



Machine Learning the TV Consumption: A Basis for a Recommendation System

Bernardo Cardoso^(✉) and Jorge Abreu

University of Aveiro, Aveiro, Portugal
{bernardoc, jfa}@ua.pt

Abstract. With the continuous growth of channels and content available in a typical interactive TV service, viewers have become increasingly frustrated, struggling to select which programs to watch. Content recommendation systems have been pointed out as a possible tool to mitigate this problem, especially when applied to on-demand content. However, in linear content, its success has been limited, either due to the specificities of this type of content or due to the little integration with normal consumption behaviors. Despite that, recommendation algorithms have undergone a set of enhancements in order to improve their effectiveness, particularly when applied to the world of linear content. These improvements, focused on the use of the visualization context, paired with machine learning techniques, can represent a big advantage in the quality of the suggestions to be proposed to the viewer. The area of user experience (UX) evaluation, in interactive TV, has been also a subject of ongoing research, extending beyond the traditional usability evaluation, pursuing other dimensions of analysis such as identification, emotion, stimulation, and aesthetics, as well as distinct moments of evaluation. This paper presents the proposal for the development of a recommendation system, based on the viewing context, and a methodology for evaluating the way this system influences the UX of the viewers.

Keywords: TV · UX evaluation · Recommendations · Visualization context

1 Introduction

In the last decades, the quantity and quality of the content available in a typical living room have increased considerably. Television, the main distributor of this content had, concurrently, a clear evolution. Not only have the number of available channels improved significantly, particularly on paid TV platforms and in more developed countries, but also, as a result of digitalization and interactivity, a number of additional features have appeared, granting a considerable increase on the available content. These new functionalities include, for example, video on demand (VOD), automated catch-up TV, digital video recorder (DVR), and over-the-top (OTT) media services like Netflix and Hulu, which can be classified in a taxonomy that distinguishes between linear content, i.e., traditional television, transmitted in broadcast mode and nonlinear content, which include services that provide movies and television programs by request [1].

All this increase in content and functionality can add up really quick, resulting in a much wider choice for the end users. Accompanying this, there is also a progressive complexification on the interaction modes. In one hand, the amount of content is so great that the viewer has difficulty in selecting a proper program to see, attaining a so-called “Paradox of Choice” [2, 3]. On the other hand, the tools that could help him in this task are of limited practical use, partly because of the physical limitations of the equipment itself (the user is too far away from the screen to be able to discern very detailed elements) and also because he interacts through a remote control with limited interaction features, which are not always adapted to the new interactive TV services [4].

This paper will focus on a proposal to develop and evaluate the user experience of an interactive TV application, dubbed TV Concierge, that aims to mitigate this paradox of choice, by artificially reducing the quantity of content presented at each moment to the user, through the use of personalization and, at the same time, building a minimalist user interface that further limits the number of interactions needed. In this way, in Sect. 2 we address a set of recently researched recommendation algorithms and techniques that rely on the TV consumption context and can be applied to linear TV content. In Sect. 3 we present the way we conceive the use of machine learning in the implementation of the recommender system. Then, in Sect. 4 we provide the results of our initial data analysis and also a visualization for a sampling of this data. In Sect. 5 we address the development of the correspondent interactive TV application and some of the mockups already designed. Next, on Sect. 6 we describe our user experience evaluation methodology and how we envisage executing it. Finally, in Sect. 7 the paper is wrapped up with the exposition of the results we expect to achieve in the upcoming months and some brief conclusions.

2 Personalization and Recommendations in Linear TV

In order to mitigate the nuisances caused by the content proliferation and a situation where the user expends a lot of time looking for content to watch (for instance, DigitalSmiths reports that 45.7% of respondents spend 5–20 min per day channel surfing [5]), several proposals have been made, most of them focusing on the use of recommendation and personalization systems, in order to facilitate the discovery of interesting content to watch [6]. These systems have achieved effective success in on-demand video subscription platforms such as Netflix and the traditional VOD platforms of the various pay-TV operators, where there is a significant control over the offer of content to recommend and where the user is already predisposed to a more advanced level of interaction, namely with a greater propensity to assign scores to the viewed content. However, in linear content, recommendations are still not an effective tool [6] and end up being just another functionality, lost among many others and far from being a solution that truly contributes to an effective mitigation of the problem [7].

Some particularities of linear content add greater complexity in creating an effective recommendation system, namely the content catalog being in constant change and the content being available only for a short time due to the TV channels programming being characterized by its constant renewal [6]. A recommendation system, that only has

access to the linear channels programming, can only recommend, at any moment, the programs that these channels are broadcasting or programs that will start in the upcoming minutes. Even systems that have access to a catch-up TV catalog, need to deal with the fact that fresh content is entering the collection all the time, since the system is constantly recording new programs and, similarly, removing older ones [1]. In contrast, VOD recommendation systems do not need to take these factors into account, as their catalogs have much slower turnover cycles.

Another characteristic of linear TV consumption that should be emphasized is that it normally follows a very regular pattern [6]. Contrary to a VOD system where the viewer usually wants to find a new movie to watch, when watching TV the consumer has habits associated with certain times of the day and follows specific recurring programs on a small number of available channels [5]. This regularity together with other contextual aspects of TV watching was also identified by other authors [3, 8, 9], who take them as the basis for proposing new approaches in recommending linear television content.

In [3], the authors looked for a better understanding of the contextual aspects involved in TV and video consumption on a domestic environment through the execution of two ethnographic studies. In the first, which involved 12 households, they observed how the household structure determines the viewing situations and the relations between the amount of attention, type of content and viewing behavior – planned, routine, or spontaneous. They also observed the different ways in which people discover new content. In a second, multi-method study, comprising 7 families with children, typical situations of visualization and their contextual aspects were assessed. After combining the results of both studies, they recognized seven viewing situations: weekend mornings; when the children are sleeping; family quality time; relaxing after school; a free moment: men and sports; and lazy afternoons. In each case they identified the contextual aspects: mood; content; viewers; time; content delivery time (linear or on-demand); and viewing mode (attention level associated with consumption). With the results of these studies, they make several proposals for recommendation systems, algorithms and user interface designs that could take into account these contextual aspects. Unfortunately, some of these contextual aspects are not easily assessed in an automated way, mood and viewing mode for instance, and cannot be used as-is in a non-intrusive solution.

Other authors found out that linear TV consumption patterns are strongly conditioned by time context and channel preferences [6]. In this sense, they propose that one way to go beyond the state of the art on the current recommendation systems for linear content is to explore and integrate this visualization context (time) in the user consumption modeling. Through empirical evaluation with a broad set of linear TV data, they demonstrate a significant improvement in the quality of the recommendations when the time context is taken into account. This usage of time context can be improved with the addition of implicit feedback, taken from consumer data analysis, and taking into account not only the linear TV services, but also the catalog of catch-up TV available today from most operators. Comparing this approach with current algorithms it has been shown it can be superior to these in accuracy, while maintaining good levels of diversity and serendipity [9].

In addition to the temporal context, an attractive idea, that has been shown to provide considerable precision gains in the recommendations, is the usage of a sequential context [8], which takes into account the last viewed program at the time of the recommendation to influence it.

An important aspect of TV consumption is that it is often shared by several different users, whose tastes can vary widely. Typical recommendation systems do not handle this situation very well, since visualization data is typically collected and modeled at the device level, aggregating all users and obscuring their individual tastes. The ideal contextual information would be for the system to know in each moment who is watching and their level of attention. That level of information is hard to attain today, without the introduction of additional hardware in the living room. A further layer of additional contextual information, that can be obtained automatically, must still be devised to mitigate this issue.

3 Technical Approach

Most recommender systems fall into two basic types: collaborative filtering and content-based filtering. In the first, a profile of a user is built with information from past behavior, i.e., books or movies purchased or ratings given. Thereafter, this model is used for comparison with models from other users, looking for similar profiles. The new/different items are then used as recommendations. Content-based filtering, on the other way, uses intrinsic characteristics of the items, like subject, color or length to find similarities between items, and use the related outcomes as recommendations. The difference between these two basic types can be demonstrated in an online bookstore like Amazon: collaborative filtering produces recommendations like “Customer that purchased this book also purchased these and these” and content-based filtering generate suggestions like “Books similar to this one are x and y”. These approaches are often combined in so-called hybrid recommender systems [10].

These standard approaches focus on the most relevant items to recommend to a user but do not take into consideration any contextual information, such as time, place, and the company of other people. They just deal with two basic types of entities - users and items - not putting them into a context when providing recommendations.

However, as we presented in the last section, context can be a fundamental entity for personalization in linear TV. Context can also impact situations already tackled by common approaches, for example, the rating a viewer gives to a movie also depends on where and how he viewed the movie, at what time, and with whom. Also, the type of movie a person recommends to a colleague will differ significantly depending on whether he is going to see it on a Monday with his parents or on a Friday night with his girlfriend [11]. This awareness leveraged the introduction of context-aware recommendation capabilities in different applications (namely in mobile platforms where location information provided by GPS has a huge impact on the acuity of the recommendations) and the development of a new research area, related to context-aware recommendation systems. For example, in [11], a notion of context is presented along with how it can be modeled in recommender systems. The major approaches to modeling contextual

information are also discussed through three main algorithmic paradigms: contextual pre-filtering, post-filtering, and modeling. The TV Concierge interactive application will make good of this research and will apply it to the recommender system being developed.

At this moment, the technical approach for the development of the recommender system is to base it on decision trees, one of the most used building blocks in current machine learning approaches. The choosing of decision trees for the building of the recommendation system was based on several benefits they offer: efficiency, interpretability, and flexibility in handling a variety of input data types (ratings, demographic, contextual, etc.) [12].

Decision trees are a very simple notion. At the core, they are nothing more than a cascade of if-then-else(s). Using the tree concept, the nodes represent questions, being the leafs the final decisions. Materializing a little the concept within the TV Concierge interactive application, the questions could be like “Is today a weekday?”, “Is the current time noon?”, “Was the previously watched channel MTV?”, and so on. A sample tree like this can be seen in Fig. 1. After answering the sequence of questions, the system will reach a leaf where there will be one or more TV programs that will be used as the recommendations.

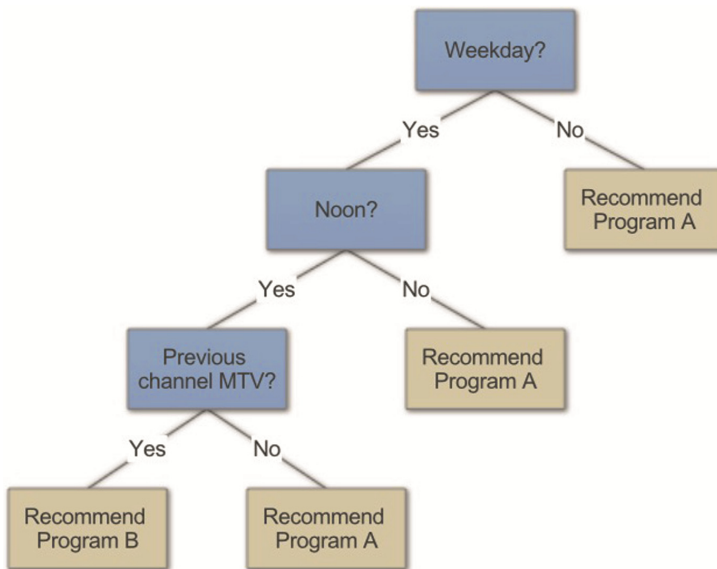


Fig. 1. Sample decision tree for program recommendation

Efficiency is very important because the system will need to create a personalized recommender system for each individual STB, and will use real-time data like the previously watched program or the current time of day as the basic input to generate a recommendation. In this framework, it needs to be very efficient to be able to generate a new program recommendation in a timely fashion. Decision trees, being a simple cascade of if-then-else(s), are very fast to use and relatively easy to construct.

Interpretability is a real plus in the TV Concierge interactive application because in some of the current recommendation systems, namely the ones based on neural networks, is very difficult to understand how the system arrived at some recommendation [10]. In the case of decision trees, interpreting the result is as simple as backtracking the tree and at each node evaluating what was the deciding attribute.

Finally, the capability to handle a lot of different data types is also a must, because TV Concierge will need to process different things like time-of-day, day-of-week, current channel, previous program, etc.

Obviously, decision trees can be written by hand, in an ad hoc way, actually, that is what most computer programmers do most of the time, since a computer program is a decision tree with some additional concepts on top of it. The use of machine learning allows for the computer to essentially write the program itself, based on the recollected usage data, in this case, from each STB. There are different algorithms to construct decision trees from a dataset, but the main idea is to uncover the most significant attribute within the data (the one with most information gain) and root the tree with it. Then, iteratively repeat the process with the remaining unused attributes, until there are no more attributes or all the elements are indistinguishable, at which point the identified items are converted in leafs representing the decision [13].

The advantages of having the computer automatically generate the decision trees are evident, since it allows each and every STB to have a personalized recommender (the decision tree), but also because sometimes the machine learning approach reveals hidden associations in the data that are not easily perceived at first sight.

The main limitation of this approach is the so-called “cold start” problem. This kind of recommender, based on historical data, is incapable of providing any recommendation until a minimum dataset is gathered. How to overcome this issue in TV Concierge is still being researched at the moment.

4 Initial Data Analysis

In the preparation for this proposal, to partially validate our assumptions and technological approaches, we have been collecting usage logs from a set of consenting users on a commercial IPTV platform (based on Ericsson Mediaroom [14]). This allowed us to understand the challenges associated with processing these events, which in real life are very noisy and need a lot of work before being in a way that can be used in a practical way. The same challenges have been found before with the same platform [15] and we resort to some of the same solutions in the pre-processing of the data, like only taking in account visualization events that spanned more than 5 min or preemptively ending watching periods that took more than 4 h without any user interaction.

However, our main interest in this initial data analysis was to find out if there were indeed easily identifiable patterns within the television usage of real users. For that, and because patterns are more easily spotted visually than analytically, we took one month of usage events from the STBs and plotted them in an agenda like view, grouping the events by day of the week (e.g. all the views of the different Mondays were grouped in a line titled Monday, etc.). The events were also color-coded by channel and titled with

the channel name and the viewing mode (live, catchup, VOD, DVR, etc.). An example plotting for two weekdays can be seen in Fig. 2. In the full plot we can also see that weekdays are much more regular than weekends. It is also noticeable that the period from 20:00 to 0:00 is where there is more turbulence. Despite that, in this example, the timeslot from late-night crossing Monday to Tuesday is very regular.

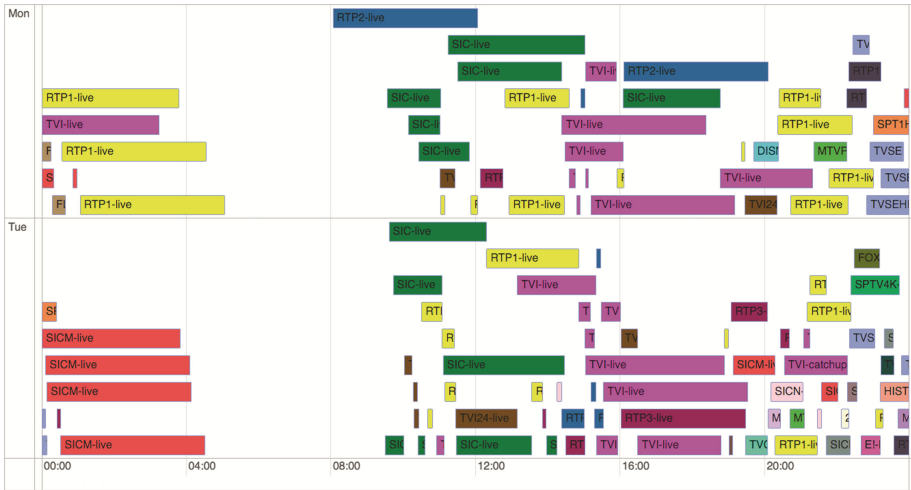


Fig. 2. Mondays and Tuesday events for one STB through the period of one month (Color figure online)

This was in line with our expectations and we found that there are indeed relevant patterns in the watching data, and even if the patterns are less evident in the primetime timeslot (the most prized time in TV) this was also anticipated, representing perhaps a lot of channel surfing and the endless pursuit of content that TV Concierge aims to mitigate. We also expect that the machine learning algorithms will catch some of the hidden, but latent, patterns still present in the data.

Anyhow, we assume that the introduction of the TV Concierge interactive application, if successful, can mitigate appreciably this turbulence. This will be actually one of the ways in which we could validate the impact of the TV Concierge, by comparing the variability of watched programs before and after the introduction of the system.

5 Prototype Design

With new recommendation algorithms, that take into account diversified contextual aspects of TV consumption, that are more focused on linear content problems, we have the foundations to the most important backend component of the TV Concierge interactive application. However, the introduction of these features cannot be dissociated from the way the viewer can access it. It is important to note that, although there is a significant research specifically regarding interfaces to interactive TV recommendations systems, some authors

report that many users preferred to have few or no interactions with the system [16]. Nevertheless, current implementations rely mainly on solutions where the recommendation is only realized after some kind of user activity [17, 18], sometimes forcing the response to a series of questions before any recommendation can be made.

This mismatch between the user expectations and the way the features are implemented shows that the relationship between the viewer and the TV goes far beyond the user interface, covering a set of other dimensions in what can be called, generally, the user experience (UX). In the ISO 9241-210 standard, UX is considered as the “person’s perceptions and responses resulting from the use and/or anticipated use of a product, system or service” [19]. In the notes associated with this definition, this standardization body stresses that UX includes all the emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and achievements associated with the user, whether they occur before, during or after use.

In “The Paradox of Choice: Why More Is Less” [2], the psychologist Barry Schwartz, explains why a lot of choices can be detrimental to people’s psychological and emotional well-being. It explains how a culture that grows with an infinite availability of constantly evolving options can also foster deep dissatisfaction, which can lead to paralysis in decision making and, in some cases, depression. It also describes the difficulty of choosing a program to view when there are so many available, and that in this sense the (traditional) recommendations systems still make the situation worse, since they always propose something new, never seen.

Combining these insights, our prototype design took an approach that tries to minimize the number of interactions the user needs to perform with the system. Simultaneously we aim to bring the recommendations to the upfront of the platform – in this respect, the system will not just wait for the user to ask for some recommendation but will preemptively suggest them. This will start from the very beginning, i.e., when the user turns on its Set-top Box (STB) it will be presented with the program that makes sense to play at the time it is turning the box on, rather than with the last channel tuned in the previous night. In addition, making use of the time-shifting capabilities of modern TV platforms, the suggested program will start from the beginning. For instance, if a potential viewer, who usually watches the TV newscast when he gets home about 8:00 pm, turns on the STB around this time, the recommendation system will automatically start playing the newscast that the viewer usually sees and not a spurious channel, kept from the last viewing session. This use case is illustrated in the left side of Fig. 3.

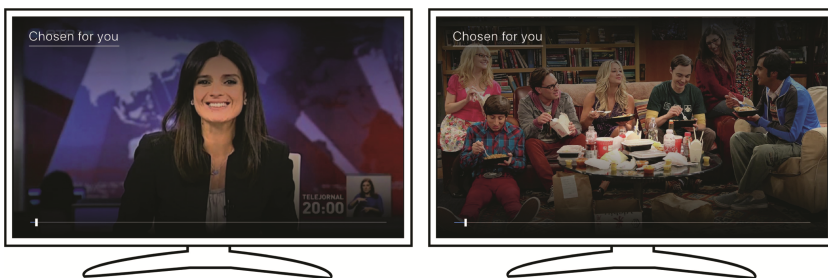


Fig. 3. Mockup for the automatic program suggestion

It is also intended that when a program finishes and whenever the system has a high degree of confidence, it will automatically start playing the next suggestion, without any intervention from the viewer. This situation is illustrated in the right side of Fig. 3 – at the end of the newscast the system suggested an episode from a series and began its reproduction automatically, from the beginning.

When the system does not have a sufficient degree of confidence, it will select a reduced number of proposals for the viewer. In this case, the playback will still start automatically, but the viewer will get, for a few seconds, the opportunity to choose another program to watch, a concept that is visible in Fig. 4, mimicking a typical binge-watching scenario [20], which is somewhat the brand mark of the current OTT video-on-demand systems from the internet.

It is also possible that the system has no suggestion to propose at a certain time, for example when no pattern has been identified, in this case, the system needs to propose a new program. Our approach to this situation is that the system will suggest a new release (e.g. a new season of a TV series) between a limited set of the most watched channels on that STB. The idea behind this methodology is that, usually, if a series that a viewer normally watches has ended, a new one will be released to take its timeslot, and the system will offer that. If that does not happen, some other channel that the viewer also watches will have something starting to promote. We already know that the user likes the channel, so there is a higher probability that it also likes a new show from that channel. This is the purpose of the interface shown on the right side of Fig. 4 – in this case, we opted for the binge-watching interface for the user to have an opportunity to actually select the new program.

Although the interactive application will start automatically when the STB is turned on, and will keep providing suggestions and playing content in an automated way, this does not mean that the user relinquishes all the control of the STB. The user can, at any moment, use his interactive TV system as usual and TV Concierge will disappear. It can be summoned again by the standard means of the interactive platform, for instance with a menu item or from a dedicated button in the remote, but it will also restart offering suggestions automatically if it detects that a program the viewer is watching just ended or upon detection of what we call a “mindless zapping”, that is, when the user starts zapping in a pattern that appears to be just hunting for something to watch.

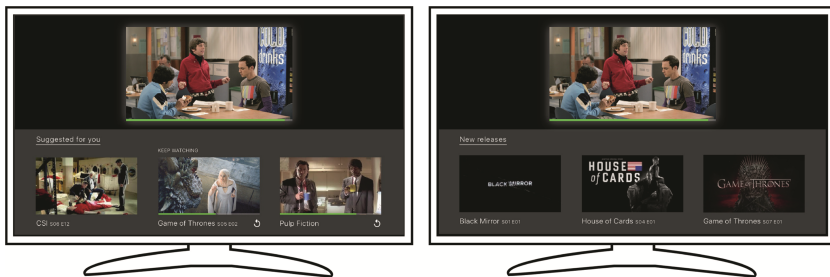


Fig. 4. Mockup of the system allowing the selection of the following suggestion

6 UX Evaluation

As stated in the introduction, one of the key aims of the TV Concierge interactive application is to evaluate the way that an interactive TV recommendation system, based on the viewing context as described in the previous section, influences the UX of its viewers. To achieve this objective, we will address several dimensions of the UX, namely: usability, aesthetic, emotional, stimulation and identification, through the use of a methodology based on a framework that uses previously endorsed and validated tools [21], and is highly aligned with some previous similar evaluations [22, 23].

In our context of evaluation, the usability dimension can be understood “as being a general quality of the appropriateness to a purpose of any particular artefact” [24]. The aesthetics dimension portrays how visually pleasing or beautiful something is perceived. The emotion dimension portrays the emotional response to the outcome of the interaction with the product or service. The stimulation dimension describes the way a product addresses the human necessity for innovative and interesting functions, interactions and content. Finally, the identification dimension indicates how a certain product allows the user to identify himself with it, by using or owning a specific product the user can reach a chosen self-presentation [22].

This framework proposes different tools for each of the distinct UX dimensions. Thus, for the usability/pragmatic dimension, the System Usability Scale (SUS) [24] will be used. This is a questionnaire with 10 items on a five-point Likert scale, which presents a final result in a range between 0 and 100. For this dimension, we will also use Pragmatic Quality (QP) component of the AttrakDiff questionnaire [25, 26]. Each of the AttrakDiff components has a value between -3 and 3 and represents an average of 7 items on a semantic differential scale of bipolar adjectives. For the aesthetics dimension, we will use the AttrakDiff Attractiveness component.

For the emotional dimension, the pictorial assessment tool Self-Assessment Manikin (SAM) [27] will be considered in its 3 indicators: pleasure, arousal, and dominance. This non-verbal tool measures the results for each of the indicators on a scale from 1 to 9. For the stimulation and identification dimensions, the corresponding AttrakDiff Hedonic Quality (HQ-S and HQ-I) components will be used.

To better understand the model that we intended to use for the operationalization of the research project, it is important to point out that the UX does not happen in a single instant, in reality, it happens in a set of distinct and conjugated moments. Thus, even before the user comes into contact with a new product or service he already has the experience of using other solutions, similar or not. When he receives information about a new system, it creates a set of expectations and a mental model of how it will work, in what Roto et al. call of anticipated UX [28]. In the first experience with the system and with every additional interaction, there are moments of what these authors call momentary UX, i.e. UX experienced, sometimes viscerally, during use. After a usage session and reflecting on it, the user achieves the called episodic UX that recovers the memory of the sensations and emotions he had during that previous use. Over time and after several periods of use (interleaved by times of non-use), a cumulative UX is reached where the various individual experiences merge to create a whole experience, which may be the most important. As

Roto et al. refer, the importance of a negative reaction during a specific use can be diminished by successive subsequent uses and the problematic use can be, in the end, remembered differently. Each of these different UX moments will be carefully studied using different tools. To address the anticipated UX, we intend to use semi-structured interviews, which will be carried out with a limited set of experts and end users. For the evaluation of the momentary UX, the ideal methodology would be direct observation, in order to capture the behaviors at the precise moment they occur. However, since the normal usage will occur in a residential environment by a large number of viewers simultaneously, this would be impractical. The alternative that this research intends to implement is the well-timed use of in-app questions, which will be presented to the viewer directly on the TV, in a way similar to the one depicted in Fig. 5. The use of this approach will also allow for the evaluation of the suggestions system since the experienced results are easily transformable in simple questions, with a very direct response from the viewer, and can be very tuned to the respective functionalities.



Fig. 5. Mockup of an in-app question being made directly on the TV

The assessment of the episodic UX will be realized carried out shortly after the installation of the prototype and will be performed in an internet application using the three instruments previously mentioned SUS, SAM and AttrakDiff, instructing the viewer that he should respond by recalling the most recent prototype usage. The evaluation of the cumulative UX, which will use the same three instruments, will be applied in the same way, three months after the system setup, along with a semi-structured interview.

The data collected by these instruments will be tabulated and compared to understand the evolution of the UX perception towards the system. A set of indicators will be collected automatically by the interactive TV platform, to serve as a control and as a potential triangulation of the information collected in the questionnaires. This will allow us to better evaluate scenarios where the user reports that he strongly agrees to “I think that I would like to use this system frequently”, the first question in SUS, and then seldom using the system.

7 Expected Results

With the insights obtained through the data analysis, we expect to achieve a rich understanding of the influence that a recommendation system, based on the viewing context, and relying on machine learning techniques, has on the UX of the viewers. Furthermore, we will also assess the relationship between the reduction of the number of content options offered to the viewer and their respective UX. As already mentioned, the excessive content can become a source of frustration for the viewer. Following this reasoning, we expect that reducing the number of options will lessen this frustration [2]. We also intend to evaluate the usage of in-app questions, directly on the interactive TV interface, in the UX evaluation context. This inquiring model has a set of advantages, since it enables a real continuity between the moment of UX and the moment of the evaluation and allows an automatic data collection. However, it interferes directly with the interaction flow which in itself will change the perceived UX [29].

The enormous growth in content available in most pay-TV platforms has the potential for creating a “Paradox of Choice” in the viewer. Taking advantage of new developments in recommendation algorithms, based on the viewing context and much more tailored to linear content, we propose to implement a recommendation system using machine learning and apply this recommendations on a preemptive interactive TV application that will act upon these recommendations and actually plays automatically the content, minimizing the viewer interactions and decisions. From the initial data analysis, we have already found that there is an effective regularity in the sampled TV usage, and that it can be leveraged into the creation of the proposed robot concierge. We need now to evaluate how such a system can affect the UX of its viewers, to get a real understanding of its potential in the linear content context.

References

1. Abreu, J., Nogueira, J., Becker, V., Cardoso, B.: Survey of Catch-up TV and other time-shift services: a comprehensive analysis and taxonomy of linear and nonlinear television. *Telecommun. Syst.* **64**, 57–74 (2017). <https://doi.org/10.1007/s11235-016-0157-3>
2. Schwartz, B.: *The Paradox of Choice: Why More is Less*. HarperCollins, New York City (2004)
3. Vanattenhoven, J., Geerts, D.: Contextual aspects of typical viewing situations: a new perspective for recommending television and video content. *Pers. Ubiquit. Comput.* **19**, 761–779 (2015). <https://doi.org/10.1007/s00779-015-0861-0>
4. Bernhaupt, R., Obrist, M., Weiss, A., et al.: Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Comput. Entertain.* **6**, 5:1–5:23 (2008). <https://doi.org/10.1145/1350843.1350848>
5. Digitalsmiths: Q4 2015 Video Trends Report - Consumer Behavior Across Pay-TV, VOD, PPV, OTT, TVE, Connected Devices, and Content Discovery (2015)
6. Turrin, R., Condorelli, A., Cremonesi, P., Pagano, R.: Time-based TV programs prediction. In: 1st Workshop on Recommender Systems for Television and Online Video at ACM RecSys (2014)
7. Churchill, E.F.: Putting the person back into personalization. *Interactions* **20**, 12–15 (2013). <https://doi.org/10.1145/2504847>

8. Aharon, M., Hillel, E., Kagian, A., Lempel, R., Makabee, H., Nissim, R.: Watch-it-next: a contextual TV recommendation system. In: Bifet, A., May, M., Zadrozny, B., Gavalda, R., Pedreschi, D., Bonchi, F., Cardoso, J., Spiliopoulou, M. (eds.) ECML PKDD 2015. LNCS (LNAI), vol. 9286, pp. 180–195. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23461-8_12
9. Gonçalves, D., Costa, M., Couto, F.M.: A flexible recommendation system for cable TV. In: 3rd Workshop on Recommendation Systems for Television and online Video, RecSysTV 2016 (2016)
10. Portugal, I., Alencar, P., Cowan, D.: The use of machine learning algorithms in recommender systems: a systematic review. *Expert Syst. Appl.* **97**, 205–227 (2018). <https://doi.org/10.1016/j.eswa.2017.12.020>
11. Adomavicius, G., Tuzhilin, A.: Context-aware recommender systems. In: Ricci, F., Rokach, L., Shapira, B. (eds.) *Recommender Systems Handbook*, pp. 191–226. Springer, Boston (2015). https://doi.org/10.1007/978-1-4899-7637-6_6
12. Zhang, T., Iyengar, V.S.: Recommender systems using linear classifiers. *J. Mach. Learn. Res.* **2**, 313–334 (2002)
13. Quinlan, J.R.: Induction of decision trees. *Mach. Learn.* **1**, 81–106 (1986). <https://doi.org/10.1023/A:1022643204877>
14. Ericsson AB Mediaroom—Ericsson Media Solutions. <https://mediasolutions.ericsson.com/products/mediaroom/>. Accessed 8 Jan 2018
15. Zibriczky, D., Petres, Z., Waszlavik, M., Tikk, D.: EPG content recommendation in large scale: a case study on interactive TV platform. In: 2013 12th International Conference on Machine Learning and Applications, pp. 315–320 (2013)
16. Buczak, A.L., Zimmerman, J., Kurapati, K.: Personalization: improving ease-of-use, trust and accuracy of a TV show recommender. In: *Proceedings of AH 2002 Workshop on Personalization in Future TV* (2002)
17. Chang, N., Irvan, M., Terano, T.: A TV program recommender framework. *Proc. Comput. Sci.* **22**, 561–570 (2013). <https://doi.org/10.1016/j.procs.2013.09.136>
18. Swearingen, K., Sinha, R.: Interaction design for recommender systems. *Des. Interact. Syst.* **6**, 312–334 (2002)
19. International Organization for Standardization: ISO 9241-210: ergonomics of human–system interaction - human-centred design for interactive systems. *Int. Organ. Stand.* **2010**, 32 (2010). <https://doi.org/10.1039/c0dt90114h>
20. Jenner, M.: Binge-watching: video-on-demand, quality TV and mainstreaming fandom. *Int. J. Cult. Stud.* **20**, 304–320 (2017). <https://doi.org/10.1177/1367877915606485>
21. Abreu, J., Almeida, P., Silva, T.: A UX evaluation approach for second-screen applications. *Commun. Comput. Inf. Sci.* **605**, 105–120 (2016). https://doi.org/10.1007/978-3-319-38907-3_9
22. Bernhaupt, R., Pirker, M.: Evaluating user experience for interactive television: towards the development of a domain-specific user experience questionnaire. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8118, pp. 642–659. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40480-1_45
23. Drouet, D., Bernhaupt, R.: User experience evaluation methods: lessons learned from an interactive TV case-study. In: Bogdan, C., et al. (eds.) HCSE/HESDD -2016. LNCS, vol. 9856, pp. 351–358. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44902-9_22
24. Brooke, J.: SUS-a quick and dirty usability scale. *Usability Eval. Ind.* **189**, 4–7 (1996)

25. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Szwillus, G., Ziegler, J. (eds.) 2003 Interaktion Bewegung Mensch & Computer, pp. 187–196. Vieweg + Teubner Verlag, Wiesbaden (2003)
26. Hassenzahl, M.: The interplay of beauty, goodness, and usability in interactive products. *Hum.-Comput. Interact.* **19**, 319–349 (2008). https://doi.org/10.1207/s15327051hci1904_2
27. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**, 49–59 (1994). [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
28. Roto, V., Law, E., Vermeeren, A., Hoonhout, J.: User experience white paper. Bringing clarity to concept user experience, pp. 1–12 (2011)
29. Ferraz de Abreu, J., Almeida, P., Beça, P.: InApp questions – an approach for contextual evaluation of applications. In: Abásolo, M.J., Almeida, P., Pina Amargós, J. (eds.) *jAUTI 2016*. CCIS, vol. 689, pp. 163–175. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-63321-3_12