TrackerControl: Transparency and Choice around App Tracking

Konrad Kollnig and Nigel Shadbolt

1 Department of Computer Science, University of Oxford

Summary

Third-party tracking allows companies to collect users' behavioural data, track their activity across digital devices, and potentially share this data with third-party companies. This can put deep insights into private lives into the hands of strangers, and often happens without the awareness of end-users. In light of this, we have developed TrackerControl, which aims to provide interested individuals with real-time evidence of tracking.

Statement of need

In the past, the analysis of app tracking often used off-device network analysis, e.g., with Charles Proxy or mitmproxy (Kollnig, Shuba, Binns, et al., 2022; Ren et al., 2016; Van Kleek et al., 2017). Such analysis usually comes with the limitation that background system-level communication may be wrongly assigned to an app and taint the analysis results.

Other past studies used static analysis, which tries to gain insights into apps' data practices without executing them (Egele et al., 2011; Han et al., 2013). Such static analysis has enabled the analysis of apps in the millions (Binns, Lyngs, et al., 2018; Viennot et al., 2014; Wang et al., 2018), but does usually not generate evidence of actual data transmission to trackers taking place, since apps are never run.

In response, past research has developed on-device network analysis tools (Le et al., 2015; Razaghpanah et al., 2018; Shuba et al., 2018; Song & Hengartner, 2015). Most of these operated at a domain-level and provided limited insights into what companies actually receive data from end-users, or to what countries data is sent and for what purpose. Furthermore, none of the previous on-device network analysis tools had been deployed at scale and most remained research prototypes.

TrackerControl

To improve the quality of tracking analysis and make it available to a wide audience, we developed the Android app TrackerControl (TC). This app provides users with real-time evidence of app tracking. TC analyses the network traffic of other apps by establishing a local VPN on the Android phone and matching all observed network traffic against a database of known tracking domains. This allows the generation of factual evidence of what companies apps share data with, and can support research (both academic and non-academic) on app privacy.

The tracking database behind TC is a unique feature of the app. The core of this database is the X-Ray 2020 database that is the product of significant research efforts over the past years (Binns, Zhao, et al., 2018; Kollnig, Binns, Dewitte, et al., 2021; Kollnig, Shuba, Binns, et al., 2022; Van Kleek et al., 2017). This database has been created from analysing more than...
2 million Android apps. The X-Ray 2020 is complemented by the Disconnect.me database, which is the foundation for tracker blocking in Mozilla Firefox on the web. We further integrate the commonly used StevenBlack hostlist for tracking in apps, as a fallback. Overall, these databases provide information on 1) the companies behind tracking on the web and in apps, 2) the countries in which these companies are based, 3) and the purposes for which tracking is conducted (e.g., analytics or ads). The visualisation of tracking inside TC loosely follows the work by Van Kleek et al. (Van Kleek et al., 2017, 2018).

The core of TC builds on the NetGuard app, which is in active use by millions of users worldwide (Bokhorst, 2021). The high maturity of NetGuard ensures the reliability of the tracker analysis while minimising battery impact and supporting the long-term maintainability of TC. In addition to providing insights into app tracking, users of TC can also block unwanted network transmissions, which has contributed to building a vibrant community of tens of thousands of users. This community has helped make TC available in 20 languages.

**Use in past and ongoing research**

TC has facilitated and inspired academic research at the intersection of policy and privacy technology (Geier & Herrmann, 2021; Kollnig, Binns, Dewitte, et al., 2021). It has also been used by the Finnish innovation fund Sitra for its ‘Digipower investigation’ into apps’ data practices. As part of this study, leading Finnish politicians and journalists used TC to analyse the practices of Android apps. The results of this investigation will be presented at the Finnish and European Parliament over the course of 2022.

TC is also part of the PlatformControl toolkit (available at https://www.platformcontrol.org/), which forms the foundation of ongoing research efforts into app privacy as part of our research group at the University of Oxford (Kollnig, Binns, Dewitte, et al., 2021; Kollnig, Shuba, Van Kleek, et al., 2022; Kollnig, Binns, Van Kleek, et al., 2021; Kollnig, Shuba, Binns, et al., 2022, 2022).

**Screenshots**

The left-hand screenshot shows the app overview. The middle screenshot shows the observed tracking companies, domains and purposes for one app. The right-hand screenshot shows the destinations of tracking data, obtained from the contacted IP addresses.

Acknowledgements

Konrad Kollnig was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/R513295/1.

The underlying network analysis functionality is provided by the NetGuard Firewall, developed by Marcel Bokhorst (Bokhorst, 2021).

TrackerControl would not have been possible without the help of many outstanding individuals, including Max Van Kleek, Katherine Fletcher, George Chalhoub, and numerous app testers and friends.

References


