

SCI3027.kand research plan:

Utilisation of viewing statistics in detecting NPVR closing credits

Saskia Kivistö
Aalto University
`saskia.kivisto@aalto.fi`

January 26, 2022

Title: Utilisation of viewing statistics in detecting NPVR closing credits

Author: Saskia Kivistö

Advisor: Wanchote Jiamjitrak

1 Abstract

Network personal video recorder (NPVR) is a service for recording broadcast TV programs for later viewing. Instead of storing recordings on the users local device, NPVR stores recordings on the content providers server. For every program listed in the Electronic program guide (EPG), a single recording is created and stored on the server. The users who record the same program will receive the same video from the server.

The recording start and end times are determined by the scheduling information given in the EPG. However, it is common that programs are not broadcasted exactly according to the EPG schedule. Thus NPVR recordings typically have some non-program content at the beginning and end. The goal of this thesis is to study whether viewing statistics can be used in determining where the actual program content ends and where the closing credits begin in a NPVR recording.

2 Goals and perspectives

3 Research material

The literature review part can be divided into three kind of subjects:

1. Previous research on methods for video content detection and labeling (closing credits, ads, etc.), to give more context on the thesis topic
2. Research on approaches to find the location of a specific pattern from data (in this case the drop in viewer count during closing credits). Needed for choosing a good model for deciding where the end credits begin based on the viewer count.
3. Research on the best practices of model validation. Needed for choosing a reasonable method for validating the results given by the labeling model.

4 Research methods

5 Challenges

The validation of the experimental work relies on collecting the end times of program recordings manually. This is somewhat time consuming. Part of the literature review could maybe be finding papers that discuss what is a sufficient amount of data for validation.

Viewing statistics are stored for 100 days. This shouldn't be a problem if I follow the bachelor's thesis course schedule, but if writing the thesis takes longer this might become a problem. In order to have access to all viewings of a recording during the course of the entire writing process, I'm using only recordings that have been broadcasted in 2022.

6 Resources

The thesis topic is from a company. I have a full time contract for this spring, and I have agreed with my line manager that I can focus mainly on this thesis during the spring. The thesis will include an applied part, and the company will provide me access to the data needed for it.

7 Schedule

Course schedule:

8.2.	1st draft ready, 2-3 pages
1.3.	2nd draft ready, ~10 pages
22.3.	3rd draft ready
11.4.-14.4.	presentation at the final seminar
17.4.	submission of the final version

Manually collecting program end times for validation:

From now on until the end of March. A menial, but time consuming task.

Finding, reading and taking notes on relevant research:

From now on until the middle of February.

Using the methods discussed in previous research for this use case:

From the middle of February to the end of March. I'm aiming to have some results ready for the 2nd draft and to have almost complete results for the 3rd draft.

Finishing the thesis:

From 3rd draft until the submission of the final version.

8 Structure of the thesis

1. Introduction

Terminology, the goal and the scope of the thesis

2. Background (literature review)

- 2.1. Previous research on methods for video content detection and labeling
(closing credits, ads, etc.)

- 2.2. Research on approaches to find the location of a specific pattern from data
(in this case the drop in viewer count during closing credits)

- 2.3. Research on the best practices of model validation

3. Research methods and data (own research)

Trying out one or several of the approaches discussed in section 2.2,
validating the results according to section 2.3

4. Results

5. Conclusions