Comments on Competition and Consumer Protection in the 21st Century Hearings,

Project Number P181201[1]

To Whom It May Concern:

We are writing in response to the FTC's solicitation for public comments regarding the upcoming hearings on competition and consumer protection.

We are writing this letter to share our personal experiences and perspectives on the current state of consumer protection with respect to the privacy risks caused by **1)** third-party tracking services embedded in mobile apps, **2)** pre-installed Android software installed by the OS vendor, **3)** commercial VPN services, and **4)** intrusive web-based user tracking.

For the last two years, our research activity focused on studying these technologies in depth due to their broad user adoption, their opacity and their potential harm to customer's privacy and security.

As detailed below, this research experience has allowed us and our collaborators to find empirical evidence of concerning and deceptive behaviors in these ecosystems, as well as instances of personal data collection performed without user consent. At the core of the problem is user's poor understanding of the underlying technologies and the lack of transparency of these products. Most users are not equipped with the required technical background and tools needed to foresight the long-lasting and severe impact that these products can cause on their privacy.

THE THIRD-PARTY ADVERTISING AND TRACKING ECOSYSTEM IN ANDROID

Mobile technologies have consolidated the data-driven economy present in the Web. Modern smartphones are ubiquitous, and they store a plethora of user and contextual data. These features make smartphones a perfect platform for collecting personal and contextual information from mobile users. This data is typically used for online advertising.

Over the past two years, Prof. Vallina-Rodriguez and his collaborators have investigated the third-party service ecosystem for mobile apps. As with the Web, many mobile app developers integrate third-party services (typically in the form of SDKs) in their apps for a variety of purposes including app maintenance (i.e., crash reports), analytics services, user engagement, user acquisition, A/B testing, social network integration, and advertising.

Third-party services inherit the set of application permissions requested by the host app, allowing them to access a wealth of valuable user data, often beyond what they need to provide the expected service to the app developer or the end-user. Unfortunately, app developers may get economic incentives to use a given third-party service in their apps, without verifying whether they can cause any harm to the end users of if they follow industry standards such as using TLS to upload customer data to the cloud. In fact, most app developers do not disclose the list of third-party services embedded in their apps in their privacy policy.

As a result, it is arguable whether users of a given mobile app are fully aware of the fact that, besides the app developer, an undisclosed number of third parties may actively collect sensitive data about them. Moreover, the dynamic nature of the ad-bidding and ad-mediation process impedes knowing a

[1] This research has been supported in part by the National Science Foundation under grants CNS-1351058, CNS-1409868, CNS-1405886 and DGE-1069311.

priori the whole list of third-party services eventually loaded in the apps at runtime, services which can also potentially collect sensitive data from users.

This lack of transparency leaves average mobile users with no guarantees to know which third-party services actively collect their data and for what purpose. Most third-party services operate in the background and do not provide any visual clues inside the apps about their presence: they can track users without their knowledge or consent over space and time, while remaining virtually invisible.

The general lack of transparency in mobile systems leaves users unable to identify the third-party services embedded by their apps, let alone know to which extent these services can collect, correlate, and aggregate their personal data and online activity across apps, devices, and platforms, and if they further share (or sell) it with other third parties, including affiliated advertising services and even data brokers.

In order to understand the ecosystem of third-party advertising and tracking service and its actors, we develop an automated approach that leverages traffic traces collected from over 14,599 mobile applications. We could find over 2,000 advertising and tracking domains of which 233 were previously unreported in popular domain blacklists such as EasyList or hpHosts. We see that large companies such as Alphabet Inc. and its subsidiaries, if aggregated, are present in over 73% of apps in our dataset.

Malicious third-party services also implement deceptive mechanisms to harvest sensitive data using side channels. We have found third-party SDKs bypassing Android's permission model to collect protected sensitive data without user consent, including the unprotected "getprop" command to obtain unique user IDs such as the device MAC Address. Other SDKs link persistent user identifiers such as the IMEI with resettable ones such as the Android ID, hence allowing them to perform longitudinal user monitoring and impeding any privacy benefit of a resettable ID. This practice is in fact against Google's Terms of Service.

Advertising and tracking services seek new mechanisms to track users across devices and platforms (i.e., cross-device tracking). We find that there is a high proliferation of cross-device advertising and tracking services, with 39% of our tracking domains being also present in the top 1000 most popular websites according to Alexa: 17 of the top-20 largest third parties have a presence both on the Web and in the mobile ecosystem.

Our analysis of the privacy policies of the most dominant mobile advertising and tracking organizations (aggregated, they are present in over 80% of all apps in our study) reveals that eight of the top-10 organizations reserve the right to sell or share data with other parties, while all of them reserve the right to share data with their subsidiaries. These findings demonstrate that a small number of companies have a privileged market position by controlling a significant portion of the ecosystem and that they can track users and share the tracking data with other entities, all with little to no transparency.

Full Paper:

1. **Apps, Trackers, Privacy, and Regulators: A Global Study of the Mobile Tracking Ecosystem**
   A. Razaghpanah, R. Nithyanand, N. Vallina-Rodriguez, S. Sundaresan, M. Allman, C. Kreibich, P. Gill NDSS Symposium 2018

PRE-INSTALLED ANDROID SERVICES

Thanks to the openness of the Android ecosystem, mobile device vendors can build and sell smartphones and other mobile devices with their custom versions of Android. Most of these custom versions deviate significantly from Google’s official Android Open Source Project (AOSP): in addition to various visual and functional changes to the base OS, vendors add proprietary apps to their firmware, and sometimes even add custom (often unknown) certificates to the system’s root certificate store.

Pre-installed services have privileged system-level access to user's information. However, these apps have been overlooked by the research community as this software is typically unavailable in app stores such as Google Play. By running with system-level privileges, hardware vendors can build detailed customer dossiers thanks to their ability to track users longitudinally across networks, apps, and devices.

In this preliminary study, we manually analyzed 15 different Android devices using static and dynamic analysis whenever it was possible. We covered eight different manufacturers including both high-end and low-end vendors like Samsung and Wiko, respectively. None of the over 1,000 pre-installed APKs present in these devices was listed on Google Play.

Our empirical analysis suggests that all these vendors actively collect user information. However, only one of them informed users during the first launch about such practices, even mentioning that minors under the age of 13 should not use the device as they are not COPPA compliant. This suggests that most Android users are potentially tracked, without explicit consent, by hardware manufacturer and also by the third-party services embedded in pre-installed apps for unknown purposes. Furthermore, we have also identified custom permissions defined by Android OS vendors which may allow third-party app developers to bypass Android's official permissions by entering into partnership agreements with the vendor.

Full Paper:

1. **An Analysis of Pre-Installed Android Software (Extended abstract)**
   J. Gamba, M. Rashed, A. Razaghpanah, N. Vallina-Rodriguez, J. Tapiador, JNIC 2018

COMMERCIAL VPN SERVICES

VPN usage has grown dramatically in recent years. According to a recent market research report, commercial VPNs are currently a 15-billion dollar[2] industry expected to increase 20% by 2022. Originally developed as a technology to privately send and receive data across public networks, VPNs are now marketed broadly as a privacy-preserving technology that allows Internet users to obscure not only their traffic but also their personal information, such as their web browsing history, from third parties including Internet service providers (ISPs) and governments alike.

Unfortunately, while many services claim to operate a robust and secure VPN infrastructure and ensure users' privacy by not logging data, the reality is the VPN ecosystem is highly opaque. The lack of practical tools or independent research that systematically audits these security and privacy claims, especially concerning information leakage and traffic manipulation, aggravates the state of affairs. Anecdotally, some VPN providers are known for selling customers' data to third-party data brokers or manipulating customer traffic. Furthermore, our previous study of the Android VPN ecosystem revealed that multiple mobile VPN providers acted maliciously by redirecting traffic to affiliate companies and injecting JavaScript for tracking purposes without user awareness. Our research findings were followed by investigations by the Center for Democracy and Technology on Hotspot Shield[3], a VPN provider allegedly intercepting and manipulating user traffic.

Unfortunately, VPN users have limited means to verify the privacy, security even geographical presence claims made by VPN services. The absence of peer-reviewed evaluation of VPN services leaves privacy- and security-conscious users with little choice but to resort to blogs, word of mouth, and review websites when shopping for VPN services. Unfortunately, we find that many of these websites are supported by affiliate programs, suggesting they are unlikely to provide users with an unbiased opinion of the technology.

---

[2] https://www.statista.com/statistics/542797/worldwide-virtual-private-network-market-by-type

[3] https://cdt.org/files/2017/08/FTC-CDT-VPN-complaint-8-7-17.pdf

To bridge this gap, we performed an exhaustive study of 200 popular commercial VPN services. In addition to widespread issues with transparency and marketing, and systematic traffic leakage due to inadequate security defaults, we observed approximately 10% of VPNs that we tested intercepting and manipulating user traffic. It is important to note that purely passive monitoring is difficult to detect from the standpoint of an end user: hence, our findings regarding traffic tampering and monitoring are at best a lower bound on its actual prevalence.

In addition to factors that compromise user privacy, we also find that many VPNs fail to deliver on their promises of geographical diversity. VPN providers frequently advertise that their service can make user traffic appear to originate from a selection of distinct vantage points, typically spread across different countries. Our study shows that the practice of `virtualizing' vantage points—i.e., physically placing VPN endpoints in one (presumably more accessible) country and then working to trick IP geolocation services into believing that the vantage point is in another country—is far more prevalent than physically distributing VPN endpoints. We find that 10% of the providers in our study appear to misrepresent the location of one or more of their vantage points, with between 5% and 30% of all vantage points located in a different country than advertised. In the most extreme case, we find a VPN provider claiming vantage points in more than 190 nations yet hosting servers in what appears to be fewer than ten distinct data centers.

While we did observe one instance of traffic manipulation, its sole purpose was to incentivize users to pay for a subscription. This class of traffic manipulation likely provides greater revenue to the VPN provider than merely injecting ads, which was the most common injection activity in prior studies. Further, VPN providers can passively inspect all unencrypted traffic passing through the VPN, and no measurement will be able to detect that; hence, our findings are likely an under-count. The most significant indication, if circumstantial, of questionable behavior is the ubiquitous use of affiliate marketing to advertise their services. Because of this, much of the information available online regarding the relative quality of VPN services is dominated by publishers participating in affiliate programs, honest evaluations of these services are hard to come by.

Finally, we find that VPN providers of all stripes often have poor default configurations, resulting in unintentional data leakage—especially in the case of a tunnel connection failure. Even in 2018, using a VPN `safely' remains a task mostly beyond the amateur user, and no VPN, at least that we were able to find, is perfect.

Full Papers:

1. **An Empirical Analysis of the Commercial VPN Ecosystem**
   M. T. Khan, J. DeBlasio, G. M. Voelker, A. C. Snoeren, C. Kanich, N. Vallina-Rodriguez
   (To appear) ACM Internet Measurements Conference 2018
   Pre-print version can be shared under request.
2. **An Analysis of the Privacy and Security Risks of Android VPN Permission-enabled Apps**
   M. Ikram, N. Vallina-Rodriguez, S. Seneviratne, M.A. Kaafar, V. Paxson
   ACM Internet Measurements Conference 2016

INTRUSIVE WEB-BASED USER TRACKING

While originally designed to deliver inert hypertext documents, the web has evolved into a system which is used both for mass media style communication and cutting-edge software delivery. While this “one size fits all” communication medium has enabled incredibly rapid innovation, this software delivery model, when coupled with modern web browsers, gives web authors the capability to run incredibly powerful and invasive software on users’ machines without their knowledge or consent. This software can be used to track users against their will using behavioral and device-based fingerprinting.

Our research in this space has focused on a holistic view of the relationship between the capabilities provided by the web browser and the use of those features to provide experiences which are actually desired by the end user. Through a novel website functionality measurement technique, we were able to infer to what extent the features of the web browser are used to deliver desirable features, vs. being used primarily in service of delivering ads and tracking code. We found that in many cases, features which can be used to subvert user privacy or even take over their machine was rarely used to deliver meaningful experiences: these features are available by default to every website, and in many instances, they are only used to create a nuisance or privacy invasion for the end user.

While the economic ecosystem of the web is currently enabled by tracking and advertising, the automatic execution of potentially dangerous code by default in all major browsers presents a significant risk to user privacy and security. We have built a browser extension which can disable many of these features by default with little or no negative impact for the user experience, but further cooperation from both the browser vendors and website authors is needed if we are to fully respect autonomy and privacy on the web.

Full papers:

1. **Browser Feature Usage on the Modern Web**
   P. Snyder, L. Ansari, C. Taylor, C. Kanich
   ACM Internet Measurement Conference 2016
2. **Most Websites Don't Need to Vibrate: A Cost–Benefit Approach to Improving Browser Security**
   P. Snyder, C. Taylor, and C. Kanich
   ACM Computer and Communications Security Conference 2017

CONCLUSIONS AND RECOMMENDATIONS:

Mobile applications, VPNs, and websites all expose users to deceptive actors who may have economic incentives to collect, stockpile, analyze, and monetize massive volumes of personal data. Typically, these goals are achieved by deceptive means that take advantage of users' limited technical knowledge and the lack of clear indicators of data flow and retention presented by applications and operating systems.

Besides technical solutions such as re-designing mobile application permissions and building auditing tools, regulation is necessary to address the market failures described above. In this section, we provide a concise and realistic list of regulatory and policy actions which could help to protect mobile, VPN, and web customers from deceptive and predatory practices.

- **Regulate affiliate programs.** Deceptive VPN providers and misleading app developers commonly implement affiliate programs to promote themselves in the market and gain an advantageous position over their competitors. While not illegal in and of itself, this approach is ripe with abuse in practice. It is imperative that the FTC reevaluate its interpretation of "Truth in Advertising" laws to investigate deceptive practices and protect users who are unable to distinguish legitimate reviews from biased and incentivized ones. Likewise, it may be necessary to identify and take legal actions against VPN providers and app developers with false claims about their infrastructure or security guarantees.

- **Enforce transparency.** Consumers do not have access to meaningful and accurate information about the technology that they use, their security guarantees, the organizations collecting personal data, and the purpose of such data collection. Moreover, the average user has a limited technical background so that they cannot fully understand the potential harms of using a particular technology and foresight the associated long-term privacy risks. As a result, the majority of users are unable to make informed choices concerning their privacy, or even when shopping for products such as VPN services and mobile apps. The new regulatory framework in the European Union, i.e., the GDPR Directive, has made positive steps towards this goal by defining a concise and clear set of privacy principles. It may be possible to take advantage of the positive inertia of this European regulatory

effort to bring transparency to mobile applications and VPN services alike. A side effect of a stricter regulatory framework it will provide sufficient incentives for promoting a privacy-by-design culture across developers and service providers.

- Education and digital literacy. The best solution for protecting customers is educating them. The FTC has a privileged position to reach out to interested audiences and run campaigns to educate users while engaging in regulatory activities. The combination of regulatory actions with these outreach campaigns will be more effective to protect users than isolated regulatory efforts. Especially, by targeting the most vulnerable populations such as teenagers and minors.

Signed,

Narseo Vallina-Rodriguez

Prof. Vallina-Rodriguez is a research assistant professor and director of the Internet Analytics Group at IMDEA Networks Institute, a public research institute based in Madrid, Spain. He is also a research scientist at the International Computer Science Institute (ICSI), a non-profit research institute affiliated with the University of California, Berkeley.

Chris Kanich

Prof. Kanich is an assistant professor of Computer Science at the University of Illinois at Chicago.