

Web Quality of Experience from Encrypted Packets

Alexis Huet, Zied Ben Houidi, Shengming Cai, Hao Shi, Jinchun Xu, Dario Rossi

Huawei Technologies Co. Ltd

{alexis.huet,zied.ben.houidi,caishengming,shihao19,xujinchun,dario.rossi}@huawei.com

ABSTRACT

Pervasive encryption makes it hard for ISPs to manage their network. Yet, to avoid user churn at times of shrinking revenues, ISPs must be able to assess the quality of experience they are delivering to their customers. The case of the Web is particularly complex, with a plethora of recently proposed in-browser metrics that aim at capturing the page visual rendering quality (e.g. Above the Fold and SpeedIndex). In this demo, we showcase that such metrics can be estimated quite accurately just from streams of encrypted packets, using classic supervised learning techniques.

1 OVERVIEW

The Web resurfaced as one of the Internet killer applications, as a gateway to search, news, shopping, and social activities. Understanding user browsing experience is essential for enhancing customer engagement and satisfaction. Yet, while a plethora of metrics has been introduced in recent years to capture the visual rendering process (such as Above the Fold or SpeedIndex [1]), collection of these application-level metrics is only possible from a browser. Indeed, the prevalence of encrypted traffic nowadays makes such information completely opaque for ISPs, who consequently struggle to manage their networks and provide their users with a satisfying browsing experience.

In this demo, we bridge this gap by extrapolating high-level fine-grained application-level QoE metrics (L7) directly from the raw low-level stream of encrypted packets (L3). In particular, our demo¹ provides:

- (1) An inspection tool opposing detailed and directly comparable views of the loading timeline, from the application and network perspectives.

¹A live preview together with a short tutorial video are available online [2]

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

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

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6886-5/19/08...\$15.00

<https://doi.org/10.1145/3342280.3342297>

- (2) An approximation of multiple application-level metrics, such as Page Load Time and Speed Index, from encrypted network traffic using supervised learning (such as XGBoost and Convolutional Neural Networks).

2 DEMO HIGHLIGHTS

We now briefly describe the demonstration workflow. We refer the readers who are not familiar with Web Quality of Experience metrics (QoE) to the illustrative introduction in our short video [2]. For clarity, we here focus on single sessions: for isolating concurrent web sessions, we can leverage techniques such as those involved in PAIN [3]. At the same time, while PAIN [3] is limited to just measuring an approximation of the Page Load Time statistics, our technique is more accurate as it approximates *any* L7 QoE indicator, as we show next.

2.1 Dataset

We instrument WebPagetest to collect simultaneously HTTP Archive files (HAR) from the browser and PCAP traces from the network. The HARs allow us to measure detailed L7 performance indicators, which we will approximate using raw packet level information extracted from PCAPs.

We perform 20 runs on non-landing pages from the Alexa top-500. After discarding failed runs, we obtain a total of about 10,000 experiments. We also include different sub-pages of the same Alexa top-500 domain, as well as different network conditions (latency, loss, etc.), to ensure generality and diversity in the collection process.

Due to space limitation of the hosting platform, the online live version of the demo [2] currently portrays a representative subsample of the overall dataset (i.e., a single run of the top-100 pages, removing adult and offensive content). The demo will instead show the full dataset, which we also plan to release upon acceptance of the demo.

2.2 Inspection tool: L7 vs L3 Web progress

For each Web site and set of network conditions, we collect simultaneously L7 browser and L3 encrypted packet traces. These complementary views are shown, at a glance, in the top (L7) and bottom (L3) portions of Fig 1 respectively, for one example website. For the sake of clarity, the plots are annotated with vertical reference lines for Time to The First Byte (TTFB), Document Object Model (DOM), Page Load

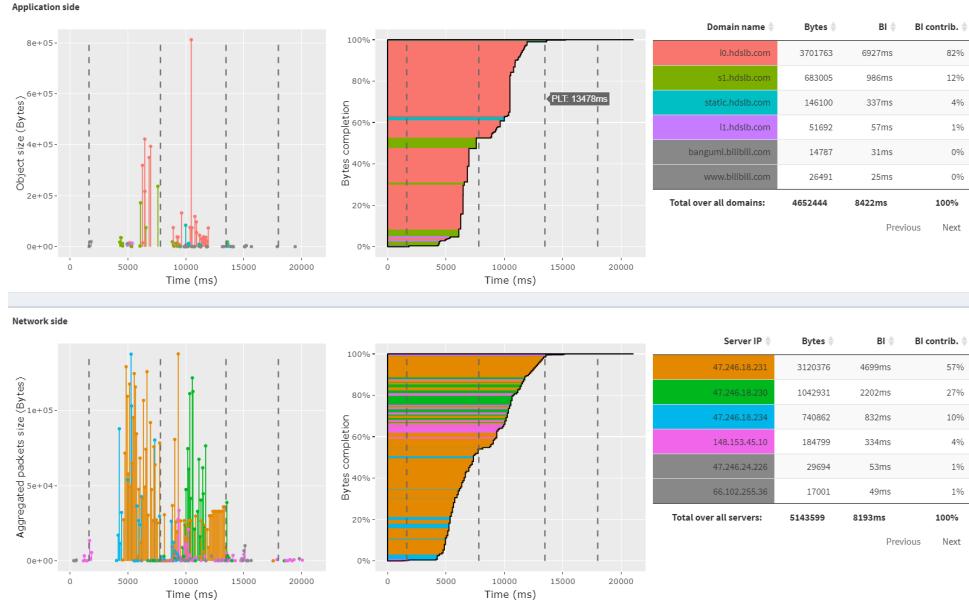


Figure 1: Illustration of the inspection tool: L7 (top) vs L3 (bottom) view of Web page progress. Screenshot is taken from the interactive demo interface accessible online [2] (a short video of the demo is accessible at the same website).

Time (PLT) and Above the Fold (ATF) times. Each color represents either an L7 domain name or an L3 IP address (colors are not related).

The top-left plot reports the object waterfall measured by the browser, whereas the bottom-left one reports the same events as seen by the network. The shapes of the time series are remarkably different due to the nature of the process at L7 compared to L3: notice for instance the 800KB object (top left plot, red spike) that the browser finishes downloading at about 12sec, but whose download started earlier in multiple packets, as the L3 time-series show (bottom left plot, green flow).

Middle plots report the cumulative view of the same events: interestingly, these visualizations are strikingly more similar. The 800KB object is visible as a sharp vertical step (from roughly 65% to 85%) in the cumulative L7 page completion plot as seen from the browser (top center plot), whereas the download progresses smoothly in L3 (bottom center plot) as the linear progress between 10 and 12 seconds shows. Finally, the right table reports details (e.g., byte volumes) pertaining to L7 domains or L3/4 TCP connections, which are the “atomic” entities of the complementary browser and network perspectives.

2.3 QoE from Encrypted Packets

We use machine learning tools to approximate application-level QoE metrics such as SpeedIndex, ByteIndex, PageLoad-Time and ATF. We train a single regression model for *all pages*, to estimate a specific QoE metric. We used XGBoost, RandomForest and CNN and obtained similar results. We perform 5-fold validation by training over over 80% of the data set, and predicting over the remaining 20% subset.

As supervised techniques require fixed-size homogeneous input, we aggregate PCAP packet traces by simply counting the per-flow bytes at a $S = 100\text{Hz}$ resolution, and limit the observation window to $T = 10\text{s}$. In other words, each packet level time-series is first transformed to 100 samples of number of bytes exchanged over 100ms intervals. We then further aggregate all flows of a session into a single time series.

As it can be seen through the demo (the panel is not shown in this extended abstract due to lack of space), supervised learning techniques accurately estimate advanced QoE indexes: for instance, median error for SpeedIndex is less than 500ms (Q1: 164ms, Q3: 1015ms), with a relative error of 18%.

REFERENCES

- [1] [n. d.]. <https://sites.google.com/a/webpagetest.org/docs/using-webpagetest/metrics/speed-index>. ([n. d.]).
- [2] [n. d.]. <https://huawei-webqoe.shinyapps.io/sigcommdemodemo/>. ([n. d.]).
- [3] Martino Trevisan, Idilio Drago, and Marco Mellia. 2019. PAIN: A Passive Web performance indicator for ISPs. *Computer Networks* 149 (2019), 115–126.

DEMO REQUIREMENTS

Equipment to be used. We will use one laptop, which requires internet connection, and one monitor. We will also display a poster detailing the motivation and theory behind our work.

Space needed. The default space of one table space and poster board is enough for our equipment and poster.

Setup-time required. Few minutes are needed to setup the demo.

Additional facilities needed. While internet access is important, the default is enough. We can fallback to a locally hosted version of the demo in case of internet failure.