White Mirror: Leaking Sensitive Information from Interactive Netflix Movies using Encrypted Traffic Analysis

Gargi Mitra
Indian Institute of Technology Madras
gargim@cse.iitm.ac.in

Prasanna Karthik Vairam
Indian Institute of Technology Madras
pkarthik@cse.iitm.ac.in

Patanjali SLPSK
Indian Institute of Technology Madras
slpskp@cse.iitm.ac.in

Nitin Chandrachoodan
Indian Institute of Technology Madras
nitin@ee.iitm.ac.in

Kamakoti V
Indian Institute of Technology Madras
kama@cse.iitm.ac.in

ABSTRACT

Privacy leaks from Netflix videos/movies are well researched. Current state-of-the-art works have been able to obtain coarse-grained information such as the genre and the title of videos by passive observation of encrypted traffic. However, leakage of fine-grained information from encrypted video traffic has not been studied so far. Such information can be used to build behavioral profiles of viewers.

Recently, Netflix released the first mainstream interactive movie called 'Black Mirror: Bandersnatch'. In this work, we use this movie as a case-study to develop techniques for revealing information from encrypted *interactive* video traffic. We show for the first time that information such as the choices made by viewers can be revealed based on the characteristics of encrypted control traffic exchanged with Netflix. To evaluate our proposed technique, we built the first interactive video traffic dataset of 100 viewers; which we will be releasing. Our technique was able to reveal the choices 96% of the time in the case of 'Black Mirror: Bandersnatch' and they were also equally or more successful for all other interactive movies released by Netflix so far.

CCS CONCEPTS

• **Networks** → **Network privacy and anonymity**; • **Security and privacy**;

KEYWORDS

Netflix interactive videos, encrypted video traffic analysis, privacy leak

ACM Reference Format:
Gargi Mitra, Prasanna Karthik Vairam, Patanjali SLPSK, Nitin Chandrachoodan, and Kamakoti V. 2019. White Mirror: Leaking Sensitive Information from Interactive Netflix Movies using Encrypted Traffic Analysis. In *SIGCOMM '19: ACM SIGCOMM 2019 Conference (SIGCOMM Posters and Demos '19), August 19–23, 2019, Beijing, China.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3342280.3342330

SIGCOMM Posters and Demos '19, August 19–23, 2019, Beijing, China

ACM ISBN 978-1-4503-6886-5/19/08…$15.00
https://doi.org/10.1145/3342280.3342330

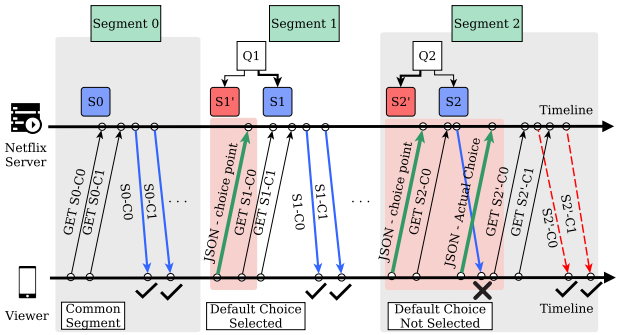

Figure 1: The streaming process of Bandersnatch

1 INTRODUCTION

Recently, Netflix released the first mainstream interactive movie called *Black Mirror: Bandersnatch*, which is the most popular among the six interactive movies on Netflix [1]. In this movie, viewers are allowed to create their own story-line by choosing one of the on-screen options (e.g., an option to throw tea over the computer or to shout at Dad) presented to them. Depending on the viewer's choice, the corresponding segment of the movie gets played. Interestingly, the choices made and the path followed can potentially reveal viewer information that ranges from benign (e.g., their food and music preferences) to sensitive (e.g., their affinity to violence and political inclination). Although this information is available to Netflix, they are bound by legal clauses that prevent them from misusing it. Today, to prevent this information from leaking to unauthorized parties, Netflix uses end-to-end encryption. Recent works [4, 5] in the domain of encrypted network traffic analysis were able to infer information such as the title of videos watched by a viewer. In this work, we raise the following research question: *'Do interactive movies leak fine-grained information about the viewers to passive eavesdroppers (compromised network devices) even when encrypted?'*.

Contributions. (1) We present the first traffic analysis technique for interactive videos that can leak more information than non-interactive videos. **(2)** We present the first dataset and identify future directions for encrypted traffic analysis on interactive videos.

2 RELATED WORK

The objectives of existing research works [3–5] have so far been limited to discerning video content, identifying the titles of videos,

Conditions	Attribute	Value
Operational	Operating System	Windows, Linux, Mac
	Platform	Desktop, Laptop
	Traffic Conditions	Morning, Noon and Night
	Connection Type	Wired, Wireless
	Browser	Google-chrome, Firefox
Behavioral	Age-group	< 20, 20-25,25-30, > 30
	Gender	Male,Female,Undisclosed
	Political Alignment	Liberal, Centrist, Communist, Undisclosed
	State of Mind	Happy, Stressed, Sad, Undisclosed

Table 1: Attributes of the dataset

and the genre preferred by a viewer from encrypted traffic of *conventional* videos.

Existing techniques from the literature are not suitable for encrypted interactive video traffic analysis due to the following fundamental reason: inter-video features cannot be used to differentiate between segments from the same video. For instance, prior works [4] have used bitrate as a feature to differentiate between two video streams. However, in the context of an interactive video, the bitrate of chunks pertaining to each choice will be the same and hence cannot be used to distinguish between two video segments. In this work, we identify an intra-video side-channel and show that it holds across various operating conditions.

3 TRAFFIC ANALYSIS TECHNIQUE

In this section, we first describe the streaming process of interactive Netflix videos. Due to the non-linearity of the script, the streaming process is check-pointed at each choice-question in the movie unlike conventional videos that stream continuously. The content of interactive videos is divided into several segments, each corresponding to one path segment in the script, with each segment consisting of multiple video chunks. Figure 1 shows an example of the control and data traffic exchanged when a viewer makes choices. The first segment of the movie (i.e., Segment 0) is common for all viewers at the end of which the choice-question $Q1$ is presented. At this point, the viewer's browser sends a JSON file (type-1 JSON file) to the server indicating that the viewer has encountered $Q1$. The viewers are then given ten seconds to choose one of the options. Our experiments revealed that Netflix considers one choice to be the default and prefetches video chunks corresponding to this choice. We denote the default choice for question Qi as Si and the non-default choice as Si'. If the viewer chooses Si, then the streaming continues uninterrupted. However, if the choice Si' is chosen, the prefetching for Si stops and a request for Si' is sent.

In the example, the viewer selects the default choice $S1$ for $Q1$. The streaming continues uninterrupted until $Q2$ appears. Like before, a type-1 JSON file is sent from the viewer to Netflix. We assume that the viewer selects $S2'$ for Q2. In this case, a type-2 JSON file will be sent subsequently. Although some chunks corresponding to $S2$ are pre-fetched, they will be discarded and the chunks corresponding to $S2'$ will be streamed. Therefore, we can conclude that the number and type of JSON files sent indicate the choice made by the viewer. However, identifying the two types of JSON files from encrypted network traffic is challenging.

Our analysis revealed that the packets carrying the encrypted type-1 and type-2 JSON files can be distinguished from other packets by their SSL record lengths. This observation was found to be consistent across various operating systems, browsers, devices, connection media, network conditions, and for all interactive movies released so far. Hence, we use *'SSL record lengths of client packets'* as the side-channel. The JSON file types can be used to infer if the choice is the default or the non-default one. The actual on-screen choices can be inferred by re-creating the sequence.

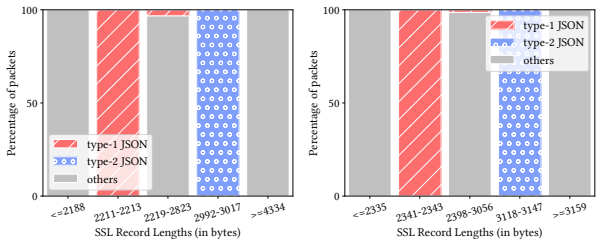

Figure 2: SSL record lengths for (Desktop, Firefox, Ethernet, Ubuntu) and (Desktop, Firefox, Ethernet, Windows)

4 RESULTS AND DATASET

We conducted our experiments on all the six interactive videos released by Netflix. Due to space constraints, we present the results for 'Black Mirror: Bandersnatch' but our results were consistent across all movies. For evaluation, we collected the encrypted network traces corresponding to viewing sessions of 100 viewers. The network traces were collected from the nearest gateway. The ground truth for evaluation was generated by asking the viewers to note down the choices they made. The viewers were free to choose the device, operating system, browser, and the geographic location from which they could watch the movie. We also collected some behavioral parameters of the viewers to highlight the implications of this privacy leak. Table 1 summarizes the behavioral and operational conditions. We have released our dataset to interested researchers [2]. Figure 2 shows the SSL record lengths of packets carrying type-1 JSON, type-2 JSON, and other traffic. Note that the two JSON files have distinctive record lengths most of the time (96%), based on which the viewer choices can be inferred.

5 CONCLUSION AND FUTURE WORK

In this work, we show the first side-channel attack on interactive videos and release a dataset for other researchers.

Countermeasures. An easy fix for the problem would be to either split the JSON file or to compress it so that it becomes indistinguishable. However, there could be timing side-channels that may still exist even after this fix. Despite disclosing this vulnerability to Netflix several months beforehand, it has not been fixed perhaps due to performance reasons.

Advancing privacy research. The video streaming engine used by Netflix is proprietary and hence we could not implement and evaluate our countermeasures. Practical implementation of even simple countermeasures may require significant changes to the software and may result in overheads and poor viewing experience. Further, these countermeasures may in-turn expose stronger unobvious side-channels. To implement and evaluate countermeasures, there is a need for an open source model of the interactive video control packet exchange, which we plan to develop in a future work.

REFERENCES

[1] 2019. *Interactive content on Netflix*. Retrieved May 20, 2019 from https://help.netflix.com/en/node/62526

[2] Gargi Mitra, Prasanna Karthik Vairam, Patanjali SLPSK, Nitin Chandrachoodan, Kamakoti V. 2019. Netflix interactive video traffic dataset. Github link: https://github.com/Gargi-Mitra/SIGCOMM2019-NetflixInteractive.git.

[3] Feng Li, Jae Won Chung, and Mark Claypool. 2018. Silhouette: Identifying youtube video flows from encrypted traffic. In *NOSSDAV*. ACM, 19–24.

[4] Andrew Reed and Michael Kranch. 2017. Identifying https-protected netflix videos in real-time. In *CODASPY*. ACM, 361–368.

[5] Roei Schuster, Vitaly Shmatikov, and Eran Tromer. 2017. Beauty and the burst: Remote identification of encrypted video streams. In *USENIX Security*.