Watching TV with the Second-Party: A First Look at Automatic Content Recognition Tracking in Smart TVs

Gianluca Anselmi*
University College London
London, United Kingdom
gianluca.anselmi.22@ucl.ac.uk

Yash Vekaria*
University of California, Davis
Davis, United States
yvekaria@ucdavis.edu

Alexander D'Souza
University of California, Davis
Davis, United States
aledsouza@ucdavis.edu

Patricia Callejo
Universidad Carlos III de Madrid,
uc3m-Santander Big Data Institute
Madrid, Spain
pcallejo@it.uc3m.es

Anna Maria Mandalari
University College London
London, United Kingdom
a.mandalari@ucl.ac.uk

Zubair Shafiq
University of California, Davis
Davis, United States
zubair@ucdavis.edu

Abstract

Smart TVs implement a unique tracking approach called Automatic Content Recognition (ACR) to profile viewing activity of their users. ACR is a Shazam-like technology that works by periodically capturing the content displayed on a TV's screen and matching it against a content library to detect what content is being displayed at any given point in time. While prior research has investigated third-party tracking in the smart TV ecosystem, it has not looked into *second-party* ACR tracking that is directly conducted by the smart TV platform. In this work, we conduct a black-box audit of ACR network traffic between ACR clients on the smart TV and ACR servers. We use our auditing approach to systematically investigate whether (1) ACR tracking is agnostic to how a user watches TV (e.g., linear vs. streaming vs. HDMI), (2) privacy controls offered by smart TVs have an impact on ACR tracking, and (3) there are any differences in ACR tracking between the UK and the US. We perform a series of experiments on two major smart TV platforms: Samsung and LG. Our results show that ACR works even when the smart TV is used as a "dumb" external display, opting-out stops network traffic to ACR servers, and there are differences in how ACR works across the UK and the US.

CCS Concepts

• **Information systems** → **Content match advertising**; *Online advertising*; *Traffic analysis*; • **Social and professional topics** → **Privacy policies**; *Corporate surveillance*; • **Security and privacy** → **Privacy protections**; *Social aspects of security and privacy*.

Keywords

Smart TV, ACR, Fingerprinting, Advertising, Tracking, Privacy

---

*Both authors contributed equally to this research.

1 Introduction

Smart TVs, which can connect to the Internet and stream content, have become widely popular. The smart TV penetration has reached almost a three-fourth of households today [15, 51], with a vast majority of globally sold TVs being smart [28]. In fact, it is challenging to buy a "dumb" TV now [43]. A number of different smart TV platforms exist, led by Samsung Tizen and LG WebOS [13].

The research community has examined privacy issues in the smart TV ecosystem, particularly third-party tracking in smart TV apps [37, 58, 62, 66], but has not looked at second-party tracking conducted directly by the smart TV platform. Smart TV platforms use a unique tracking approach dubbed Automatic Content Recognition (ACR) [42] to profile viewing habits of smart TV users. Unlike traditional online tracking in the web and mobile ecosystems that is typically implemented by third-party libraries/SDKs included in websites/apps, ACR is typically directly integrated in the smart TV's operating system. At a high level, ACR works by periodically capturing the content displayed on a TV's screen and matching it against a content library to detect the content being viewed on the TV. It is essentially a Shazam-like technology for audio/video content on the smart TV [38].

ACR is implemented by all major smart TV manufacturers, including Samsung [9] and LG [55]. There has been public and regulatory scrutiny of ACR tracking. Most notably, the FTC sued Vizio and Inscape in 2017 for their use of ACR tracking in smart TVs without user consent [12]. However, prior research lacks an in-depth measurement and analysis of ACR tracking in smart TVs despite (a) being known to exist for many years, (b) its unique tracking approach as compared to web or mobile, and (c) its deployment in the vast majority of smart TVs today.

Our research aims to bridge this gap in the prior literature. While it would be ideal to study ACR tracking in a white-box setting, it requires reverse-engineering and/or jailbreaking a smart TV's operating system, which is challenging. An alternate auditing approach,

commonly used in the measurement community [26, 46], is to analyze network traffic flows between devices at the client-side and tracking endpoints at the server-side. In this work, we adopt this auditing approach to analyze the network traffic between ACR clients on the smart TV and ACR servers for two major smart TV manufacturers: Samsung and LG.

We use this auditing approach to answer three research questions about ACR tracking in smart TVs.

First, we investigate if ACR tracking is agnostic to how a user watches the TV. Specifically, we compare ACR tracking across scenarios where a user (a) watches linear TV (e.g., via antenna), (b) streams content from a smart TV platform's app such as Samsung TV Plus [48], (c) streams content from a third-party app such as Netflix, (d) uses the TV as an external display for laptop or gaming console via HDMI, (e) screen casts content via Wi-Fi from an external mobile phone or laptop, or (f) stays on the TV's homepage. ACR tracking across these scenarios raises unique concerns. For instance, ACR tracking when the TV is being used as a "dumb" external display raises privacy concerns. Similarly, ACR tracking of copyrighted third-party content raises intellectual property concerns [59]. We find that: (1) ACR network traffic exists when watching linear TV and when using smart TV as an external display using HDMI, (2) ACR network traffic is not present when streaming content from third-party apps such as Netflix and YouTube.

Second, we investigate whether privacy controls offered by smart TV manufacturers have an impact on ACR tracking. Prior research has shown that privacy controls do not always work [31, 32, 50]. We compare ACR tracking before and after exercising the offered privacy controls. We find that: (1) opting-out stops ACR network traffic, (2) login status does not impact ACR network traffic.

Third, we investigate whether ACR tracking differs across smart TVs bought and operated in different jurisdictions: the UK and the US. Both the UK and the US have distinct privacy laws and regulations. Moreover, data transfers between the UK and the US are regulated by the UK-US Data Bridge, which may impose geographic constraints on data transfers between smart TVs and ACR servers [14]. We study whether there are any differences in ACR tracking across the UK and the US and whether they use different geographically located ACR servers. We find that: (1) smart TVs in the UK and the US contact distinct ACR domains, (2) unlike the UK, ACR is active in the US even when streaming content from platform streaming app.

To the best of our knowledge, our work presents the first in-depth measurement and analysis of *second-party* ACR tracking in smart TVs. For reproducibility, our code and data is available at: https://github.com/SafeNetIoT/ACR

2 Background and Motivation

Smart TV tracking can be broadly classified into second-party and third-party tracking. Third-party tracking on smart TV is similar to traditional web and mobile tracking. Smart TV app developers include a tracking library or SDK that collects and shares app usage data and user/device identifiers with third parties. In contrast, second-party tracking refers to the tracking conducted directly by the smart TV platform via its operating system. Unlike third-party tracking that is limited to the app that includes the tracking library,
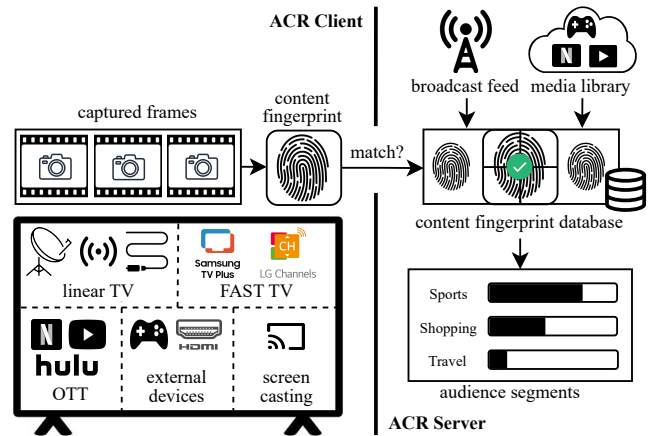

Figure 1: Overview of ACR tracking in smart TVs.

second-party tracking is potentially agnostic to how a user watches TV (i.e., whether watching linear or OTT, streaming app used, etc.).

Automatic Content Recognition (ACR) [42] is widely used for second-party tracking in smart TVs. As shown in Figure 1, ACR periodically captures frames (and/or audio), builds a fingerprint of the content, and then shares it with an ACR server for matching it against a database of known content (e.g., movies, ads, live feed). When the fingerprint matches, ACR server can determine exactly what piece of content is being watched on the smart TV. This enables smart TV platforms like Samsung and LG to profile users into *audience segments* [10, 30], which are then used to target personalized ads. Fingerprints in ACR are essentially hash of the content, which can be matched at the server-side to identify the content. However, the fact that the hash of content rather than raw content is sent to ACR servers does not necessarily make the data "anonymous" [57]. Moreover, the viewing habits of a user is potentially identifying [40].

Since its inception in 2011, with roots in Shazam song identification, ACR tracking has been adapted to identify other modalities of content. In 2012, DirecTV and Viggle expanded ACR into the TV ecosystem [53], while Samsung partnered with a content recognition tech company – Enswers to integrate ACR into their smart TVs [52]. LG smart TVs incorporated ACR in 2013 with a partnership with Cognitive Networks [21]. The same year, Sony also partnered with Samba TV to use its own ACR [54]. Moreover, Vizio and Roku adopted ACR in 2014 [17].

ACR tracking has raised privacy concerns. Most notably, Vizio was sued by the FTC for selling customer data to third parties, who then used it for personalized ads. This lawsuit was settled in 2017 with Vizio agreeing to provide clearer disclosures and opt-out mechanisms [19]. However, opting out is typically not straightforward, often requiring navigation through various settings in multiple subsections, with no universal off switch [64]. It is also unknown whether these privacy controls actually work as intended.

3 Design & Methodology

3.1 Design

To investigate ACR tracking, we setup a dedicated infrastructure that facilitates data collection and experimentation on smart TVs.
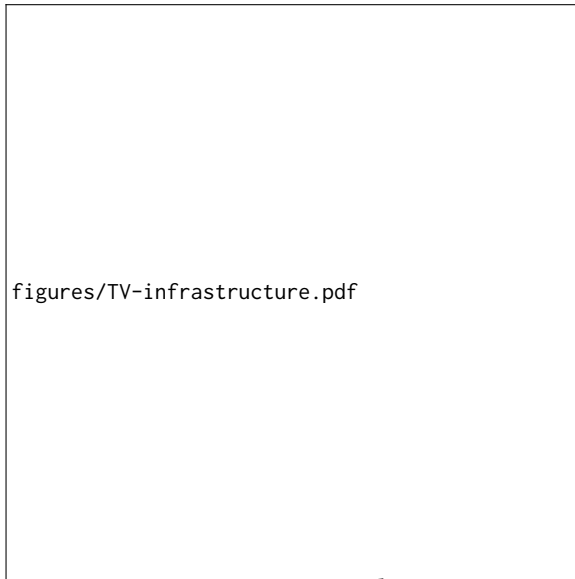
Figure 2: Experimental setup.

Figure 2 shows the design of our setup. We consider two smart TVs: Samsung and LG. We deploy our infrastructure in both the UK and the US. Each component of our infrastructure is described below.

Access Point Server. The servers are the core of the infrastructure. Each server, one per TV, works as an access point for the smart TVs, using dedicated Wi-Fi cards or adapters. In addition, they have one wired network interface connected to the Internet, through which ACR domains are contacted. The servers have been configured with the installation of the *Mon(IoT)r* software [39, 46], a powerful tool to capture network traffic from IoT devices. These servers store all the captured traffic and the extracted data for further analysis.

Smart TVs. We select two smart TVs, specifically Samsung and LG (with 23% and 18% market share, respectively [16]), making them two of the leading brands in the smart TV market, due to their widespread popularity and in-built integration of FAST channels.

Scripts. The experimentation process comprises a set of scripts, running directly on the servers, designed to automatically control the smart TVs while running different tests and analyzing their network traffic. In particular, we use support scripts for interacting with smart TVs and triggering a specific function, e.g., opening Netflix app (*Trigger Scripts*) and verifying the correct execution of the experiments (*Validation Scripts*). These scripts are entirely automated. *Trigger Scripts* rely on Android mobile phones, physically connected to the servers. We use the capabilities of Android Debug Bridge (ADB) [1] to establish remote control over the Tuya Smart app [23], effectively transforming mobile phones into remote controls for the smart TVs. Finally, we use *Analysis Script* for analysis of the network traffic.

3.2 Methodology

Network Traffic Collection. Our servers are able to capture the encrypted traffic during the entire duration of our experiments. Our analysis is focused at extracting traffic patterns from the data captured by *Mon(IoT)r* without decrypting it. This tool allows us to design dedicated experiments tailored to each TV and scenario under investigation. The capture contains exclusively the traffic

transmitted to and received from the smart TV. The experimental workflow is as follows. After the initiation of traffic capture, we automatically power on the smart TV, using server-controlled smart plugs. This initial power-on phase is crucial, as the majority of DNS requests are typically sent within the first few seconds after device activation [61]. This is essential to identify the domain names associated with the contacted IP addresses. After that, the core experiment starts, for a duration of one hour. The specific content of the experiment varies based on the scenario. Finally, the experiment concludes with the smart TV being powered off and the network traffic capture being terminated. The entire process is automated.

Scenarios. The term "scenario" refers to a distinct experimental setup designed to examine a specific functionality of the smart TV. Our experiments cover six scenarios.

- **Idle**. The smart TV is powered on and remains on its home page for the entire duration of the experiment.
- **Linear**. The experiment involves watching a single linear channel broadcasted via the TV antenna.
- **FAST**. This involves streaming a single channel from the FAST platform of the TV manufacturer (Samsung TV+, LG Channels). Free Ad-supported Streaming TV (FAST) is essentially linear broadcast TV that is streamed over the Internet.
- **OTT**. An over-the-top (OTT) app, streaming app that provides streaming content over the Internet (Netflix or YouTube), is used to stream content.
- **HDMI**. A separate laptop (browsing and watching YouTube videos) or gaming console (playing popular games) is connected to the TV via HDMI.
- **Screen Cast**. This scenario investigates the screen cast feature by mirroring YouTube content streamed on a separate phone or laptop onto the smart TV screen.

Phases. As shown in Figure 3, we delineate four distinct phases for executing each scenario, each characterized by a unique configuration determined by two key factors: the linkage of a user account and the option to accept or reject advertising/tracking settings on the TV. In two of these phases, we are logged in using a TV account, while in the other two, we are logged out. Additionally, in two phases, we actively opt-out of all advertising/tracking options available directly on the TVs, thereby declining such services. Table 1 in Appendix B lists all the selected opt-out options. Across all settings, ACR is specifically disabled by turning off *viewing information services* [69]. Conversely, in the remaining two phases, we opt-in to such settings. It is important to note that without accepting the ToS and privacy policy we are unable to watch or access most of the Smart TV content. So, our opt-in or opt-out always assumes that ToS and privacy policy are opted-in. Our four phases are as follows:

- **LIn-OIn**. Logged In-Opted In
- **LOut-OIn**. Logged Out-Opted In
- **LIn-OOut**. Logged In-Opted Out
- **LOut-OOut**. Logged Out-Opted Out

Identifying ACR traffic. We employ the *Analysis Scripts* to extract relevant information and statistics from the captured network traffic. We primarily focus on domains potentially associated with ACR tracking. To this end, we filter the list of contacted domains, identified with DNS captured packets, retaining only those containing the string "acr". To the best of our knowledge, no official
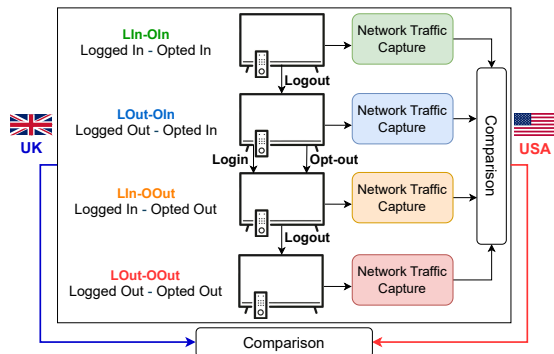
Figure 3: Methodology. The methodology is repeated for each scenario and TV in the two countries.

documentation exists on the specific domain names used for ACR by Samsung or LG, nor is there a standard requiring all domains with "acr" in the name to be related to ACR on smart TVs. But we use this approach to identify ACR traffic for the following reasons:

- Identified domains with the "acr" string were classified as tracking-related by sources like Netify [24] and Blocada [7]. Blocada.org [7], an open-source privacy suite, lists all such domains for LG and Samsung as tracking-related, matching the patterns we observed [7].
- Numbered ACR domains we observed suggest a consistent naming scheme by LG and Samsung, likely used to distinguish ACR servers by region or other factors.
- We further validated our approach by comparing presence of these domains before and after opting-out of ACR on the Smart TV and analyzing contact frequency across scenarios. These domains showed regular contact patterns, unlike other ad/tracking domains like samsungads.com.

This approach however, may not generalise to all TV manufacturers. Future research can build on our findings by following the heuristic in the last point. Our automated methodology and released scripts support further exploration in this area.

4 Results

The findings within the first two subsections pertain exclusively to the UK. The final subsection compares these results against the US.

4.1 Comparison Across Scenarios

To understand ACR tracking, we compare the network traffic flows to ACR domains during different scenarios, focusing particularly on phase LIn-OIn. LIn-OIn likely represents the most common configuration in the wild, since smart TV users need to be logged in (e.g., LG TVs require login for app downloads) and opted in (default option when setting up the TV) to access most of its functionalities.

Our analysis reveals different behaviors between Samsung and LG regarding their use of ACR domains. When ACR is enabled on LG TVs, a single domain is contacted (eu-acrX.alphonso.tv, where X is an arbitrary number that changes periodically). This domain belongs to Alphonso, a technology company that manages LG Ad Solutions [3]. On the other hand, Samsung contacts multiple ACR domains (acr-eu-prd.samsungcloud.tv, acr0.samsungcloudsolution. com, log-config.samsungacr.com, log-ingestion-eu.samsungacr.com). All these domains belong to Samsung, aligning with the fact that

Samsung Ads offer its own proprietary ACR tracking [49]. Furthermore, we geolocate the IP addresses using MaxMind [33] and IP2Location [25]. Due to known limitations and inaccuracies of GeoIP databases, we perform additional validation using RIPE IPmap [47] to accurately map the observed ACR domains to their server locations. We first perform traceroute from a location in the US or UK, then use RIPE IPmap for geolocation. In case of discrepancies, we rely on RIPE IPmap because – (1) It offers multiple geolocation engines, each with unique techniques. (2) Its latency engine quickly computes measurements using RIPE Atlas probes with known locations. (3) It uses a reverse DNS engine that leverages geographical identifiers in PTR records to estimate IP locations. Our analysis reveals that all LG ACR domains resolve to Amsterdam, Netherlands. Network traffic between the UK and the EU raises no cross-jurisdictional regulatory concerns [11]. For Samsung, acr-eu-prd.samsungcloud.tv and log-ingestion-eu.samsungacr.com are both located in London, UK, while acrX.samsungcloudsolution.com locates to Amsterdam, Netherlands, and log-config.samsungacr.com in New York, USA. This raises concerns about UK TV users' viewership data being stored in the US, where different privacy regulations apply. However, both Alphonso (for LG) and Samsung are on the DPF (Data Privacy Framework) List [14, 45], allowing data transfers between the UK and the US under the UK-US Data Bridge.

Figure 4 shows the frequency of network traffic directed towards ACR domains in all scenarios. The data is presented in a packet-per-millisecond format, where each spike corresponds to a single millisecond slot. For scenarios in which similar behaviors are observed, only one plot is presented. For both LG (a) and Samsung (b) TVs, the scenarios with the highest ACR traffic are Linear and HDMI. During the remaining scenarios, ACR traffic is significantly less – peaks get reduced by up to 12× – suggesting that ACR client on the TV may not be sending fingerprints. LG's official documentation mentions that its ACR captures frames every 10ms [56]. However, the fact that we observe network traffic every 15 seconds suggests that LG likely batches the frames captured every 10ms and sends the resultant content fingerprint every 15 seconds. The remaining scenarios for LG show a lower amplitude of communication occurring every fifteen seconds, with peaks every minute. For Samsung, we observe a consistent amount of traffic across various scenarios for the ACR domains (acr0.samsungcloudsolution.com, log-config.samsungacr.com, log-ingestion-eu.samsungacr.com). Based on their names, we hypothesize that communication with these domains primarily consists of logging information or maintaining connection persistence ("keep-alives"). We assume that the domain acr-eu-prd.samsungcloud.tv transmits fingerprints, as it exhibits the highest network traffic during Linear and HDMI. Communication occurs once per minute, with peaks observed approximately every five minutes. Interestingly, the remaining scenarios exhibit consistent peak values occurring every minute, accompanied by additional smaller traffic one minute following each peak. Samsung's official documentation [8] mentions that its ACR captures the frames every 500ms, suggesting that Samsung batches the captures as well and sends the fingerprints every minute. The differences in ACR capture frequency explains the different network behavior across the two brands.

We assume that, during OTT, ACR may not collect screenshots of third-party owned streamed content due to copyright issues.
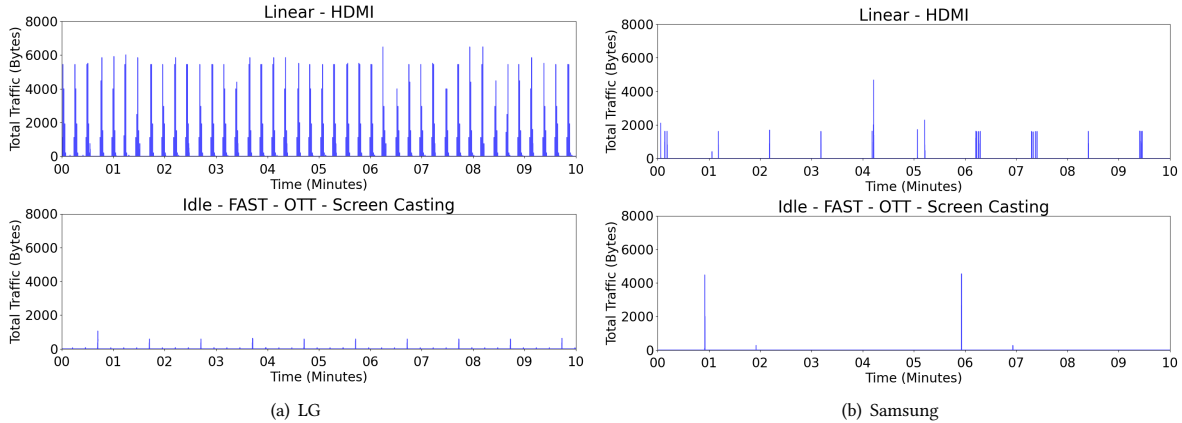
(a) LG   (b) Samsung

Figure 4: 10 minutes of ACR traffic in different scenarios during LIn-OIn in UK.

Another explanation could be that the third-party app wants to preserve the privacy of its users, for example Netflix prefers to have ACR deactivated during its streaming "in order to preserve the integrity of its subscribers viewing experiences and maintain sole control over measurement of its viewership" [67]. The same reasoning applies to FAST channels, which LG and Samsung consider to be "aggregator apps" [70], where providers may have agreements restricting ACR usage.

Linear and HDMI do not seem to present the previous restrictions. Linear, as FAST, may feature content from multiple providers, but the broadcasting network typically holds the rights to the content and has agreements in place with content owners and advertisers. These agreements include provisions for the use of ACR.

4.2 Comparison Across Phases

This section describes the methods we use for understanding impact of privacy controls, specifically in terms of the influence of user login status and opt-out settings on ACR network traffic.

LIn-OIn differs from LOut-OIn by only that the user is logged in/out on the TV. We first understand the differences during these phases across TVs by plotting the CDF of data transferred to ACR domains (in bytes) in each scenario during the LIn-OIn and LOut-OIn phases as shown in Figure 5 (see Figure 7 for USA). We observe distinctions in the data transfer periodicity amongst LG and Samsung measurements as also captured in Figure 4. Overall, Samsung transmits upto $2X$ more data at a higher frequency to ACR domains as compared to LG. Interestingly, LG sends the most data to ACR endpoints when content is streamed via HDMI and screen casting, while Samsung does so during linear TV streaming in the logged-in phase and FAST in the logged-out phase. These findings suggest implementational differences in the content fingerprint generation and transmission algorithms used by the two TVs.

Next, we look at the differences between the logged-in and the logged-out status for the same TV manufacturer. Analysis reveals that the set of ACR domains contacted across scenarios in LOut-OIn remains identical to those observed in LIn-OIn. Traffic volume and frequency patterns also exhibit a high degree of similarity between these phases for the same TV manufacturer as also represented in Figure 5. Hence, we conclude that although differences exist across TV manufacturers, for a given TV, user login status appears to have no material impact on the ACR network traffic behavior. We also

assume that ACR tracking may be relying on the Advertising ID of the TV and/or the IP address rather than the user account ID. For completeness, we add more details in Tables 2 and 3 in Appendix C.

Similarly, LIn-OOut and LOut-OOut phases exhibit identical behavior. They differ from the previous two in that we have opted out of all advertising/tracking. Interestingly, once opt-out is exercised (Table 1), there is a complete absence of communication with any previously identified ACR domains, and no new ACR-related domains are observed. These findings suggest that the opt-out mechanisms implemented on LG and Samsung smart TVs are working.

4.3 Comparison Across Countries

To investigate the differences across the US and the UK, we now analyze the results of our experiments conducted in the US. The comparison reveals some key differences in the ACR tracking across the US vs. the UK. This comparison is particularly relevant due to the differing data protection laws: the GDPR (General Data Protection Regulation) in the UK [60] and CCPA (California Consumer Privacy Act) in the US [41], particularly in California where we conducted out US experiments. It's possible that these variations in data protection laws may influence how ACR tracking operates in each region. More details on network traffic to US-specific ACR domains are provided in Table 4 in Appendix C.

Both LG and Samsung TVs utilize domain names consistent with those identified in the UK, with some differences in the names (e.g. the term "EU" in the UK domains and "US" in the US domains related to ACR). LG contacts a single ACR domain, tkacr**X**.alphonso.tv (where **X** is a number that changes periodically). Samsung, however, contacts the same three domains (acr-us-prd.samsungcloud.tv, log-config.samsungacr.com, and log-ingestion.samsungacr.com), but omits the fourth domain that it uses in the UK. Analyzing the geolocation of the US domains, their IP addresses all belong to servers that are physically located in the US.

As shown in Figure 6, in the US – similar to the UK – both smart TV platforms exhibit a significantly higher network traffic with ACR domains during Linear, FAST, and HDMI scenarios. Idle, OTT, and Screen Casting display considerably less traffic with ACR domains. Interestingly, the FAST scenario deviates from the UK findings. Watching LG Channels and Samsung TV+ in the US results in ACR traffic levels comparable to linear channel viewing, possibly because FAST platforms have different agreements with the content
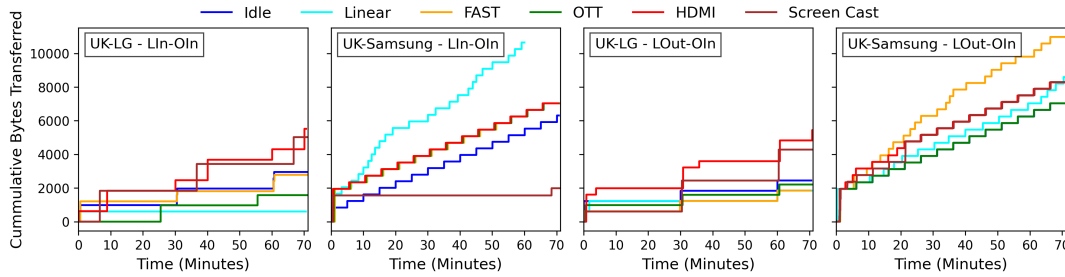
Figure 5: Cumulative distribution of bytes transmitted to ACR domains over the time during different opted-in phases in UK.
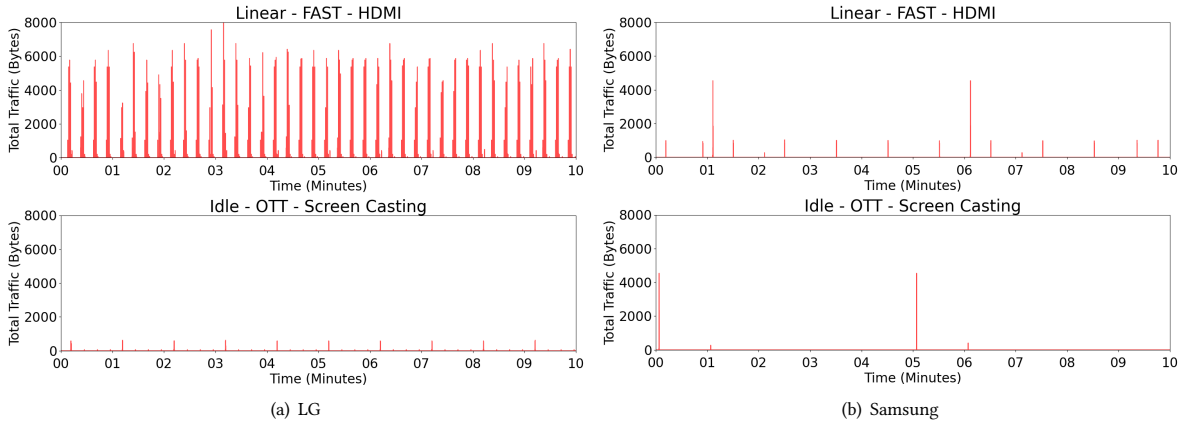

(a) LG

(b) Samsung

Figure 6: 10 minutes of ACR traffic in different scenarios during LIn-OIn in US.

providers than the ones in UK, allowing ACR to operate. All other assumptions made for the UK are valid also for the US, due to the similarity of the behaviors.

For all four phases, the observations are identical in both countries. In both LIn-OIn and LOut-OIn phases, user login status seems to have no influence on ACR tracking. All traffic patterns for the first two phases in both the countries during all scenarios are shown in Figures 8, 9, 10 and 11 in Appendix C, for completeness. Conversely, opting out of advertisement/tracking (LIn-OOut and LOut-OOut) disables ACR in both the countries. There is a complete absence of communication with any ACR domains during these phases.

5 Related Work

Online advertising and tracking pervades web, mobile, and IoT ecosystems. Research community has investigated advertising and tracking in IoT devices [2] emphasizing on either the IoT traffic [6, 22, 34] or security and privacy implications of such tracking [4, 5, 18, 26, 27, 44, 63, 71].

Advertising and tracking in smart TVs by third-parties has also garnered attention from the research community. In the last few years, researchers have attempted to study advertising and tracking in the smart TV ecosystem [20, 29, 36, 37, 58, 62, 65, 66]. Varmarken et al. [66] developed a tool to collect and analyze network traffic from top-1000 apps on smart TVs – Roku and Amazon Fire TV. They also show ineffectiveness of DNS-blocklists in blocking advertising and tracking traffic and PII-exposure by the ecosystem. Mohajeri et al. [37] studied the same two OTT streaming devices by analyzing their network traffic. They found that tracking involved collection and sharing of unique identifiers such as device IDs, serial numbers, MAC addresses, and SSIDs. Tileria et al. [62] showed the existence of similar tracking ecosystem for Android TVs. Tagliaro et al. [58] studied security and privacy issues for the Hybrid Broadcast Broadband TV (HbbTV) standard that allows broadcasters to improve their offered content to the broadcast signal as well as OTT streaming app users. While smart TV third-party tracking has been extensively studied, second-party tracking (e.g., ACR) has received little attention. We fill the gap in the literature by analyzing ACR technology in smart TVs.

6 Conclusion

We present a first look at second-party Automatic Content Recognition (ACR) tracking in smart TVs. Using a black-box auditing approach, we tested two major smart TV brands (LG and Samsung) in various scenarios and experimental setup in two different countries (UK and US). Our findings indicate that (1) ACR operates even when it is used as a "dumb" display via HDMI; (2) opt-out mechanisms stop ACR traffic; (3) ACR works differently in the UK as compared to the US. As future work, we plan to explore more advanced man-in-the-middle (MITM) techniques to understand the payload of ACR network traffic. Moreover, we plan to investigate the link between ACR tracking and ad personalization in smart TVs. Finally, although different than ACR, our auditing approach can be adopted to assess privacy risks of Recall [35] – which analyzes snapshots of the screen using generative AI [68]. To foster future research, our code and data is available at https://github.com/SafeNetIoT/ACR.

Acknowledgments

References

[1] ADB 2023. Android Debug Bridge. https://developer.android.com/tools/adb.

[2] Hidayet Aksu, Leonardo Babun, Mauro Conti, Gabriele Tolomei, and A Selcuk Uluagac. 2018. Advertising in the IoT era: Vision and challenges. *IEEE Communications Magazine* 56, 11 (2018), 138–144.

[3] Alphonso ACR Software Terms of Service 2022. https://alphonso.tv/privacy/smart-tvs/smart-tv-tos/.

[4] Natã Miccael Barbosa, Joon S Park, Yaxing Yao, and Yang Wang. 2019. " What if?" Predicting Individual Users' Smart Home Privacy Preferences and Their Changes. *Proc. Priv. Enhancing Technol.* 2019, 4 (2019), 211–231.

[5] Johannes K Becker, David Li, and David Starobinski. 2019. Tracking Anonymized Bluetooth Devices. *Proc. Priv. Enhancing Technol.* 2019, 3 (2019), 50–65.

[6] Oladayo Bello, Sherali Zeadally, and Mohamad Badra. 2017. Network layer interoperation of Device-to-Device communication technologies in Internet of Things (IoT). *Ad Hoc Networks* 57 (2017), 52–62.

[7] Blocada. 2024. 1Hosts (Lite). https://blokada.org/mirror/v5/1hosts/lite/hosts.txt. Accessed: 2024-08-25.

[8] IAB Canada. 2022. A Samsung Ads guide for advertisers. https://iabcanada.com/wp-content/uploads/2022/09/Samsung-Ads-ACR-Guide-CA.pdf. Accessed: 2024-05-14.

[9] IAB Canada. 2022. Understanding ACR: A Samsung Ads guide for advertisers. https://iabcanada.com/wp-content/uploads/2022/09/Samsung-Ads-ACR-Guide-CA.pdf.

[10] Samsung DSP Help Center. 2024. https://help.dsp.samsungads.com/docs/audiences. Accessed: 2024-05-08.

[11] European Commission. 2021. Data protection: Commission adopts adequacy decisions for the UK. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_3183. Accessed: 2024-05-14.

[12] Federal Trade Commission. 2017. Case 2:17-cv-00758 Document 1. https://www.ftc.gov/system/files/documents/cases/170206_vizio_2017.02.06_complaint.pdf. Accessed: 2024-05-07.

[13] Statista Research Department. 2023. Smart TV streaming device market share worldwide as of 2020, by platform. https://www.statista.com/statistics/1171132/global-connected-tv-devices-streaming-market-share-by-platform/. Accessed: 2024-05-07.

[14] Innovation & Technology Department for Science. 2023. Factsheet for United Kingdom Organisations. https://assets.publishing.service.gov.uk/media/650c4c7efbd7bc000de54786/factsheet_for_uk_organisations.pdf.

[15] Virginie Dremeaux. 2021. Across Europe, TV viewing is already a connected experience. https://www.freewheel.com/insights/blog/across-europe-tv-viewing-is-already-a-connected-experience. Accessed: 2024-05-07.

[16] Joseph D'Souza. 2024. Smart TV Statistics By Usage, Demographics and Facts. https://www.coolest-gadgets.com/smart-tv-statistics. Accessed: 2024-08-25.

[17] Cyrus Farivar. 2018. Vizio, sued for making creepy smart TVs, will notify customers via the TVs. https://arstechnica.com/tech-policy/2018/09/vizio-smart-tv-owners-to-learn-of-snooping-settlement-via-their-snoopy-tvs/.

[18] Kassem Fawaz and Kang G Shin. 2019. Security and privacy in the Internet of Things. *Computer* 52, 4 (2019), 40–49.

[19] Chaim Gartenberg. 2018. Vizio nears $17 million settlement for Smart TV data-tracking lawsuit. https://www.theverge.com/2018/10/4/17937052/vizio-17-million-settlement-tv-data-tracking-lawsuit.

[20] Ben Gilbert. 2019. There'sa Simple Reason Your New Smart TV Was so Affordable: It's Collecting and Selling Your Data, and Serving You Ads. *Business Insider* (2019).

[21] Michael Gorman. 2013. LG partners with Cognitive Networks to make Smart TVs smarter and more interactive. https://www.engadget.com/2013-08-29-lg-parnters-with-cognitive-networks-to-make-smart-tvs-smarter-an.html.

[22] Danny Yuxing Huang, Noah Apthorpe, Frank Li, Gunes Acar, and Nick Feamster. 2020. Iot inspector: Crowdsourcing labeled network traffic from smart home devices at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–21.

[23] Tuya Smart Inc. 2024. Tuya Smart App. https://play.google.com/store/apps/details?id=com.tuya.smart.

[24] Netify Network Intelligence. 2024. Samsungacr.com - Domain Info. https://www.netify.ai/resources/domains/samsungacr.com. Accessed: 2024-08-25.

[25] IP2Location 2024. https://www.ip2location.com.

[26] Umar Iqbal, Pouneh Nikkhah Bahrami, Rahmadi Trimananda, Hao Cui, Alexander Gamero-Garrido, Daniel Dubois, David Choffnes, Athina Markopoulou, Franziska Roesner, and Zubair Shafiq. 2022. Your echos are heard: Tracking, profiling, and ad targeting in the amazon smart speaker ecosystem. *arXiv preprint arXiv:2204.10920* (2022).

[27] Haojian Jin, Boyuan Guo, Rituparna Roychoudhury, Yaxing Yao, Swarun Kumar, Yuvraj Agarwal, and Jason I Hong. 2022. Exploring the needs of users for supporting privacy-protective behaviors in smart homes. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.

[28] Federica Laricchia. 2024. Smart and connected TVs - statistics & facts. https://www.statista.com/topics/4761/smart-and-connected-tvs/#statisticChapter. Accessed: 2024-05-07.

[29] Hieu Le. 2021. *Characterizing the Fire TV Advertising and Tracking Ecosystem*. University of California, Irvine.

[30] LG. 2022. https://www.lg.com/global/newsroom/news/home-entertainment/lg-smart-tvs-get-a-new-acr-solution-legacy-technology-replaced-by-lg-ads-solutions/. Accessed: 2024-05-08.

[31] Zengrui Liu, Umar Iqbal, and Nitesh Saxena. 2023. Opted Out, Yet Tracked: Are Regulations Enough to Protect Your Privacy? arXiv:2202.00885 [cs.CR]

[32] Célestin Matte, Nataliia Bielova, and Cristiana Santos. Do cookie banners respect my choice?: Measuring legal compliance of banners from iab europe's transparency and consent framework. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 791–809.

[33] MaxMind 2024. https://www.maxmind.com/en/home.

[34] M Hammad Mazhar and Zubair Shafiq. 2020. Characterizing smart home iot traffic in the wild. In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, 203–215.

[35] Microsoft. 2024. Retrace your steps with Recall. https://support.microsoft.com/en-us/windows/retrace-your-steps-with-recall-aa03f8a0-a78b-4b3e-b0a1-2eb8ac48701c.

[36] Hooman Mohajeri Moghaddam. 2022. *Tracking and Behavioral Targeting on Connected TV Platforms*. Ph. D. Dissertation. Princeton University.

[37] Hooman Mohajeri Moghaddam, Gunes Acar, Ben Burgess, Arunesh Mathur, Danny Yuxing Huang, Nick Feamster, Edward W Felten, Prateek Mittal, and Arvind Narayanan. 2019. Watching you watch: The tracking ecosystem of over-the-top tv streaming devices. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. 131–147.

[38] Gabriel Hongsdusit Mohamed Al Elew. 2023. Your Smart TV Knows What You're Watching. https://themarkup.org/privacy/2023/12/12/your-smart-tv-knows-what-youre-watching. Accessed: 2024-05-07.

[39] Mon(IoT)r 2019. https://moniotrlab.khoury.northeastern.edu/tools/.

[40] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*. IEEE, 111–125.

[41] State of California Department of Justice. 2018. California Consumer Privacy Act of 2018 (CCPA). https://oag.ca.gov/privacy/ccpa. Accessed: 2024-05-14.

[42] Tim Peterson. 2022. WTF is automatic content recognition (ACR)? https://digiday.com/future-of-tv/wtf-is-automatic-content-recognition/.

[43] Nick Pino. 2023. Dumb TVs - here's why you can't find them anymore. https://www.tomsguide.com/features/dumb-tvs-heres-why-you-cant-find-them-anymore. Accessed: 2024-05-07.

[44] Davar Pishva. 2017. Internet of Things: Security and privacy issues and possible solution. In *2017 19th international conference on advanced communication technology (ICACT)*. IEEE, 797–808.

[45] Data Privacy Framework Program. 2023. DFP List. https://www.dataprivacyframework.gov/list. Accessed: 2024-05-14.

[46] Jingjing Ren, Daniel J Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. 2019. Information exposure from consumer iot devices: A multidimensional, network-informed measurement approach. In *Proceedings of the Internet Measurement Conference*. 267–279.

[47] RIPE. 2024. RIPE IPmap. https://ipmap.ripe.net/. Accessed: 2024-08-25.

[48] Samsung. 2024. Samsung TV+. https://www.samsung.com/us/televisions-home-theater/tvs/tvplus/. Accessed: 2024-05-07.

[49] Samsung Ads 2024. https://www.samsung.com/uk/business/samsungads/.

[50] Iskander Sanchez-Rola, Matteo Dell'Amico, Platon Kotzias, Davide Balzarotti, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos. 2019. Can I opt out yet? GDPR and the global illusion of cookie control. In *Proceedings of the 2019 ACM Asia conference on computer and communications security*. 340–351.

[51] Audrey Schomer. 2024. The state of smart TV: a special report. https://variety.com/vip-special-reports/the-state-of-smart-tv-special-report-1235841024/. Accessed: 2024-05-07.

[52] Catherine Shu. 2014. Content Recognition Tech Company Answers Receives $2M From Samsung Ventures. https://techcrunch.com/2014/04/25/content-recognition-tech-company-answers-receives-2m-from-samsung-ventures/.

[53] Catherine Shu. 2018. TV ACR – a brief history, state of play and where it's going. https://www.tvadsync.com/2018/02/22/tv-acr-a-brief-history-state-of-play-and-where-its-going/.

SmartTV	Opt-Out Option
LG	**Enable** *Limit ad tracking*
	Disable *TV membership agreement for marketing comms.*
	Enable *Do not sell my personal information*
	Edit *User agreements* in Privacy and Terms as follows:
	Disable:
	– *Viewing information agreement*
	– *Voice information agreement*
	– *Interest-based & Cross-device advertising agreement*
	– *Who.Where.What?*
	Disable *Home promotion*
	Disable *Content recommendation*
	Disable *Live plus*
	Disable *AI recommendation* (*Who.Where.What, Smart Tips*)
Samsung	**Disable** *I consent to viewing information services on this device*
	Disable *I consent to interest-Based advertisements*
	Disable *Customization Service*
	Enable *Do not track*
	Disable *Improve personalized ads*
	Disable *Get news and special offer*

Table 1: Opt-Out Options in the Smart TVs.

[54] Sara Sluis. 2019. The Marketer's Guide To ACR Technology In Smart TVs. https://www.adexchanger.com/ad-exchange-news/the-marketers-guide-to-acr-tech-in-smart-tvs/.

[55] LG Ad Solutions. 2022. How ACR solves major advertiser challenges: Transparency, Cross-Platform Advertising, Measurement, Incrementality. https://lgads.tv/insights/how-acr-solves-major-advertiser-challenges/.

[56] LG Ad Solutions. 2023. Site Privacy Policy. https://lgads.tv/site-privacy-policy/. Accessed: 2024-05-14.

[57] FTC Staff in the Office of Technology. 2024. No, hashing still doesn't make your data anonymous. https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/07/no-hashing-still-doesnt-make-your-data-anonymous. Accessed: 2024-08-25.

[58] Carlotta Tagliaro, Florian Hahn, Riccardo Sepe, Alessio Aceti, and Martina Lindorfer. 2023. I Still Know What You Watched Last Sunday: Privacy of the HbbTV Protocol in the European Smart TV Landscape. In *30th Annual Network and Distributed System Security, NDSS 2023*.

[59] The Copyright Act of 1976, USA 1976. The Copyright Act of 1976, USA. https://www.copyright.gov/reports/guide-to-copyright.pdf.

[60] U.K. The Information Commissioner's Office (ICO). 2023. The UK GDPR. https://ico.org.uk/for-organisations/data-protection-and-the-eu/data-protection-and-the-eu-in-detail/the-uk-gdpr/. Accessed: 2024-05-14.

[61] Oliver Thompson, Anna Maria Mandalari, and Hamed Haddadi. 2021. Rapid IoT Device Identification at the Edge. In *Proceedings of the 2nd ACM International Workshop on Distributed Machine Learning*. https://doi.org/10.1145/3488659.3493777

[62] Marcos Tileria and Jorge Blasco. 2022. Watch over your TV: a security and privacy analysis of the android TV ecosystem. *Proceedings on Privacy Enhancing Technologies* 3 (2022), 692–710.

[63] Rahmadi Trimananda, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, and Athina Markopoulou. 2022. {OVRseen}: Auditing Network Traffic and Privacy Policies in Oculus {VR}. In *31st USENIX security symposium (USENIX security 22)*. 3789–3806.

[64] TurnOffSmartTVFeatures 2023. How to Turn Off Smart TV Snooping Features. https://www.consumerreports.org/electronics/privacy/how-to-turn-off-smart-tv-snooping-features-a4840102036/.

[65] Janus Varmarken, Jad Al Aaraj, Rahmadi Trimananda, and Athina Markopoulou. 2022. FingerprinTV: Fingerprinting Smart TV Apps. In *Proceedings on Privacy Enhancing Technologies (PoPETs)*, Vol. 2022. 606–629.

[66] Janus Varmarken, Hieu Le, Anastasia Shuba, Athina Markopoulou, and Zubair Shafiq. 2020. The tv is smart and full of trackers: Measuring smart tv advertising and tracking. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (2020).

[67] Verance. 2020. Verance Develops Solution to Prevent Emerging Issues With Netflix & Other Streaming Services, Avoiding Unforeseen Challenges of ACR Technology. https://www.verance.com/verance-develops-solution-to-prevent-emerging-issues-with-netflix-other-streaming-services-avoiding-unforeseen-challenges-of-automatic-content-recognition-acr-technology/.

[68] Tom Warren. 2024. Windows AI feature that screenshots everything labeled a security 'disaster'. https://www.theverge.com/2024/6/3/24170305/microsoft-windows-recall-ai-screenshots-security-privacy-issues.

[69] Bryan Westover. 2022. Stop your snooping smart TV — how to turn off data collection for every brand. https://www.tomsguide.com/how-to/stop-your-snooping-smart-tv-how-to-turn-off-data-collection-for-every-brand.

[70] Alan Wolk. 2023. Why FAST Channels Are Not the Same as Cable Networks. https://www.nexttv.com/news/why-fast-channels-are-not-the-same-as-cable-networks-wolk. Accessed: 2024-05-14.

[71] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User perceptions of smart home IoT privacy. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–20.

A Ethics

In our experiments we do not collect any data from real users on the Internet. All experiments are contained within our own testbed. When conducting the experiments, we fully respected the ethical guidelines defined by our affiliated organization, and we received approval.

B Opt-Out Options

All the opt-out options selected directly on the two smart TVs are shown in Table 1.

C Amount of Bytes Towards ACR Domains

Tables 2, 3, 4 and 5 quantify the amount of data (kilobytes) exchanged with LG and Samsung ACR destinations across various scenarios.

Figures 8, 9, 10 and 11 show the details of ten minutes traffic for each presented scenarios, respectively for UK LIn-OIn, UK LOut-OIn, US LIn-OIn and US LOut-OIn.
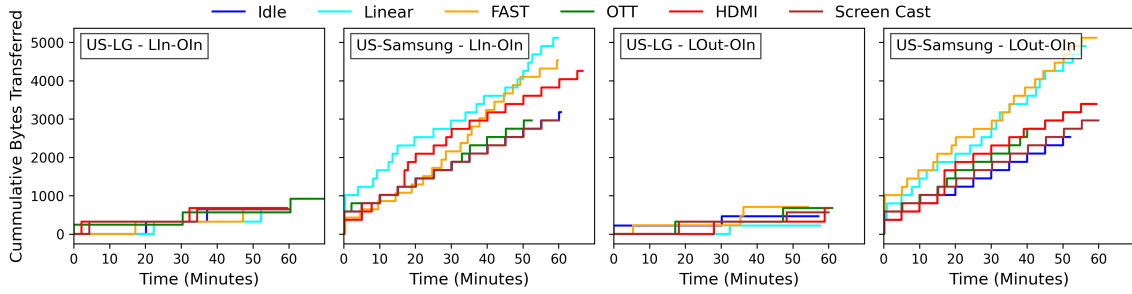
Figure 7: Cumulative distribution of bytes transmitted to ACR domains over the time during different opted-in phases in USA.

Domain Name	Idle	Antenna	FAST	OTT	HDMI	Screen Cast
eu-acrX.alphonso.tv	264.7	4759.7	262.8	264.3	4296.5	266.2
acr-eu-prd.samsungcloud.tv	-	440.9	8.5	8.6	204.8	30.3
acr0.samsungcloudsolution.com	-	-	11.1	11.3	11.0	11.7
log-config.samsungacr.com	9.5	10.8	9.2	8.9	9.3	10.0
log-ingestion-eu.samsungacr.com	176.9	298.4	125.4	161.6	162.3	-

Table 2: Number of kilobytes sent/received to/from ACR domains in different scenarios during LIn-OIn in UK.

Domain Name	Idle	Antenna	FAST	OTT	HDMI	Screen Cast
eu-acrX.alphonso.tv	258.0	4801.9	255.5	250.6	4229.5	272.8
acr-eu-prd.samsungcloud.tv	8.6	463.9	8.6	8.5	184.0	16.1
acr0.samsungcloudsolution.com	11.1	11.1	11.0	11.1	11.0	24.3
log-config.samsungacr.com	9.2	9.1	-	9.1	9.2	10.4
log-ingestion-eu.samsungacr.com	159.9	232.3	-	169.8	170.6	195.3

Table 3: Number of kilobytes sent/received to/from ACR domains in different scenarios during LOut-OIn in UK.

Domain Name	Idle	Antenna	FAST	OTT	HDMI	Screen Cast
tkacrX.alphonso.tv	215.3	4583.2	4948.3	214.9	4125.0	240.4
acr-us-prd.samsungcloud.tv	-	184.4	176.6	-	148.5	-
log-config.samsungacr.com	10.5	10.5	-	9.7	19.7	10.1
log-ingestion.samsungacr.com	143.5	253.2	237.4	156.1	224.8	172.1

Table 4: Number of kilobytes sent/received to/from ACR domains in different scenarios during LIn-OIn in US

Domain Name	Idle	Antenna	FAST	OTT	HDMI	Screen Cast
tkacrX.alphonso.tv	236.3	4612.4	4832.5	191.3	4633.5	222.0
acr-us-prd.samsungcloud.tv	-	153.5	166.1	-	160.2	-
log-config.samsungacr.com	9.6	9.6	9.6	10.4	10.4	9.6
log-ingestion.samsungacr.com	112.7	216.3	247.5	187.5	146.9	157.9

Table 5: Number of kilobytes sent/received to/from ACR domains in different scenarios during LOut-OIn in US
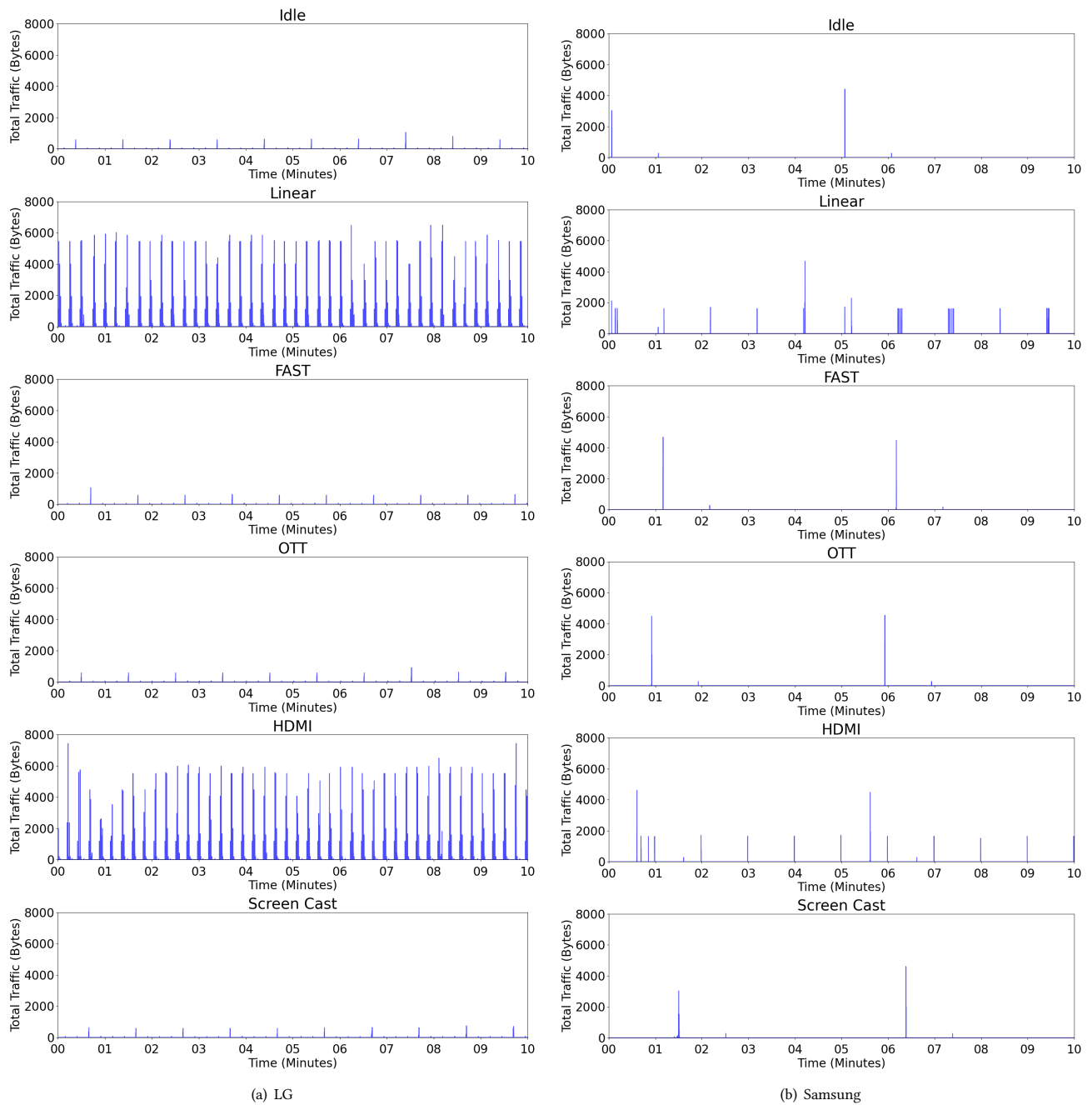
(a) LG

(b) Samsung

Figure 8: 10 minutes of ACR traffic in different scenarios during LIn-OIn in UK.
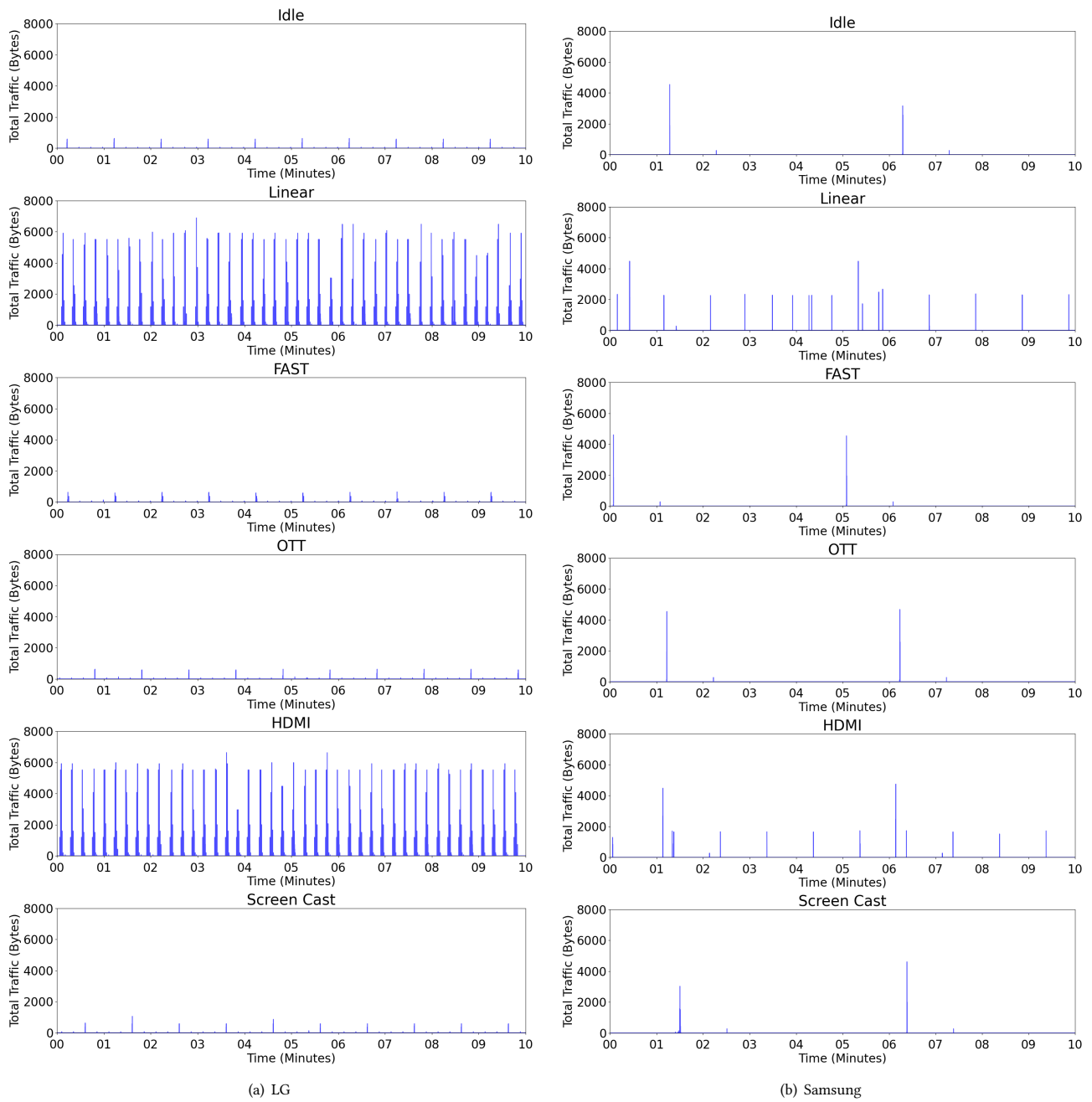
(a) LG

(b) Samsung

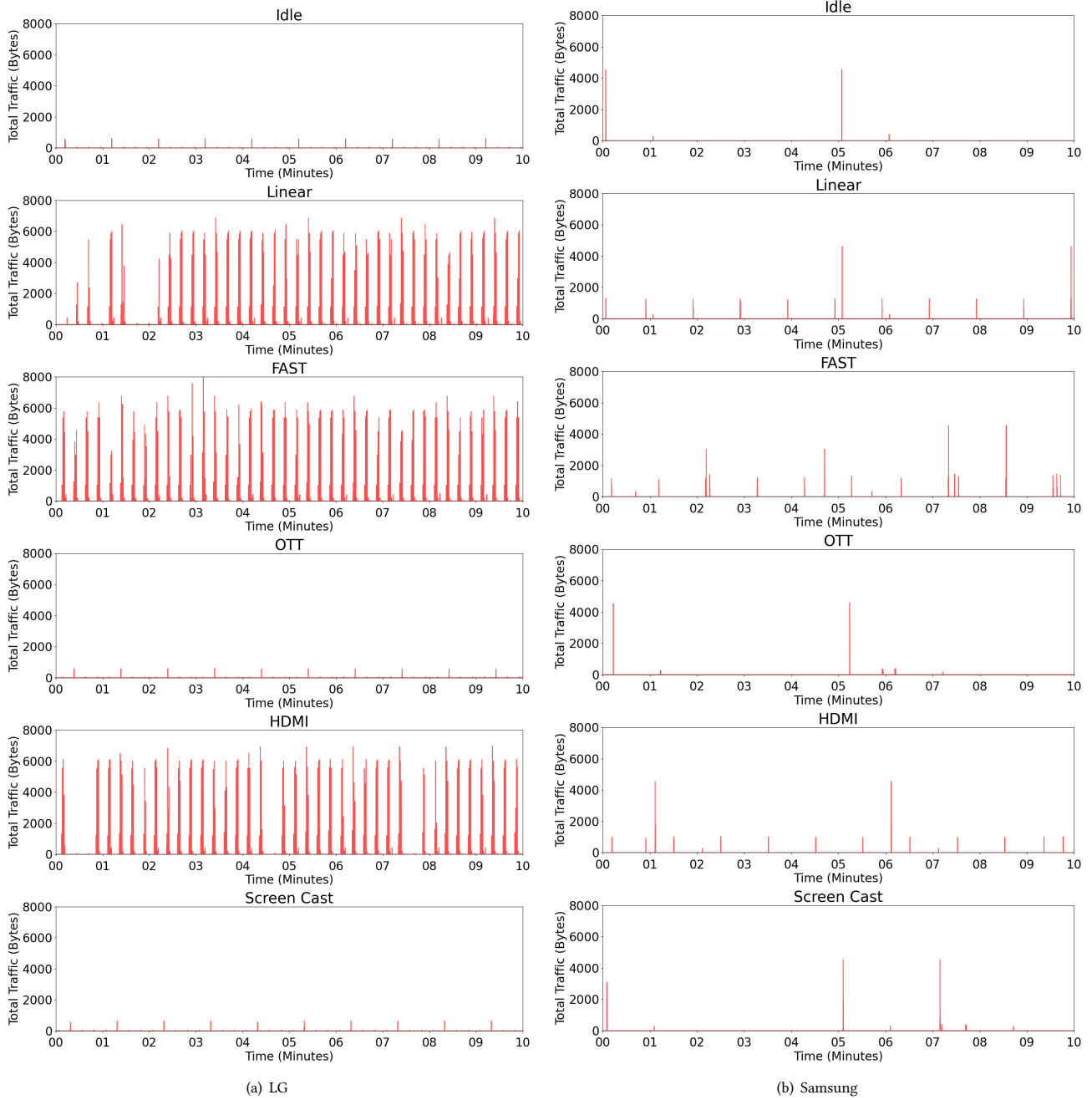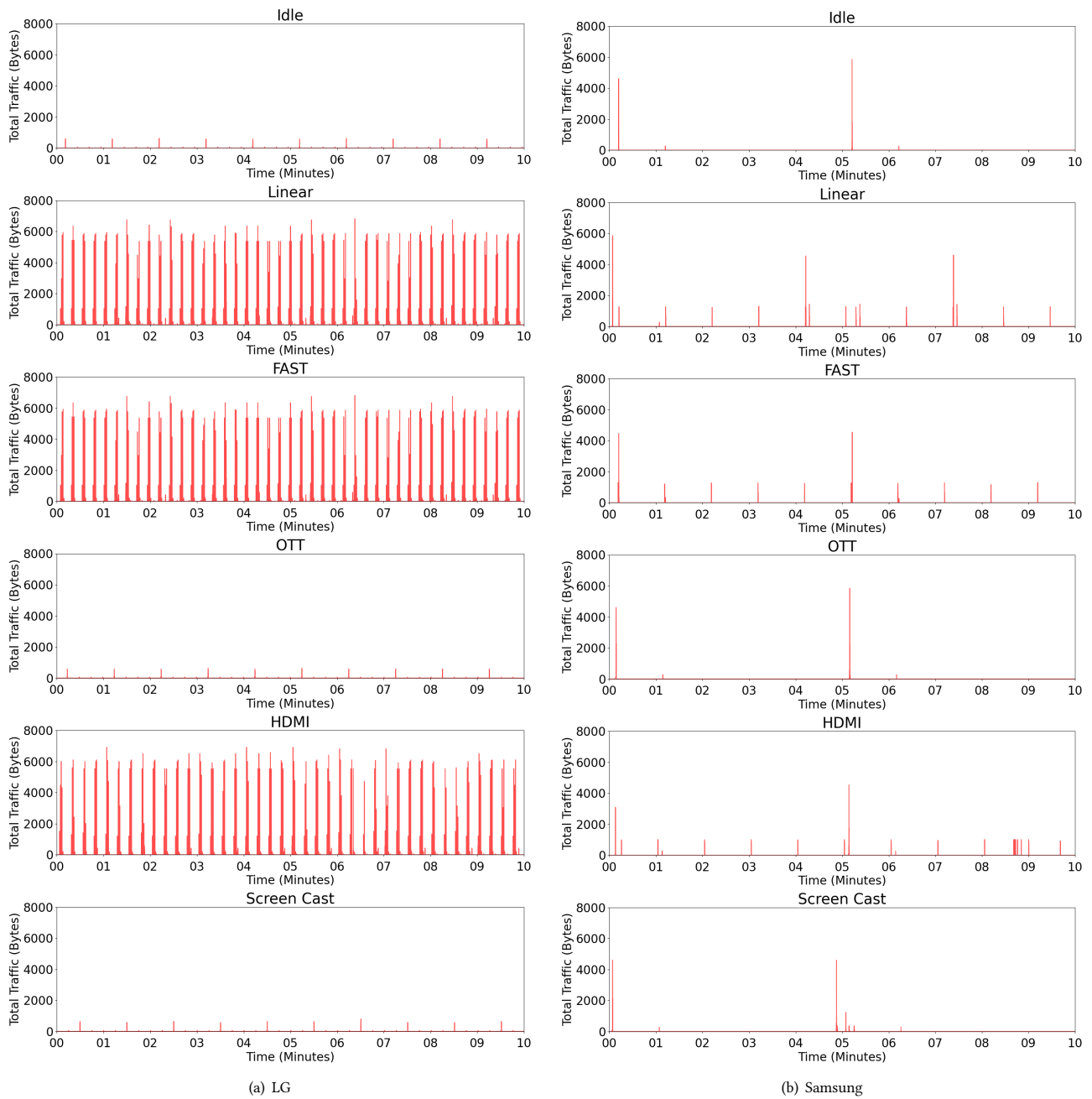**Figure 9: 10 minutes of ACR traffic in different scenarios during LOut-OIn in UK.**

(a) LG

(b) Samsung

Figure 10: 10 minutes of ACR traffic in different scenarios during LIn-OIn in US.

(a) LG

(b) Samsung

Figure 11: 10 minutes of ACR traffic in different scenarios during LOut-OIn in US.