De-anonymizing Encrypted Video Streams

Master's Thesis

7 September 2018

Stefano Peverelli
`pstefano@student.ethz.ch`

supervised by
Melissa Licciardello

Systems Group
Department of Computer Science
ETH Zürich

Abstract

In the last recent years streaming services such as Netflix, Youtube, Amazon Prime Video, Hulu and others, have become the main source for video content delivery to the public. With the effort of private companies and of the AOM consortium [1], various coding formats and streaming techinques have been refined and have gained popularity. *Adaptive Bitrate Streaming*, between others, enables high quality streaming of media content over HTTP, and represents nowadays the industry's standard.

DASH *Dynamic Adaptive Streaming over HTTP* is an instance of Adaptive Bitrate Streaming originally developed by MPEG. In DASH each media file gets encoded at multiple bitrates, which are then partitioned into smaller segments and delivered to the user over HTTP. Netflix's use of DASH services is no mistery, indeed it is already five years that each title on Netflix sits with its own different bitrate copies on a CDN, waiting to be served to clients in a particular area of the planet. [2]

Reed et Al. [3] have shown how, despite a recent upgrade in Netflix infrastructure to provide HTTPS encryption to video traffic, it is possible to recover unique fingerprints for each title, due to the adoption throughout the entire Netflix library, of *per-title encoding*. They make use of adudump [4] a command-line program that can run on passive TAP device [5] or on a live network interface, and uses TCP and ACKS sequences to infer the sizes of application data unit *ADUs* transferred over each TCP connection.

Our approach reiterates parts of Reed et Al.'s work, but cannot rely on their every assumption and discovery, due to constant changes that Netflix is brining to their enconding and streaming algorithms. Moreover our intent is focused on finding out if, by analyzing coarse-grained traffic data, we are able to identify a video based on its *bitrate-ladder*.

Contents

1 Introduction

According to the latest Cisco's VNI [6], video will account for 82% of all IP traffic in Europe by 2021; in addition, the overall IP traffic per person will triplicate from 13*GB* to 35*GB*. These forecasts clearly picture the growth of the streaming industry, posing, at the same time an important question on the present and future states of the final user's privacy.

As shown by Reed et Al. [3] anonimity of user's viewing activity is at risk. Not for the use that Netflix or other streaming services do of user's session data, but because of the risk of a man-in-the-middle attack *MITM* carried by an *evil* party that has control over the flow of packets over a network.

In particular, they have shown how the adoption of HTTPS to protect video streams from Netflix *CDN*s to user's end devices, does not hold against passive traffic analysis.

1.1 Motivation

The goal of this project is to replicate part of the work conducted by Reed et Al. and to investigate the possibility of identifying a Netflix stream solely based on the observed average bandwidth. This, follows from the intuition that *per-title encoding* embeds the nature and the complexity of video frames in a unique way, that may reveal the identity of the content being streamed.

1.1.1 *Per-Title Encoding*

In December 2015 Netflix announced [2] that it was introducing a new method to analyze the complexity of each title and find the best encoding recipe based on it. Their goal with the adoption of per-title encoding was to provide users with better quality streams at a lower bandwidth.

Before then, each title was encoded with a *Fixed Bitrate Ladder*; their pipeline returned a list of {*Bitrate, Resolution*} pairs that represented the sufficient bitrate to encode the stream at a certain resolution (Table 1.1), with no visible artificats.

Bitrate (kbps)	Resolution
235	320×240
375	384×288
560	512×384
750	512×384
1050	640×480
1750	720×480
2350	1280×720
3000	1280×720
4300	1920×1080
5800	1920×1080

Table 1.1: Netflix original's Fixed Bitrate Ladder.

This "one-size-fits-all" ladder, as reported, achieved good results in the encoded video's perceived quality (**PSNR** [7]) given the bitrate constraint, but, would not perform optimally under certain conditions. For instance, high detailed scenes with sudden changes of light, or rapid transitions of camera shots, would require more than 5800*kbps;* in contrast, more static frames, as in animated cartoons, may be encoded at higher resolutions mantaining the same bitrate level.

In summary they noticed how in certain cases, the produced encoding would either present some small artifacts (*e.g.* complex scenes), or waste bandwidth, (*e.g.* static, plain scenes). For this reason, they came up with per-title encoding.

In order to find the best fitting bitrate ladder for a particular title, there are several criterias that they took into account, the principal ones being:

- How many quality levels should be encoded to obtain a *JND* between each of them.
- Best {*Resolution, Bitrate*} pair for each quality level
- Highest bitrate required to achieve the best perceivable quality

As aforementioned, each title's perceived video quality, gets computed as a measure of *Peak signal-to-noise ratio*. The comparison is performed between the produced encode, upsampled to 1080*p*, and the original title in 1080*p*, and the best {*Bitrate, Resolution*} pair is assigned to that specific quality level, as depicted in Table 1.2.

In Figure 1.1, we can see the impact of per-title encoding on the original bitrate ladder: in order to achieve the same perceivable quality level (point **B** and **C**), it requires a lower bitrate to be encoded to (point **A**). Moreover, with around the same bitrate, one can see

Resolutions	Fixed Bitrate Ladder (kpbs)	Per-Title Bitrate Ladder (kbps)
320 × 240	235	150
384 × 288	375	200
512 × 384	560	290
512 × 384	750	
640 × 480	1050	
720 × 480	1750	440
720 × 480		590
1280 × 720	2350	830
1920 × 1080	3000	1150
1920 × 1080	4300	1470
1920 × 1080	5800	2150
1920 × 1080		3840

Table 1.2: Comparison between the two different approaches for the same title: note how different titles may have different numbers of quality levels. For each movie, the minimum number of quality levels gets computed to produce a just-noticeable-difference (JND), when switching bitrates during playback.

how per-title encoding can achieve a higher resolution compared the fixed case (point **A** and **D** respectively). It follows obviously that, holding to a high-quality stream while maintainig or lowering the used bandwidth is key: the end user will get same or better quality then before, at a lower bandwidth.

1.1.2 User's Privacy

As of 2018, data began the most valuable commodity on the planet [8], and with its value rising, society starts to question how to legislate to protect both industry's and user's rights.

In 2016, Netflix announced the introduction of HTTPS to protect the content being streamed to users. With the addition of TLS on top of HTTP, Netflix aim, was to avoid the risk of eavesdropping on unsecure connections, protecting themselves and users from third-party applications and governments potentially collecting viewer's data and streaming habits.

Avoiding deep packet inspection from potential eavesdropper certainly adds another layer of security, but given the narrow relationship between Netflix and ISPs, one might conclude that the chance that IP packets do not get inspected by ISPs (especially in the United States, after the reclassification of ISPs as Title I *Information Services* [9]), is still a matter of mere trust given to the platform-provider relationship.

Cases of user's data breaches as the Kanopy one [10], are a testimony of how easy would be for an attacker to extract information to the point at which users could become

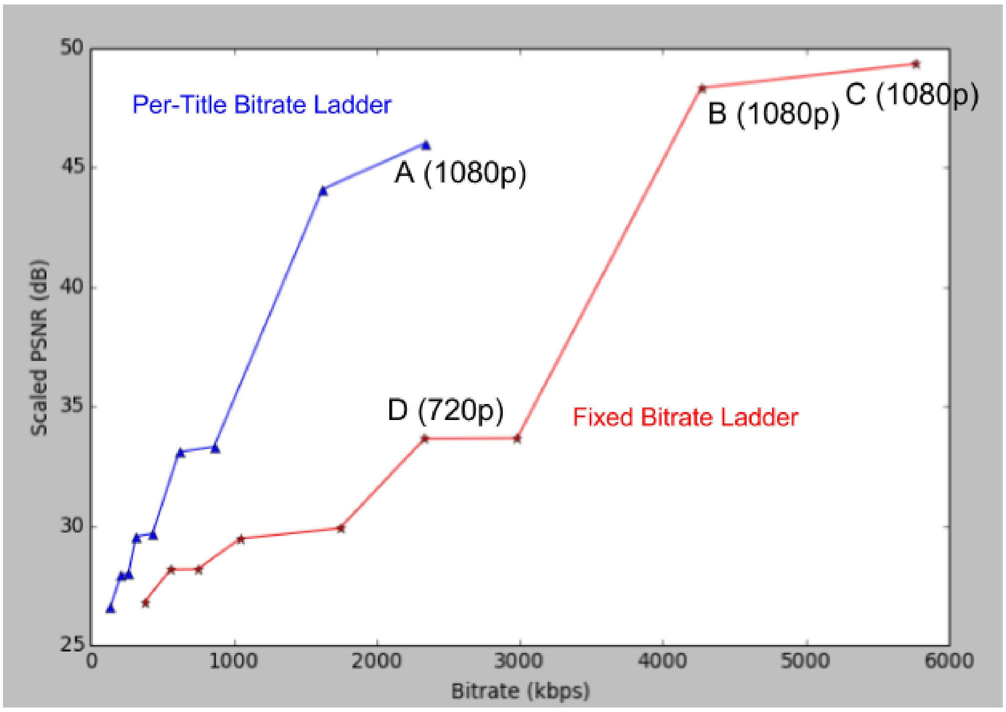

Figure 1.1: Difference between per-title vs. fixed bitrate ladders.

identifiable by the solely information the platform was collecting in their internal log files. The content of which, include between others: timestamps, geo-location data, client-device informations and IP addresses.

The ability to cluster people based on just their video-streaming habits, poses a potential threat for how the data could be processed and used by parties with access to it. One could imagine how government agencies could easily get in possess of sensitive information about the nature of the content a particular user is interested about, or how ISPs could profit from selling data profiles to advertisement companies that in turn would exploit their information to improve per-user recommendation algorithms.

Considering this trend, it is for us crucial to investigate how parties that can have access to transport-layer information, (more details in **??**), could exploit per-title encoding to identify video traffic.

1.2 Related Work

As previously mentioned, our work is mainly insipired by Reed et Al.'s [3] research paper, in which they presented a novel method to de-anonymize encrypted netflix stream in real-time with limited hardware requirements. Their system was able to identify a video using uniquely TCP/IP headers, by making use of **adudump**, a program built on top of **libpcap** [11], a powerful C/C++ library for network traffic capture. They acquired for each video, metadata information with a tampering tool, and then matched adudump traces against with. The evaluation of their method revelaed that they could identify majority of videos by recording only 20 minutes of traffic each.

An earlier proof on how bandwidth analysis could reveal the content of ecnrypted traffic, was given by Saponas et Al. [12], in their "SlingBox-Pro" case study, that exposed how, by recording network traffic and producing and combining different trace levels, they were able to identify 98% of 40 minutes video traces.

Similar work has been conducted also by Moser [13], who analyzed how bitrate ladders could uniquely embed the identity of Netflix titles. In his work he further studies the impact of the aggregation of each video's segment bandwidth on the overall accuracy of his system.

1.3 Main Objective

The main goal of this project is to build a system that can manipulate network bandwidth and observe video traffic from Netflix, to verify the intuition that each title could be identified by its own bitrate ladder.

Furthermore we investigate and discuss possible countermeasures that streaming providers could adopt to preserve users privacy. A detailed explanation of our approach is presented in **??**.

1.4 Structure of this Report

In this section, we outline of the contents and the structure of this report.

Chapter 1 introduced our motivation behind the project, presented featured work that influenced and inspire our approach, and it outlined the goals we would like to ultimately achieve.

Chapter 2 presents the attack scenario from an ISP perspective, (the ISP acting as the adversary), the infrastructure needed to acquire the data, the nature of the information that could be infered, and the consequences that might arise.

Chapter 3 shows our version of the attack, in a different context, but up to some extent, with similar conditions to the one depicted in Chapter 2. It also higlights the similiarities and differences from the featured approaches we followed.

In Chapter 4 we evaluate our system, we present results, and discuss the relevance of such a method.

Chapter 5 tries to summarize and draw conclusion based on the claims made at the beginning of the project.

Bibliography

[1] Wikipedia. Alliance for open media, 2019. URL https://en.wikipedia.org/wiki/Alliance_for_Open_Media.

[2] Netflix TechBlog. Per-title encode optimization, 2015. URL https://medium.com/netflix-techblog/per-title-encode-optimization-7e99442b62a2.

[3] Andrew Reed and Michael Kranch. Identifying https-protected netflix videos in real-time. In *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy*, CODASPY '17, pages 361–368, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4523-1. doi: 10.1145/3029806.3029821. URL http://doi.acm.org/10.1145/3029806.3029821.

[4] Jeff Terrell, Kevin Jeffay, F. Donelson Smith, Jim Gogan, and Joni Keller. Passive, streaming inference of tcp connection structure for network server management. IEEE International Traffic Monitoring and Analysis Workshop, 2009.

[5] Wikipedia. Network tap, 2019. URL https://en.wikipedia.org/wiki/Network_tap.

[6] Business Insider Intelligence. Video will account for an overwhelming majority of internet traffic by 2021, 2019. URL https://www.businessinsider.com/heres-how-much-ip-traffic-will-be-video-by-2021-2017-6?r=US&IR=T.

[7] Wikipedia. Peak signal-to-noise ratio, 2019. URL https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio.

[8] The Economist. The world's most valuable resource is no longer oil, but data, 2018. URL https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

[9] Wikipedia. Net neutrality in the united states, 2019. URL https://en.wikipedia.org/wiki/Net_neutrality_in_the_United_States.

[10] @xxdesmus. Kanopy.com leaking api and website access logs, 2019. URL https://rainbowtabl.es/2019/03/21/kanopy-data-leak/.

[11] tcpdump. URL https://www.tcpdump.org/.

[12] Scott Saponas, Jonathan Lester, Carl Hartung, Sameer Agarwal, and Tadayoshi Kohno. Devices that tell on you: Privacy trends in consumer ubiquitous computing. In *Proceedings of the 16th USENIX Security Symposium*. USENIX, August 2007. URL https://www.microsoft.com/en-us/research/publication/devices-tell-privacy-trends-consumer-ubiquitous-computing/.

[13] Florian Moser. Identifying encrypted online video streams using bitrate profiles, 2018. URL https://github.com/famoser/network-experiments.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

De-anonymizing encrypted video streams

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Stefano	Peverelli

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zurich, 07.09.2019	

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.