Binge-watching Data Analysis Tool

José Maria Reis Simões d'Assis Cordeiro
josemcordeiro@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

November 2019

Abstract

Binge-watching is an uprising social phenomenon widely understood as the consumption of media content for long periods of time. Due to the lack of research that combines both objective and subjective data to understand series, movies and documentaries viewing habits, we proposed a novel methodology to explore viewers' emotions and engagement at home.

This thesis is part of a wider study, and aims to work on a set of tools that are pivotal for the collection and interpretation of the data related to these events. The creation of Binge-watching Data Analysis Tool, which detects viewers' interaction on an internet-distributed video service (i.e. Netflix), collects their physiological data, prompts forms and displays graphically the different data gathered for analysis are some of the goals of this project.

The technology developed allowed to cover more than ∼90% of physiological data during binge-watching sessions at home. The data gathered empirically supported previously theorized results by registering, for example, an increased average heart rate while binge-watching.

Keywords: Binge-watching; Physiological data; Chrome Extension; Engagement; Standalone Smartwatch Application; Netflix;

1. Introduction

The increasing popularity of internet-distributed TV services, such as Netflix, has changed media consumption behaviour in recent years. This was done by giving their subscribers the possibility to customize their TV diet regarding content, time, screen and location [16]. According to Netflix's 2017 annual report, over 117 million subscribers distributed amongst 190 countries watched 140 million hours of TV shows and movies per day [27].

The release of full (and even new) seasons at once together with the affordable prices that internet-distributed video services offer foster television consumption [1] and has recently popularized the so-called phenomenon *binge-watching* (BW).

Despite the number of publications on the topic, there is not a standardized definition for the term BW [25]. Recently, Netflix defined binge-watching as watching 2 to 6 episodes of a series in a single sitting [28] whilst some scholars considered in their studies binge-watching as watching two complete episodes of the same series [5, 14, 30, 31]. Merikivi et al. (2019), conducted a deep analysis on previous publications and defined it as the "consumption of more than one episode of the same serialized video content in a single sitting at one's own time and pace" [25]. The co-existence of multiple definitions is due to the fact that people binge in various ways [22].

In the context of this project, we understand binge as the consumption of personally selected content for long periods of time in a row, in number of episodes, we consider three or more episodes, same as Sarah et al. (2019) [34].

Steiner and Xu (2019) described catching up, relaxation, sense of completion, cultural inclusion, and improved viewing experience as the main motivations for BW [10]. Recently, engagement and parasocial relationships were also linked to BW and stronger effects on media consumers [34]. Deborah et al. (2019), confirmed, in an empirical way, previously suggested by Pittman et al. (2015) [31], the relationship between binging sessions and a higher level of engagement and immersion [31].

The consumption of TV shows for long periods of time has been related to factors such as loneliness or depression, and may lead to heart disease [37]. Also, poor sleep quality is more associated to binge-watchers, since most of the sessions occur between 9PM and 11PM [33], creating a pre-sleep arousal [19]. Therefore, it is essential to study this social phenomenon to better understand the correlation between the viewers and the series.

A growing number of academic publications aim to analyse the motivations, attitudes and psychological factors behind this mode of TV consumption and its effects on health. Most of them are based on questionnaires or user's report/e-diaries [25]. However, there is a lack of studies that combine both objective and subjective measurements, and in particular, an absence of research that explores viewers' emotional states from a physiological (and objective) point of view whilst binge-watching.

This project is the continuation of a pilot study conducted by M-ITI, and this is the bridge between that project and a novel study "Excessive Media Use in Time of Netflix. "Binge-watching": Motives, Experiences, and Effects on Sleep" [36] funded by Swiss National Science Foundation. Its aim was to develop reliable technology able to understand the interactions with a streaming media service provider of a group of participants and their physiological reaction while binge-watching, before sleeping.

Particularly, this thesis involved developing software tools that collect three different types of data: physiological data (e.g. heart rate through a smartwatch), self-reported data (e.g. through an e-diary) and behavioural data (e.g. user's actions on Netflix interface collected by a browser application, for example, Chrome Extensions). It was also important to define a good interface to display, in a correct and clear way, all the the data collected.

To the software developed was given the name of Binge-watching Data Analysis Tool (BWDAT).

2. Background

To draw a novel approach on engagement measurement on BW sessions, based on the heart rate (HR) variability, an extensive review was done, not only on previous BW studies, but also on video games immersion studies and home health monitoring to evaluate reliable methods and wearable devices.

2.1. Binge-watching studies

Over the past years, international scholars have studied the binge-watching phenomenon by exploring both its causes (e.g. motives), time of the day to engage in this activity (e.g. morning, afternoon), location (e.g. bedroom, living room) of the users, and consequences (e.g. in people's life and health).

Several studies collected information through online surveys, focus groups, interviews, or self-reporting forms before and after binge-watching.

A study conducted by Deloitte in 2015 showed that TV drama is the most popular TV fiction genre to binge-watch, since a continuous narrative lends itself well to multi-episode viewing [7].

A study conducted by Erickson et al.19 analysed the hypothesis of binge-watching being related with the unconscious and unidirectional bonds created between the viewers and some characters, also known as parasocial relationships [34].

Some habits, while engaging in this phenomenon, were also studied, revealing participant's intentions and awareness while watching series. It was shown that BW can be intentional or not [17], and that viewers are not able to reproduce a clear image of their own binge-watching activity [9].

Some interviews were conducted by Stein and Xu (2018), revealing that viewers tend to rewind and rewatch scenes to better understand them [10], such behaviour has became more common on computers [3]. The main reasons to BW were found to be catching up, relaxation, sense of completion, cultural inclusion, and improved viewing experience [10].

Some habits relating BW and health issues and consequences were studied. For example, Sung et al. (2015) analysed the relationships between psychological variables concluding that the more depressed and lonelier viewers were, the more likely was for them to BW [37].

In Germany, Mikos (2016) showed that BW is a habit that became part of people's culture, and that it tends to be done on the weekend [26]. In the Netherlands, people aged between 18 and 34 tend to binge-watch alone at home, preferably in the living room and bedroom. Also, in around half of the sessions, viewers used their phones or ate while watching series [9]. In the United Kingdom, a study revealed that BW sessions lasted on average 02:07:41, and that 31% used mobile devices [33].

Horvath et al. (2017) conducted a laboratory study in which 3 groups of 17 participants would watch each episode of the "The Game" series in a different rating. That study concluded that viewers that binge-watched had a less accurate memory regarding the content watched, when answering questionnaires 140 days later [15].

A group of interdisciplinary researchers from the Madeira-ITI [5] developed a novel method that combines both self-reported and biosensing data to analyse viewers' motivations, emotional states and engagement when binge-watching at home.

The framework of that project enabled the prototyping of several tools. First, a B-Watch app [12], developed by Vasanth, was installed in a smartphone and smartwatch, paired via Bluetooth, allowing the acquisition of heart rate (HR) and inertial data from the viewers in a non-intrusive manner. A timestamp was also recorded in order to synchronize sensor data to media content. That information was written in different table sheets and stored in specific DropBox accounts. Second, Cabral and

Rigby developed a Chrome Extension to gather participants' viewing actions for specific episodes on Netflix's interface (e.g. name of the episodes watched, pauses, skipped content).

Additionally, participants had to answer a Self-assessment Manikin (SAM) [4] and a video questionnaire, that contained eight 6-second clips from each episode watched. The combination of HR and the SAM values gathered allow the observation of participants' emotional changes [23] while watching Netflix.

This study revealed that night and evening time were the most common times to binge-watch, and the bedroom was more used than the living room, mostly in night sessions (12 out of 15). Also, the computer was the preferred used device to BW [5].

On average, a binge-watching session (more than 2 episodes watched in a row) lasted 02:10:40 (HH:MM:SS), the longest session lasted 6:08:11, and the shortest 00:47:38 [5].

Since both Chrome Extensions and smartwatch data have one second time stamps, and the actual time that each sample was collected, it was possible to plot the data using the same Y axis along a single X axis. The plot of that data would give an output shown in Figure 1.

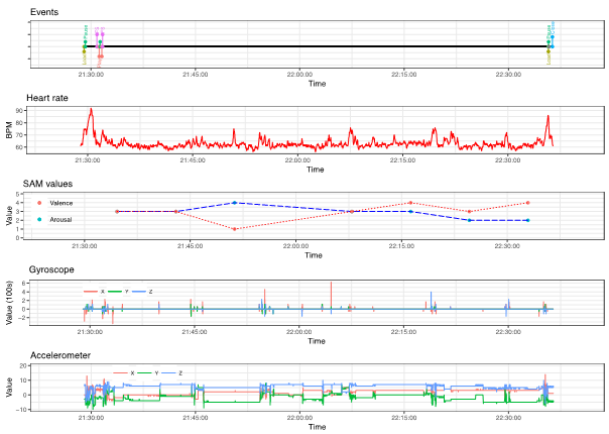

Figure 1: Data plot in the BWatch study

2.2. Engagement

Many factors were found appealing to BW, one, pointed by Pittman and Sheehan (2015), was the level of engagement of the viewers with the content watched [31]. Melanie et al. (2004) defended that stronger media effects are also related with a higher level of engagement and immersion [24]. Such levels can be created by increased levels of transportation and parasocial relationships [34] with the series being watched. That makes engagement one of the pivotal factors to be taken into consideration when studying BW.

It is possible to measure engagement and deduce emotional states in an indirect way, through biosensing data [23], this methodology is denominated as psychophysiological research [8].

The key factors described by Martin et al. (2019) are cardiovascular measures, HR, galvanic skin response, unable to distinguish valence for Electromyography (EMG), respiration rate and respiratory amplitude [23].

In this extended abstract we will focus on the HR data collection as an attempt to measure valence and arousal, related with engagement, during BW sessions in an indirect non-intrusive way. The HR can reflect the emotional state of a person. Surprise, negative emotions (i.e., anger, anxiety, embarrassment) and positive ones (e.g., happiness, joy) can lead to higher HR values [18].

Heart rate is the number of times a person's heart beats per minute.

Earliest studies applying physiological data to study video-games are from Hazlett (2003) [13] and Mandryk et al (2005) [21], which used electrodes to measure HR and other vital signals.

Some companies, such as, Valve and Ubisoft use O.Zen sensors (formally known as Innergy) for their studies when developing new video-games or features [29].

An academic study tried to measure participant's affect to a military training video-game used as sensors, a Microsoft Kinect, to detect movement, and an Affectiva Q-Sensor, to collect physiological responses to events during the study [6].

However, Q-Sensor logged data did not fully cover the whole gaming session of several participants, because of that, the values collected of HR were not used.

Another study, attempting to correlate emotions with 2 detectors, ProCompInfiniti system and sensors, and BioGraph Software from Thought Technologies, measured the galvanic skin response, electromyography and HR [20]. They successfully managed to associate some emotions to those physiological data, but in an laboratory environment.

They used the HR values when they were unable to distinguish valence for EMG alone, using high HR values to move valence from neutral to high, and low HR values to move valence from neutral to low [20].

Nogueira et al. (2015) used Blood Volume Pulse (BVP) readings measured at the thumb using a clip-on sensor to calculate the HR for biofeedback games [29].

That knowledge allowed to, not only understand user's engagement, but to improve user's experience, by using biofeedback to change the game experience.

2.3. Home health monitoring studies

Due to the lack of previous at home BW studies, it is advised to review other in the wild studies that collected physiological data with reliable devices. For that, here we will analyse some home health monitoring studies devices, that need to use very reliable, since they aim to monitor people's health, and how to properly choose one.

More recently, smartwatches are becoming more accurate for physiological data measurements, even being used to some Home Health Monitoring studies [2], due to its non-invasive methods.

On the master's thesis written by André Manso (2019), he used the Samsung Gear S3 Frontier. Despite the battery fully covered the exercises performed, the values registered had many errors, because the movement caused problems on detecting HR variability [2].

These problems were very often reported on other studies that compared measured values with a smartwatch and a reliable device. The main cause was pointed as the quick movements of the activities would hinder the detection of HR variability [32].

An application was developed in 2017 by Diogo Cabral for the M-ITI at home study, to collect physiological data while its users would watch series, it was the Bwatch App [12].

The BWatch App was developed for the Madeira-ITI study for mobile and smartwatch [12], being both connected via Bluetooth. The smartphone writes to a file every second the last values received, updating them when it gets a notification from the smartwatch when one of the values changes. If the connection is lost it does not update the values, so it is vital to make sure that both devices are near to accurately register the data.

During the study some usability problems emerged, mostly related with the battery life. Since each session can last more than 4h, and that there is constant communication between the smartwatch and the smartphone, the battery can sometimes run flat before the session.

However, to study engagement with video content is essential for BW studies to have the information of the content watched and the participant's behaviour. To do so, it is necessary to analyse the current streamer platforms and ways of obtaining that content data.

3. Streaming Media Services

In the past years, many platforms appeared to provide media streaming services, such as, YouTube, Netflix, Hulu, Vimeo, Amazon Prime, etc.

Netflix is one of the biggest on-demand content provider, with 158 million paid memberships in over 190 countries. Netflix works across multiple devices, from Smart TVs, smartphones, laptops and smartphones.

An analysis to 3.5 million logs of historic data of 174 users, on average covering 138 days of internet usage, from Google Chrome and Mozilla Firefox showed that YouTube and Netflix are used by a larger number of people [11].

Two main web browsers are Mozilla Firefox and Google Chrome. Both of them allow developers to develop their own scripts to run on user's browsers. Mozilla has Add-Ons and Google Chrome has Chrome Extensions.

Chrome Extensions can interact with different elements of the browser, such as, window, tabs (open, close, change) and cookies. It uses different sets of tools like JavaScript, html, CSS and Chrome Extensions libraries, which allow a non-static interaction with the browser. In addition, Chrome Extension update automatically, without the need of the user's permission, it is more reliable in terms of security, and it does not decrease the browser performance, compared with Mozilla Firefox [35].

The BWatch Chrome Extension was developed for the first study in Madeira-ITI related with this topic [5]. This Extension performs by loading an initial form in a new tab when the user accesses the Netflix in the Chrome browser and by attaching a tab with the list of all the video questionnaires to the tab bar. During the session the Chrome Extension detects specific events, such as, skip intro, pause/play video, new episode, audio selected, subtitles selected for certain episodes of series specified in the code. That information is registered in a data sheet per session, uploaded to a Dropbox account, it is coded using html, JavaScript, jQuery and Chrome Extension libraries. The Chrome Extension contains all the links for the forms and also the names of the different series and episodes to be analysed written down in a static way.

The BWatch Chrome Extension revealed few problems. In some computers it did not load the proper forms or it did not collect the data properly, e.g., it was not clearly registered the end of an episode or session. That lead to an inaccurate data collection of the events.

4. System design and implementation

A first version of the BWDAT was developed, followed by 3 technical and usability studies, each providing useful information, leading to 3 cycles of improvements. This present methodology represents the final outcome, simplified in Figure 2.

4.1. Streamer and Browser

The streamer selected was Netflix, and, as laptop is the most popular device to be used to binge-watch during the night, it was the chosen device.

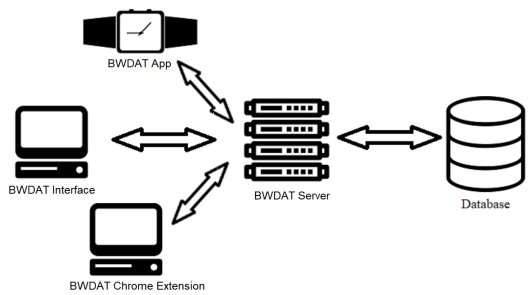

Figure 2: BWDAT Chrome Extension functional and communication model.

Google Chrome was the browser used, given the previous work done on Chrome Extensions development and the information available online.

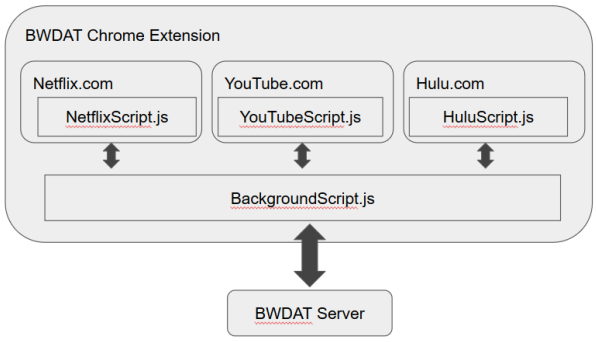

Figure 3: BWDAT Chrome Extension functional and communication model.

Unlike the previous studies, since no official definition of binge-watching has been delivered [25], and because it varies from person to person [22], it was given the freedom to the user to define the beginning and end of a session. A session starts when the user opens a Netflix tab, and ends when the user closes all the Netflix tabs, on Google Chrome browser.

Since it is expected for different streamers interface interactions to be registered in the future, for example, YouTube and Hulu, it was implemented an independent communication system between the different scripts run, on each tab, and the background script, shown in Figure 3. On a first stage, only the Netflix script was developed and implemented.

To display the forms for the pre and post study, and before and after each session, a trigger system was created. After installing the BWDAT Chrome Extension the user is prompted with the login interface.

When the user logs in the Chrome Extension, and submits its ID, if defined, the pre-study form will be displayed and the log in page closed.

The session starts when the participant opens a Netflix tab, and no more Netflix tabs are currently opened. If there is a predefined pre-session form a new tab is opened and the user taken to that tab. Once the form is complete the user will go back to the Netflix tab and start the session.

The actions on the video, such as, pause, play, forward, rewind were detected using the JavaScript functions of the video tag, this way, even if the buttons and interface changes, they are still functional. Skip intro and skip credits were triggered when the participant would click on the interface, and the target compared with both tags and they match. It is also registered the remaining time and the total length of the episode for each action.

When the participant opens a Netflix tab, it is register in a vector. If the vector is previously empty a starting session is triggered, and a request is done to the server. Every time a Netflix tab is closed, the action is registered, and, if it is the last Netflix tab open it closes the session.

The search was registered by looking for the p variable value on the URL, when the '/search' was detected.

After the trials an alarm system was implemented that verifies if any physiological data was collected 5 seconds before the beginning of the session when a session starts. If it is not found, a blocking pop-up message is prompted to ask the user to verify its device.

When the user closes all the Netflix tabs, if registered, a post session form is displayed on a new window.

To prevent no replied questionnaires, a pinned fix tab is added the browser with all the forms' URLs. Also, a value p was added to the URL to keep track of who is replying.

When analysing the sessions it was noticed some empty sessions, some were just to Login into the Netflix. In order to distinguish empty sessions from the Login ones, the Login detection was here implemented by reading the URL.

4.2. Physiological Data

For the physiological data acquisition Polar m600 was the chosen device, because it has Wear OS 2.0 and its 6 LEDs and reviews reveal a good battery lifespan, with an acceptable commercial cost. Android Studio was chosen to develop the application, which uses Java, very popular among developers, and Android covers a wider range of devices.

A novel approach was taken, compared with previous studies. In order to limit the need of hardware interactions, it was developed a standalone type of Smartwatch application. This way, first, the user only needs to interact with the smartphone one time, to insert the Wi-Fi password on the. Secondly, despite Bluetooth is more battery friendly, by

using a direct Wi-Fi connection between the smartwatch and the router, as it has a wider range, it allows the participant to have more freedom.

Despite only the Polar m600 model was chosen for this study, not only a square but also circular screen shape were developed to allow all the other models to support this application in the future.

To identify each smartwatch, since restrictions are applied on accessing the smartwatch hardware identification number, an incremental number is generated by the server and stored on the smartwatch application's memory cache. Without this credential no data can be collected.

If during the initial acquisition of the ID there is no Wi-Fi connection the user will be prompted with a warning. Also, when the Bluetooth is on, it will be switched off when the participant presses 'Start!', because the smartwatch will primarily use Bluetooth for http requests to save battery, and we want direct WiFi connection.

Every time that a value on the sensor changes, that value is updated in the memory cache of the Smartwatch. The last registered values are sent to the server every second.

To test the data reliability of the Polar m600, 4 healthy people used the smartwatch and Empatica E4 to compare the HR values collected by both of them.

The tests showed that, despite User 1 had an 11% mean absolute error, on average it is below 10%, so we can consider that Polar m600 has an acceptable heart rate accuracy.

4.3. Back-end

A server created using IST Virtual Machine technology maintained a good level of reliability and security.

For this project a Virtual Machine from IST was provided. Its Operating System is Debian/Ubuntu, with 2 vCPUs, 4GB of RAM, with 20GB of disc memory.

The language chosen to develop the server was Node.js, since it is based on JavaScript and it is well accepted for the development of applications by big companies, such as, NASA, Walmart, Intel, mostly due to its scalability power.

It is very important to always have the application running, therefore, it was taken the decision that all the errors were logged into a log file instead of crashing it.

A connection to the database is created when a request is made and closed after sending the query response in order to keep a single DB connection open per request.

A DB from IST was provided to register all the data collected. It is MySQL based and has the capacity to store 10GB of information.

The design of the tables was focused on allowing a good scalability, since this project is open for future upgrades. It was defined that a user always needs to be associated with a project, and every Chrome Extension data needs to be associated with a user and project.

The devices work in an independent way. Each device has a unique number that will register the data when 'Start!' is pressed and stop when 'Stop!' is pressed.

The association of the physiological data of the device and the participant's sessions is done through a correlation of a device and a user and project, with an initial and an optional end time.

This way, we allow studies with and without devices, and, in the future, to be associated with other type of input data, not only the Chrome Extension.

4.4. Graphical interface

To manage and visualize all the data a web interface was developed using JavaScript, HTML and CSS since they are more dynamic and actual web development languages.

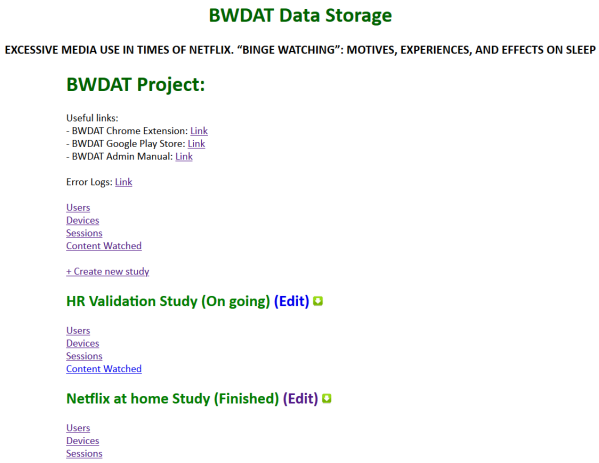

Figure 4: Main BWDAT Admin interface

The main interface (Figure 4) has all the important links and access to the different projects.

To create a project it is required a unique project name. It is possible to choose to have a study with or without devices, specific content to be registered or all the forms' URLs mentioned before.

It is possible later edit the information predefined for each project. The project state (Not started, On going and Finished) informs the state of the project to help keep tracking of the current projects running.

The User's interface display all the different users and it allows to edit the end time or remove them. When creating a new user a new random unique user code is generated in the server and it dynamically adds to the table, allowing to create many users in a row easily. To study the evo-

lution of the participant throughout the study, on the User's interface, when clicking Visualize, a series of graphics is displayed, see Figure 6. For more details of each session, when clicked on the graphic, it redirects to the page with the full session graphical and tabled values description.

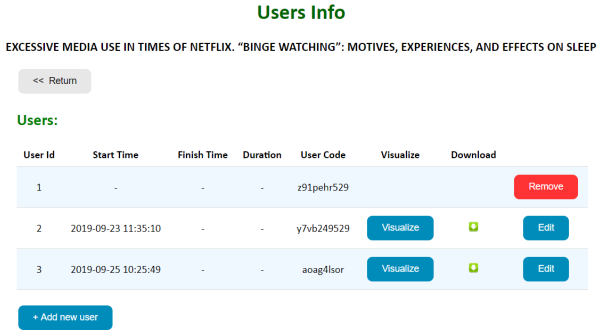

Figure 5: Users interface

Figure 6: User 3 Sessions Interface

The project's owner has to record when a participant receives and returns the devices. To help keep tracking a table displays which user has each device and when they are available.

The longest session registered had 6h, when the participant fell asleep while BW [5]. Therefore, it was proven that Polar m600 battery, with the current implementation, lasted enough for the required and desired time stamps of the study.

Also a simple interface was developed to compare the different user's actions for the same episode (content watched).

5. Results

39 participants were recruited through an online platform. All of them residents in Switzerland, with ages between 21 and 52. From those only 34 had valid sessions.

Some users were asked to watch 3 full episodes in a row before sleep, and the others were asked to watch only 1 episode before sleep, for 3 days. This way, we consider the first group as the BW group and the second one the control group. The data acquisition lasted 4 months, from June 2019 to September 2019.

Only 9 smartwatches were used on this study, since it was decided to connect previously the smartwatches to mobile devices, so participants did not have to connect to theirs, for privacy issues. The devices used were Nokia Smartphones (Nokia 5) and Samsung Tablets (Galaxy Tab E). To deliver the devices, an appointment was scheduled and usability instructions were given to the participants. Also, an instruction form was sent to the participants with the links for the different questionnaires.

To analyse the data the BW sessions were divided from the control group. 15 participants made part of the control group and the remaining 19 were the BW group. We will focus our analysis on the second group.

From the 19 sessions, 6 had 100.00% coverage and 17 had more than 90% coverage. We understand coverage as the % of the session that has HR registered, for example, a 1h session should have 3600 HR values. 3 of the other 13 sessions with less than 100.00% coverage had a 100.00% data transmission, in which, since the participant pressed 'Start!', on the smartwatch, until the end of the session no data was loss. But, as users 23, 24 and 25 only pressed the button few seconds after some content was being watched, it didn't covered the whole session. On average, each session had a 89.52% coverage, with a 89.99% data transmission.

The missing gaps of physiological data, that managed to proceed after the collection, lasted between 5 seconds to 1 minute. All the sessions with less than 81% coverage had a stop in a certain point in the session which never recovered until the end, see as an example Figure 7.

Some of the sessions with a lower HR coverage had some reported issues. User 9, which had a 62.55% coverage, reported being interrupted and problems with the device. Number 16, with a 34.38% coverage, reported internet connection problems. Participant 18 answered the phone, which may had led to only a 80% coverage, and participant 37 reported low technical problems with the smartwatch battery, leading to a 90% coverage and not 100%.

Regarding the Chrome Extension, no major faults were reported. Some users did not close the Netflix right after the session, replying first to the post-session form sent by e-mail. However, is easy

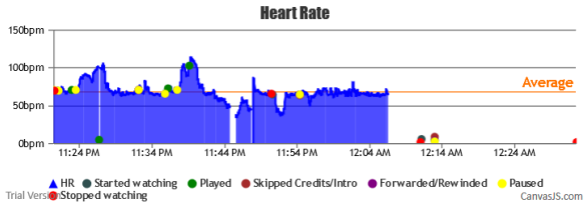

Figure 7: Example of a data collection fail in a session

to see that viewers quickly learnt that the post-form would be triggered by closing the tab, since the upcoming sessions had their participant ID added by the system to the URL. The forms replied using the e-mail didn't have that component.

On average, 1 pause and 1 forward were done per episode, both on BW and non BW sessions. The majority of the pauses lasted less than 1 minute. Only few pauses were longer than 1 minute, but shorter than 5 minutes.

On average a session lasted 01:48:31 (HH:MM:SS), despite the average sum of the episodes per session was 01:57:00. That can be explained because some participants skipped the intro, and others did not fully watch all the episodes, forwarding some of the parts. Those values are similar to the study run by Deborah et. al (2019) which registered a 02:10:40 average time of BW sessions [5], and the predicted of 02:07:41 by Rigby et. al (2019) [33].

No correlation was found between the duration of the session and the data coverage, see Figure 8 and Figure 9.

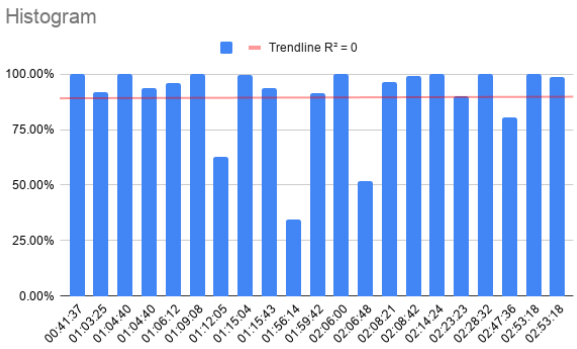

Figure 8: Coverage of BW sessions

The sessions started on average at 21:45:38 and ended at 23:34:10. When comparing BW sessions with the control group, we can find that non BW sessions started later, on average at 22:44:31, but ended almost at the same time as the BW group, 23:24:04. This is interesting, since the values are very close to the ones predicted by Rigby et al. on which BW sessions would occur between 9PM and 11PM.

Also, it was found that, unlikely of what was expected in some papers, users found the played content, not much by searching keywords, but mostly from the main screen. On the other hand, it showed that the average time of each episode was 39minutes, similar to the median value registered by Frank et. al (2019) [11] of 44minutes.

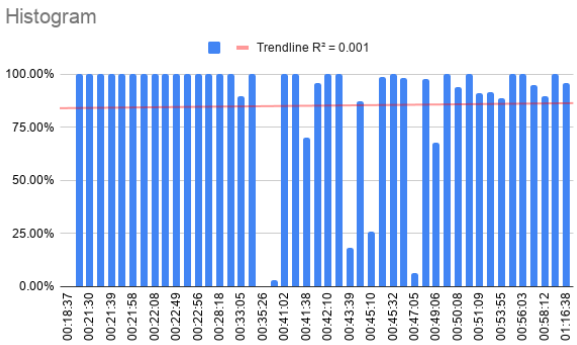

Figure 9: Coverage NBW sessions

When comparing the average HR values from both groups we can see that both had similar minimum values (NBW = 55bpm and BW 56bpm), but the average HR of BW sessions were higher (NBW = 64.60bpm and BW = 72.63bpm) and higher values were registered on BW sessions (NBW = 85.86bpm and BW = 99.14bpm). This proves what was predicted previously, that BW sessions create more valence and arousal on the viewers, therefore, registering higher values of HR.

6. Conclusions

Taking into consideration everything that was written before, here we draw our conclusions and design the desired future work.

The present project is part of a novel study that proposes a new and interdisciplinary approach to better understand the binge-watching phenomenon. In particular, the study uses new tools that allow the collection of physiological data in a non-intrusive manner, and correlates it with self-reported data.

Due to the complexity of the information gathered and the need to synchronize it in a very accurate fashion, the development of a set of tools was pivotal and this, precisely, is the contribution of this master thesis.

First, a Chrome Extension that records users' interactions on Netflix. Second, a smartwatch app that collected heart rate and inertial data with a lifelong battery that managed to covered the BW sessions with more than 89% rate. Third, the server and DB delivered a good performance to the flux of data registered.

The interface provided was a major help on giving independence for scientist create new studies and manage to analyse the data collected in a clear

way.

This technology allowed to confirm in and empirical way, for the first time ever in a publication, the behaviour and engagement expected in BW sessions with physiological data in the wild.

6.1. Future Work

As stated before, this project is the beginning, rather than the end, of an ongoing journey to develop a limitless set of tools to explore binge-watching and other phenomenons in the wild.

Along the way, few limitations were found to this technology. For example, it is hard to distinguish from when the user is actually watching while the content is being played, or if it is performing other activities, for example, texting on WhatsApp.

As future development, it is suggested as the main feature required to be implemented is a backup system on the smartwatch, for when the Wi-Fi connection is lost. It should also be implemented on the Chrome Extension on a next stage.

A similar system to BWDAT developed for smart TVs should also be further explored.

Since the usage of the smartphone was the most common activity that generated interruptions, it would be useful to develop a smartphone App that would register phone activity during the sessions, running in the background.

Since skin temperature can give valuable information about user's emotions, it would be interesting to adapt one device to collect that information, using for example an iButton.

Despite this study focusing on the effects of BW on the sleeping behaviour, a more extended study, to analyse viewers engagement throughout several months would be interesting, since the technology proved capable of supporting such kind of projects.

Despite Polar m600 performed well, further testing should be done using other smartwatches that support the BWDAT App.

References

[1] Amanda Lotz. Portals: A Treatise on Internet-distributed Television. *Ann Arbor: University of Michigan Press.*, 2017.

[2] André Manso. Long Term Real-Time Pervasive Monitoring System for non-Hospitalized Patients. *Master's Thesis, Instituto Superior Técnico*, July 2018.

[3] F. Bentley and J. Murray. Understanding video rewatching experiences. *ACM Conference on Interactive Experience for TV and Online Video (TVX'16)*, pages 69–75, 06 2016.

[4] Bradley, M. M., Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behaviour Therapy and Experimental Psychiatry, 25(1), 49-59.*, 1994.

[5] Deborah Castro, Jacob Rigby and Diogo Cabral and Valentina Nisi. The Binge-Watcher's Journey. Investigating Motivations, Contexts, and Affective States Surrounding Netflix Viewing. *Convergence*, 2019.

[6] J. DeFalco, J. Rowe, L. Paquette, V. Sherry, K. Brawner, B. Mott, R. Baker, and J. Lester. Detecting and addressing frustration in a serious game for military training. *International Journal of Artificial Intelligence in Education*, 28, 09 2017.

[7] Deloitte. Digital Democracy Survey. A multi-generational view of consumer technology, media and telecom trends., 2015.

[8] C. DeWall, C. Anderson, and B. Bushman. *Handbook of Psychology, Second Edition*. 09 2012.

[9] Dimph de Feijter, Vassilis-Javed Khan, Manix S. van Gisbergen. Confessions of a 'Guilty' Couch Potato. Understanding and Using Context to Optimize Binge-watching Behaviour. *In proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, Chicago, USA, June 22-24, 2016, pp. 59-67.*, 2016.

[10] Emil Steiner, Kun Xu. Binge-watching motivates change: Uses and gratifications of streaming video viewers challenge traditional TV research. *Convergence, 1-20*, January 2018.

[11] Frank Bentley, Max Silverman and Melissa Bica. Exploring Online Video Watching Behaviors. *TVX 2019, June 4–6*, June 2019.

[12] Harry Vasanth. BWatch., 2017.

[13] R. Hazlett. Measuring emotional valence during interactive experiences: boys at video game play. pages 1023–1026, 01 2006.

[14] Helena Govaert. How is the concept of "binge-watching" of TV shows by customers going to impact traditional marketing approaches in entertainment sector? *Universiteit Gent*, 2014.

[15] J. Horvath, A. Horton, J. Lodge, and J. Hattie. The impact of binge watching on memory and perceived comprehension. *First Monday*, 22, 09 2017.

[16] Jacob Rigby, and Duncan Brumby, Sandy Gould and Anna Cox. "I Can Watch What I Want": A Diary Study of On-Demand and Cross-Device Viewing. *the 2018 ACM International Conference*, pages 69–80, June 2018.

[17] Karyn Riddle, Alanna Peebles, Catasha Davis and Fangxin Xu. The Addictive Potential of Television Binge Watching: Comparing Intentional and Unintentional Binges. *Psychology of Popular Media Culture 4: 589–604*, 2017.

[18] S. Kreibig. Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84:394–421, 04 2010.

[19] Liese Exelmansa, Jan Van den Bulckb. Binge Viewing, Sleep, and the Role of Pre-Sleep Arousal. *Journal of Clinical Sleep Medicine*, 2017.

[20] R. Mandryk and M. Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *Int. J. Hum.-Comput. Stud.*, 65:329–347, 04 2007.

[21] R. Mandryk and K. Inkpen. Physiological indicators for the evaluation of co-located collaborative play. pages 102–111, 01 2004.

[22] Mareike Jenner. Binge-watching: Video-on-demand, quality TV and mainstreaming fandom. *International Journal of Cultural Studies 2017, Vol. 20(3) 304–320. doi 10.1177/1367877915606.*, 2015.

[23] Martin Čerticky, Michal Čerticky, Peter sinčak, Gergely Magyar, Ján Vaščák and Filippo Cavallo. Psychophysiological Indicators for Modeling User Experience in Interactive Digital Entertainment. *Sensors (Basel)*, February 2019.

[24] Melanie C. Green, Timothy C. Brock, Geoff F. Kaufman. Understanding Media Enjoyment: The Role of Transportation Into Narrative Worlds. *Communication Theory 14: 311–27*, November 2004.

[25] J. Merikivi, J. Bragge, E. Scornavacca, and T. Verhagen. Binge-watching serialized video content: A transdisciplinary review. *Television & New Media*, xx:1–15, May 2019.

[26] L. Mikos. Digital media platforms and the use of tv content: Binge watching and video-on-demand in germany. *Media and Communication*, 4:154, 07 2016.

[27] Netflix. Netflix Annual Report., 2017.

[28] Netflix survey. Netflix & Binge: New Binge Scale Reveals TV Series We Devour and Those We Savor., 2016.

[29] P. Nogueira, V. Torres, R. Rodrigues, E. Oliveira, and L. Nacke. Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game. *Journal on Multimodal User Interfaces*, 10:1–32, 12 2015.

[30] Ofcom. The communications market report 2017., 2017.

[31] M. Pittman and K. Sheehan. Sprinting a media marathon: Uses and gratifications of binge-watching television through netflix. *First Monday*, 20, 10 2015.

[32] R. Reddy, R. Pooni, D. Zaharieva, B. Senf, J. El Youssef, E. Dassau, F. III, M. Clements, M. Rickels, S. Patton, J. Castle, M. Riddell, and P. Jacobs. Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise (preprint). *JMIR mHealth and uHealth*, 6, 03 2018.

[33] J. Rigby, D. Brumby, A. Cox, and S. Gould. Old habits die hard: A diary study of on-demand video viewing. *2018 CHI Conference on Human Factors in Computing Systems - CHI '18 Montreal, Canada 21-26 April*, pages 1–6, 04 2018.

[34] Sarah E. Erickson, Sonya Dal Cin and Hannah Byl. An Experimental Examination of Binge Watching and Narrative Engagement. *Social Sciences 8: 19*, January 2019.

[35] Search Engine Journal. FireFox Addons vs Google Chrome Extensions.

[36] Université de Fribourg. Excessive Media Use in Times of Netflix. "Binge-Watching": Motives, Experiences, and Effects on Sleep, 2019.

[37] Yoon Hi Sung, Eun Yeon Kang and Wei-Na Lee. A Bad Habit for Your Health? An Exploration of Psychological Factors for Binge-Watching Behavior. *65th Annual International Communication Association Conference, San Juan, Puerto Rico, May 21-25 2015.*, 2015.