Swipe Along: A Measurement Study of Short Video Services

Shangyue Zhu, Alamin Mohammed, Aaron Striegel

Introduction

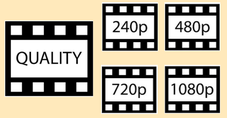
Why is Short Video?

Short videos have emerged recently as a popular form of short-duration User Generated Content (UGC) within modern social media. With increasing popularity, short video services use a large portion of Internet bandwidth. For example, TikTok has now exceeded 65 million active daily users in the U.S. alone, while the average number of watched videos is over 50 per day for each user. The explosive growth has bought billions of dollars in annual economic benefits to short video platforms.

What are we concerned with?

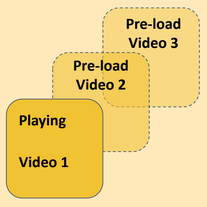
Background pre-loading aims to eliminate start-up delay or wait time between videos, which is now prioritized higher in Quality of Experience (QoE) objectives, given the application design that now facilitates 'swiping' to the next video. Thus, perceptual quality analysis, pre-loading strategies, and the impact of bandwidth requirements will be potential challenges for short video research.

The scope of our work is to characterize and experimentally compare four of the most popular short video services: TikTok, YouTube Shorts, Instagram Reels, and Facebook Watch.



Video perceptual quality

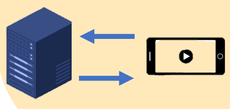
Investigate and compare the difference between the uploaded video and the source video by VMAF.



Video pre-loading

Investigate the pre-loading policy on different services.

Impact on Bandwidth



Explore the performance of short video services in different network environments

Services

Service Design Characterization

	Samples (#)	Labels (#)	Resol. (#)	VP9 (%)	AVC (%)
Service1	2,472	1	8+	0	100
Service2	3,583	8	12+	98.3	100
Service3	2,264	18+	42+	88.7	69.4
Service4	1,821	57+	215+	86.1	67.3

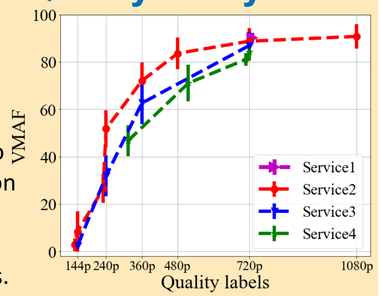
Table1: Encoding information per service.

Table 1 shows the resolutions associated with a smaller number of labels in most services, suggesting precise video quality tuning of variants. The "+" instances present they may carry an even larger number of labels and resolutions in the wild.

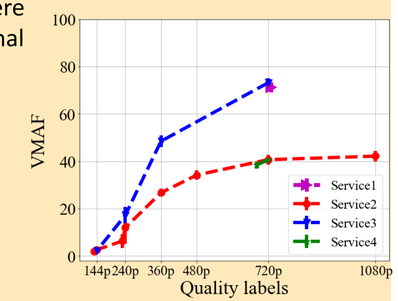
Perceptual Quality Analysis

VMAF scores

The VMAF score is obtained by comparing the transcoded video to the original. Degradation to the original video comes from reduced resolutions and bitrates. The VMAF score has a range of 0 to 100, where 100 equal to the original content, whereas 0 indicates the lowest possible quality.



Low motion videos



High motion videos

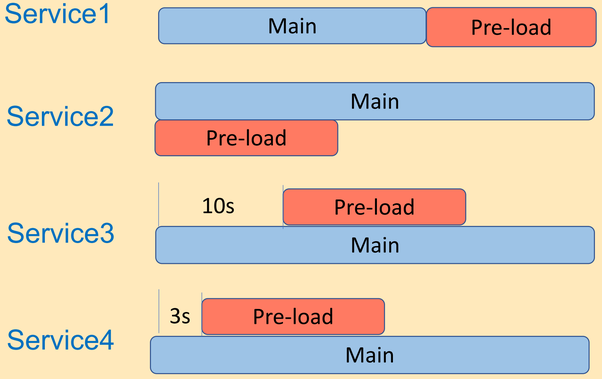
Analysis

Generally, for all services, VMAF drops as motion increases, with the strongest drop seen for high-motion videos. On the high end, most services opt to not go past the video quality where returns significantly diminish, which corresponds to 720p, since all these services being smartphone-oriented. On the low end, there are 2 services offer variants at extremely low quality (VMAF of 20 and below), which may be due to either default encoding settings or attempt to be able to play even at the harshest network conditions.

Pre-loading Policy

What is Pre-loading?

This example provides a brief glimpse into one of the pre-loading policies under such conditions where network connectivity is disabled just as the first video completes playing. Despite the network being disabled, portions of the next three videos are still able to be played (via swiping) with eventually the fifth and sixth videos having no content.



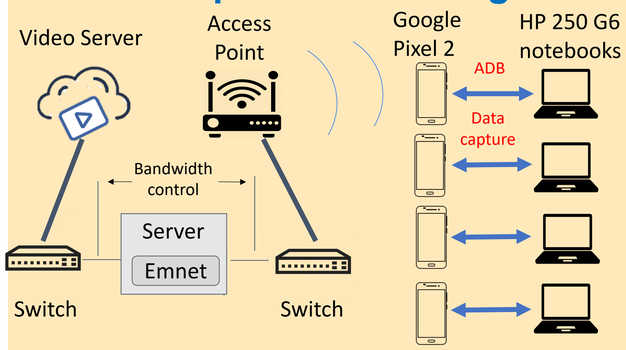
We conclude the pre-loading policy analysis by comparing the differences in relative positioning of the pre-load, in reference to the main video download. We summarize all characteristics of pre-loading policies from the current short video services in Table 2.

Attributes	TikTok	Shorts	Watch	Reels
# of Pre-load	3(iOS) 5(Android)	1	4	16-17
Pre-load size	0.5-1.8 MB	0.1 - 6 MB	0.1– 5 MB	0.1 – 1.6 MB
Pre-load duration	1-25s	20s (iOS) 30s (Android)	2- 4s (iOS) 2- 10s (Android)	1 - 3s
Average video duration	25 s	28 s	101 s	14 s
Average video size	4. 9 MB	1.7 MB	19 MB	1.1 MB
Protocol	TCP	QUIC/TCP	QUIC/TCP	QUIC/TCP
Quality level	1+	Up to 8	Up to 6	Up to 6

Table2: Attributes of studied video services.

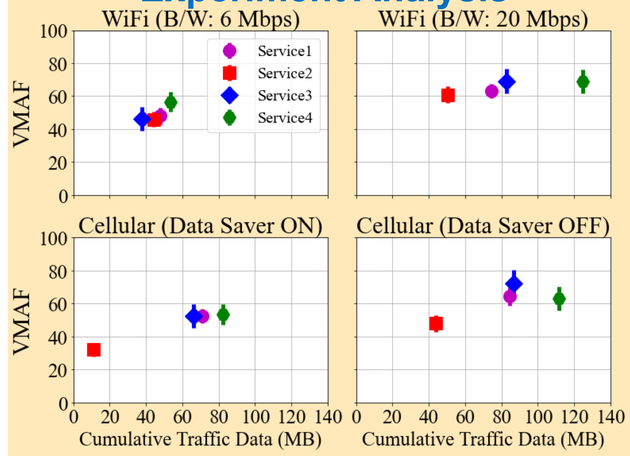
Evaluation

Experiment Setting



To analyze overall data consumption and QoE under different network configurations, we explore 4 distinct experimental scenarios. Initially we set-up a competitive network scenario on WiFi, where four clients running the same service simultaneously stream a sequence of 15 videos, share a bandwidth of 6 Mbps and then of 20 Mbps. Following, we conduct single client measurements on a cellular network for exploring the 'Data Saver' mode.

Experiment Analysis



Service1 and Service3 manage to strike a better trade-off between quality and data consumption, with Service4 using the most data but failing to provide significantly better quality than other services. The 'Data Saver' reduces data consumption by a minimum of 15% for most services, reaching a striking 60% for Service2. Such data savings are associated with quality reduction of about 15-25% for Service1, 3 and 4, but more (~33%) for Service2