I Know What You Watched Last Sunday A New Survey Of Privacy In HbbTV

Marco Ghiglieri
Security in Information Technology
Technische Universität Darmstadt
Darmstadt, Germany
marco.ghiglieri@ec-spride.de

***Abstract*—Smart entertainment devices like Smart TVs, game consoles and Blu-ray players are getting more and more popularity in households. They have at least one interface that can establish a connection to the Internet. Especially Smart TVs and set-top boxes with network connection enable many third parties to provide real-time information and content directly to the consumers, for example additional information about a current TV program. HbbTV is one standardized way to deliver content to the devices via Internet. However, each vendor has additionally an unstandardized system to bring interactive content to the consumers. Most of these techniques are based on current web technologies. In case of HbbTV the content is presented on a transparent web site overlay over the current channel. As we have shown in former publications tracking techniques well-known from web sites can be used to gather the consumers' viewing behavior on Smart TVs without actively using an HbbTV application.**

In this work, we show how the current state in HbbTV has changed compared to our first survey in 2012. A new scenario where the listening behavior of consumers' while listening to a radio channel can be gathered is introduced. Furthermore, we explain our open source protection implementation which covers both scenarios effectively. Finally, we discuss some threat scenarios with regard to the wrong usage of this data.

I. Introduction

A. Smart Entertainment Devices

In the near future, entertainment devices with network functionalities like Smart TVs or set-top boxes will be in every household. For example, the worldwide shipments of Smart TVs are rising from 52 million in 2011 to expected 141 million units in 2015, which will be around 55 percent of global television shipments [1]. All devices are wireless or cable-bound network enabled. Thus, the easiest way to connect them to the home network is via Wi-Fi, where the consumer has only to enter the Wi-Fi key. The home network is often connected to the Internet. This leads to highly interactive functionalities, e.g., video on demand services, additional information to the current program and online games.

Hybrid Broadcasting Broadband TV (HbbTV) is a standardized technique covering video on demand and information services for Smart TVs and set-top boxes. It uses web technologies like HTML[1], CSS[2] and JavaScript to bring the content to the consumers. Technically, a web site with transparent background over the current channel is delivered to the devices. According to [2] a worldwide usage of HbbTV is discussed. Europe has the highest coverage as of today. In Germany, 89% of the sold Smart TVs have the HbbTV functionality in 2013.

As we have shown in former publications [3] and [4] well-known web tracking methods can be implemented in HbbTV to gather the viewing behavior without knowledge of the consumers, i.e., consumers are not aware of the tracking methods used in background while watching TV. The collected data can be analyzed and used to get better viewing statistics or for the delivery of personalized advertisements directly to the devices. Conventional TVs are not able to communicate with the broadcasting stations since they have no network interface that connects to the Internet. This interface can be used as a back channel in Smart TVs, i.e., data from the TV can be sent back to the broadcasting stations.

B. Overview and Our Contribution

In section II, we show shortly that this paper is the third paper on the topic privacy in HbbTV with new results. We begin with a short introduction in HbbTV (see section III) and follow with an overview of our analyzed scenarios. We show two scenarios, a Smart TV environment and a radio environment where tracking methods has been found. In both scenarios the use of tracking methods is not expected by the consumers and however in the latter scenario the Internet functionality while listening radio is not necessary. We discuss in detail which techniques are used and show differences between our first survey in 2012 and in the first two months of 2014. Section V covers an explanation of our extended protection system. Finally, we discuss threat scenarios for the devices analyzed in this paper, show some out-of-scope vulnerabilities and give a short outlook of future work.

II. Related Work

A document from the working group of the German TV Platform published in 2013 states that the issues found in former publications are not real and the architecture of Smart TVs is not comparable with other computer devices [2]. However, this section summarizes some issues of Smart TVs which have been seen on other computer devices like laptops as well.

A. Privacy Issues in HbbTV

To the best of our knowledge, we are not aware of any other academic work related to security and privacy in HbbTV

[1]Hypertext Markup Language
[2]Cascading Style Sheets

except our own articles published in 2013 [3] and 2014 [4]. In 2013, we described the entire issue and our results based on a detailed analysis of German and Austrian broadcasting stations. Later, we introduced a privacy protection system, which prevents consumers being tracked by broadcasting stations. Many national magazines and newspapers reported about this privacy issue.

Moreover, there are some other publications related to HbbTV in general. Lukac et al (see [5]) discuss a technical implementation of HbbTV. In [6] and [7] is shown how the HbbTV functionality can be ported to Android. [8] shows a way to test an HbbTV application properly. [9] and [10] are describing methods to synchronize HbbTV applications and other home network devices. Another paper explains how to record and deliver HbbTV applications, so that the user may watch a program repeatedly with all data available [11] at replay time.

B. Vulnerabilities in Smart Entertainment Devices

Numerous publications about vulnerabilities in smart entertainment devices have been published in the last years. For Smart TVs, Michéle et al [12] found a vulnerability in the media players of many vendors, which can be used to inject malicious software in a Smart TV. Many other security issues have been found in Smart TV subsystems. Auriemma reported a vulnerability where a Smart TV is not usable when it receives invalid data packets at a specific network port [13]. More hacks for the underlying hardware or software were published by SeungJin researchers [14] and Mulliner and Michéle [15]. A manual was published outlining how to evaluate the security of your home entertainment system [16].

III. Hybrid Broadcasting Broadband TV

A. HbbTV Standard

ETSI approved the HbbTV standard in November 2012 [17]. The standard specifies the technical framework for HbbTV applications and how it should be implemented in Smart TVs and set-top boxes. Basically, an HbbTV application is a website, which can be displayed as a transparent overlay over the current program. The development of the HbbTV standard is done by the HbbTV consortium. It is a pan-European initiative with the aim to harmonize the broadcast and broadband delivery of entertainment to the end consumer through Smart TVs and other set-top boxes [18].

The current version of the HbbTV standard does not provide any information about how to develop an HbbTV application privacy friendly.

B. General Functionalities of HbbTV

According to the standard an HbbTV supporting device needs an (for the consumer) invisible web browser that can open HbbTV applications from the Internet. The HbbTV application is requested when the device receives an URL in the DVB[3] stream sent by the broadcasting station. The URL is then extracted from the DVB stream, and the site will be loaded in background by the web browser. The content of this URL can be any web site on the Internet written with standard web techniques like HTML, CSS and JavaScript with an additional HbbTV JavaScript library that enables the page to be displayed. The broadcasting stations typically provide a kind of landing-page URL of their HbbTV applications, which displays a notification message that an HbbTV application is available and ready to be activated by pressing the *Red Button* on the remote.

Figure 1 outlines the typical process of activating an HbbTV application from consumer's perspective. Different possibilities to start an HbbTV application on the devices exist (see HbbTV standard for the full list). The most common way is that the HbbTV notification, which tells the consumers that an HbbTV application is available, is shown on the DVB program (2). If a consumer presses the *Red Button*, the HbbTV application starts in full screen mode (3).

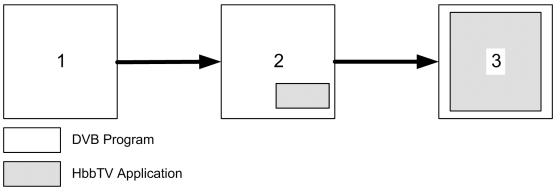

Fig. 1: Activation Process of HbbTV Application

In our analysis, we have seen devices which have different HbbTV activation modes. All devices analyzed have an option to disable the HbbTV service. Moreover, some of them have an HbbTV option called notify. That mode does not show any notification to the consumer, but if the consumer knows that an HbbTV application is available, he/she can activate the HbbTV service by pressing the *Red Button* on the specific channel. Afterwards the HbbTV application sent by the broadcasting station is executed. From a privacy perspective it is a very useful function, but not useable since the user is not notified and does not know if a HbbTV application is available.

The HbbTV applications differ from channel to channel. For one broadcasting station it is only a replacement for teletext, for another it is a portal to deliver a variety of additional services such as media libraries or the possibility to watch previous content. The broadcasting stations are in all cases responsible for the content they provide via HbbTV.

In the remainder we are only focusing on the first two steps (see figure 1). Thus, we collected data before the consumer intentionally activated the full screen application by pressing the *Red Button* on the remote. The started HbbTV application after pressing the button is not in scope of this paper.

IV. HbbTV in Real World

In this section, we compare our results from 2012 [3] and [4] with data collected in January and February 2014. We shortly summarize the conditions for the old and new survey:

- We focus on broadcast dependent HbbTV applications, i.e., HbbTV applications that are loaded by

[3]Digital Video Broadcast - The DVB project is an Alliance of many companies worldwide. It defines specifications for digital media delivery [19]. In this context, a DVB stream is a data stream that transports radio or TV signals via satellite (DVB-S), cable (DVB-C) or terrestrial (DVB-T) to the end devices.

turning to a channel with HbbTV signal in the DVB stream.

- Channels without HbbTV were not analyzed since there is no extended privacy risk.
- We used standard hardware for capturing the packets, e.g., a laptop and a Wi-Fi router.
- The software Wireshark[4] and standard routing technologies for capturing the packets are used.
- The traffic of the HbbTV channels was captured with different broadcasting methods: satellite (DVB-S) and terrestrial (DVB-T). This time we did not capture HbbTV on cable (DVB-C). The position of the satellite was Astra 19.2E and the terrestrial signal was in the Frankfurt Rhine-Main region of Germany.
- We divided the time where packets are exchanged in different time phases (see figure 2):
  Phase 1: Time between switching to a channel and the HbbTV notification is shown,
  Phase 2: Time between displaying the HbbTV notification and when the *Red Button* is pressed by the consumer and
  Phase 3: Executing the HbbTV application after pressing the *Red Button*. This phase has not been analyzed, since the consumer has actively started the HbbTV application by pressing the button and data exchange between the Smart TV and the Internet is expected.
- The results focus on phases 1 and 2, in which the consumer has not intentionally activated the HbbTV application. Phase 3 is out-of-scope.

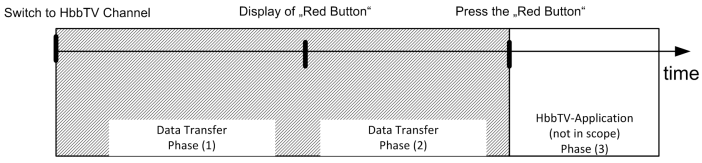

Fig. 2: Phases

We describe two environments where HbbTV is used: Smart TV and digital satellite radio. The Smart TV environment was the basis in the 2012 survey as well. The digital satellite radio environment is introduced in this work for the first time.

In both configurations we also spotted no network activity between HbbTV applications and device vendors this time. Furthermore, we will not show how the packets are captured in detail since it has been already explained in former publications. Instead, we will discuss some requests in more detail.

Statistically, we found a higher number of channels using HbbTV and extended our list from [4] accordingly. In 2012 we analyzed HbbTV channels from Germany and Austria. This time we added further channels from Germany, Austria, France and Spain. All HbbTV channels analyzed are listed in table II. New in this article are digital satellite radio channels implementing the HbbTV technology (see table III).

[4]http://www.wireshark.org/

A. Characteristics and Grouping of HbbTV Channels

In this subsection we explain our criteria we used for categorizing the channels. We use these characteristics in section IV-C and IV-E

1) HbbTV: A channel has an HbbTV signal if we gathered data while on a channel and a notification is shown to the consumer. If the notification is not shown and data has been collected, we do not assume that a HbbTV application is available.

2) Start-up Requests: When switching to an HbbTV channel (Phase 1), the broadcasting station provides an URL within the DVB stream to load the first HbbTV notification showing that an HbbTV application is available. In background, a variety of scripts, images and other resources are loaded. We call this action *start-up request*. If no application is available, no URL in the DVB stream is provided and no further action is performed by the device.

3) Periodic requests: In addition to the start-up requests, we measured on some channels periodic requests after the HbbTV notification has been displayed (Phase 2). The time between each request differed from 1 to 15 minutes depending on the channel. These requests are made before the user actively started the HbbTV application by pressing the *Red Button* (Phase 3), so the user does not expect any data to be transferred to some other party, e.g., broadcasting stations.

4) Counting Pixel: A counting pixel is a request of an image or other resource that is used to count visitors and page impressions of a web site. In smart entertainment devices it could be used to count the consumers on a specific channel with HbbTV. If a counting pixel is only be used for counting the start time of an HbbTV channel, a periodic request is not needed. If the counting pixel should be used to measure the time a consumer is on a channel, a periodic request can be performed.

5) Tracking: If a tracking script is found, we assume that the data is being used for other purposes than for the functionality itself. A tracking script is developed for collecting data of a consumer to get a more accurate viewing/listening behavior.

6) Data transferred to third parties: Transferring data to third parties is for the functionality of HbbTV in most cases not necessary. If the broadcasting station uses a datacenter for its services, we do not mention that as third party.

7) Personalized ads: The HbbTV notification for the consumers can be modified so that it can deliver personalized ads. The consumer is not in an HbbTV application at this moment.

8) Cookies: Cookies can be set on smart devices as well. On some channels we found cookies, which save values on the smart device. We do not judge them as bad or good.

In table I we listed the characteristics and assigned them to groups. The characteristics of A,B,C,D and Z from [4] have not changed. We extended the groups by A^+,A^- and B^-. Further information about the groups are in section IV-C.

As already mentioned, we clearly state that the characteristics are measured before the consumer activates the HbbTV application by pressing the *Red Button* (Phase 1 and 2).

Characteristics/Group	A+	A	A-	B	B-	C	D	Z
HbbTV	(x)	x	x	x	x	x	x	x
Start-Up request	x	x	x	x	x	x	x	x
Periodic requests				x	x	x	x	?
Counting Pixel			x	x	x	x	(x)	?
Tracking					x			?
Data transferred to third parties						x	(x)	?
Personalized Ads							x	?

TABLE I: Groups of Channels

B. Smart TV Environment

The Smart TV environment has three components: a Smart TV, Internet connection via home network and a DVB signal coming from different sources. As shown in figure 3 the DVB signal is received by satellite or terrestrial and is cable-bound connected to the built-in receiver of the Smart TV. The Smart TV is connected to the Internet and the home network via LAN or Wi-Fi. In this scenario the test devices were Samsung UE40D6200/UE40ES6300 and LG 42LN5758.

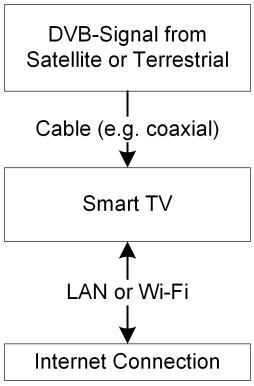

Fig. 3: Smart TV Configuration

C. Comparison and Results - Smart TV Environment

In this section, we compare the results from our publication [3] with new results. In summary the HbbTV traffic of our newer survey is much more privacy friendly on already analyzed channels. Some channels have not changed and others are the first time in our analysis. A list of analyzed channels can be found in table II. The list of HbbTV channels is not complete and only represents our results.

All data in this comparison was obtained in the Smart TV environment explained above. It has been captured directly without an encryption. This time, we found differences on the different signal sources, which we marked in the channel list separately. As we have done in our former work, we split the channels in different groups based on table I with common characteristics. The column old group represents the group assignment in [4] and the column group reflects the assignment of our new results.

HbbTV applications, which are partially transferred by DVB (DSM-CC[5]), have been found. DVB has no back channel, so it is a very privacy friendly way to transfer an HbbTV application to a smart device. However, we observed that this method can be used in different ways: backup for Internet connection or for notifying the consumer. When it is used as a backup for Internet, we measured that with an Internet connection, HbbTV data was transferred via Internet. On one channel we detected that this method is used to transfer the HbbTV notification to the device, which is the most privacy friendly way we have seen so far.

[5]Digital storage media command and control - can be used for transporting packet data

Group	Old Group	Channel
A+	-	Super RTL (HD) (S), RFO, France24(en/fr)
A	A	ZDF Group (+), 3Sat (HD)
A	-	HSE 24 (HD), ORF1 (HD) (C), ORF2 (HD) (C), Local TV portal (*), Local TV channels (+)
A	C	ARTE (HD),
A-	B	ARD Group (+,C)
B	C	Pro Sieben (HD), Kabel 1 (HD), Sat 1 (HD), Puls 4 Austria (HD) (C)
B	-	Sixx (HD)
C	Z	RTL (C,S), VOX (C,S)
C	-	RTL2 (C,E,S), sonnenklar.tv (HD), QVC (Plus) (HD) (E), RTVE
D	D	Anixe (HD), Anixe iTV
-	A	Bibel TV (HD)

TABLE II: TV Channels

E - encrypted traffic found
S - HbbTV signal available in DVB-S, not on DVB-T
$*$ - HbbTV application was started when turning to channel
C - Cookies found
+ - Channel group with same characteristics; more than one channel
HD - Channel is available in high definition

The channels in the new introduced sender group A+ technically do only a start-up request, i.e., a request is performed to the server of the broadcasting station without receiving an HbbTV application. We assume that these channels are testing an HbbTV implementation, which may soon be transformed in a complete HbbTV application. The only exception in this group is *France24* that sends the notification via DVB. This means that the HbbTV notification does not perform any Internet request. Nevertheless, one request without content is made to the servers of *France24*. This could be a check if an Internet connection is available. Thus, the resulting group A+ has three channels. The probability of accurate tracking in this group is very low.

In the groups A-D the following traffic flow is common. The Smart TV is triggered by the initial HbbTV URL in the DVB stream. The URL is then extracted and a request to the Internet is performed. In return a HbbTV notification with different resources is delivered; the start-up request. The resources differ from channel to channel and can be a variety of HTML files, scripts and images. If the previous requests are performed successfully, the notification that the HbbTV application is ready and can be activated by the *Red Button*, is displayed to the consumer (see figure 4).

The probability of consumer tracking is higher in group B-D. We have seen periodic requests before the consumer pressed the *Red Button*. Some channels deactivated the periodic requests compared to our first survey. However, the periodic requests to the server of the broadcasting stations are this time in an interval from 70 seconds up to 15 minutes.

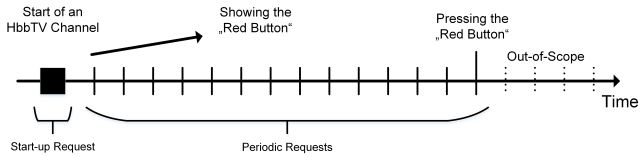

Fig. 4: HbbTV Traffic Flow

The shorter interval time is found in group B and the longer time in group D. These requests can be used to pre-load content or to track consumers very accurately. With such kind of requests, the broadcasting station is able to check if a consumer is still watching a specific channel. Periodic requests stop when the consumer is turning to another channel since the specification requires to stop all open connections when switching. Additionally, in group C we found tracking scripts at different channels than last time. The purpose of these services is clearly defined, they can be used to get more information about the consumer and the used devices.

In group A are channels where the start-up request is performed to show the HbbTV notification. Here, we list some channels that have not changed their characteristics compared to the old survey. One channel has been moved to this group: *ARTE*. The tracking service Google Analytics has been deactivated. The two channels *ORF 1* and *ORF 2* are encrypted and cannot be watched without an appropriate smart card, but the HbbTV application is notified and can be started. The channel *local TV portal* starts its HbbTV application in full screen when turning to the channel. The consumer does not need to press the *Red Button*.

A$^-$ summarizes channels which are privacy friendly, but have counting pixels. The pixel of that sender group is explained in detail in section IV-E. The old group assignment of the *ARD group* was B. In this survey the periodic requests, which enabled the station to create detailed viewing behavior of a consumer, has been disabled. This broadcasting station is notifying an HbbTV application even if the Internet connection is not available. It uses the DVB stream (DSM-CC) to transfer necessary data.

Group B has all characteristics of group A. Additionally, periodic requests before the consumer has pressed the *Red Button* could be measured. The requests seem to be counting pixel since they contain data about the watching time and the channel. The time period is every 70 seconds. The request looks like

```
http://hbbtvserver.com/?c=SAT1&seq=1
&open=1&sid=3729871
```

where seq and open are parameters that count up every request and sid seems to be a session identifier. Parameter c is the name of the channel. That request returns a JavaScript command that does nothing. With this technique it is easily possible to determine the time a consumer is on a specific channel. Google Analytics and other tracking services have been removed from the channels (*Pro Sieben, Kabel 1, Sat 1, Puls 4 Austria*), so they moved from group C to B. *Sixx* is added to group B.

The channels of group C are sending data to third parties like INFOnline[6], IVW[7], Google Analytics[8], etracker[9] and ScorecardResearch[10]. These services are tracking services, which can be used to track a customer accurately. As in all other groups the consumer is not able to avoid sending data to such services. In this group *QVC* uses HTTPS to transfer data to its servers. We managed to break that encrypted connection with the mitmproxy[11] and could see a request to Google Analytics.
The two channels *RTL* and *VOX* that were formerly in group Z moved to group C. They are now providing an HbbTV application to all consumers of HbbTV supporting devices. Periodic requests could not be measured in this group.

Group D has the most privacy invasive channels which are showing an advertisement every 15 minutes over the current running program (see figure 5). The customer has no

Fig. 5: Advertisements of Group D

option to disable that unwanted overlay. Even if the channel is broadcasting an ad, the HbbTV ad is additionally shown. We did not recognize any changes since 2012. At this point, we are not aware if the ads are personalized. Interestingly, we got different ads on different devices.

Bibel TV has disabled the HbbTV application. All channels from the former group Z have move to another group. That group represented channels that could not be received by our test devices last time.

Since our results are reflecting which data flow is going from the device to the broadcasting stations, we do not know how the gathered data is processed by the broadcasting stations. But even just the possibility to collect a variety of data is a high privacy risk for consumers. The data can be used to gather consumer behavior in real time with more accuracy than ever before. In group D, we could already see how the HbbTV technique can be used in the wrong way. A consumer does not expect background Internet transfers and even the displaying of advertisements from the Internet is totally unexpected.

Most Internet service providers assign an IP address to an household for 24 hours, which makes it unique and identifiable within that period. Even afterwards devices can be identified by using cookies with IDs. The Samsung devices did not send back any cookies. The other (LG and invento) devices did.

[6]https://www.infonline.de/
[7]http://www.ivw.eu/
[8]http://www.google.com/analytics/
[9]http://www.etracker.com/en.html
[10]https://www.scorecardresearch.com
[11]http://mitmproxy.org/

With such data it is easily possible to link a device with an already existing dataset even after an IP renewal.

As of now, we do not see a direct link to the consumer, only to the device. New services which requires to enter credentials or billing information can, however, be used to link the data directly to the consumer. On the home shopping channel *HSE24* we already had the possibility to log-in to an home shopping account. Unfortunately, our entered data was transmitted without encryption, so we stopped at this point.

Broadcasting stations which provide more than one channel may track the consumers viewing behavior over these channels, i.e., it is possible to determine the consumer's preferences or political orientation over these channels. Furthermore, data to identify a Smart TV is transferred, for example screen resolution, vendor, IP address and cookies. For the future, it is likely that more services for HbbTV will implement consumer logins for user identification.

D. Digital Satellite Radio Environment

The digital satellite radio configuration has four components: a set-top box with Internet connection, DVB satellite signal, connection to a Hi-Fi system. The set-top box is connected with the satellite and has a connection to the Internet and the home network. The Hi-Fi system gets its audio signal directly out of the set-top box. Sometimes the set-top box is additionally connected to a (conventional) TV, which is however not relevant for our configuration. We used a set-top box from invento model Scena 6n IDL6651N.

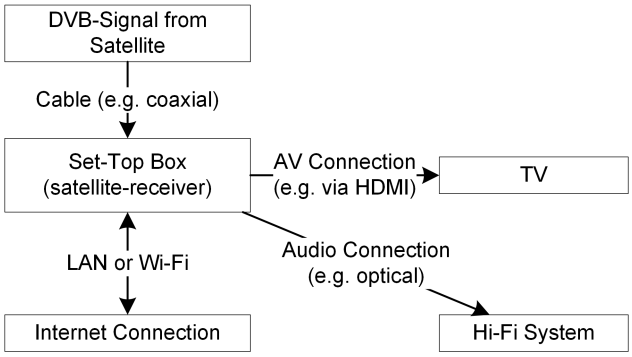

Fig. 6: Set-Top Box Configuration

The digital radio signal with HbbTV is only transmitted by DVB-S, so other broadcasting methods have not been analyzed.

E. Results - Digital Satellite Radio

In this work, we found the first time HbbTV in over 20 digital satellite radio channels. They are provided by the same radio station. The characteristics of the traffic can be assigned to group A^-, which means that a start-up request and a counting pixel have been measured. Furthermore, an image with the radio channel logo is requested. However, on all radio channels we observed that the privacy friendly method via DVB (DSM-CC) is enabled. Therefore, if an Internet connection is not established, the HbbTV application is available via DVB. In DVB HbbTV notifications the radio channel logo was not customized, i.e., just the logo from the radio station has been shown on all radio channels.

However, nobody would expect any data transfer to the Internet while listening radio. In the scenario explained in section IV-D, the consumer will be counted automatically if an Internet connection is available. It is not necessary to turn on a TV. If you disconnect your Internet from the receiving set-top box, the consumer will probably lose functions like Internet radio or software updates of the box. If the box is also used for receiving TV, the HbbTV applications are not available anymore.

In this context, even if the analyzed HbbTV traffic is only in the more privacy friendly group A^-, we could not think of a useful application without a TV. The only scenario, in which this functionality is useful is if the radio channel is turned on on a Smart TV, then it provides information about the current running program, but we assume that few people use a Smart TV to listen to radio. The radio channels that use HbbTV are listed in III.

Group	Radio Station
A^-	Bayern 1, Bayern 2, Bayern 3, Bayern Plus, BR Klassik, BR Puls, HR 1, HR 2 Kultur, HR 3, HR 4 Rhein-Main, HR Info, MDR 1 Radio Sachsen, MDR Figaro, MDR Info, MDR Jump, MDR Klassik, MDR Radio Sachsen-Anhalt, MDR Sputnik, MDR Thüringen, NDR 1, NDR 2, NDR 90.3, NDR Blue, NDR Info, NDR Infor Spezial, NDR Kultur, SWR 1, SWR 2, SWR 3, SWR 4, SWR info, WDR 1, WDR 2, WDR 3, WDR 4, WDR 5, WDR Event, You FM

TABLE III: Radio Channels

In detail we have found requests like

```
http://hbbtvserver.com/stat/p.png?redir=1
&app=1&sid=28479&sub=-1&delivery=11
&uid=5aab3ecd0cad34153b2c83cd1031d0ab
&d=84305.1393622531075
```

These requests can be categorized as counting pixel since the response is an image with 1x1 pixels. We could not clarify the meaning of the parameters redir, app, sub, delivery and uid. Redir, app, sub and delivery never changed in our analysis. However, uid changed from one device session to another, i.e., if you turn off the device and restart the uid changes. Thus, this session uid can be used to create a history of one device session over all channels the broadcasting station is providing. In this case, it provides many radio and tv channels and is so enabled to track a device in one session, e.g., a device watched channel 1, channel 2 and switched to radio channel 2 and so on. This data is at this point not directly assignable to a device.

However, we found cookies, which have a long life time – over 20 years –. These cookies can be concatenated with the log requests to profile a consumer. We do not know if this is made at server level. The cookies seem to be set before the consumer has the possibility to disable it. An option to disable it is available after starting the HbbTV application with the *Red Button*.

We found no periodic requests, so the time a consumer is on a channel cannot be recorded accurately. Only the turn on time is reported.

Our protection system has been updated to protect the consumers against that transmission.

V. HbbTV Privacy Protector

We developed the HbbTV Privacy Protector to give consumers the option to control an HbbTV application on a specific channel. The protector disables an HbbTV application per default. As outlined in figure 7 the implementation consists of a transparent HTTP proxy server, a web server and a whitelist for the HbbTV URL patterns. The dynamic HbbTV detection method can find new HbbTV channels and add them to the whitelist.

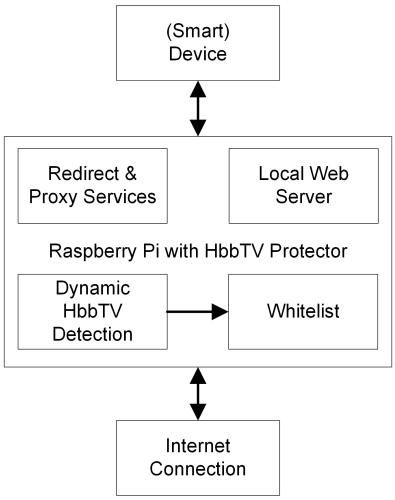

Fig. 7: Abstract Components of HbbTV Privacy Protector

Basically, the proxy intercepts the traffic between the device and the Internet uplink. When an HbbTV start-up request is detected (for example when changing a channel and thereby autostarting an HbbTV application), the request is intercepted and forwarded to the local webserver. It delivers an HbbTV application which waits until the consumer actively enables HbbTV by pressing the *Green Button* on his/her remote. We have not chosen the *Red Button* to not confuse consumers. When that happens, the interception rule is removed from the proxy server and the device is forwarded to the broadcaster's HbbTV application.

The software and the pre-packed Raspbian are available on our website[12].

A. Technical Background

We developed the tool compatible with the *Raspberry Pi* platform, a small ARM based computer that is quite popular for building TV media centers. In this case, we implemented the tool on top of the Raspbian Linux operating system.

To operate properly the smart device Internet traffic needs to be routed through the *Raspberry Pi* for example using the Internet connector and can then be passed on to the local Wi-Fi using a USB Wi-Fi adaptor (see Figure 8 as an example with a Smart TV). Due to the HDMI connector with HDMI-CEC support, the *Raspberry Pi* can be controlled and configured from a Smart TV using the TV remote. The small device size makes it possible to mount it on the TV backside, out of sight. Even the USB port of the Smart TV can be used to power the *Raspberry Pi*. As an advantage it starts automatically when switching on the TV.

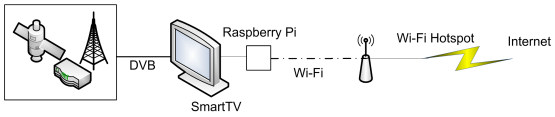

Fig. 8: Architecture of the HPP

B. Redirect and Proxy

The implementation is based on mitmproxy[13], a lightweight HTTP and HTTPS interception proxy written entirely in python. We developed a custom python script to add the HbbTV interception functionality to *mitmproxy*. Traffic redirection can be done transparently on a Linux device between the smart device and the Internet uplink using the Linux *iptables* tool. *Iptables* is configured to transparently forward incoming traffic on port 80 (HTTP) to port 8080 and port 443 (HTTPS) to port 8443. On port 8080 resp. 8443 the mitmproxy is waiting for traffic. This redirection of traffic cannot be seen by the smart device.

C. Dynamic HbbTV Detection and Whitelist

In the first version we only had a hardcoded list of HbbTV URLs which turned out to be problematic since a change of an URL is possible at any time. We extended the system by a method to auto detect new HbbTV applications: Dynamic HbbTV Detection. This method fills the list of HbbTV URLs. Therefore, we can provide a start list and it grows automatically.

The auto detection mechanism uses the fact that all HbbTV applications on Samsung devices start with an HTTP HEAD request. Unfortunately, we noted that the HEAD request is only done by our Samsung test devices. Samsung is one of the biggest players in the Smart TV markets, so we leave this method as an useful function integrated.

Additionally, the current version implements a method to detect an HbbTV application with searching an specific HTML object containing

```
application/oipfApplicationManager
```

in the requested web page. All HbbTV applications need this fragment because this object is used to enable the HbbTV browser in the smart device to display the content.

D. Local Web Server

The function of the local web server is to provide our HbbTV application. This application shows the consumer a message that an HbbTV application is available and can be activated by pressing the *Green Button*. In order to enable the

[12]http://www.smarthome.sit.tu-darmstadt.de

[13]http://www.mimproxy.org

web server, here Apache Web Server, to provide an HbbTV application the mime type of HbbTV has to be added. The HbbTV mime type is

```
application/vnd.hbbtv.xhtml+xml      hbbtv
```

E. Configuration and first Start-up

The Raspbian image is pre-configured to work with the USB Wi-Fi adaptor EDIMAX EW-7811UN. The system boots and starts directly the mitmproxy with our script. The on-board LAN port has to be connected to the device and the USB adaptor connects to the home network. To simplify the configuration process we implemented a small web site that can be used to configure the Wi-Fi and the language. At the moment only WPA and WPA2 networks are supported. As language we have a German and an English version. For using the graphical configuration interface the LAN port has to be connected to a computer. The user interface can then be requested by entering http://192.168.10.1 in a web browser. On the LAN port the DHCP server is assigning IP addresses, so a direct connection to a home network is not recommended.

VI. Security and Privacy Issues

In this section we will summarize how the privacy issues in HbbTV can affect consumers. The listed attacks are passive, which means that the attacks can be issued without a device modification.

A. Loss of Sensitive Data

The value of sensitive data is often not known until it is lost or leaked. In our work, we have seen that an HbbTV application can send viewing and listening behavior data to third parties and the broadcasting station. For broadcasting stations the value of this data is high, because they can (1) see in real time in which topics the consumer is interested, (2) profile consumers to show them only interesting content or (3) profile consumers to sell the data to an advertisement company for creating (personalized) ads for smart entertainment devices. In all cases, the data should not be automatically sent and processed without consent of the consumer. Third parties like Google have more possibilities to profile a consumer since they get a lot more data. They can concatenate data from different sources: laptop, smart phone, tablets, smart entertainment devices and many more.

B. Eavesdropping Neighbors

We have shown in our publication [3] that a wireless network can be captured and then be processed in such a way that a profile of a consumer is creatable. This attack is even possible when the network is secured with the current WPA2 standard.

C. Burglary

An even more problematic attack is the possibility to eavesdrop a wireless network to gather data to determine if people are not at home. For example the Smart TV is usually turned on at 8pm, however this time the device is off. The burglar can infer that an apartment is empty if some Wi-Fi signals are not measurable. Even in a WPA/WPA2 Wi-Fi network, some meta data is not encrypted and can be read by an attacker, e.g., size of sent packets, mac address of sender and receiver. An intelligent algorithm could be developed to analyze the wireless network traffic fast and accurate. The best way to prevent such an attack is to use cable-bound connections since eavesdropping of cable connections is much harder than wireless connections.

VII. Security Issues not in Scope

We shortly summarize issues found while analyzing the HbbTV traffic.

- HTTPS certificates have not been validated correctly on the web browser in our Samsung UE40ES6300. This issue could lead to a severe phishing attack since all HTTPS certificates were marked as trustworthy. We communicated that to Samsung in October 2012 and the patch has been distributed in April 2014.
- Unidentified Google requests were found on the LG device. The requests are encrypted. However, this problem is outside of the responsibility of the broadcasting stations, so we did not analyzed these requests further.
- The set-top box sends a request to the vendor if a channel is changed. The content of this request is the channel logo that is displayed in the channel information. This leads to the possibility that the vendor can make statistics which box are watching which channel. However, the consumer cannot stop this transmission. The only way is to disconnect the box.

VIII. Conclusion & Future Work

In this paper we showed that the HbbTV technique can still be used to track consumers by broadcasting stations. We expected that more broadcasting stations removed tracking mechanisms due to the press and media coverage we have reached.

In our new analysis some channels reduced the probability of tracking. They removed Google Analytics and another sender group removed the periodic requests. But we detected four different third party tracking scripts on new HbbTV channels. The implementation of such services clearly shows the intention of the broadcasting stations. For the consumer, it is not easily possible to disable the tracking services and remain the HbbTV functionality, so we discussed our extended implementation of a protection system that helps to prevent sending data without consent of the consumer. It is obvious that HbbTV applications must be more controlled to prevent further privacy issues. At the moment for example, we are not sure why digital satellite radio channels need an HbbTV application with counting pixel. These counting pixels delivered by the HbbTV application can be used to count radio listeners easily.

The findings out of scope resulted in a need to have a steady control of new devices and a posibility to give the consumer a control option for the devices. It seems that many implementations on the devices, which we have been seen so far, are only proof of concepts. As a future work, we will be developing an entire protection system for the smart home.

Not only smart entertainment devices are aimed, smart home devices like smart plugs are risky as well.

ACKNOWLEDGMENT

We thank Florian Oswald and Erik Tews for their support. Florian wrote his Bachelor's thesis in the context of this project and Erik implemented a proof of concept of the protection system. Moreover, we thank Michael Waidner for his support to make this project possible. This work was supported by the European Center for Security and Privacy by Design (EC SPRIDE), funded by the German Federal Ministry of Education and Research (BMBF), and the Center for Advanced Security Research Darmstadt (CASED), funded by the LOEWE program of the Hessian Ministry for Science and the Arts (HMWK).

REFERENCES

[1] Greg Tarr, " IHS: Smart TVs Rise To 27% Of TV Shipments ," http://www.twice.com/articletype/news/ihs-smart-tvs-rise-27-tv-shipments/105108, accessed on 23.02.2014.

[2] Working Group Smart TV of the German TV-Platform, "Marktanalyse Smart-TV - Eine Bestandsaufnahme der Deutschen TV-Plattform ."

[3] M. Ghiglieri, F. Oswald, and E. Tews, "Hbbtv – i know what you are watching," in *Informationssicherheit stärken – Vertrauen in die Zukunft schaffen*. Bundesamt fr Sicherheit in der Informationstechnik, 05 2013, pp. 225–238.

[4] M. Ghiglieri and E. Tews, "A privacy protection system for hbbtv in smart tvs," in *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, 01 2014, pp. 648–653.

[5] Z. Lukac, M. Radonjic, B. Veris, T. Maruna, and N. Kuzmanovic, "The experience of implementing a hybrid broadcast broadband television on network enabled tv set," in *MIPRO, 2011 Proceedings of the 34th International Convention*, May, pp. 840–844.

[6] M. Milosevic, S. Markovic, N. Jovanovic, B. Mlikota, and M. Balac, "Porting and integration of functionality required by hbbtv standard on android based dtv platform," in *Consumer Electronics ?? Berlin (ICCE-Berlin), 2013. ICCEBerlin 2013. IEEE Third International Conference on*, Sept 2013, pp. 1–4.

[7] M. Milosevic, S. Markovic, B. Mlikota, M. Zivanovic, and B. Prtvar, "Porting of hbbtv functionality on dtv platform based on android os," in *Telecommunications Forum (TELFOR), 2012 20th*, Nov 2012, pp. 1353–1356.

[8] Z. Lukac, V. Zlokolica, B. Mlikota, M. Radonjic, and I. Velikic, "A testing methodology and system for functional verification of general hbbtv device," in *Consumer Electronics (ICCE), 2012 IEEE International Conference on*, Jan 2012, pp. 325–326.

[9] M. Zorrilla, I. Tamayo, A. Martin, and I. Olaizola, "Cloud session maintenance to synchronise hbbtv applications and home network devices," in *Broadband Multimedia Systems and Broadcasting (BMSB), 2013 IEEE International Symposium on*, June 2013, pp. 1–6.

[10] C. Ziegler, "Second screen for hbbtv 2014; automatic application launch and app-to-app communication enabling novel tv programme related second-screen scenarios," in *Consumer Electronics ?? Berlin (ICCE-Berlin), 2013. ICCEBerlin 2013. IEEE Third International Conference on*, Sept 2013, pp. 1–5.

[11] J.-C. Dufourd, S. Thomas, and C. Concolato, "Recording and delivery of hbbtv applications," in *Proceddings of the 9th international interactive conference on Interactive television*, ser. EuroITV '11. New York, NY, USA: ACM, 2011, pp. 51–54. [Online]. Available: \url{http://doi.acm.org/10.1145/2000119.2000129}

[12] "Watch and be watched: Compromising all smart tv generations," in *2014 IEEE 11th Consumer Communications and Networking Conference (CCNC)*, 01 2014.

[13] L. Auriemma, "Endless restarts," Apr. 2012, http://aluigi.altervista.org/adv/samsux_1-adv.txt.

[14] L. SeungJin, "Dirty note on Samsung Smart TV Security," Dec. 2012, http://beistlab.files.wordpress.com/2012/12/samsung_smart_tv_attack_surfaces2.pdf.

[15] C. Mulliner and B. Michéle, "Read it twice! a mass-storage-based tocttou attack." in *WOOT*, 2012, pp. 105–112.

[16] Rikke Kuipers, Eeva Starck & Hannu Heikkinen, "Smart TV Hacking: Crash Testing Your Home Entertainment," http://www.codenomicon.com/resources/whitepapers/codenomicon-wp-smart-tv-fuzzing.pdf.

[17] IRT GmbH, "HbbTV Specification (approved by ETSI as ETSI TS 102 796 v1.2.1 in November 2012)," http://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.02.01_60/ts_102796v010201p.pdf.

[18] IRT GmbH , " HbbTV = More entertainment at your command ," http://hbbtv.org, accessed on 02.04.2013.

[19] DVB Project, "History of DVB," https://www.dvb.org/about/history, , accessed on 15.04.2014.