervasive Systems
- Junehwa Song: Social and Culture Computing

MARC LANGHEINRICH is currently a professor with the Faculty of Informatics, Università della Svizzera Italiana, Lugano, Switzerland, where he heads the Research Group for Ubiquitous Computing. His main research interests include usable privacy, pervasive displays, and ubiquitous computing. He received a PhD degree in computer science from ETH Zürich. Contact him at langheinrich@ieee.org.