Aalto University
School of Science

Utilisation of viewing statistics in video recording credits detection

Saskia Kivistö

Puhepaja 2

April 5, 2022

1. Introduction

1. Introduction
2. User viewing behaviour data
3. Problem formulation and methods
4. Results

1. Introduction

Goal of this thesis:

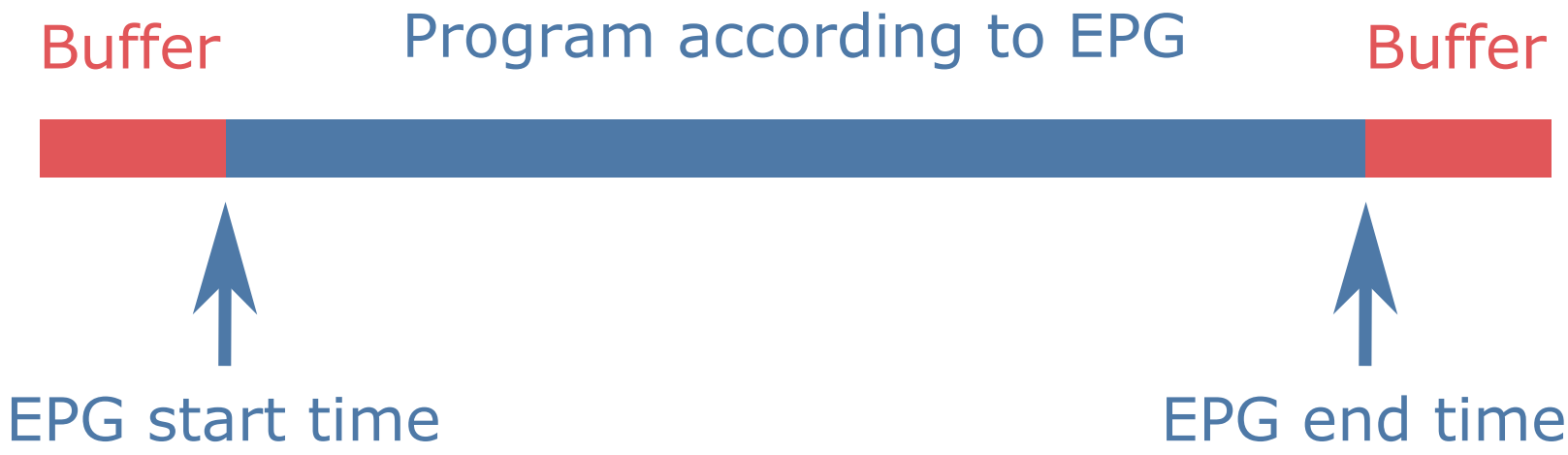
- Detect where recorded TV programs actually start and end
- by analysing patterns of user viewing behaviour

1. Introduction

The issue with recording TV programs:

- TV programs are not usually broadcasted exactly according to EPG
- Timing recording by EPG can lead to program being recorded only partially
- Buffer can be added to ensure that whole program is recorded

1. Introduction



1. Introduction

1. Introduction

1. Introduction

- Goal: identify extra content

1. Introduction

- Goal: identify extra content
- Why?
  - Its annoying to fast forward over extra content
  - Extra content consumes storage space

2. User viewing behaviour data

1. Introduction
2. User viewing behaviour data
3. Problem formulation and methods
4. Results

2. User viewing behaviour data

Where does the user behaviour data come from?

- The NPVR service provider company I am working for
- NPVR $\approx$ a normal video recorder
  - but the TV programs you record are stored in a cloud server instead of a device in your home

2. User viewing behaviour data

Why is the data collected?

- monitoring the popularity of programs
- monitoring the user experience quality
  - smoothness of streaming etc.
  - with data related to the above, it can be calculated which parts of a recording were watched during a view, and which parts were fast forwarded

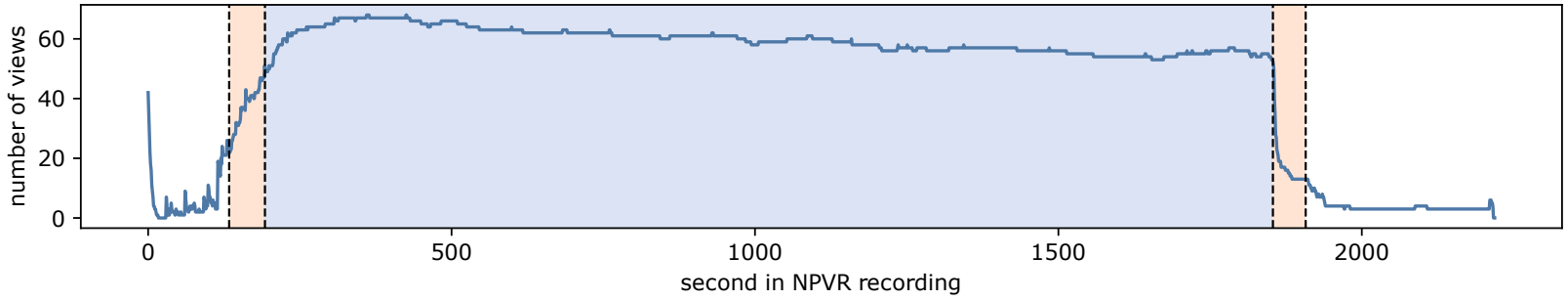
2. User viewing behaviour data

What the examined data actually is?

- For each view of any recording:
  - it can be calculated which parts of a recording were watched during a view, and which parts were fast forwarded
- The data examined in the thesis is an aggregation of the above

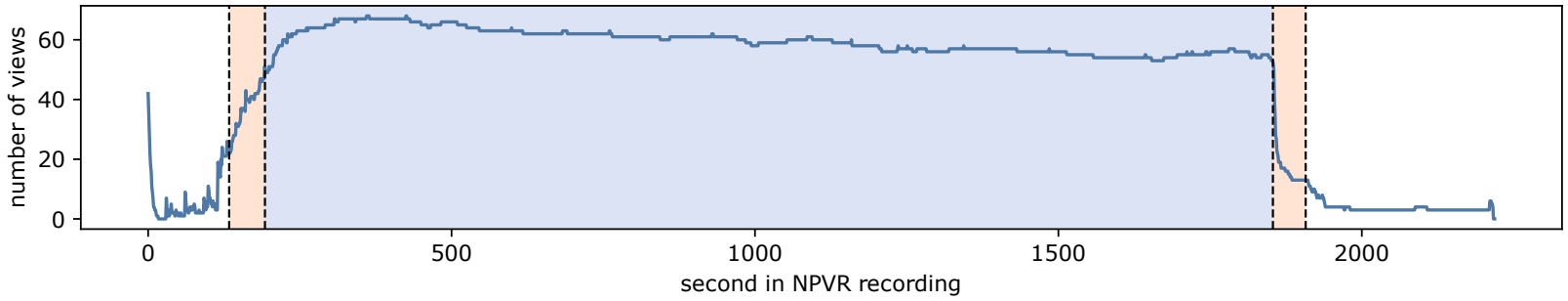
2. User viewing behaviour data

An example of the aggregated data



2. User viewing behaviour data

An example of the aggregated data



opening credits

program content

closing credits

- Could creidts be detected solely with this data?

3. Problem formulation and methods

1. Introduction
2. User viewing behaviour data
3. Problem formulation and methods
4. Results

3. Problem formulation and methods

Problem formulation

- Signal processing, offline change point detection
- Minimisation problem for the sum of costs of each segment (segments are divided by change points)
  - cost function
  - search method

3. Problem formulation and methods

Methods

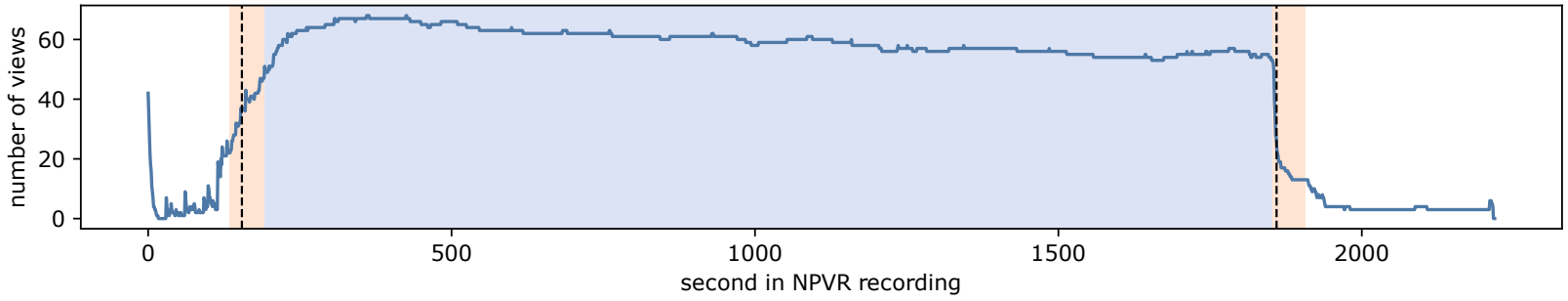
- cost function: variance, detects distribution mean shifts well
- search method; dynamic programming, produces optimal segmentation
- Python scientific library `ruptures` implements the above

4. Results

1. Introduction
2. User viewing behaviour data
3. Problem formulation and methods
4. Results

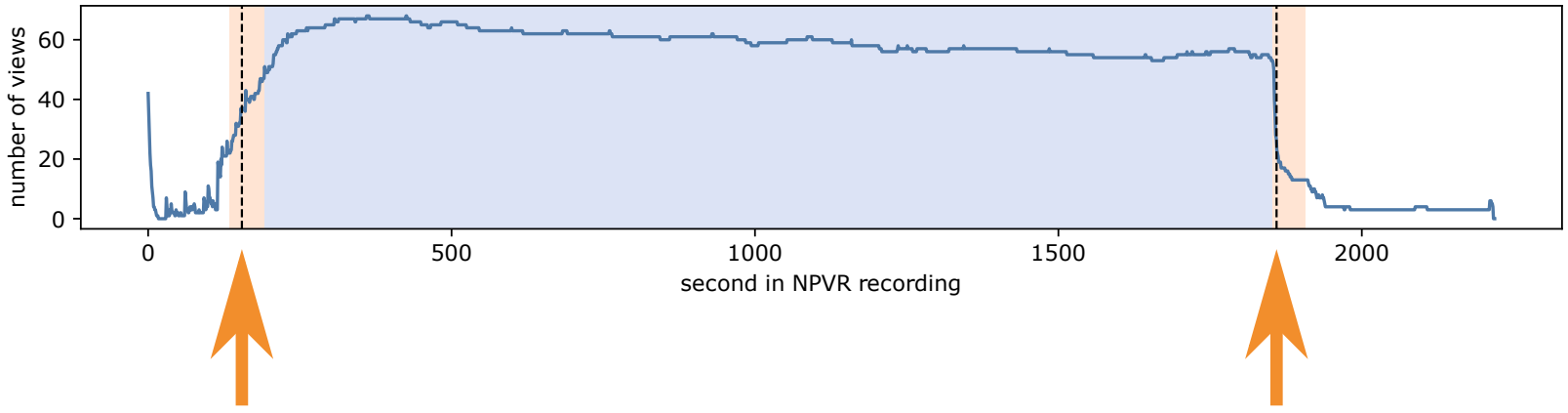
4. Results

Typical output



4. Results

Typical output

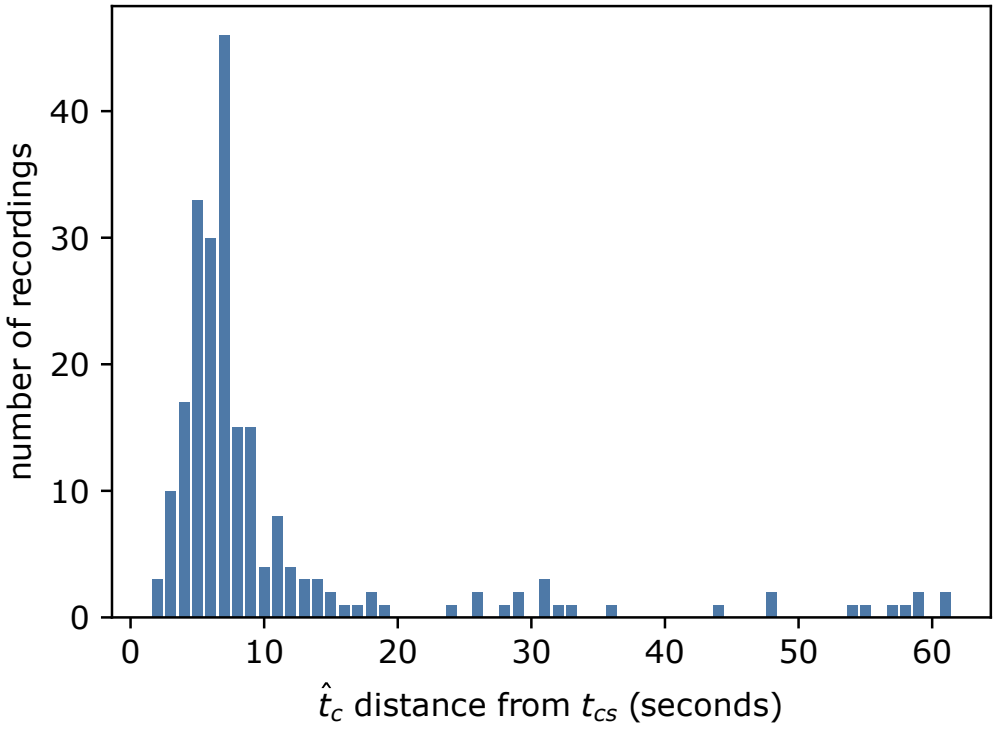


change point

change point

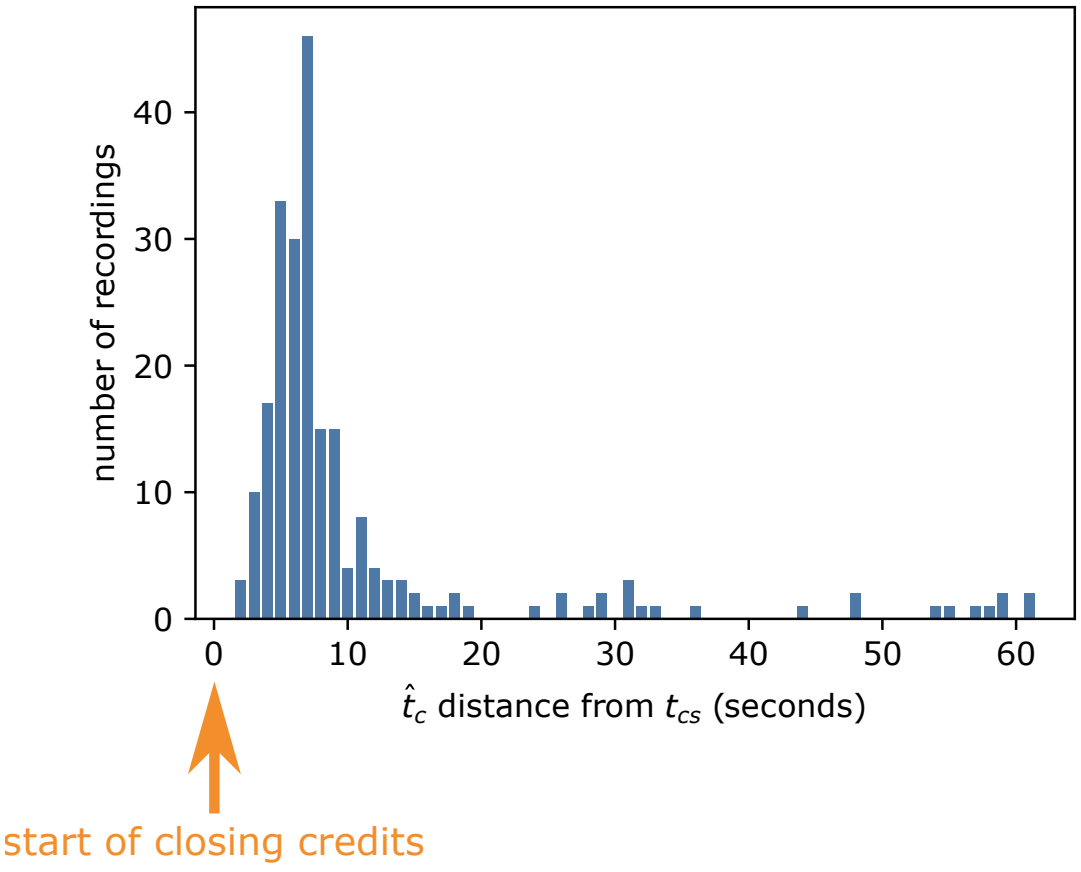
4. Results

Predicted change point distance from closing credits start



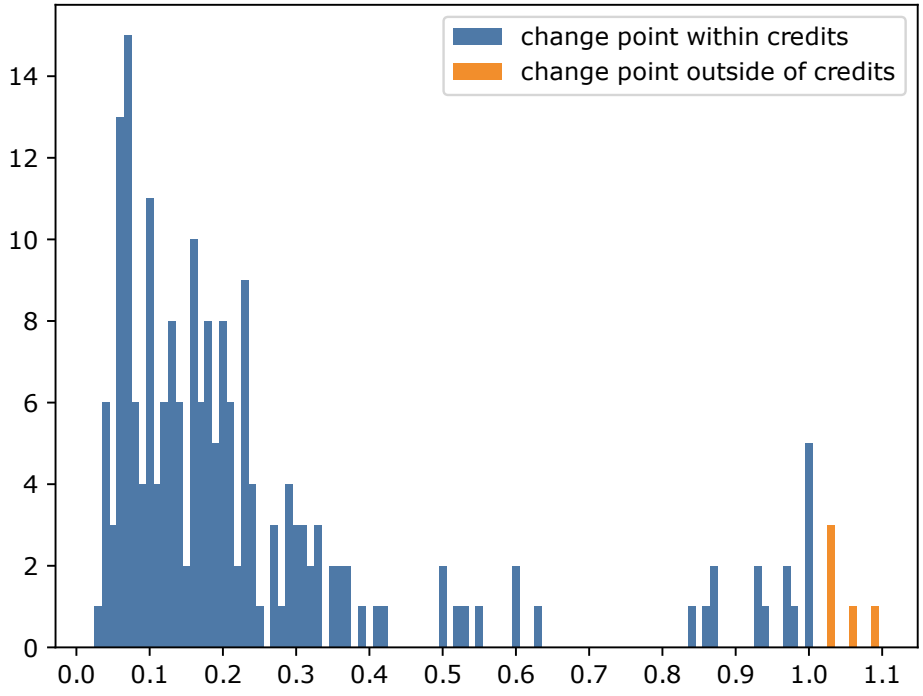
4. Results

Predicted change point distance from closing credits start



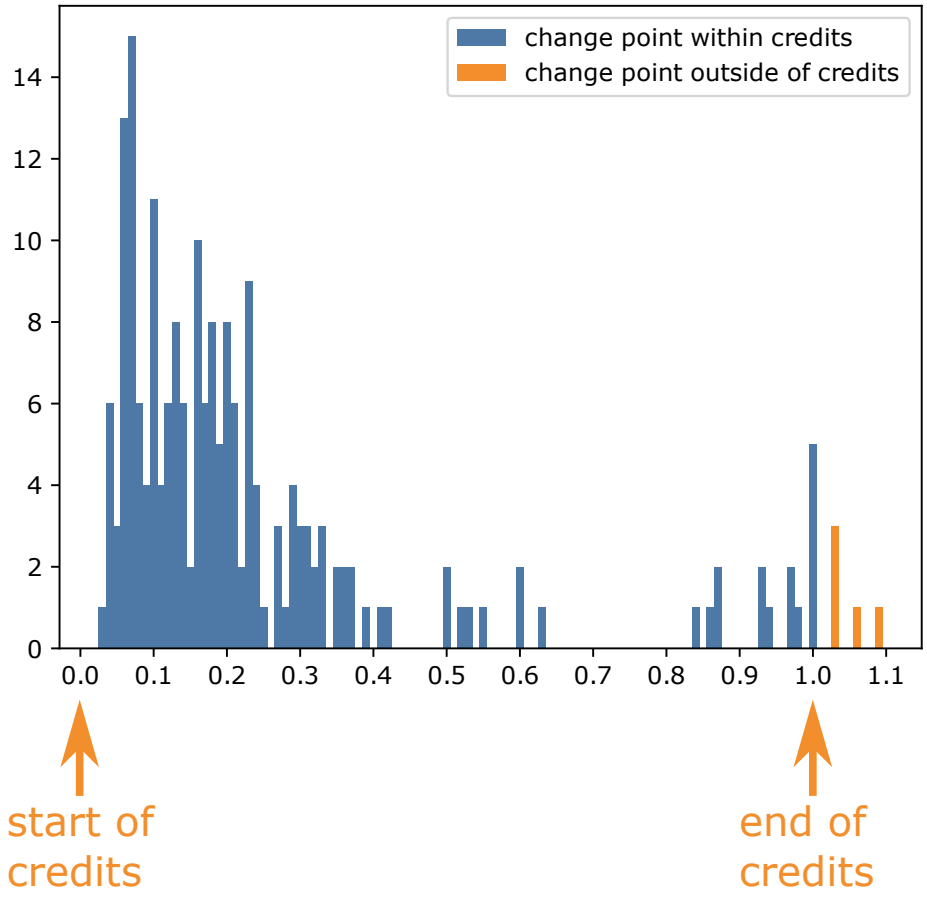
4. Results

Change point position relative to closing credits



4. Results

Change point position relative to closing credits



4. Results

Can credits be detected based solely on user viewing behaviour?

- Yes, for approximate location of credits
- No, for exact start or end of credits

Thank you!