Aalto University
School of Science
Bachelor's Programme in Science and Technology

Utilisation of viewing statistics in video recording credits detection

Bachelor's Thesis

April 17, 2022

Saskia Kivistö

Aalto University	ABSTRACT OF
School of Science	BACHELOR'S THESIS
Bachelor's Programme in Science and Technology	

Author:	Saskia Kivistö
Title of thesis:	Utilisation of viewing statistics in video recording credits detection
Date:	April 17, 2022
Pages:	
Major: **Code:**	Computer Science SCI3027
Supervisor: **Instructor:**	Assistant Professor Parinya Chalermsook Wanchote Jiamjitrak, M.Sc. (Tech) (Department of Computer Science)
Abstract	
Keywords:	change point detection, segmentation, statistical signal processing, NVPR
Language:	English

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknistieteellinen kandidaattiohjelma

KANDIDAATINTYÖN TIIVISTELMÄ

Tekijä:	Saskia Kivistö
Työn nimi:	Videotallenteen alku- ja lopputekstien paikannus katsomistilastojen perusteella
Päiväys:	17. huhtikuuta 2022
Sivumäärä:	
Pääaine:	Tietotekniikka
Koodi:	SCI3027
Vastuuopettaja:	Apulaisprofessori Parinya Chalermsook
Työn ohjaaja:	DI Wanchote Jiamjitrak (Tietotekniikan laitos)

Tässä kandidaatintyössä tutkitaan onko katsomiskäyttäytymistilastojen avulla mahdollista erottaa ohjelmasisältö ylimääräisestä sisällöstä televisiolähetyksissä, jotka on nauhoitettu verkkopohjaisella digitaalisella mediatallentimella. Nauhoitteiden alkuun ja loppuun jää ylimääräistä sisältöä, koska nauhoituksen tallennusaika on muutaman minuutin pidempi kuin ohjelman ohjelmaoppaan mukainen kesto. Tällä varmistetaan että ohjelma tallentuu kokonaisuudessaan, vaikka sen lähetysaika poikkeaisikin hieman ohjelmaoppaasta. Kannuste ylimääräisen sisällön tunnistamiselle on, että se vie tallennustilaa ja sen ohi kelaaminen on puuduttavaa.

Tutkimuskysymystä lähestytään signaalinkäsittelyn kautta segmentointiongelmana. Mahdollisten segmenttien homogeenisyyttä mitataan neliöllisellä tappiofunktiolla. Tappiofunktion minimoivan segmenttiyhdistelmän laskemiseen sovelletaan kahta algoritmia, joista ensimmäinen löytää optimaalisen ratkaisun kun segmenttien määrä on määritelty ennalta ja toinen kun segmenttien määrä ei ole tiedossa.

Alku- ja lopputekstien sijainti löytyy oikein suurimmassa osassa 279 tallenteen otoksesta. Lopun muutoskohta paikantuu tyypillisesti muutaman sekunnin päähän lopputekstien alkukohdasta, mutta jos lopputekstien aikana tai päätteeksi on vielä varsinaista ohjelmasisältöä muutoskohta paikantuu lähemmäs lopputekstien loppua. Alun muutoskohdan sijainnissa alkutekstien sisällä on enemmän vaihtelua.

Alku- ja lopputekstien tarkkaa alku- ja päätöskohtaa ei pysty määrittämään pelkkien katsomistilastojen avulla. Katsomistilastot voisivat kuitenkin olla hyödyllisiä tarkan alun ja lopun paikantamisessa yhdistettynä muuhun analytiikkaan ja metatietohin ohjelmien tyypillisestä rakenteesta.

Avainsanat:	muutospisteiden havaitseminen, segmentaatio, tilastollinen signaalinkäsittely, NVPR
Kieli:	englanti

Contents

1 Introduction

Watching television used to be a very time-sensitive activity. If one wanted to see a TV program, they had to watch it when it was broadcasted. Videocassette recorder brought the freedom to store programs and to choose freely when to watch them. Even though newer technologies have replaced the videocassette recorder, one major inconvenience still remains. Programs are not broadcasted strictly according to the schedule of the Electronic program guide (EPG). Sometimes programs start earlier or end later than what is stated in the EPG. To ensure that the entire program is recorded, recording must start before the EPG start time and end after the EPG end time. This usually results in some non-program content being included in the recording. Skipping over the non-program content can be frustrating for the person watching the recording.

Knowing when the program truly starts and ends would solve the issue, but this information is not generally available. This leaves the option to detect the start and end of the program from the recording. Various artificial intelligence solutions have been devised to solve the problem, for example [2] [3] [5] [6], but as program content can be quite varied it is difficult to find a universal solution.

Network personal video recorder (NPVR) is a type of service for recording broadcast TV programs for later viewing. Instead of storing recordings on the users' local device, NPVR stores recordings on the content provider's server. The start and end times of recordings are determined by the scheduling information given in the EPG, and a fixed amount of margin is added on both ends to ensure that the entire program is recorded. This leads to non-program content being included in recordings, which the users typically skip. However, statistics of which parts of the recording users watch and which parts they skip can be collected and analysed.

The goal of this thesis is to study whether user viewing behaviour can be used to detect when the actual program content begins and ends in an NPVR recording. I am writing this thesis for an NPVR service provider company. From the perspective of an NPVR service provider, detecting the location of core program content is useful for the following reasons. Firstly, less storage space is needed if the surplus content is discarded. Secondly, it is convenient for the customers if the relevant content of a recording is pre-identified and they do not have to search for it.

This thesis is largely based on the Selective review of offline change point detection methods by Truong et al. [7]. The program content detection task is formulated as a offline signal change point detection problem, and the suitable algorithms for solving the problem are chosen with the help of the typology established in the research paper. The empirical calculations are done with the Python library `ruptures`, which is associated with the aforementioned research paper.

The thesis is structured as follows. Section 2 gives an overview on the characteristics of the viewing behavioiur data. Section 3 discusses the theoretical background of signal change point detection from the perspective of this specific use case. Section 4 discusses the Python scientific library `ruptures` and how it is used to detect change points in this thesis. The results are evaluated in section 5. Section 6 considers the viability of using user viewing behaviour for change point detection, based on the previous sections. Lastly, the main points of this thesis are summarised in section 7.

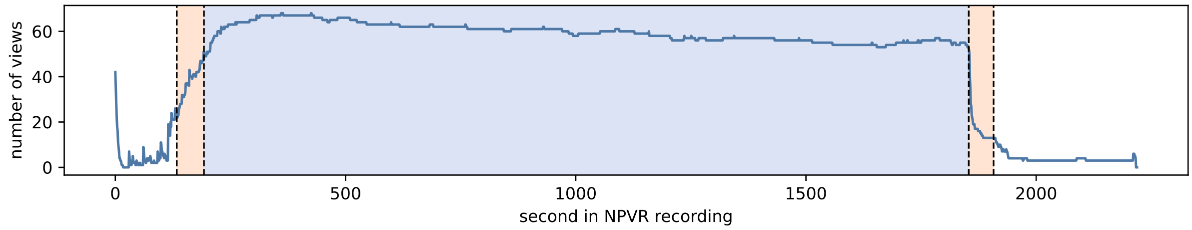
2 User viewing behaviour data

Whenever an NVPR video is watched by a user, certain metrics about the viewing event are saved. The main reason for collecting viewing metrics is the monitoring of the amount of views and the user experience quality. It can be calculated from the metrics of one view which parts of the video the user watched and which parts they skipped. Aggregating this data from multiple views of the same recording allows acquiring an overview of what parts of the recording users typically watch. This type of recording specific aggregated view count is referred to as user behaviour data in this thesis.

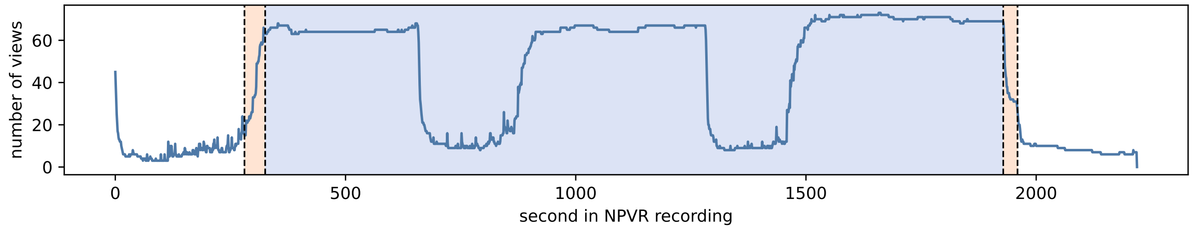
The content of an NVPR recording can be divided into program and non-program content. Non-program content is the surplus material at the very beginning and end of a recording, that is a byproduct of ensuring that the entire program is recorded even if its broadcast time deviates sligtly from the EPG schedule. Program content can be categorised into opening and closing credits, core program content and advertisement breaks, although not all programs have them.

User interest for the content categories varies. Core program content is what the users watch the recordings for, and non-program content is irrelevant for the users. This is reflected in the viewing behaviour. Users also tend to skip over advertisement breaks. Viewing behaviour regarding the opening and closing credits is not as clear cut, but generally users are content with starting watching during the opening creidts. Closing credits do not typically interest users, given there is no extra content included in them, such as a preview of the next episode.

Two examples of user viewing behaviour data are illustrated in Figure 1. The data in both figures consists of a sample of 100 user views, but the views are from two different TV programs. Figure 1a views are from a sitcom episode and Figure 1b views are from an episode of a soap opera. The episodes are divided into one second-long segments, and calculated for each segment is the number of user views in which the segment was watched. The segments are plotted on the horizontal axis, and the number of user views per segment is plotted on the vertical axis. For example, 63 users from the sample of



(a) 30 min sitcom episode without advertisements, 100 views



(b) 30 min soap opera episode with two advertisement breaks, 100 views

Figure 1: Visualisation of user viewing behaviour for two example recordings

100 users watched the part of the sitcom episode between 0:10:00 - 0:10:01 (600 on the horisontal axis in Figure 1a).

The background colour in the figure represents the content type of the time segment: non-program content is white, core program content is blue and beige indicates the opening and closing credits. The change points are marked with a vertical dashed line. There is a distict increase in the number of views during the opening credits in both figures. Likewise, there is a decrease in the number of views during closing credits. The program visualised in Figure 1b has two advertisement breaks, which appear as recesses in the number of views during the core program content. Figure 1a episode has no advertisement breaks.

2.1 Data cleaning

This section explains how the sample views are chosen. An uniform sample size is used for all recordings in order to simplify result evaluation and comparison. The characteristic pattern for viwing behaviour described in the end of the previous section is usually visible from a dozen views. Sample size of 100 views was chosen because it should be large enough to ensure that it is very unlikely that the characteristic pattern does not emerge.

As the first step, all views of a recording are sorted from oldest to newest. The views will be added to the sample in this order, but potentially incorrect or irrelevant views are filterd out. This means that the sample views are not simply the 100 earliest ones. Firstly, views which end before the recording process ends are discarded. Also views with no skips are discarded, since the change point detection relies on inspecting which parts the viewers skip, so views without skips are useless. Lastly views with implausible player source duration are removed. This includes the cases where player source duration does not exist, it is negative or close to maximum value for unsigned 32-bit integer.

3 Theoretical background

3.1 Signal change detection

Locating video content changes from user viewing behaviour time series data can be formulated as a signal processing problem, more precisely as a change point detection problem. The signal in this case is the timeseries data of user viewing behaviour, as described in section 2. The assumption in signal change point detection is that the signal consists of segments with different statistical properties, for example a different central tendency. Change points are the transition points between segments. However, signal processing is not the only possible approach for time series change point detection. Statistical approaces, such as Bayesian methods, are also widely used, but in order to narrow down the scope of this thesis they wont be examined.

Signal change point detection problems can be divided into two main categories, depending on whether the change detection must be done for incoming data in near real-time, or not. Methods solving the former case are referred to as online algorithms. Offline algorithms solve the latter case, and they differ from the online algorithms by getting the entire dataset as input and typically being more computationally complex, but also by detecting the changes more accurately. The user viewing behaviour data is aggregated from multiple views after the views have ended, meaning that all of the data to be analysed is available from the start, allowing offline methods to be used for change point detection.

Literature review by Truong et al. [7] compared different offline change point detection methods. It classified the methods according to how the homogenuity of a signal segment is measured, and how the segments where to evaluate the homogenuity are chosen from the signal. The measure of homogenuity is referred to as *cost function* and the way to choose the segments is referred to as *search method*. The objective of change point detection algorithms is to find the segmentation that minimises the sum of the cost function outputs of the segments. The function to be minimised can be expressed as follows:

$$\sum_{k=0}^{K} c(y_{(t_k . t_{k+1})}) \tag{1}$$

where $y_{(t_k . t_{k+1})}$ is a segment of the signal from index $t_k + 1$ to t_{k+1}, t_{k+n} is the index of the nth change point, c is the cost function and K is the number of change points.

Offline change point detection methods can be divided into two categories, based on whether the number of change points in the signal, K, is known beforehand, or not. If K is unknown, a constraint on the number of change points needs to be added. This is refferred to as the *complexity penalty*. Similar to cost function and search method, there are multiple alternative complexity penalties that can be used. Examining the data in section 2, based on human observation it would be reasonable to assume that a single change point can be found quite reliably for the opening and closing credits. However, advertisement breaks can also produce notable drops in the number of views which may be interpreted as change points, so if the recording might contain them the number of change points cannot be known in advance.

3.2 Cost function

What is a good cost function depends on what kind of change occurs in the the signal on a change point. The literature review by Truong et al. [7] divides cost functions into two categories, parametric and non-parametric. Parametric models make some assumptions about the statistical parameters of the data where the sample comes from, whereas non-parametric do not.

Maximum likelihood estimation is a commonly used mathematically simple parametric model for change point detection. There are multiple cost functions based on this model for different probability distributions.

Looking at the user viewing behaviour data discussed in section 2, it seems that what best describes what happens to the signal in change points is a shift in the distribution mean. A cost function that works well for mean shift detection is least squared deviation c_{L_2}, which is a maximum likelihood model. It measures the variance and is defined as follows:

$$c_{L_2}(y_{a.b}) = \sum_{t=a+1}^{b} \|y_t - \overline{y}_{a.b}\|_2^2 \qquad (2)$$

where $y_{a.b}$ is a segment from signal y from index a to b, and $\overline{y}_{a.b}$ is the empirical mean of $y_{a.b}$.

Another possibly suitable cost function for the data is the least absolute deviation c_{L_1}. It detects changes in the median and can be used also for detecting shifts in other central tendencies, such as mean and mode. It is defined as follows:

$$c_{L_1}(y_{a.b}) = \sum_{t=a+1}^{b} \|y_t - \overline{y}_{a.b}\|_1 \qquad (3)$$

3.3 Search method

Truong et al. [7] divided search methods into two categories, by whether they provide always an optimal solution for change point location with the chosen cost function, or if they provide an approximate solution. This thesis will examine only optimal methods. Optimal methods produce inherently more reliable and accurate results that approximate ones, with the tradeoff of typically higher computational complexity. The approximate methods are excluded on the basis that the goal of this thesis is to study whether it is at all possible to detect credits from solely viewing behaviour statistics, not determining the most viable approach to do it in practice. Thus high reliablity and accuracy outweights low computational complexity.

Two optimal search methods were presented in the literature review, `Opt` and `Pelt`. `Opt` was first introduced by Bellman [1] for an unrelated problem. `Pelt`, short for Pruned Exact Linear Time, was first indroduced by Killick et al. [4]. The main difference between the methods is that `Opt` requires the number of change points K to be known beforehand, and `Pelt` can be used when K is unknown.

Both of the methods use dynamic programming to find the change points. Dynamic programming is a mathematical optimisation method, where the problem is solved by recursively dividing it into smaller subproblems. In this case the problem is minimising equation 1. The observation that enables a dynamic programming approach is that:

$$\min\left[\sum_{k=0}^{K} c(y_{(t_k.t_{k+1})})\right] = \min\left\{c(y_{(0.t_1)}) + \min\left[\sum_{k=1}^{K} c(y_{(t_k.t_{k+1})})\right]\right\} \quad (4)$$

Equation 4 means that the minimisation problem for the sum of costs for all segments in an optimal segmentation, equation 1, can be reformatted into a minimisation problem for a sum of two elements, where the first element is the first segment at the beginning of the signal and the second element is the optimal set of segments for the rest of the signal. Analogously, the second element can be expressed as the minimised sum of its first segment and the optimal set of segments for the rest of the signal, and so forth.

As mentioned in section 3.1, advertisement breaks in the recording may produce recesses in the viewer count and thus appear as change points. The number of advertisement breaks per program varies, meaning that K is unknown if the recording might contain advertisement breaks. However, programs in noncommercial channels never contain advertisements, and for them it can be assumed that there should be exactly one change point at the beginning and one at the end of the program. This implies that `Opt` should be a viable search method for recordings from non-commercial channels with $K = 2$. The `Opt` algorithm follows equation 4 and proceeds as follows for $K = 2$:

1. For every possible subsample of the signal, calculate C_1, which is the value of

homogenuity within that subsection with the chosen cost fuction.

2. For every possible subsample of the signal that has at least two values, calculate C_2, which is the two C_1 that cover the entire subsample and minimise the cost function.

3. Divide the signal in two so that the cost function is minimised by the sum of C_1 of the first section and C_2 of the second section.

A more thorough pseudocode for the algorithm is available in Truong et al. literature review [7].

In principle, finding the optimal set of change points works similarly when K is unknown. It would be possible to choose a sufficiently large maximum value $K = K_{\max}$ and iterate `Opt` for all $K = 1$, ..., $K_{\max}$ and choose the result which produced the best minimisation. However, the computational complexity of this method is very high. A more viable solution is to add a complexity penalty term into equation 1 to restrict the number of predicted change points. With the complexity penalty β equation 1 becomes:

$$\sum_{k=0}^{K} \left[c(y_{(t_k \cdot t_{k+1})}) + \beta \right] \tag{5}$$

When β is linear the `Pelt` algorithm can be used. In principle, it follows the dynamic programming approach of equation 4, but there is a pruning rule which discards some indices which cannot be change points, reducing the computational complexity.

3.4 Result evaluation

Standard statistical methods for univariate analysis will be used for result evaluation. The statistics will include five-number summary first devised by Tukey et al. [8], which includes the median, the 1st and 3th quartiles and the minimum and maximum.

4 Methods

4.1 Sample data and ground truth

In order to evaluate how well a method detects change points, a ground truth is needed for comparison. Ground truth can be obtained by having a person look at a video recording and having them write down the timestamps of the change points.

I have collected the start and end times of the credits from 279 NPVR recordings by hand with a margin of error of ± 1 seconds. The sample consists of episodes from 16 different TV series. 138 of the sample recordings were broadcasted on non-commercial channels, and thus do not contain advertisement breaks. Some general information about these recordings is listed in Table 1. The 141 recordings with advertisement breaks are listed in Table 2. Every recording in the sample has at least one hundred user views.

# series	**# episodes**	**episode length**	**genre**
1.	6	30 min	reality television
2.	10	30 min	comedy
3.	32	30 min	comedy
4.	9	40 min	drama
5.	4	45 min	drama
6.	16	45 min	drama
7.	30	45 min	drama
8.	8	50 min	drama
9.	20	50 min	drama
10.	3	90 min	drama

Table 1: Sample recordings without advertisement breaks categorized by series

# series	**# episodes**	**# ad breaks**	**episode length**	**genre**
11.	25	1	30 min	soap opera
12.	35	1	30 min	soap opera
13.	31	2	30 min	soap opera
14.	26	3	50 min	drama
15.	15	1, 2 or 3	30-60 min	reality television
16.	9	4	90 min	reality television

Table 2: Sample recordings with advertisement breaks categorized by series

4.2 Change point detection with `ruptures` library

Change points are detected from the sample data with the Python library `ruptures`. The library is based on the findings of the Truong et al. literature reviw [7]. Selected algorithms examined in the literature review are implemented in `ruptures`.

Choosing the most suitable algorithm for this use-case can be done by considering the three aspects of change detection methods discussed in the literature review: cost function, search method and constraint. Accuracy is more important than low computational complexity, so optimal method is preferable to an approximate one. Two optimal search methods , `Opt` and `Pelt`, are implemented in `ruptures`, as described in section 3.3.

The `ruptures` library has three different cost functions that can be used with `Opt` and `Pelt`. These are least absolute deviation c_{L1} as described in Equation 3, least squared deviation c_{L2}, also known as variance or quadratic error loss, as described in Equation 2 and kernel cost function c_{rbf}.

Following the convention used by Truong et al. the actual change points will be denoted by t and the change points produced by `ruptures` will be denoted by $\hat{t}$ [7]. Each change point has a subscript, where the first letter denotes the credits where the change point belongs to (o: opening credits, c: closing credits). If there are separate change points for the start and end timestamps of the credits, the subscript will have a second letter that denotes which one of those the change point represents (s: start, e: end). This notation is illustrated in Figure 2.

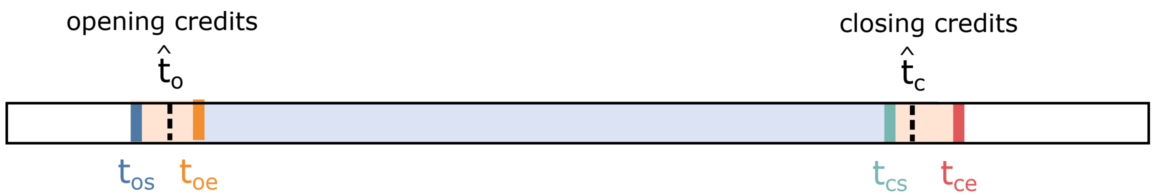


Figure 2: Notation for the actual (t) and predicted ($\hat{t}$) change points

5 Results

5.1 Non-commercial channels with `Opt` as search method

An example of `Opt` output with cost function c_{L2} and number of change points $k = 2$ is visualised in Figure 3. The viewing behaviour data is the same as in Figure 1a. The difference between the figures is that in Figure 3 the vertical dashed lines mark the change points determined by `ruptures` instead of the actual change points checked by hand. Both of the predicted change points align with the credits. The change point for the end of the program is close to the beginning of the closing credits and the change point for the start of the program is in the middle of the opening credits.

Running `Opt` c_{L2} $k = 2$ for Table 1 recordings, it turns out that the predicted change points usually land somewhere on the credits. Out of the 138 recordings, the first change point landed outside of opening credits in 23 recordings and the second change point was not on the closing credits in one recording. More statistics of the result are listed in Table 3.

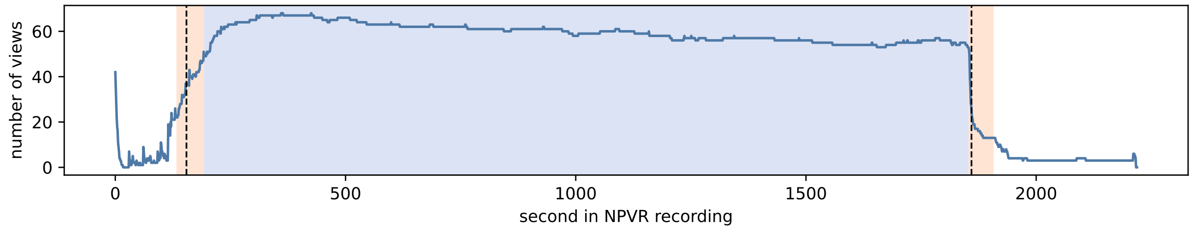


Figure 3: `Opt` c_{L2} output for Figure 1a data

statistic	$\hat{t}_o - t_{os}$	$\hat{t}_o - t_{oe}$	$\hat{t}_c - t_{cs}$	$\hat{t}_o - t_{os}$
minimum	-22	-265	3	-107
1st quartile	9	-17.75	5	-50.75
median	16	-5	7	-35.5
3rd quartile	45,75	-5	8	-28
maximum	157	152	36	5
variance	676	1380	17.2	873
standard deviation	26.0	36.2	4.15	29.5

Table 3: Five-number summary and other statistics for `Opt` c_{L2}, $k = 2$

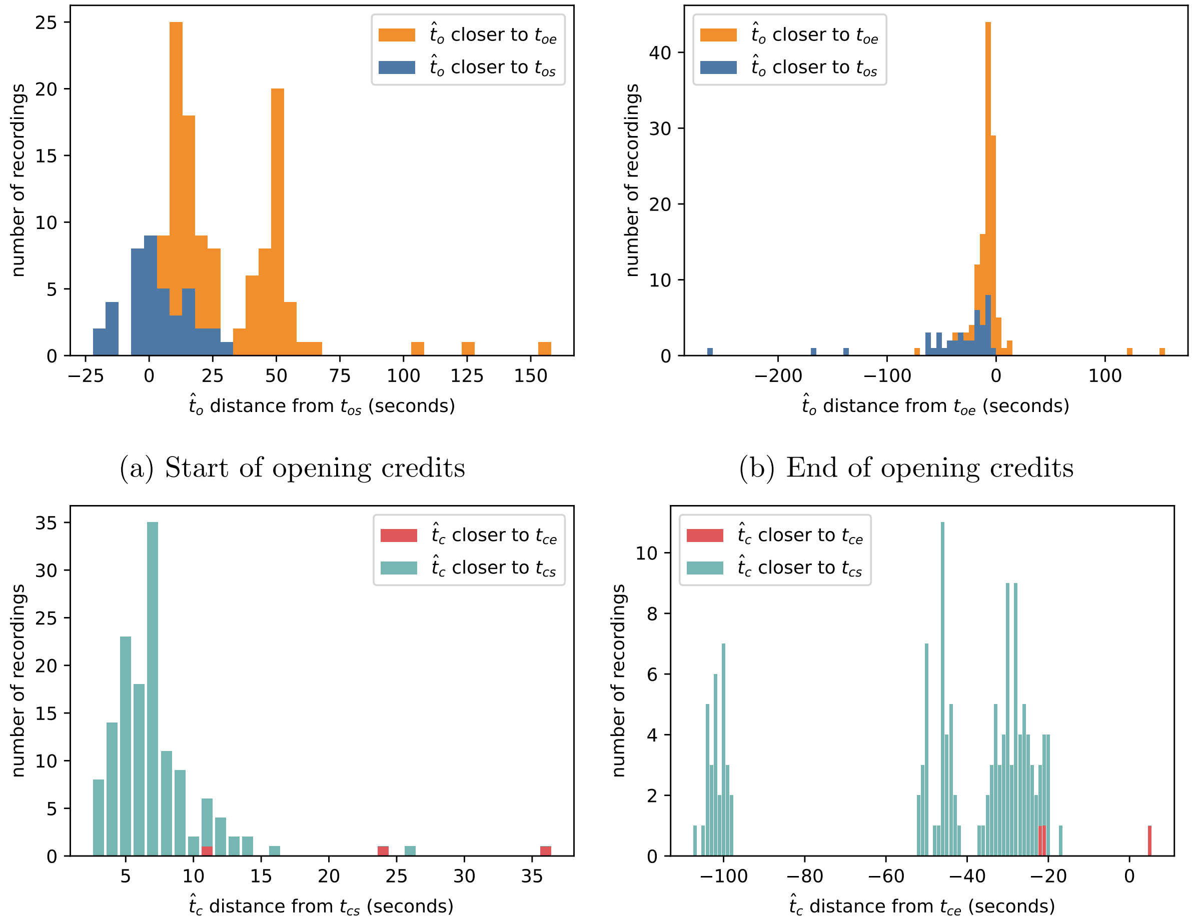
On some episodes of series 6. in Table 1, the opening credits occur few minutes after the beginning of the program. For those episodes I set the ground truth opening credits change points t_{os} and t_{oc} to the beginning of the program instead of the actual start and end time of the opening credits. The reasoning behind this was the assumption that an average viewer does not want to miss the first minute or two of the program. However, for two of these recordings `Opt` c_{L2} $k = 2$ predicted the first change point closer to the opening credits than the start of the program. This explains the large maximum value for $\hat{t}_o - t_{oe}$ in Table 3. Without these two outliers the maximum value for $\hat{t}_o - t_{oe}$ would be 11.

The five-number summary of the quartiles in Table 3 gives some insight about the output, but the results can be understood more intuitively by plotting the data. Plotted in Figure 4 are the predicted locations for opening and closing credits, $\hat{t}_o$ and $\hat{t}_c$, compared to the actual start and end times of the credits. Figure 4a has $\hat{t}_o - t_{os}$ values plotted, Figure 4b has $\hat{t}_o - t_{oc}$ values plotted, et cetera.

In order to gain more insight about the results, the output change points were divided into two populations, depending on whether the predicted change point $\hat{t}$ is closer to the start or the end of the credits. For the opening credits it seems that $\hat{t}_o$ typically falls in the middle of the credits and aligns closely to neither t_{os} or t_{oc}.

For the closing credits, it is visible in Figure 4c and 4d that most often $\hat{t}_c$ is very close to

t_{cs}, although with a delay of few seconds. Another detail worth noting is that although three fourths of the sample $\hat{t}_c$ results are within 8 s from t_{cs}, for all samples it holds that $\hat{t}_c > t_{cs}$, meaning that $\hat{t}_c$ is never predicted to be before the closing creidts.



(a) Start of opening credits

(b) End of opening credits

(c) Start of closing credits

(d) End of closing credits

Figure 4: `Opt` c_{L2} output compared to actual change points of credits

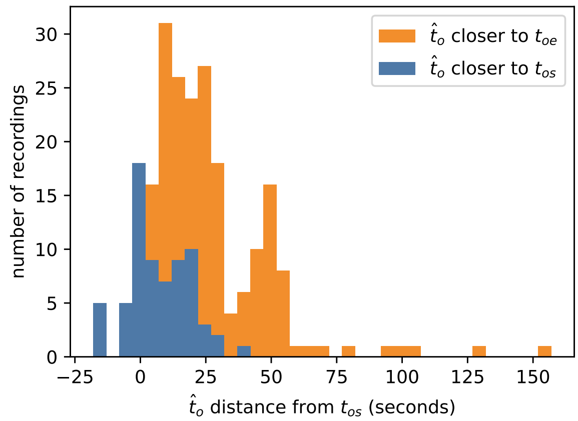
5.2 All channels with `Pelt` as search method

Both `Opt` and `Pelt` are optimal search methods, meaning that with the same cost function the output produced by the methods is identical, if the number of change points predicted by `Pelt` happens to be the same that was given as a parameter k to `Opt`. For example, if `Pelt` c_{L_2} happened to predict $k = 2$ change points for all Table 1 recordings the results would be identical to what is presented in section 5.1.

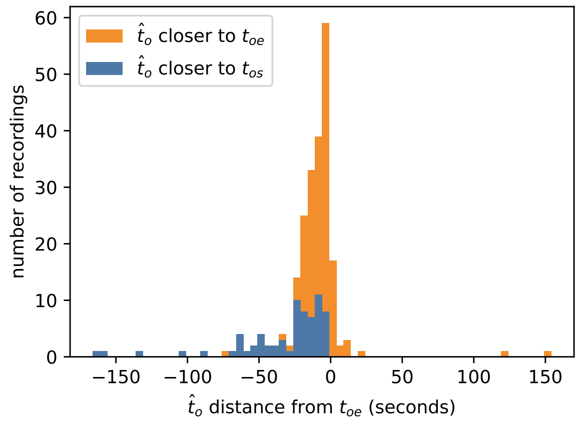
Change points were calculated with `Pelt` as the search method and c_{L_2} as the cost function for all of the 279 recordings listed in Table 1 and 2. Statistics of the results are listed in Table 4. The results are plotted in Figure 5 in a similar way as the `Opt` results in Figure 4. The results appear to be quite similar to the `Opt` results.

statistic	$\hat{t}_o - t_{os}$	$\hat{t}_o - t_{oe}$	$\hat{t}_c - t_{cs}$	$\hat{t}_o - t_{os}$
minimum	-18	-166	2	-107
1st quartile	9	-20	5	-46
median	20	-9	7	-30
3rd quartile	38	-5,25	9	-23
maximum	157	152	136	14
variance	497	726	227	896
standard deviation	22,3	27,0	15,1	29,9

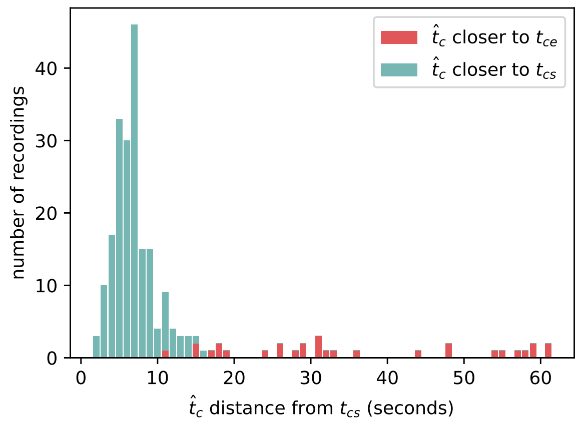
Table 4: Five-number summary and other statistics for `Pelt` c_{L2}



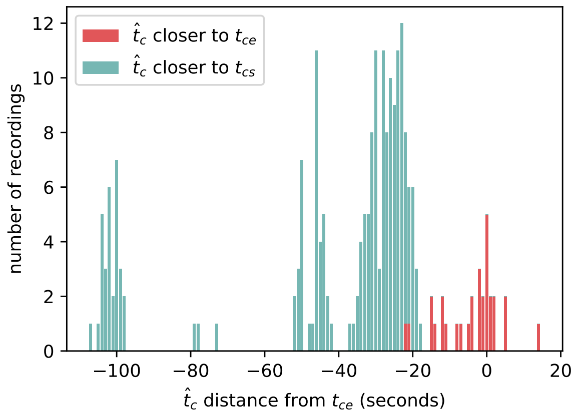
(a) Start of opening credits



(b) End of opening credits



(c) Start of closing credits



(d) End of closing credits

Figure 5: `Pelt` c_{L2} output compared to actual change points of credits

6 Discussion

There seems to be some differences in the results between different kinds of programs. The approach of using user behaviour for change segment detection most likely works best for recordings with a conventional structure with clearly separated interesting and uninteresting content. Structural things that are likely to cause problems for this method are for example opening credits that are not located at the very begining, credits embedded to core program content, recapitulations of previous episodes, sneak peeks of future episodes, especially at the end of closing credits.

Another thing that would benefit from more consideration, is choosing how to clean the data and picking the sample size.

7 Conclusions

The goal of this thesis was to study whether the opening and closing credits and advertisement breaks can be identified from an NVPR recording based solely on user viewing behaviour data. The verdict is that in most cases the general location of an change segment can be found with a reasonable accuracy. However, identifying more exact locations of change points is more difficult. It seems that locating the start of closing credits and the change points for advertisement breaks might be viable, but the end of closing credits and exact start and end for opening credits cannot likely be detected solely from user viewing behaviour.

The accuracy of the results might be improved by considering the differences between different programs and channels. User viewing behaviour is more useful for programs with a typical and predictable structure. For better results it might also be worth considering, if user viewing behaviour based change point detection could be combined with other methods for change point detection.

References

[1] Bellman, R. *On a routing problem.* Quarterly of Applied Mathematics. 1958, vol. 16 :1. P. 87–90. Available at: doi:10.1090/qam/102435.

[2] Berrani, S.-A. & Manson, G. & Lechat, P. *A non-supervised approach for repeated sequence detection in TV broadcast streams.* Signal Processing: Image Communication. 2008, vol. 23:7. P. 525–537. Available at: doi:10.1016/j.image.2008.04.018.

[3] Ibrahim, Z. A. A. & Gros, P. *TV Stream Structuring.* ISRN Signal Processing. 2011, vol. 2011. P. 1–17. Available at: doi:10.5402/2011/975145.

[4] Killick, R. & Fearnhead, P. & Eckley, I. A. *Optimal Detection of Changepoints With a Linear Computational Cost.* Journal of the American Statistical Association. 2012, vol. 107:500. P. 1590–1598. Available at: doi:10.1080/01621459.2012.737745.

[5] Kompatsiaris, Y. & Merialdo, B. & Lian, S. *TV Content Analysis: Techniques and Applications.* CRC Press, 2012. ISBN 978-1-4398-5562-1.

[6] Manson, G. & Berrani, S.-A. *Automatic TV Broadcast Structuring.* International Journal of Digital Multimedia Broadcasting. 2010, vol. 2010. P. e153160. Available at: doi:10.1155/2010/153160.

[7] Truong, C. & Oudre, L. & Vayatis, N. *Selective review of offline change point detection methods.* Signal Processing. 2020, vol. 167. P. 107299. Available at: doi:10.1016/j.sigpro.2019.107299.

[8] Tukey, J. W. *Exploratory Data Analysis*, volume 2. 1977. ISBN 0-201-07616-0.