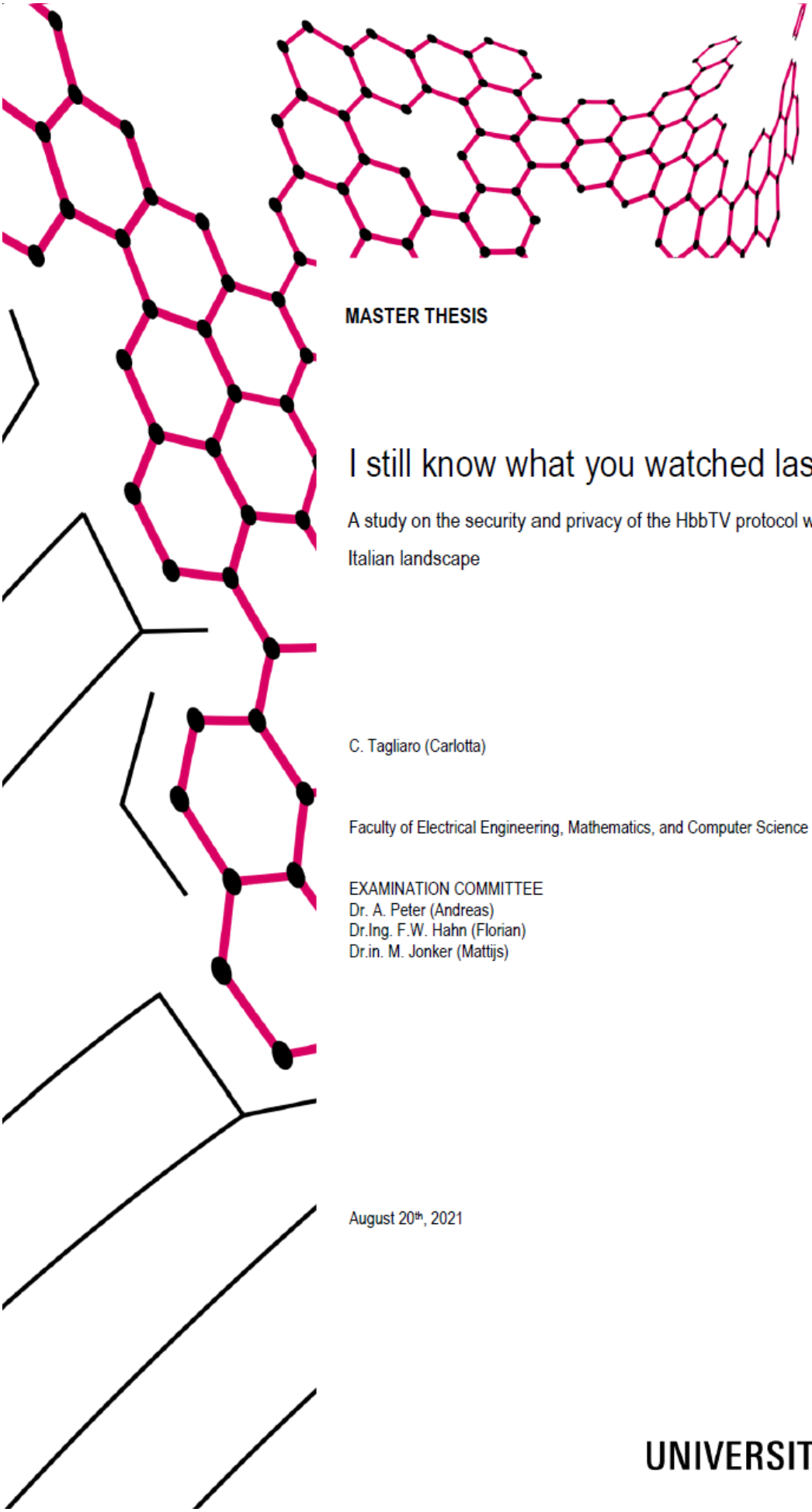

MASTER THESIS

I still know what you watched last Sunday

A study on the security and privacy of the HbbTV protocol with a focus on the Italian landscape

C. Tagliaro (Carlotta)

Faculty of Electrical Engineering, Mathematics, and Computer Science

EXAMINATION COMMITTEE
Dr. A. Peter (Andreas)
Dr.Ing. F.W. Hahn (Florian)
Dr.in. M. Jonker (Mattijs)

August 20th, 2021

UNIVERSITY OF TWENTE.

I still know what you watched last Sunday

A study on the security and privacy of the HbbTV protocol with a focus on the Italian landscape

Carlotta Tagliaro
c.tagliaro@student.utwente.nl
EIT CyberSecurity Master Student
University of Twente, the Netherlands

***Abstract*—The ever-increasing support for the HbbTV standard in Smart TVs and Set-Top boxes allows broadcasters to enrich contents offered to consumers via the standard broadcast signal with Internet-delivered apps, e.g., the possibility to rewatch a show. It works using standard web technologies as transparent overlays over a TV channel. Despite numbers of HbbTV-enabled devices are rapidly growing, little or no security studies are present in the literature and no standard protective measure is in place.**

This works aims at showing the current state of HbbTV in the Italian landscape and discuss its implications for consumers' privacy and security. We describe some of the techniques used by different broadcasters to measure users' (viewing) preferences and show how their wrong implementation causes severe risks. To complement the analysis, we carried out an online survey to assess the level of awareness of Smart TV and HbbTV related risks. Results show a low level of perception of the possible threats users are exposed to. Finally, an open-source security mechanism is presented so to ensure a safe experience for the user when watching TV and reduce the privacy issues that HbbTV may pose.

***Index Terms*—HbbTV, Hybrid Broadcast Broadband TV, Security, Privacy, Smart TV, DVB, Network, Risks, Human Factor**

I. Introduction

STATISTICS reveal that, up to date, 1.7 billion TV households exist worldwide[1] having a huge impact on our society as a whole. In Western Europe, the average television viewing time per person amounts to 240 minutes per day. However, some difficulties arise when trying to keep television at pace with new digital media being developed. In recent years, a major shift towards on-demand, streaming services, as Netflix, has been witnessed.

To combine the broadcast content delivery typical of standard TVs with the powerful digital content delivery of the new platforms and improve the video user experience for consumers, the Hybrid Broadcast Broadband TV (HbbTV) [5] initiative was started in 2009 by an industrial consortium comprising the industry leaders. It sets a standard for a broadcast/broadband hybrid protocol to deliver content to smart TVs, Set-Top Boxes, and other kinds of connected multiscreen devices in an interconnected environment.

Few studies have been conducted on the security and privacy posture of the HbbTV protocol showing that little or no security is provided to ensure consumers' safety. Additionally, little or no control is given to the viewer; she has no means to detect whether a connection is secured, which data are transferred, and how those are then used. Users' privacy is even at greater risk when combining such powerful tools with tracking and data analytics.

The security issues that HbbTV can open to are varied and with different levels of severity. They can range from a simple echo request from the broadcaster to check if the user is still watching the channel to a phishing attack that replaces the application's URL and induces the user to insert credit card details in a fake check-out page. Some solutions to protect users' privacy in the smart TVs domain are already present in the market [26]; however, those are related to specific TV models and vendors and do not take into consideration the extra interaction via the HbbTV protocol carried on by the single individual broadcaster.

Ghiglieri et al. [10]–[13] moved the first steps in the direction of assessing HbbTV's security posture highlighting the severe risks consumer' privacy was exposed to. Users' awareness of HbbTV's risk was also assessed resulting in a lack of such. Additionally, a HbbTV Privacy Protector was developed to let consumers decide whether a specific HbbTV-enabled channel can or cannot load Internet data.

Most of the studies present in the literature date back to 2016. Recently, a major shift towards HbbTV 2.0 has been witnessed and with it, new security measures are introduced (e.g., increased use of HTTPS over HTTP) together with the more widespread adoption of HbbTV.

In this paper, we study how the landscape has changed and if broadcasters have become more careful with users' security and privacy by looking at the traffic between a Smart TV and the servers offering the HbbTV applications. Additionally, it will test through a survey if consumers, since being more frequently exposed to such features, have also become more security-aware. The main focus of this research will be Italian broadcasters, given the geographical setting of the collaborating company, with a small digression on German and French ones.

This paper is structured as follows: Section II gives some background and technical information on HbbTV with a focus on its security issues. Section III introduces the complementary work that has been done on the security and privacy aspects of the HbbTV protocol.

Section IV presents the traffic analysis for nine Italian TV channels together with its results. The outcome shows that little progress has been made over the years and consumers'

[1] https://www.statista.com/statistics/268695/number-of-tv-households-worldwide/

are still exposed to privacy and security risks, e.g., being tracked before expressing consent. To complete the picture, a reduced version of traffic analysis is discussed for some French and German broadcasters that offer HbbTV applications showing that consumers' still experience a lot of profiling. Section V aims at investigating security awareness among users' showing, again, a general lack of knowledge of HbbTV's risks. However, it also shows that when consumers are confronted with potential risks linked to the use of Smart TVs and HbbTV, they show high concern for their data.

Section VI discusses the problematic immaturity of HbbTV's security together with the main results of the present study and how they could impact consumers.

Section VII describes the proposed solution to ensure users' privacy, carefully describing why each design and architecture choice was made. Differently from the Privacy Protector, such a tool allows for higher customization and more flexible security measures.

Section IX summarizes the main findings and proposes further steps that could complement the presented study.

II. Background

The *Hybrid broadcast broadband TV (or "HbbTV")* - as stated by the HbbTV Association [5] - *is a global initiative aimed at harmonizing the broadcast and broadband delivery of entertainment services to consumers through connected TVs, set-top boxes, and multiscreen devices*. In other words, it represents both a widely adopted industry standard, the ETSI Technical Specification 102 796 [6], and a driving force to promote a unified hybrid TV delivery across different platforms [3]. A hybrid TV offers both broadcast and broadband content to the viewer.

The initiative dates back to 2009 when a group of industry leaders, lead by the German broadcaster RTL, introduced a different form of Teletext using the HbbTV standard and the CE-HTML interface language, an XHTML-based standard for webpages with remote user interfaces typically used in consumer electronic devices.

The HbbTV standard works either via broadcast or via IP link; however, it is most powerful in an Internet-connected environment where a combination of broadcast and broadband networking can deliver additional content to the consumer. For HbbTV to work, the TV or Over The Top (OTT) device must have support for it, and then, the broadcaster must provide at least one HbbTV application for the user to interact with. When such an application is delivered to the consumer, she is typically displayed with a relevant icon informing that some extra HbbTV content is available. Such additional information is typically in the form of program guides, viewer interaction (e.g., with some quizzes during a show), lyrics of viewed music videos, better advertising and customized content.

To interact with such extra content, until HbbTV version 1.5, the user could use the remote control of the Smart TV, more specifically through the colored buttons. Instead, with HbbTV 2.0, the possibility to connect different devices, such as smartphones and tablets, was added hence allowing multi-device interactions.

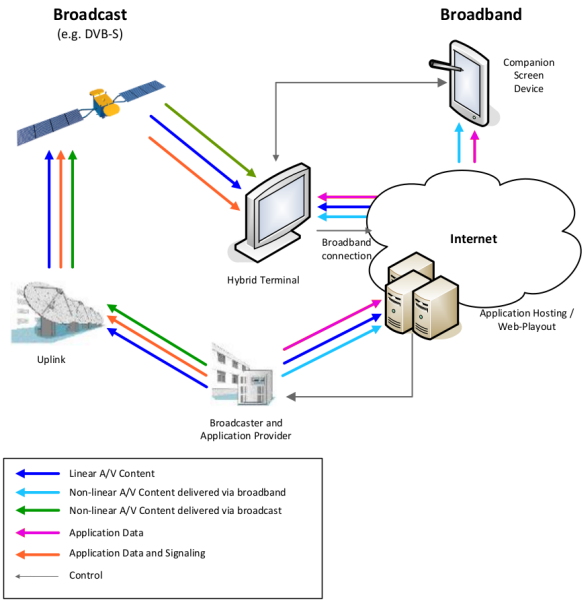

Fig. 1: HbbTV System Overview [7]

A. HbbTV Specifications

A hybrid terminal can support two different connections in parallel; on one side it is connected to a broadcast Digital Video Broadcasting (DVB) network, while on the other side it is connected to the Internet via a broadband interface. Through the first one, the terminal receives standard broadcast Audio/Video (A/V) content and allows for the signaling of stream events to an application. The Internet connection, instead, allows for bi-directional communication with the provider and can receive non-linear A/V content. The broadband interface may also connect with other HbbTV terminals or Companion Screen Devices (e.g., smartphones and tablets) on the same local network. The interaction between the different actors can be seen in Figure 1.

Through the Broadcast interface, the terminal also receives application data and stream events that are transferred using Digital Storage Media - Command and Control (DSM-CC) object carousels. Non-realtime content is transferred using the File Delivery Protocol (FDP) protocol. The recovered data is sent to the Runtime Environment of the terminal composed by the Application Manager, the Browser, and the Companion Screen Interface.

Via the Broadband Interface, the hybrid terminal has a connection to the Internet. This connection provides a way to request application data from the servers of a provider. Data collected in this way is again transferred to the Runtime Environment [7].

The Internet-delivered HbbTV applications are embedded as a link in the DVB stream sent by the broadcaster, which will be then extracted and loaded in the background of the browser. The content can be any website written with standard web techniques such as HTML, CSS, and JavaScript. When the application is loaded, the user is typically displayed with a notification overlay to show the app being ready to be activated through the remote control (via the standard *Red Button*).

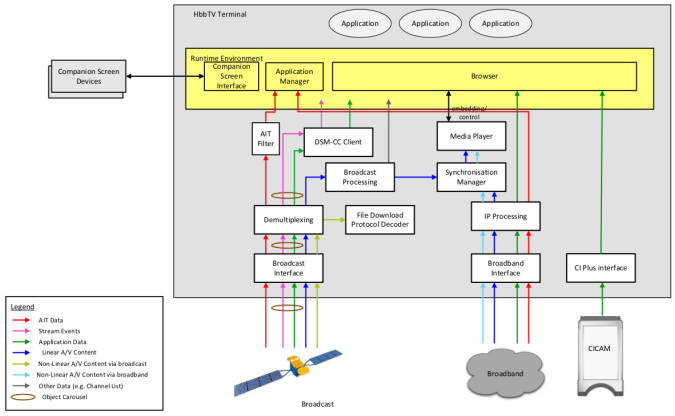

Fig. 2: Functional Components of a Hybrid Terminal [7]

B. HbbTV Adoption Rate

The adoption rate and support of HbbTV are growing steadily over the years; in 2014, already 92 percent of Germany's smart TVs support HbbTV [13].

In the following, adoption rates and relevant data from 2020 will be reported with a particular emphasis on three countries, Italy, Germany, and France being them respectively the country where the present study is conducted, the one with the highest adoption rate, biggest support for HbbTV and where part of the research is conducted, and the one where, again, part of the study was carried out [1], [21].

a) ***Italy****:* Totally, 25 million TV households[2] are present in Italy; Among them, 4,05 million represent HbbTV connected devices.

b) ***Germany****:* Germany represents the leading country when looking at HbbTV statistics being also the first place where such a standard was deployed. 38.52 million TV Households are present; Over 90% of Smart TVs that are sold support the HbbTV standard;

c) ***France****:* In December 2020, 28.8 million households are estimated and French broadcasters report, by October 2020, 1.75 million HbbTV connected devices.

C. Security Concerns with HbbTV

Despite statistics showing an ever-increasing adoption of HbbTV, little or no literature is present on the security issues that it might entail. Up to now, the main focus of security researchers has been vulnerabilities linked to physical access to such devices either through the USB port or local network access [14], [22], [24].

However, smart TVs that support HbbTV can access online content and web pages through the integrated web browser. These features open up a plethora of different attacks.

Before delving into the description of some possible attacks, it is worth mentioning that the HbbTV specification presents a security-related chapter. It is stated that only broadcast-related applications shall be trusted and broadcast-independent applications shall not be trusted. However, it is not mentioned how to perform such a control, and additionally, the user has the power to bypass such a restriction by making broadcast-independent applications trusted. On the broadband side, it is mentioned that security is provided through the adoption of the Transport Layer Security (TLS) protocols. Concerning the adoption of TLS, several requirements are presented such as the supported cipher suites, the minimal key length, and the forbidden use of compression algorithms. At the same time, the adoption of TLS (or more specifically HTTP over TLS (HTTPS)) is strongly suggested; whether it is implemented depends both on device manufacturers and on the actors that deliver content to the end-user. Additionally, the validation of server certificates, as described in the official document, appears to be weak and insufficient; only the match between hostname (or IP address) contained in the server certificate against the requested one is carried out. Both the SSL certificate expiration date and the identity of the Certificate Authority that released the certificate are not checked.

Finally, it is important to make some considerations privacy-wise. The standard allows the end-user to specify its tracking policy. Two alternatives are given, Do Not Track (`DNT`) set to 1, no tracking consent, or set to 0, tracking consent. The `DNT` parameter is included in every outgoing HTTP request to explicit user's tracking preferences. However, again, in this case, it is up to application developers and device manufacturers to correctly implement privacy. Several problems might arise if tracking websites are allowed, especially in autostart applications (i.e., applications that run without the user knowing and without the need for her consent) or even if persistent cookies are stored. Persistent cookies remain until the expiry date and as reported in Section III can be extremely problematic since the date set is far in time allowing tracking over a long period. Other issues might be generated from allowing third-party cookies to be stored [7].

As previously described, most of the HbbTV applications run inside a built-in browser that allows to both display HTML content and runs JavaScript code. To manifest the end-user the option to access extra content, a small hint is displayed on the TV screen. This is implemented as a semi-transparent HTML layer that overlays the actual TV program and it contains an URL encoded in the DVB stream (retrieved from a specific web server). In such a way, the TV becomes visible to the broadcaster even before the user consents to it hence possibly breaching the user's privacy. More details on such a problem are reported in Section III.

As mentioned above, a problem is represented by third-party tracking. A study conducted in 2013 over 66 different German stations showed that 13 among them used Google Analytics functionalities to track users; others used other services or their scripts (however, these latter cannot be determined with their approach since it would require access to server-side code) [14]. This could not only cause damages on the end-user side, but the attacker might exploit such a feature to spam fake analytics via proxy networks simulating real TVs and influence broadcasters' decisions, e.g., to discontinue a certain show.

Another family of attacks can be referred to as content-based attacks; the malicious actor can replace the URL or the content that the user will be displayed with. Several opportunities open up for the attacker. They could exploit DVB/DSM-

[2]TV Households are the share of households with a television set.

CC injection to replace content into stream carousels, directly specifying the URLs pointing back to their malicious content. In case of connections not secured with TLS, attackers could potentially spoof content being transmitted and perform a Man-in-the-Middle (MITM) attack replacing the original content. As pointed out by Herfort [14], none of the 66 stations were implementing TLS, but (hopefully) now broadcasters should be more security-aware. Interestingly, due to poor server configurations, also Watering Hole attacks were potentially successful with the attacker being able to directly access the server and replace content.

What do such content-based attacks entail? The end-user can be displayed with any arbitrary content, might be tricked into clicking a malicious link, and JavaScript code can be run without the user knowing. Attackers might exploit such a weakness in the system to insert fake news banners using a partly transparent HTML overlay (similar to the one used to signal the presence of extra content). This can lead to misinformation and can have potentially dramatic consequences. Especially in the pandemic period we are living in, the power and influence that fake news can have on people are in the eyes of everybody eventually leading to no-mask demonstrations that can hamper everyone's health [4], [19]. Additionally, miners can exploit such a possibility to use several TVs' CPUs to mine bitcoins using JavaScript-based code [9], [23].

Besides, the TV can be used to attack further devices in the user LAN. Using a timing approach, attackers can scan users' private networks searching for connected devices by exploiting the XmlHttpRequest object in JavaScript. This could lead to the reconfiguration of components in the local network facilitating further attacks (e.g., by reconfiguring the home router). Using this technique, the attacker can eventually transfer all the gained information to Internet drop-zones hence exposing the victim's IP address.

III. Related Work

Few studies have been conducted on the security posture of HbbTV deployments, mainly due to its recent widespread adoption, lack of (apparent) security relevance in the TVs domain, and the manufacturer/developer-dependency that causes a plethora of different implementations and solutions. Especially concerning the topic of dynamic advertisement, such a lack is predominant given the standard has only been introduced from mid-2020.

Some researches are to be considered extremely relevant to the topic of the present document: a collection of papers on the privacy of HbbTV [10], [11], [13], a 2019 hijack attack on the broadcast communication [8] and a survey on the lack of security and privacy awareness of consumers [12]. In the following, the results are reported.

Before delving into privacy, it is important to mention a 2014 attack performed by Oren and Keromytis that was able to manipulate an HbbTV URL at the DVB level (or even the entire application by manipulating the DSM-CC object carousel) and caused several devices to receive malicious URLs or content. Since HbbTV supports graphic overlays over the actual content of the screen, this might lead to phishing attacks entirely covering it [20].

Nowadays, much sensitive and personal information is processed by web applications and particular care must be put into ensuring security and privacy. Standard security mechanisms must be adopted even when dealing with HbbTV apps, for example, an online shop that allows the end-user to add items in the chart by simply pressing the red button. Clearly, in this example, data about the buyer (e.g., name, credit card number) must be protected. However, in 2014, a German channel transferred a user's login without HTTPS thus allowing potential attackers to record the complete login process and later exploit it [13]. Additionally, even when HTTPS is used, some smart TV models incorrectly implement certificate validation allowing MITM attacks.

Ghiglieri and Waidner [13] conducted three different tests in 2012, 2014, and 2015 to analyze the HbbTV data flow from the smart TVs to the broadcasters and vice-versa. They analyzed the dataflow also before consumers pressed the red button to explicitly launch the HbbTV apps and found several privacy issues. First, periodic requests (spanning from every 1 second to 15 minutes) were made to allow broadcasters or other third parties to measure how long consumers remain on a specific channel. Those do not launch or open the HbbTV apps but still collect information through the "counting pixel" technique, e.g., screen resolution, device vendor, or other.

In 2012, many broadcasters deployed traffic measurement methods without considering the legal aspects and without presenting a privacy policy to consumers. In 2015, many channels switched to HTTPS for securing HbbTV applications indicating that some steps towards security and privacy maturity were being made. However, as previously mentioned, some home shopping channels still did not adopt such a protocol. At the same time, heavy exploitation of cookies was found in 2015; their expiration dates ranged from 30 days to 1 year thus remaining on the devices for long periods and without letting the user delete them. Several "invisible" tracking scripts, such as webtrekk, were also found. Despite some progress, still, no governing rules were present and such technology did not completely enforce users' privacy [17].

For consumers, the best way to protect from such leaks would be to disconnect their device from the Internet, but this is not acceptable. Broadcasters should stream HbbTV notifications over DVB (before the user presses the red button) to ensure that her information is not leaked through the Internet. However, this would turn it into a trust problem. Ghiglieri and Tews propose an interesting solution, called the Privacy Protector, that allows end-users to control their data by barring channels to load Internet data unless the consumer presses the green button [11]. Additionally, to ensure a secure authentication and authorization mechanism for consumers, Matejka et al. proposed an architecture for a Security Manager [16]. Such a tool should verify the user being actually who he claims to be and later enforce some access control policies; in the context of a Smart TV, this translates to the possibility of users logging safely to different accounts, eventually through multi-factor authentication, via such a manager.

In 2019, at a conference talk, Massimo Bozza showed the feasibility and extreme easiness of hijacking HbbTV DVB connections. Through the use of the HiDes UT-100c, a modu-

lator (transmitter)[3] and the C++ library TSDuck[4] it is possible to modify the Application Information Table (AIT) (containing the HbbTV related information) in the stream, force the kill of all "legit" HbbTV applications by specifying the `0x04 KILL` code, add the malicious application redirecting to the target URL and automatically start it by specifying the code `0x01 AUTOSTART`. The flow of the inline hijack of the DVB signal is simple; the stream (from the antenna) is captured through a tuner, it is then processed with TSDuck and its plugins to modify it in the way described above, lastly, the stream is re-modulated and sent to the TVs. This weakness in the DVB architecture allows an attacker to possibly replace the HbbTV application with arbitrary and/or malicious content. The opportunities for an attacker are multiple: a user being displayed with fake news banners, redirected to a malware-download website, or a scam/phishing one [8]. Such attack was similarly conducted by Michéle et al. using a Terratec TStick+ as modulator and, on the software side, different libraries such as *tzap* [17].

Even more worryingly, the survey conducted by Ghiglieri et al. [12] reveals that only a small percentage of the interviewed are aware of those privacy and security risks, and even fewer can mention a concrete consequence of such. At the same time, as previously mentioned, even when confronted with the risks, almost no one is willing to fully disconnect their device from the Internet since it would mean losing all the added functionalities of a smart TV. Clearly, this shows that consumers underestimate the potential harm that can arise from Smart TVs-related issues and highlights the need for enhancing security awareness. Lastly, the survey shows that when users are made aware of the risks, they are willing to adopt (and pay for) a privacy protector solution as long as it does not block Internet-dependent additional features.

IV. Traffic Analysis

In the following sections, the different contributions to assess the security and privacy of the HbbTV protocol will be outlined. The contributions are organized following a logical order; first, traffic analysis for Italian channels is performed to understand what type of communication happens between the Smart TV and the several broadcasters' servers and what data are transmitted via the HbbTV protocol. Results are then analyzed to spot potential security and privacy issues. Then, the analysis of the Italian scenario is complemented with the same testing methodology replicated over some French and German broadcasters to have a more comprehensive picture of HbbTV adoption and maturity in different European countries.

The following nine Italian channels have been selected: *Rai 1, Canale 5, Spike, RealTime, SportItalia, RDS (Radio Dimensione Suono), RTL, La7, Radio Kiss Kiss*; note that these channels belong to different broadcasters either public or private. Those have been selected with several parameters in mind: average audience share, broadcaster, offered content and, of course, enabled support for HbbTV. All the tests have been performed between February and May 2021.

[3]http://www.hides.com.tw/product_cg74469_eng.html
[4]TSDuck website: https://tsduck.io/

Such a contribution can be split into two main tests and procedures that were carried out, represented in Figure 3:

1) Listen and capture the traffic between the Smart TV and the servers to later analyze it to check what domains are contacted and in search of cookies and/or consumers' data.
2) Extract the initial URLs contacted by the Smart TV to launch the HbbTV application by replicating the DVB hijack attack explained in [8] for each of the nine Italian channels that are analyzed. Such URLs are opened on a Chrome browser while a Transparent Proxy is listening.

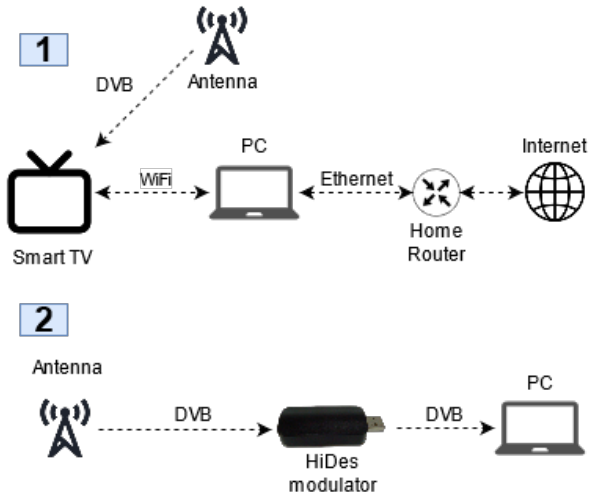

Fig. 3: Testing design for the two traffic analysis

In the following, more details for the two tests are presented.

A. Traffic Capture with Wireshark

The testing environment is composed of three main devices:

- Sharp Aquos LC-32Bi6E Smart TV with Android 9;
- PC with Ubuntu 20.04 operating system with Wireshark[5] installed;
- Samsung M5500 Smart TV.

The laptop is connected through the Ethernet interface with the home router and its Wi-Fi hotspot is enabled. The Smart TV is then connected to it. Since two of the analyzed channels (i.e., Mediaset and La7) do not receive any HbbTV app on the Android device probably because of a compatibility issue, it is necessary to perform the analysis for such channels on a Samsung device (where the HbbTV app is available and usable).

An instance of Wireshark is started on the Wi-Fi interface to which the Smart TV is connected to capture the traffic generated by and directed to the Smart TV. Traffic is collected for one hour using the following methodology adapted from the work by Ghiglieri et al. [11]:

1) Listen 15 minutes without any interaction to spot potential information being transmitted before user consent or user explicit action to enable the HbbTV app;

[5]Wiresharkwebsite:https://www.wireshark.org/

2) Give consent and interact for 20 minutes with the suggested buttons, different for each channel, to see what type of data is sent and if HTTP connections are spotted in the extra features offered by the HbbTV app;
3) Revoke user consent (if possible) and listen 10 minutes without any interaction;
4) Restore consent, change the channel, re-tune to the channel and listen for 15 minutes without any interaction.

For each channel taken into analysis, a factory reset of the TV is done so that there is no interference in the captured traffic.

Traffic analysis is automated, only for the Android device, through a bash script to have precise timings between the different periods aforementioned. Automation is possible by connecting the Android TV via the Android Debug Bridge (ADB) to the laptop and issuing key events directly from the script.

Traffic is collected in *.pcap* files which are then analyzed. To make the analysis process faster and more efficient, using Tshark[6], the command line version of Wireshark, the *.pcap* file is converted into a *.csv* file taking out the useful parameters for the analysis, e.g., *ip.src_host* (please refer to the Wireshark documentation[7]).

Given the *.csv* file, the following information is extracted:

- "domain": the contacted host;
- "occurrences": the number of requests to a specific host;
- "protocols": can be either HTTP and/or the different TLS versions depending on whether the communication is encrypted or not;
- "consent_status": tags identifying the four periods the testing phase has been divided into;
- "packet_number": the numeric ID of HTTP packets that will later be manually inspected (since unencrypted).

Additionally, the intent and common usage of the contacted service have been manually added. HTTP packets are inspected using the Wireshark interface looking for cookies, parameters, and API calls. This allows to see what type of service is requested and their respective answers.

All the domains that have the purpose equal to "Tracking" and that are found either before the consent is given or after it being revoked can signal misbehavior of the provider that does not wait for explicit user's consent before delivering some tracking and targeted content. At the same time, even the non-tracking domains found before accepting the privacy notice are not in line with the HbbTV protocol, i.e., no further communication should take place unless the user agrees.

Additionally, a time analysis complements this step. Such an analysis shows whether there is a certain frequency of requests while the other measure shows if such a pattern is reliable (high standard deviation means that the time window between requests varies a lot).

B. TSDuck Extraction & Transparent Proxying

The second test consisted of replicating a DVB/DMS-CC hijack attack using the TSDuck library and the UT-100c HiDes modulator. First, a complete scan of the Ultra High Frequency (UHF) channels was performed to identify to which one each of the nine analyzed Italian TV channels belonged. This task was carried out using the `tsscan` function of TSDuck.

```
|  Service: 0x218C (8588), TS: 0x0004 (4), Original Netw: 0x013E (318)          |
|  Service name: Rai 1 HD, provider: Rai                                        |
|  Service type: 0x01 (Digital television service)                              |
|  TS packets: 533,296, PID's: 11 (clear: 11, scrambled: 0)                     |
|  PMT PID: 0x01AC (428), PCR PID: 0x01B6 (438)                                 |
|-------------------------------------------------------------------------------|
|     PID  Usage                                          Access        Bitrate |
|   Total  Digital television service ..................... C    7,762,295 b/s |
|  0x01AC  PMT ............................................ C       15,923 b/s |
|  0x01B6  AVC video (1920x1080, main profile, level 4.0  C    6,799,402 b/s |
|  0x01C1  AC-3 Audio (ita, AC-3, 3/2 (L,C,R,SL,SR), @48  C      460,588 b/s |
|  0x01C2  MPEG-1 Audio (eng, Audio layer II, 128 kb/s,   C      137,853 b/s |
|  0x024C  Teletext (ita, Initial Teletext page) ........ C      112,803 b/s |
|  0x028A  MPEG-1 Audio (Oth, Audio layer II, 64 kb/s, @  C       75,192 b/s |
|  0x07D1  MPEG-2 Private sections (AIT) ................ C+       4,453 b/s |
|  0x07D2  MPEG-2 Private sections (AIT) ................ C+       4,453 b/s |
|  0x0BB9  DSM-CC U-N (MHP Object Carousel) ............. C+     100,082 b/s |
|  0x0BBA  DSM-CC U-N (HbbTV) ........................... C+      50,041 b/s |
|  0x0C1D  DSM-CC Stream Descriptors ..................... C+       1,499 b/s |
|          (C=Clear, S=Scrambled, +=Shared)                                     |
=================================================================================
```

Fig. 4: Parsed DVB Stream of Rai 1 channel

As mentioned in Section II, the URLs of the HbbTV applications are included in the DVB stream thus, by analyzing the broadcast stream, it is possible to extract them (and, eventually, alter them). Thanks to the `tsp` function, it is possible to capture the DVB stream passing through the HiDes modulator in Transport Stream (TS) format. The file is then converted into a *txt* format (for ease of reading) and the segments related to the AIT are extracted in binary form by specifying their respective Program IDs (PIDs). As described in Section III, the AIT contains the HbbTV information together with the start-up links of the applications. In Figure 4, a parsed DVB stream table for Rai 1 is presented. It shows the different substreams with their respective PIDs. It is important to note that, together with the standard Audio/Video streams, also the AIT is present.

The `application_type` parameter should equal to `0x0010` since it means that the information is related to HbbTV and the URLs for the HbbTV applications can be found. This is the file that an attacker should modify to replace the original application URLs with their (malicious) ones.

The extracted links for each of the nine Italian TV channels can be found in Table I. Such URLs are then opened in a Chrome browser following two approaches to mimic the Smart TV environment:

- Use an extension that emulates the built-in browser of Smart TV, nominally RedOrbit HbbTV Emulator 13[8];
- Manually change the User-Agent (UA) of the request using the UA of a real Smart TV, e.g., *HbbTV/1.4.1 (+DRM+MEDIA360;Samsung;SmartTV2017;T-KTSDEUC-1290.3;;)+TVPLUS+SmartHubLink Chrome*.

This was necessary since some of the analyzed links opened only in one of the two modalities. Each time, the browser's data and cookies are deleted.

To collect such traffic, a transparent proxy using mitmproxy[9], an interactive HTTPS proxy, is set up. The proxy

[6]Tshark: https://www.wireshark.org/docs/man-pages/tshark.html
[7]Wireshark documentation: https://www.wireshark.org/docs/dfref/

[8]https://chrome.google.com/webstore/detail/redorbit-hbbtv-emulator/mmgfafehampkahlmoahbjcjcmgmkppab?hl=en
[9]https://mitmproxy.org/

TABLE I: Extracted HbbTV start up links

Channel Name	Link
Sportitalia	http://www.kineton.it/hbbtv/sportitalia/sportitaliachannel/index.html
RDS	http://hbbtv.rds.radio
RealTime	http://discovery.castoola.tv/realtime
RTL	https://cdn.rtl.it/hbbtv.rtl.it/rtlchannel/index.html
Rai 1	https://tivuon-hbbtv.tivu-alchemy.net/index.html?configuration=DTTprod https://tivuon-hbbtv-lativu.tivu-alchemy.net/index.html?configuration=prod https://www.raiplay.it/hbbtv/launcher/RemoteControl/index.html?delivery=2 https://www.raiplay.it/hbbtv/RaiPlay2020/index.html
Spike	http://www.kbbtv.tech/viacom/viacomchannel/index.html
Canale 5	http://hbbtv.mediaset.net/app/mplayhbbtvgold/backdoor.shtml http://hbbtv.mediaset.net/app/mplayhbbtvgoldzoo/dev/index.html https://infinity-tivuon.infinitytv.it/hbbtv/index.html https://mhptivu.mediaset.net/app/mplayhbbtvtivu/index.html https://tivuon-hbbtv-lativu.tivu-alchemy.net/index.html?configuration=prod
La7	https://ht.la7.it/index.php
Radio Kiss Kiss	http://www.kineton.it/hbbtv/kisskiss/kisskisschannel/index.html

was set to work on the same machine as where the extracted links were opened from. The mitmproxy CA certificates were then installed in the Chrome browser to make it work even for HTTPS traffic. After this procedure, plaintext traffic was intercepted. Traffic was saved in text files. These files were later manually inspected for analysis using the mitmdump tool available with mitmproxy.

The methodology adopted resembles the approach taken in Section IV but only 30 minutes of plain traffic were captured: listen 10 minutes without interaction, accept the privacy notice and interact for 10 minutes, revoke consent, and listen for 10 minutes. The contacted networks domains and IP addresses together with the adopted protocol (HTTP or HTTPS), the known purpose of the service, and eventual tracking cookies are extracted.

C. Results

Since the results of both tests aimed at showing what type of information is exchanged between the smart TV and the broadcasters and whether some security and privacy risks might arise, they are aggregated and then presented in the following section.

All nine Italian channels show connections to at least one tracking service (with possible profiling cookies set) even before the user has a chance to decide whether or not to accept the privacy notice. Table II, summarizes which channels connect to which known tracking service before consumers' consent and also presents the results from the analysis in a synthesized manner.

It is also noteworthy that two out of nine channels make POST requests to an AWS API "/audiencesavemessage" with the user ID, model, and device brand as parameters of the body for later stage profiling of the consumer.

In general, cookies appear to have on average long expiration dates ranging from 2021 to 2048. Given that such cookies can be used to track, with potential linkage to other data, a user's browsing behavior, such persistence could pose privacy risks. Additionally, given the adoption of HTTP by some services, such cookies are sent plaintext and if an attacker is sniffing on the communication channel where the information is sent, he will be able to intercept it. If cookies contain sensitive information, that could be later used to mount a targeted attack.

As an example, RealTime sends plaintext cookies identifying the geographical location and Internet Service Provider (ISP) of the consumer even before their explicit consent.

RDS does not present any privacy policy when accessing the HbbTV app for the first time and the user starts to be profiled without having provided any consent. Such a policy is nowhere to be found even in the sub-menus of the application. Rai, on the other hand, despite not showing any privacy policy when accessing the HbbTV app for the first time, allows the consumer to consult it at an external link that can be found in a sub-menu of the app.

Some channels offer the possibility to revoke the consent given to data processing at a later time while others do not. Since deleting cookies in the TV is not immediate (in fact, such a procedure typically requires a factory reset) and given their long expiration dates, it would be fair to give the consumer the possibility to revoke the consent to data processing. Specifically, three channels out of the nine examined do not allow the user to revoke consent.

Moreover, the channel RTL, despite presenting the possibility to revoke consent, in reality, profiling and identification cookies are not deleted so requests to any tracking services are still made.

A methodology that was found to be widely used by several broadcasters is the "tracking pixel"[10]; this technique consists of tracking consumer behavior by uploading a 1x1 pixel image when the user visits a web page or opens a certain content. Given its small size, it is invisible to the naked eye but can provide a lot of data to advertising or analytics agencies that can infer user preferences in this way. In particular, SportItalia, RealTime, and Spike adopt this methodology by returning 1x1 pixel GIF89a objects in requests.

Finally, some channels perform periodic requests to check if the consumer is still watching. This methodology was also reported by Ghiglieri et al. in 2015 when they conducted a similar study in Germany [13]. Additionally, what can be noticed are periodic requests to tracking services. Almost all channels show frequent requests to such profiling domains (on average around every minute); for example, SportItalia makes requests to smartclip around every 70 seconds while

[10]https://en.ryte.com/wiki/Tracking_Pixel

TABLE II: Result of the analysis of nine Italian Channels

Channel Name	Tracking Services	Privacy Notice shown	Possibility to Revoke Consent	Tracking Pixels	Periodic Requests
Sportitalia	Smartclip POST to "/audiencesavemessage"	Yes	Yes	Yes	Yes
RDS	Google Tag Manager Google Analytics Facebook	No	No	No	Yes
RealTime	Google Analytics discovery-log-castoola GETs to */log-audience*, */log-view* and */log-ad*	Yes	Yes	Yes	Yes
RTL	Google Analytics DoubleClick	Yes	Yes	No	Yes
Rai 1	DoubleClick	No	No	No	Yes
Spike	Google Tag Services SecurePubAds	Yes	Yes	Yes	No
Canale 5	tags.tiqcdn.com (Tealium Inc.)	Yes	Yes	No	Yes
La7	tags.tiqcdn.com (Tealium Inc.) SecurePubAds cdn.permutive.app	Yes	Yes	No	Yes
Radio Kiss Kiss	POST to "/audiencesavemessage"	Yes	No	No	No

RDS contacts Google Analytics around every 14 seconds. In all cases, a low standard deviation signifies that some recurrent pattern can be identified.

D. German and French Landscape

For further comparison, the traffic analysis was replicated in Germany and France to reveal any differences or similarities in the adoption of the HbbTV protocol in these countries.

For foreign channels, the procedure is similar but simpler. Only the first test was carried out and only half an hour of traffic was analyzed for each channel taken into consideration. The test involving the URL extraction and the transparent proxy was omitted.

1) ***Germany****:* In the context of the German analysis, five different channels are selected from a subset of the analyzed channels in the 2016 paper by Ghiglieri and Tews [11]. Such channels are: Arte, Anixe, SWR BW, HSE Live and ZDF. Based on the level of privacy invasion, they divide the analyzed channels into four groups. For our study we selected one representative channel from each group, except for HSE, to see how the situation changed over five years.

All five channels present privacy policies. However, in only two of them, nominally, Arte and HSE, the privacy policy is shown as soon as the user arrives at the respective channel while in the other three this has to be searched in the submenu of the HbbTV application. All five channels offer to the user the possibility to revoke the consent but two of them, SWR and ZDF, offer the user only the possibility to revoke consent to the "counting pixel" technique.

Unlike the Italian scenario, all five channels adopt the "counting pixel" technique and there are fewer third-party tracking services. What is observed is essentially in-house tracking, unlike the Italian channels that rely on larger services such as Google Analytics or Smartclip to collect information about consumers.

We highlight the differences between the two approaches: the first approach has the following advantages: since, being well-defined and well-established services, consumer data are presumably treated in full respect of privacy and with adequate security techniques both in the communication of profiling data and in their storage. Conversely, adopting local services could present potential security issues for consumer data caused by any bugs that are never discovered. However, the use of larger third-party services could allow for the aggregation of various information and data from different websites/applications thus leading to more accurate targeted content/advertising even possibly on different user devices (not limited to smart TV).

Greater use of HTTP, that is, of unencrypted traffic, was also found compared to what was observed in Italy. In particular, this has led to two serious problems: two channels of the five analyzed, Arte and HSE, allow users to log in using their credentials linked to an account. Associated with this account can be found sensitive information such as an address, credit card data, etc. The credentials are not encrypted but they are sent in plain text allowing an attacker to capture them and use them later for malicious purposes. For HSE the problem was already explicitly mentioned in the previously mentioned paper but no solution was put in place [11].

2) ***France****:* the adoption of the HbbTV protocol is not yet in full swing and few channels have presented the possibility to interact with such applications. Two channels were analyzed, Arte and NRJ12, with the former being in common with the analysis in Germany.

In particular, both channels analyzed show the privacy policy before starting to interact with the application and allow the withdrawal of consent. Arte supports the "counting pixel"

technique. Also, in this case, the tracking is entrusted to minor services and tends to be local, unlike what has been found in the Italian context.

As in Germany, Arte presents the possibility of logging in using a code sent to the user via email. However, this code is sent in plain text, without being encrypted, allowing possible attackers to intercept it.

V. HbbTV's Risk Awareness Survey

In the context of this project, an anonymous survey was also conducted on a sample of 100 individuals aimed at highlighting the level of awareness of users in the context of the use of smart TVs and the risks associated with them. The sample was reached by spreading the URLs of the survey to indirect contacts of the persons involved in the study in order not to introduce bias. The only requirements to participate in it were to be older than 18 years old and watch Italian television. The language used to formulate the questions was Italian.

The survey takes inspiration from the 2015 survey defined by Ghiglieri et al. [12] to assess the level of consumers' awareness concerning privacy and security issues with HbbTV. Such an initial questionnaire was conducted in Germany, the leading country for the adoption of HbbTV, and, as reported in Section III, it confirmed a generally low level of awareness of privacy-related risks in this context. It also showed that, even if being exposed to the risks deriving from the uncontrolled use of Smart TVs, consumers are not willing to fully disconnect their devices from the Internet so as not to lose the extra features offered. Thus, a solution that provides security (against ill-intentioned attackers), privacy, and functionality is needed.

The survey that is described in the following proposes again the same approach adopted by Ghiglieri et al. but after six years and in a different country. The idea is to see whether with the more widespread adoption of HbbTV and Smart TVs, the situation of users' awareness has improved and whether Italian consumers have a different approach towards their privacy than German ones.

The ethical guidelines defined by the affiliated organization when conducting such surveys were fully respected and the approval from the ethical committee was received before sending the form to participants.

The study was built using SoSci Survey[11], a German platform that allows for heavy customization of sections.

Figure 5 shows the structure of the survey. What follows is a more detailed description of what each section means and what information it should gather from participants.

1) *Introduction*: participants are generally informed about the topic of the survey. Some details are omitted not to impact their answers. Details on the questionnaire, such as its anonymity and the duration, are included;
2) *(Smart) TV Demographics*: participants are asked whether they own a TV or eventually a Smart TV. Those who do not own a Smart TV are asked if they would like to own one. Only the ones who own a Smart TV, or want to buy one, continue to the next section. The others are redirected to "Final Questions";
3) *Awareness of Security and Privacy Risks*: participants are asked whether they are aware of security and privacy risks linked to Smart TVs; if yes, they should enumerate them and list eventual measures to counteract these risks;
4) *HbbTV Demographics*: participants are asked whether they ever saw HbbTV notifications and if they are aware of how such protocol works;
5) *Risk Assessment of Scenarios*: participants are given eight different risky scenarios (one per page) in random order. For each, they are asked to give a score based on how critical they think such a scenario is. The score goes from 1 (very low risk) to 5 (very high risk). Additionally, they are asked to justify their rating. The full list of scenarios can be found in Appendix A;
6) *Privacy Policy related Questions*: participants are asked whether they read privacy policies when accessing digital services and if they were ever shown with such banners asking for data treatment consent when watching TV;
7) *Selection Grid*: participants are presented with a table showing five modalities, reported in Appendix B, of connecting the TV to the Internet with different security levels and extra functionalities available. They are asked to vote for their preferred one and also mention all the desired features that a tool to enforce security in this context should have;
8) *Final Questions*: participants are asked their age, gender, and their area of expertise to have insights on the sample who answered such a survey.

100 participants completed the survey. Out of these, 70 answered that they either possess a Smart TV or would be willing to buy one. Those continued the survey while the others were redirected to the Final Questions section. Accordingly, from now on, the total number of responses will be considered 70.

The second step, as reported in the above-mentioned list, consisted of assessing whether participants are aware of security and privacy risks. 40 participants (57%) did not mention any risk confirming an alarmingly low level of awareness; 20 (29%) participants identified only one risk; the remaining (14%) identified either 2, 3, or 4 risks per person. 44 text answers are analyzed and clustered together in macro-answers. The most often mentioned risk is linked with privacy and consists of tracking and profiling (18 participants). The second most (10 participants) is data and credential leakage due to unencrypted traffic or unreliable services. Lastly, only 16 participants (23%) were able to mention at least one security measure to prevent such risks with Firewalls being the most mentioned answer (7 participants).

39 participants (56%) reported having seen HbbTV notifications while using their Smart TVs. However, only 10 (14%) correctly mentioned that such protocol is both a combination of standard broadcast signal (with the delivery of the URL through the DVB stream) and broadband communication for the delivery of Internet-based content.

Out of the total of 70 participants, 65% stated that they never

[11] https://www.soscisurvey.de/

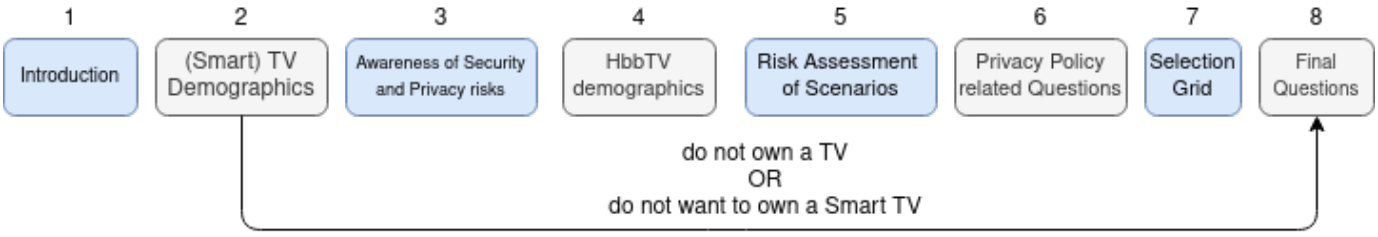

Fig. 5: Survey structure

or rarely read the privacy policy presented when accessing a digital service for the first time and 81% stated that they did not read such privacy policy presented while watching a TV channel. In addition, only 27 respondents (38%) were able to mention at least one type of data that could potentially be collected while using an Internet-connected Smart TV. The most mentioned collected type of data is ”viewing times and preferences” (21 times) while the second most common one is ”personal information”, e.g., email addresses, date of birth, etc. (5 times). Surprisingly, only 3 participants mentioned that geographical location can be collected too.

However, these results seem to be at odds with the Risk Assessment section. When presented with broadcasters being able to store and analyze usage habits (scenario n.4), participants assigned it an average risk value of 2.81. However, all it took was adding that the information is used to show personalized advertising (scenario n.2), to raise this value to 3.1. In addition, the fact that broadcasters might aggregate data from other services and that they might sell information to third parties (scenario n.8) was given a risk score of 3.47. This highlights how consumers are concerned about how their data is being used for profiling but there is little awareness and unwillingness to inform. This puts the responsibility and duty in the hands of issuers to ensure the consensual handling and collection of their consumers' data. For the complete risk scoring assigned to each scenario, refer to Table III.

TABLE III: Risk Score for the eight Scenarios

Scenario	Average Risk Score	Standard Deviation
1	2.67	1.06
2	3.1	1.14
3	3.43	1.10
4	2.81	1.19
5	3.69	1.14
6	3.99	1.16
7	3.09	1.04
8	3.47	1.01

Lastly, participants were asked how they would like to connect their Smart TV taking into consideration security aspects, functionality, required effort, and costs. For the five modalities please consult Appendix B. 26 participants (37%) preferred the cheaper solution that requires some configuration to secure the Internet communications of the Smart TV. Only 14 participants (20%) voted for the option of connecting the device without further security measures while 45 respondents (64%) would be willing to adopt some solution to improve security. The most mentioned features to keep into consideration when designing a tool to provide security to smart TVs are ”ease of use” (18 times) and ”highly customizable” (17 times) with eventually two different options for both experienced and not users.

VI. Discussion

Both the results presented in Section IV and Section V highlight a problematic immaturity in the context of HbbTV adoption in Italy on either broadcasters' and consumers' sides.

Particular care is put in trying to identify problems concerning consumer's privacy. Explicit references to General Data Protection Regulation (GDPR) articles can be found throughout the present section. GDPR has been effective since 2018 to protect European citizens' data and privacy [2]. Some references will also be made to the Italian Garante della Privacy[12] given the geographical setting of our analyses.

Results of the technical analysis show the negligent behavior of broadcasters offering HbbTV applications. Fair treatment of users' data is not always guaranteed as, for example, demonstrated with the connection to tracking services before their consent. Particularly, several violations of GDPR are found. Tracking before the user has expressed her consent is in contrast to the ”Conditions for consent”, of the European Regulation and with the guidelines on cookies by the Italian Garante della Privacy. Additionally, withdrawing consent, with the consequent deletion of data, should be possible and as easy as giving it according to the same statement of GDPR. This is not the case for some Italian channels as shown in the results that either do not allow for consent revocation or ask the user to directly contact their dedicated office. The absence of the privacy notice when accessing the HbbTV application for the first time, as for RDS, signals a violation of transparent information communication of data handling and the provision of correct information to the data subject while the ”hidden” policy, in a sub-menu of the app, as for Rai, violates again the same principles of transparency.

The incorrectness, incompleteness, and non-transparency of the privacy policies presented by the different channels, do not help consumers' in understanding what information is collected about them and how this is then processed. This is in contrast with what is stated by the principles relating to the processing of personal data, to the provision of the correct information and transparency.

[12]https://www.garanteprivacy.it/

On the other side, users seem unaware of the potentially privacy-invasive tool they have in their houses. When asked what are the risks associated with the use of Smart TVs or HbbTV, only a few were able to mention at least one. Despite this lack of awareness, when confronted with potentially risky scenarios, users seem to be extremely concerned about their data and personal information, e.g., viewing preferences, even more than what was reported by Ghiglieri with his survey on German consumers [12]. This worryingly highlights deep illiteracy on this topic that might cause security and privacy problems to unaware users.

The consequences of having such a "relaxed" approach towards security in the Smart TV context, and more specifically in the HbbTV one, could lead to several issues. In the following, two of them are mentioned. The advent of online shopping apps delivered through HbbTV is already a reality in Germany and will not take long to reach the whole of Europe. To purchase online, the consumer is required to insert sensitive data such as credit card information, billing address, name, etc. The wrong handling of this data might lead to severe security problems, as theft of personal information and credentials. As reported in Subsection IV-D, the use of plain HTTP with no encryption of data when logging in to such services poses a serious threat to the consumer. HbbTV's new functionalities need to catch up with security before being ready to be deployed Europe-wide.

As a second example, dynamic advertising is introduced. Static delivery of adverts, over standard broadcast signal, is expected to be replaced by dynamic Ad insertion (and replacement) where adverts are delivered over the Internet and dynamically inserted in commercial breaks [25]. Directly linked to this change, is the possibility of having targeted ones. Addressable Television (ATV) will combine the advantages of traditional TVs with the benefits of digital marketing [15].

Several issues arise from the adoption of dynamic and targeted advertising. On one hand, there is an ethical issue, often referred to as autonomy. Algorithms, with their persuasive power, can *nudge the behavior of data subjects and human decision-makers by filtering information* [18]. In the HbbTV context, this reflects in consumers being nudged to buy certain products deemed suitable for them thus posing their autonomy at risk.

At the same time, adverts are not typically encrypted. This might lead to MITM attacks that replace the location of the advert to load a different media. The situation is worsened by the so-called "interactive ads" that allow users to interact with them to directly purchase online. No sanitization or additional checks on the advert link is being done allowing any ill-intentioned party to replace the legit advert with malicious content.

VII. Proposed Solution: HbbTV Blocker

To mitigate the security and privacy issues described in Subsection II-C, Subsection IV-C and discussed in the previous section, a security tool, the HbbTV blocker, is designed and developed. Its implementation consists of a gateway on top of a Raspberry Pi[13] that intercepts traffic to and from the Smart TV and a graphical interface.

In the prototype, the Raspberry is directly connected through the Ethernet interface to the home router while the Smart TV is connected to the Wi-Fi hotspot of the former device. In such a way, all the traffic directed to the Smart TV passes through the Raspberry. The prototype was designed to work on the nine Italian channels mentioned in Section IV.

A python script intercepts the DNS queries using the *pyshark* library[14], a wrapper for tshark, and filters them using the documented Wireshark filters. If the contacted domain matches against specific string patterns defined for each of the nine channels from the TSDuck extracted URLs, the current channel variable is set to the name of the matched channel.

Another python script checks the current channel and enforces the corresponding blacklist. All nine blacklists (one per channel) are designed starting from traffic captures done in Section IV and they contain sub-strings of the contacted tracking and analytics domains. To enforce those lists, *iptables*, a program that permits to configure IP packet filter rules of the Linux kernel firewall[15], is used. For each of the strings contained in the blacklist, a new iptables rule is added to block specific traffic. The format of the rule is the following `iptables -A INPUT -m string --string "domain" --algo bm --to 65535 -j DROP` with "domain" being replaced with the current string.

The approach of using blacklists is preferred over whitelists. Whitelisting would require defining a specific set of allowed domains for each of the applications used by the user. The installation of a new app that is not contemplated in the list, would mean that the traffic is automatically blocked since it is not included in any whitelist. Additionally, defining whitelists for huge services, like Google or Amazon AWS, is a nontrivial task considering a large number of domains and subdomains in use. Thus the blacklist approach is deemed more scalable if new applications are installed, and easier to manage. This solution can be found in other widely used tools, for example, the Pi-hole[16], to block unwanted traffic.

The consumer has access to a simple graphical dashboard where different options are available; she can decide the "traffic behavior" of each channel, block all HTTP traffic, have a look at the number of blocked requests per channel, and, for advanced users, upload customized blacklists. The suggestions of having an easy-to-use and customizable tool, as reported in Section V, were kept in mind when designing the dashboard to make it suitable for all users. Three traffic behaviors can be selected:

- **Allow all**: all the traffic, including tracking and profiling, passes through the gateway;
- **Block tracking**: enforces the above-mentioned blacklists designed to block tracking and analytics domains;
- **Block all**: blocks all traffic independently of its nature.

[13]https://www.raspberrypi.org/
[14]Pyshark documentation: http://kiminewt.github.io/pyshark/
[15]Iptables documentation: https://linux.die.net/man/8/iptables
[16]Pi-hole website: https://docs.pi-hole.net/

The *Block all* modality and the possibility to block all HTTP traffic are enforced using specific iptables rules too. It is to be mentioned that those two options might hinder user experience since they also block traffic unrelated to HbbTV. The user is presented with an informative alert when trying to turn on such options.

There is a shift of responsibility which is both in the hands of the developer and the consumer of the product. The developer has to maintain the tracking blacklists updated and complete. However, given the previous considerations, if a user wants to block HbbTV traffic or wants to block specific domains, it will be their responsibility to possibly hinder the Smart TV functionalities or upload the correct blacklist. Again, Pi-hole was taken as inspiration since most of its available blacklists are written by the community of users.

The reason why a gateway was chosen over the proxy used by Ghiglieri and Tews in their Privacy Protector tool [11] is dual. On one side, they adopted mitmproxy as a transparent proxy which requires its CA to be installed in the Smart TV to capture HTTPS traffic. To install such certificates, root access is required. Unfortunately, gaining root privileges on a Smart TV is not an easy task since no documentation is available and every different model has its custom procedure (if any). Their approach worked fine a few years ago when mostly only HTTP was used in HbbTV communication but, with the advent of HTTPS, it requires a lot of effort.

Additionally, even with simpler proxies that do not act as Men-in-the-Middle but simply collect traffic headers, there is a problem when setting those on Smart TVs. Android proxy settings apply only to browser traffic. Other applications' traffic, including HbbTV, does not pass through the proxy for security reasons. To bypass this limitation and make all traffic flow through the proxy, root access is required. In the end, the gateway approach was deemed the best in terms of universality and adaptability to different models, brands, and operating systems of Smart TVs and considering the requirement of "ease of use" highlighted in Section V.

To complete the discussion, some performance analysis of the tool is conducted. For such tests, a Raspberry pi 4 with 8GB of RAM was used. The collected values are: number of packets exchanged per second and the number of bytes per second, CPU, and RAM usages. Such measures are collected by a script every 10 seconds to avoid the observer effect. As described above, the Raspberry is connected through the Ethernet interface to the router and the Smart TV is connected to its Wi-Fi hotspot. Then for 30 minutes, the TV is used to navigate the channels and use their respective HbbTV applications. The results are reported in Table IV both for when the Raspberry is idle, i.e., the tool was not running, and when the tool is active. Despite showing an increased usage of resources, results indicate that the tool is not highly demanding therefore can be used in the envisioned scenario without problems. Additionally, the relatively low number of exchanged packets for HbbTV applications makes the tool suitable for its intended task of blocking requests.

TABLE IV: Performances Tests' Results

.	Idle	Running
Average CPU usage (%)	3.70%	35.66%
Average RAM usage (%)	11.14%	14.66%
Average Packets per second	0	256 (max 7433)
Average KBytes per second	0	176 (max 7600)

VIII. Limitations and Future Works

In this section, the main limitations of the present work are presented. There are two main directions where the present study can be expanded and that represent the two main limitations.

On one side, the next step could be rooting one Smart TV to install the required certificates for a transparent proxy to work. This will allow for the capture of plain text traffic directly from the Smart TV thus having a more detailed picture of what communications are in place between broadcasters and consumers, i.e., what data are transferred and which domains are contacted. Such a procedure, since heavily discouraged by vendors, is not documented and only with a deep analysis of old Smart TVs' firmware versions, could it be done. Additionally, the necessary steps change for every vendor and model of devices. Because of time constraints and only one device available, the rooting of the Smart TV was not executed.

On the other hand, the security tool that has been developed could be expanded to comprise additional features to make it marketable. This would require some usability testing so that users can test whether new options are required and if the "ease of use" requirement is satisfied. As discussed above, the *Block all* feature might hinder Smart TVs' functionalities; there is no straightforward way to understand when the user exits the "standard" TV channels app switching to a different one, e.g., YouTube. For such a reason, she will still experience all traffic being blocked. However, as reported in Section V, since consumers do not want to disconnect their Smart TV not to lose the extra functionalities, it is considered probable that such an option will be only rarely selected in favor of blocking only tracking.

The description of the developed tool is to be intended as an initial step in its development and it should be considered as a prototype. Therefore, the use of different network interfaces, e.g., USB adapter, should be evaluated to make it more reliable and less impacting on performances. Additionally, the performance results reported in Section VII are not to be intended as complete but they only give a glance at the resources' usage.

A minor limitation that is worth mentioning concerns the number of Italian channels analyzed. Only nine channels out of the ones that offer HbbTV functionalities were selected based on different parameters, e.g., audience share. To complement the traffic analysis, additional channels could be included to have a complete overview of all HbbTV traffic in Italy.

IX. Conclusions

The present study shows that consumers are exposed to severe privacy issues in the context of HbbTV-enabled de-

vices through different contributions to the existing literature, both traffic analysis and a consumer survey. All the Italian broadcasting stations that have been analyzed show some negligence in handling users' data in compliance with the current regulations by adopting heavy profiling and tracking (third-party) services. The situation seems to not have evolved from five years ago when most of the related work on the security of the HbbTV protocol was published. The risk is still high also in two other European countries, France and Germany.

Additionally, consumers seem to have a worryingly low level of awareness for risks linked to Smart TVs and HbbTV. However, they show high concern when confronted with potential issues highlighting the need for a security tool that enhances the security level while improving understanding of such risks.

The present paper defines a solution to mitigate the security and privacy issues arising from the unregulated adoption of HbbTV, HbbTV blocker. It takes into consideration the need for high customization and ease of use expressed by consumers. Given the adaptability of such a tool to different contexts, a next expansion could include different modules to secure different smart-home appliances, e.g., smart refrigerators.

References

[1] Deployments | HbbTV. Last accessed: 2021-07-05. URL: https://www.hbbtv.org/deployments/.

[2] General data protection regulation (GDPR) – official legal text. Last accessed: 2021-07-05. URL: https://gdpr-info.eu/.

[3] HbbTV: What is it and how does it work? - BSG SA. Last accessed: 2021-02-17. URL: https://bsgroup.eu/hbbtv-what-is-it-and-how-does-it-work/.

[4] Hundreds gather in Madrid for anti-mask protest. URL: https://www.bbc.com/news/av/world-europe-53802226.

[5] Overview | HbbTV. Last accessed: 2021-05-26. URL: https://www.hbbtv.org/overview/.

[6] ETSI TS 102 796 V1.5.1, Hybrid Broadcast Broadband TV. Standard, ETSI, September 2018. URL: https://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.05.01_60/ts_102796v010501p.pdf.

[7] Hbbtv 2.0.3 specification. Standard, HbbTV Association, October 2020. URL: https://www.hbbtv.org/wp-content/uploads/2020/10/HbbTV-SPEC-00525-HbbTV-SPEC-00515-008-hbbtv203_2020_10_14.pdf.

[8] Massimo Bozza. Catch the wave - Hijack HbbTV. Codemotion in collaboration with LVenture Group and LUISS EnLabs, 2019. URL: https://www.youtube.com/watch?v=2yeahbhPu9o.

[9] Eric Chong. The growing trend of coin miner JavaScript infection. Last accessed: 2021-02-05. URL: https://www.fortinet.com/blog/threat-research/the-growing-trend-of-coin-miner-javascript-infection.html.

[10] M. Ghiglieri. I Know What You Watched Last Sunday - A New Survey Of Privacy In HbbTV. 2014. URL: https://www.techrepublic.com/index.php/resource-library/whitepapers/i-know-what-you-watched-last-sunday-a-new-survey-of-privacy-in-hbbtv/.

[11] M. Ghiglieri and E. Tews. A privacy protection system for HbbTV in Smart TVs. In *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, pages 357–362, 2014.

[12] M. Ghiglieri, M. Volkamer, and K. Renaud. Exploring consumers' attitudes of smart TV related privacy risks. In *Human Aspects of Information Security, Privacy and Trust*, Lecture Notes in Computer Science, pages 656–674. Springer International Publishing.

[13] M. Ghiglieri and M. Waidner. HbbTV security and privacy: Issues and challenges. 14:61–67.

[14] Martin Herfurt. Security concerns with HbbTV. URL: https://www.researchgate.net/publication/277007241_Security_concerns_with_HbbTV.

[15] S. Leffel, K. Muller, et al. Addressable TV Advertising. *smartclip*, 2020. URL: https://smartclip.tv/addressable-tv-advertising-white-paper/.

[16] J. Matejka, P. Podhradský, and J. Londák. Security manager for hybrid broadcast broadband architecture evolution. In *2016 International Symposium ELMAR*, pages 57–62. ISSN: 1334-2630.

[17] Benjamin Michéle. Broadcast. In *Smart TV Security: Media Playback and Digital Video Broadcast*, SpringerBriefs in Computer Science, pages 35–80. Springer International Publishing. URL: https://doi.org/10.1007/978-3-319-20994-4_3.

[18] B.D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi. The ethics of algorithms: Mapping the debate. 3(2). URL: https://doi.org/10.1177/2053951716679679.

[19] C. O'Connor and M. Murphy. Going viral: Doctors must tackle fake news in the covid-19 pandemic.

[20] Yossef Oren and Angelos D. Keromytis. From the Aether to the Ethernet - Attacking the Internet using Broadcast Digital Television. In *23rd USENIX Security Symposium (USENIX Security 14)*, pages 353–368, San Diego, CA, August 2014. USENIX Association. URL: https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/oren.

[21] Saint Girons R. Hbbtv country review. *HbbTV Symposium*, 2016. URL: https://www.hbbtv.org/wp-content/uploads/2017/01/Regis-Saint-Girons-Country-Review-HbbTV-Symposium-2016.pdf.

[22] REVULN. The TV is watching you: Samsung 0-day, 2013. Last accessed: 2021-01-18. URL: https://vimeo.com/55174958.

[23] Jan Rüth, Torsten Zimmermann, Konrad Wolsing, and Oliver Hohlfeld. Digging into browser-based crypto mining. In *Proceedings of the Internet Measurement Conference 2018*, IMC '18, page 70–76, New York, NY, USA, 2018. Association for Computing Machinery. URL: https://doi.org/10.1145/3278532.3278539.

[24] L. SeungJin and S. Kim.Smart. Smart tv security - #1984 in 21st century. InCanSecWest, 2013. URL: https://cansecwest.com/slides/2013/SmartTV%20Security.pdf.

[25] Eric Shiffman. HbbTV: How addressable TV is implemented in the EU. Last accessed: 2021-03-23. URL: https://www.spotx.tv/resources/blog/product-pulse/hbbtv-how-addressable-tv-is-implemented-in-the-eu/.

[26] J. Varmarken, H. Le, A. Shuba, A. Markopoulou, and Z. Shafiq. The tv is smart and full of trackers: Measuring smart tv advertising and tracking. *Proceedings on Privacy Enhancing Technologies*, 2020:129–154, 04 2020.

Appendix A
Eight Risky Scenarios

1) The channel you are watching gets information about when and how long you watch it. For broadcasters with multiple channels (for example, Canale 5 and TGCOM24), there is the possibility that the information from both channels will be merged.
2) Your usage habits (i.e. what you use your Smart TV for, when and how often) are stored by the TV broadcasters. The information collected about you is analyzed to show you personalized (i.e. tailored to you) advertising.
3) Your usage habits (i.e. what you use your Smart TV for, when and how often) are stored by the TV broadcasters. The purpose and the way such data is stored are not explicitly stated and not certain.
4) Your usage habits (i.e. what you use your Smart TV for, when and how often) are stored and analyzed by the TV broadcasters.
5) A TV broadcaster offers you the possibility to direct home shopping of the item that is being advertised by simply entering your credentials and credit card information on its website.
6) A TV broadcaster offers you the possibility to direct home shopping of the item that is being advertised by simply entering your credentials and credit card information on its website. It cannot be ruled out that such information is only received by the broadcaster itself.

7) TV broadcasters may rely on and aggregate data about you coming from bigger services, such as Google and Facebook, to better tailor their content to your preferences.
8) TV broadcasters may rely on and aggregate data about you coming from bigger services, such as Google and Facebook, to better tailor their content to your preferences. This might be also used to show you targeted advertisements. It cannot be ruled out that such information is not sold to other third parties.

APPENDIX B
SELECTION CHOICES IN SURVEY

TABLE V: Five modes of connecting Smart TV to Internet

Modality	The Smart TV is connected to the Internet without further precautions	The Smart TV is not connected to the Internet at all	The Smart TV is not connected to the Internet and is also used as an external monitor for a PC/laptop	The Smart TV is first secured by you via a protection software before you connect it to the Internet.	The Smart TV is secured via preconfigured protection software before you connect it to the Internet
Internet Features	No restrictions	None	Only standard functions of the laptop / PC incl. media library / no updates	No mandatory restriction	No mandatory restriction
Risk	Potential risks	None	None	None, limited	None, really limited
Additional Effort	None	None	Laptop/PC must be configured and connected	one-time 15 min. for configuration of the protection software	None, since preconfigured
Additional Cost	None	None	None	one time 20€	One time 40€