Analysis of Top Internet Websites Load Times

Samuel M. Garcia Murça
Dept. Statistics and Consumers
Instituto Angolano das Comunicações
Luanda, Angola
samuel.murca@inacom.gov.ao

Daniel Perdices, Jorge E. López de Vergara
Dept. Electronics and Communication Technologies
Escuela Politécnica Superior, Universidad Autónoma de Madrid
Madrid, Spain
{daniel.perdices, jorge.lopez_vergara}@uam.es

***Abstract*—Motivated by the significant surge in demand of web browsing traffic, the economic costs of minimizing delays, and the paramount importance of Web Quality of Experience (WebQoE), this paper delves into the study of website load times during user navigation. In order to answer to the premises of the study, measurements concerning the W3C timing metrics of the top Internet domains were actively and periodically obtained during a month from different geographical regions. We found that these metrics distribution is akin to the Gamma distribution with positive bias, and they vary influenced by geographical location, time, and day of the week. Finally, the popularity of these websites has been compared to their performance, showing that some actors consistently deliver the best quality while others fall behind, despite their relevance in the rankings.**

***Index Terms*—Web Quality of Experience (WebQoE), Navigation Time, Cloud Computing, top websites, longitudinal study.**

I. Introduction

Nowadays, most of the content consumed on the Internet is web-based, where the browser has become the main platform through which a considerable number of services are accessed, such as media streaming, research, or entertainment (social networking, online gaming, chat, etc.), leading to a growth of the demand for this traffic. This increase of network traffic generated by web applications [1], and consequently the performance of websites, have become subjects of discussion not only by experts in diverse areas such as networking, IT or marketing, but also by users whose Quality of Experience (QoE) might be severely affected.

The delay sensitivity [2] of those services at the top of the rankings of the most accessed websites on the Internet has further sparked interest in understanding their behavior. From the knowledge of their nature or the relationship between different metrics, scientists and engineers are able to improve and/or design better products and processes.

Therefore, following the idea of other works, such as [3], [4], in this research we have done a longitudinal study, measuring every hour the top Internet domain loading times during a month from different locations. Once we had these measurements, we have analyzed them in order to understand their behavior, and find which main factors influence the Quality of Service (QoS). These factors under analysis include the user location or the access date, both of them being relevant. Moreover, we have also found strong correlations among different metrics, such as between *Page Load Time (PLT)* and *Front End Time (FET)*. Understanding these relations among the metrics helps service providers to improve the perceived web performance. Furthermore, we have also found that the popularity of a website does not necessarily imply a good QoE: some domains included in the top ranks have exhibited a page load time higher than expected in a large percentage of measurements.

The rest of the paper is structured as follows: First, we provide some background and discuss the related work. Next, we describe how we developed the measurements. Then, we analyze the obtained results. After this, we provide a discussion of the results. Finally, we conclude the paper and outline some future research lines.

II. Background and Related Work

Nowadays, QoS and QoE are crucial for service providers and customers. In the former case, these metrics are a relevant indicator not only for problem identification and resolution, but also in the planning or design of new products and services [2]. In the latter case, customers' perception of the performance may lead them to choose one provider over other competitors, while also sharing their experience for better or for worse. Thus, it is no surprise that the topic of web performance has been extensively studied in the past. In this section, we cover the topic of performance measurements for the web, recent tools to measure it, and related work.

A. Web Performance Measurements

Generally, we define web performance as the set of metrics that represent the perception of a user of how fast or slow a website loads and functions. This is important because shorter load times improve the user experience, so users are likely to prefer them and increase the time they spend on these websites [5]. On the other hand, slow web domains test users' patience and may cause them to leave the website before they find everything it has to offer. Additionally, more and more companies are turning to ad-based monetization as a way to increase their revenues, which uses advertisement embedded into the website to financially support it. Thus, the website performance has a big impact on its profitability.

Several studies have pointed out the importance of delay and its direct relationship to business value. For example, Amazon

This research has been partially funded by the Angolan Government, by Universidad Autónoma de Madrid, by the Spanish State Research Agency under the project AgileMon (AEI PID2019-104451RB-C21) and by the Spanish Ministry of Science, Innovation and Universities under the program for the training of university lecturers (Grant number: FPU19/05678).

and Google reported losses of between 0.6 and 1.2% when the delay increased between 0.4 and 1 second, while Shopzilla reported a 12% increase in revenue for a 5-second reduction in Page Load Time [6].

Nearly half of users expect a domain to load within 2 seconds, and 40% of users will abandon the page if it takes longer than 3 seconds to load. A 1-second delay in page response can mean a 7% reduction in users staying on the domain [7]. Additionally, website performance affects not only its users, but also their ranking in Google's search results. It has been more than a decade since Google indicated that page speeds are a factor in ranking websites in its results, although content relevance remains the main factor in this process [6].

Nowadays, websites are becoming increasingly complex. They have intertwined static and dynamic resources, e.g. JavaScript, CSS, images, videos, or audio, loading both in parallel and sequentially. Consequently, web performance is not just solely based on the network delay, but it is affected by this entangled load of many resources. The W3C provides the *Web Performance Timing Interface* [8], [9] with the premise of providing a complete picture of end-to-end latency, i.e. complete user experience information regarding the time duration of each resource that makes up the web pages. It provides data that can be used to measure the performance of a website. Unlike the JavaScript-based libraries that have historically been used to collect similar information. This API can be much more accurate and reliable, and after collection it can also transmit the results to a metrics server [9].

B. Measurement problems and tools

What is observed in previous traffic studies [10], [11], and consequently corroborated by recent W3TECHS reports, is that more than 80% of websites are under the HTTPS protocol [12]. This makes the process of passively measuring the web performance difficult, due to the encryption mechanisms that make up these connections.In order to cope with this problem, it is possible to use active methodologies such as browser development tools, commercial applications (e.g. *pingdom* from Solarwinds [13] or *dotcom monitor* [14]) or freemium and open-sourced applications such as *Apache Jmeter* from Apache Foundation [14], [15], or *browsertime* and *sitespeed*, both from sitespeed.io [16]. These tools use the *W3C Navigation Timing* standard and the JavaScript *Navigation Timing API* to collect data associated with the timing performance of web applications. Sitespeed.io solutions also use the *W3C Selenium Webdriver* standard to automate and/or simulate user interaction with websites and run under Docker containers.

C. Related Work

Once the theoretical foundations of web performance measurements have been presented, the works related to the present paper are shown. Several studies have been published focusing on the core of this research, which studies and analyzes some parameters related to WebQoS and/or WebQoE.

A comparative study of methods for measuring the loading times of the resources that make up web applications was presented in [1], focusing on the use of the *Navigation Timing API* to measure HTTP response times within browsers, with the aim of encouraging the use of the browser as a platform for large-scale, representative end-to-end network performance measurements. Based on the various experiments conducted, the authors claimed that *Navigation Timing API* provides information on when a connection is established, how long it took to make the request, as well as how long it took for the request to be fulfilled, details that the other methods fall short of, and also managed to provide a good estimate of latency with the Google Chrome web browser.

In 2016, another study was conducted [6] where the authors proposed new metrics to estimate the WebQoE. Their contributions focused on two parts: first, they provided a comprehensive taxonomy of existing WebQoE metrics and tools where they classified them into 4 categories (Time Instant Metrics, Time Integral Metrics, Compound Scores, and Mean Opinion Score (MOS)); and second, they presented two metrics inspired by Google's *SpeedIndex*, which explicitly considers the delay of all events in the life of a web page, but has some limitations due to computational complexity. The proposed metrics, namely *ByteIndex* and *ObjectIndex*, were based on the percentages of bytes and objects downloaded at an instant t respectively. After testing them with the top 100 Alexa domains using the Google Chrome browser, they showed significant levels of correlation with lighter computational needs with respect to *SpeedIndex*. Besides, they also recognized the need to use geographically dispersed vantage points (e.g. PlanetLab nodes, Amazon nodes or other cloud service provider platforms).

In 2020, further research on this topic was published [17]. In this case, the approach is based on the use of some attributes of the W3C recommendation and the *Navigation Timing API* interface. Pursuing the goal of measuring the network at zero cost with active measurements and without the need of anchors, they used several *Navigation Timing API* timestamps, to estimate the state of the underlying network, i.e. *RTT* and *Throughput*, using deep learning techniques. They achieved consistent results even exceeding some of the existing solutions on the market. However, such an approach was limited to laboratory experiments, lacking a more in-the-wild measurement campaign.

In 2022, the study [18] was published. There, the authors in partnership with Cloudflare evaluated the relative accuracy of lists of the most popular websites, Count of requests from top 5 browsers, Unique IP addresses, Unique IP root page loads, Unique client IP addresses requests from top 5 browsers, Count of HTTP requests on the server side that present a diverse set of perspectives on what website popularity means. It is noted that each tool uses its own methodology for inferring popularity and such methods perform relatively poor against the study benchmark metrics, except for the Chrome user experience report (CrUX), a dataset of real-world web browsing activities captured by Google from Android and PC versions of Google Chrome and analyzed by the authors of such study. It is important to note that Cloudflare provides only a fraction of the most visited websites listed in top 1000 of other rankings, such as Umbrella (1.99%), Majestic (10.12%), and Alexa (14.97%).

It is notorious that all the studies share the idea that WebQoE is expressed in terms of the page load time (*PLT*) [5], [19], but they do not characterize or describe the parameters that make it up. In [20], the author defines $PLT = DLT + SCT + SRT + PDT$ (see section III-D). The author leaves aside the processing time of the object, the influence of the time of day, the region and the day of the week, so in our work we propose to analyze each of these parameters.

III. System Design and Development

This section presents the strategy used in our study to obtain the performance metrics data for the top Internet domains from different geographic locations.

A. Top Internet domains

Of the various synonyms and/or meanings of the word *Principal*, the following is of particular relevance: *"That has the first place in estimation or importance and takes precedence and preference over others"*. Having said that, other criteria were added to the previous idea, with the aim of listing the top domains on the Internet, i.e. to make the ranking of the 100 most visited websites on the Internet. As stated in [21], this is an arduous task.

For this purpose, we proceeded to analyze and filter the results reported by the tools most commonly used by researchers [18], such as Alexa Internet belonging to Amazon, discontinued in May 2022, Majestic Million by Majestic SEO, and Umbrella by Cisco, both from March 2022. The filtering and analysis criteria consisted of discarding aliases (e.g. *www.google.[country code]* and *www.google.com*), adult sites (to avoid usage blocking policies) and the addition of some domains from Spain (such as *elpais.com* and *marca.com*), to have some local data to compare results with local requests. The final list of studied websites is available in the project repository at GitHub[1].

B. Browser and Operating System

Google Chrome browser has been used, as it has the largest market share [22], [6]. Similarly, Windows is the dominant desktop operating system worldwide, where in June 2022 it had a market share of just over 76% [23]. Consequently, this setup has been used to imitate an average user.

C. Geographical Locations

To satisfy the need to measure and study the load times of the main Internet websites from different geographical regions over a considerable period of time, we turned to cloud computing solutions, specifically Microsoft Azure Cloud from three regions: America (Central Canada), Europe (West Europe), and Asia Pacific (Australia West), which general purpose equipment (D1V2/DS1V2) had the following characteristics[2]: 3.5 GB of RAM, 1 CPU and 50 GB of SSD and a network bandwidth of 750 Mbit/s. Note that the available network bandwidth is comparable to those presented in existing FTTH user access links. The total cost of the setup was about 250 USD during the whole experiment period. Extending the measurements in number of months, browsers, vantage points or websites is possible, leveraging cloud scalability features.

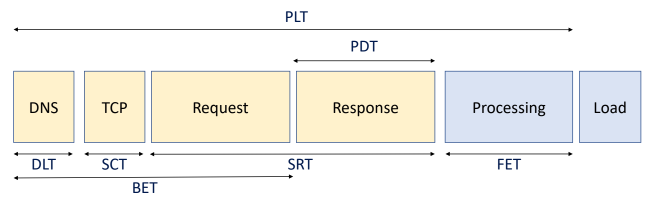

Figure 1: W3C performance metrics collected in this study.

D. Metrics to Collect

The analytical component of the study focuses on the analysis of WebQoS related data behavior and web performance metrics presented below [16], [9] and depicted in Figure 1: *Domain Lookup Time (DLT)* is the time it takes to perform the DNS lookup, *Server Connection Time (SCT)* is the time it takes to connect to the server, *Server Response Time (SRT)* is the time taken by the server to send the response, *Back End Time (BET)* is the time it takes for the network and the server to generate and start sending the requested object, *Front End Time (FET)* is the time taken by the browser to parse and create the page, *Page Download Time (PDT)* is the time it takes to download the page (the HTML code), and *Page Load Time (PLT)* is the time it takes for the page to load, from the time the page view starts (e.g. when clicking on a link) until loading is complete in the browser. They are obtained using the *Navigation Timing API*. Apart from W3C metrics, we also took these classical network metrics: *Transfer Size (TFS)* is the size of the compressed response (web page); *Round Trip Time* (RTT) is usually measured using ICMP or the TCP handshake if ICMP is disabled, as we discovered that this is frequent; finally, *Throughput* (TP) is the average throughput (bits received over time).

IV. Analysis and Results

After collecting the performance metrics data of the top websites at each hour from 18th June to 18th July 2022, from multiple vantage points across the globe, the analysis and the referred results will be presented. In a first step, all the outputs resulting from the measurements were unified, and sequentially all the necessary preprocessing was applied.

A. Dataset Analysis

First, an extreme value analysis was performed for each metric and website to ascertain the consistency of the data (data cleanliness). It is worth mentioning that *www.sky.com* and *www.arxiv.org* presented incongruent values in the *TFS*,

[1] https://github.com/sam190890/Measuring-Top-Internet-WebSites

[2] https://learn.microsoft.com/en-us/azure/virtual-machines/dv2-dsv2-series

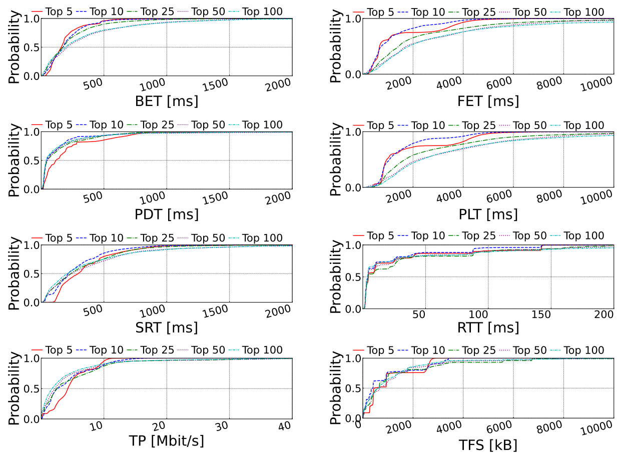

Figure 2: cumulative distribution function of the web performance metrics.

i.e. “0” as the *DLT* and *SCT* metrics in about 66% of the measurements are 0 due to the existence of persistent connections and the internal DNS cache of Google Chrome. Thus, *DLT* and *SCT* of these domains were suppressed.

B. Characterization of the collected variables

To characterize the aforementioned metrics for each website, probability distributions are the standard way to model data without predictors. However, before getting into them, some visual orientation is necessary, using the Cumulative Distribution Function (CDF), to select a small set of models that may fit your data.

In Figure 2 it is observed that the obtained distribution functions present behaviors similar to the Gamma distribution, with the peculiarity that they have long tails to the right, this is, positively skewed, which corroborates with the results of the study [1]. Such distributions have been used in the modelling of waiting times, *RTT* and other network parameters [24]. Note that the behavior is not similar among different top domains.

C. Expected values

One of the great disadvantages of point estimators is that they are almost never exactly equal to the actual values they are estimating. Therefore, the need arises to estimate an interval instead of a point value, in order to be able to guarantee a confidence of the estimate, frequently $1-\alpha = 95\%$.

Since all quantities are estimated using the mean across multiple samples, simple confidence interval of Monte-Carlo methods as a result of the Central Limit Theorem (CLT) or Lyapunov CLT can be computed. This is,

$$\mu \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \tag{1}$$

where μ is the mean, σ the standard deviation and $z_{\alpha/2}$ is the percentile $\alpha/2$ of the normal distribution. Applying these to all the metrics, we obtained the following confidence intervals:

$$\begin{aligned}
BET &= 330.4663 \pm 0.1071 \text{ ms}\\
FET &= 3\,692.7585 \pm 0.7985 \text{ ms}\\
PLT &= 4\,180.8915 \pm 0.8175 \text{ ms}\\
PDT &= 157.4926 \pm 0.0704 \text{ ms}\\
SRT &= 388.9155 \pm 0.1307 \text{ ms}\\
TFS &= 1\,174\,557.5000 \pm 211.5000 \text{ bytes}\\
TP &= 4.1069 \pm 0.0008 \text{ Mbit/s}\\
RTT &= 28.7934 \pm 0.0067 \text{ ms}
\end{aligned}$$

D. The importance of the geographical location

The influence of geographical location on the behavior of the variables is studied in this section. Figure 3 shows the distributions of the metrics for the three different regions, which seem to have different behavior. In order to respond to this question, the goodness-of-fit or non-parametric hypothesis testing methods, *Kolmogorov-Smirnov test* (K-S test), whose premise is to quantify the l^{∞}-distance of the CDFs, i.e.

$$D = \sup_{x\in\mathbb{R}} |F_1(x) - F_2(x)|, \tag{2}$$

where F_1 and F_2 are the two CDFs we want to compare. If both of the CDFs have N samples, then the hypothesis of $F_1 = F_2$ is rejected at a level of confidence $1-\alpha$ if and only if

$$D > \sqrt{-\tfrac{1}{N}\log\tfrac{\alpha}{2}}. \tag{3}$$

Applying the K-S test, we confirm that the distributions cannot be considered similar across different locations for none of the web metrics (for a significance level of $1-\alpha = 95\%$, and a p-value near 0 for all cases). Thus, based on the data, we conclude that web performance depends on the region.

E. The relevance of the hour of the day

This section studies the influence of the hour of user's location on the behavior of the variables. In order to know if they are similar or not, the same methodology applied in the section IV-D will be used, given the number of comparisons or combinations between the schedules for each metric, only the maximum of D across all possible pairs of hours is analyzed.

From the results of the *Kolmogorov-Smirnov* test, illustrated in Table I, it is observed that the time taken for the *BET*, *SRT* and *RTT* are influenced by the time the pages are accessed, and represent 50% of the time metrics, as the others present no reason to reject the null hypothesis, i.e., values generated by distributions of the same metric do not vary depending on the hour of the day.

F. Correlation Analysis

Another possible approach to web performance modeling with these many variables is to predict ones that are complex, such as *PLT*, as a function of simpler ones, such as *RTT* or *TP*. To do this, it is necessary to understand how to summarize the relationship between variables [25].

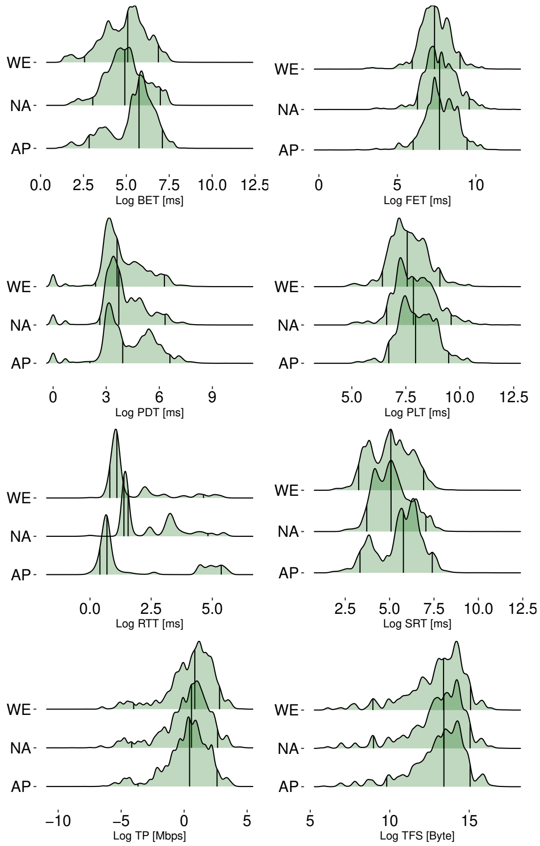

Figure 3: PDFs of the web performance metrics by region.

TABLE I
KOLMOGOROV-SMIRNOV TEST RESULTS BY HOUR

In **red**, rejected hypothesis at $1 - \alpha = 95\%$.

Metric	D	p-value
BET	0.021	**0.04**
FET	0.012	0.50
SRT	0.022	**0.03**
PDT	0.013	0.50
PLT	0.012	0.5
TFS	0.007	1
TP	0.02	0.06
RTT	0.029	**0.001**

In this section, we will proceed with the correlation analysis to analyze the relationship between variables, whether dependent (or endogenous) and/or independent (or exogenous).

For our purpose, Pearson correlation coefficient r is the most common numerical measure of the degree of linear relationship between two quantitative variables. However, this correlation coefficient does not consider non-linear relations, so, a low value of r does not necessarily mean that the variables are independent [26]. Another issue is that it is not a robust strong statistic, as it can often change considerably due to the presence of outliers, so the Spearman correlation coefficient ρ was also applied.

About the interpretation of these values, according to [26], 0.10, 0.30 and 0.50 correspond to small, medium and large effects, respectively. Besides, the authors of [18] present a similar interpretation for both coefficients, i.e. Pearson and

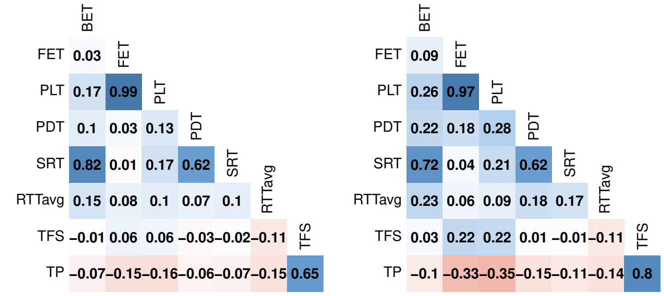

Figure 4: Pearson (left) and Spearman (right) Correlation Matrices for all regions. Darker colors show higher correlation. Blue and red for positive and negative correlation, respectively.

TABLE II
TOP 10 OF THE BEST AND WORST WEB DOMAINS

Domain	%	Domain	%
translate.google.com	100	www.espn.com	12.17
www.ebay.com	100	www.cnn.com	11.41
www.gmail.com	100	www.qq.com	10.97
www.google.com	100	www.foxnews.com	5.58
www.ikea.com	100	www.washingtonpost.com	5.52
www.linkedin.com	100	www.alibaba.com	5.22
www.myshopify.com	100	www.reddit.com	3.42
www.twitter.com	100	www.nike.com	1.13
www.wikipedia.com	100	www.nbcnews.com	0.69
www.wordpress.com	100	www.sohu.com	0.00

Spearman, classifying them as insignificant for those below 0.10; weak, between 0.10 and 0.39; moderate, between 0.40 and 0.69; strong, between 0.70 and 0.89; and very strong, for those greater or equal than 0.90.

In some contexts, even small correlations can be of great practical importance. It should also be noted that, unless the sample is large, the correlation may be very different in the sample than in the population from which the sample was selected [26]. Therefore, in the present study, correlations greater than 0.30 will be considered as relevant. According to Figure 4, both methods present four interesting correlations: moderate (SRT, PDT), strong (SRT, BET and TP, TFS) and very strong (PLT, FET).

G. Website load time behavior

Next, the behavior of the loading times of each website was studied based on the acceptance, as we defined next. The determination threshold for the acceptance is what defines the metric. It was obtained from the mean between the lower limit of the *PLT* for 40% of users and the time it must take to load for the user to pay attention, resulting in 6.5 seconds, i.e. $PLT \leq 6.5$ is *Acceptable*, and $PLT > 6.5$ is *Not acceptable*.

Based on this threshold, Table II shows the top 10 domains that had best and worst behavior during the study. The percentage shows the number of samples that are below *Not acceptable* threshold.

V. DISCUSSION

In this section, we discuss the obtained results shown above. With 95% confidence, the average time taken for a page to load is less than 4.8 seconds. Of the domains studied, despite all them are included in the top 100 web domains in the world, only ten always had load times in the acceptable range, set at 6.5 seconds, during the month the measurements lasted.

There is a very strong linear correlation between the *PLT* and the delay in processing the front end (*FET*). The European and American regions show strong similarities and better WebQoE compared to the Asia Pacific region. It has been observed that hours 14, 15, and 16 are the ones that provided the highest PLT values, i.e. higher delays. Half of the time metrics studied are influenced by the time the pages are accessed, and all of them are influenced by the region from which the pages are accessed. The probability density functions of the metrics studied present characteristics or behaviors that resemble Gamma distributions, with the peculiarity of positive skewness.

It is worth mentioning that these top domains cover a large percentage of everyday users' traffic [27], so the results here provided, being a small set of the whole Internet, can be useful for understanding the behavior of the main services to which users connect on a daily basis.

For reproducibility, the collected and processed dataset, used to carry out this work, is available at GitHub[1].

VI. Conclusion

Throughout this article, we have studied the load time of websites during one month from different geographical locations. The analysis of the obtained data has been presented in order to be able to characterize the variables measured, their correlation, the existence of the influence of geographic location and time on the variables measured, and the overall website behavior.

Based on the results obtained during the study, the following improvements are proposed for future work. It would be interesting to increase the period for carrying out the measurements, e.g. one year to extract long-term trends, or extend the geographical horizon or geographical locations from which the measurements are carried out. Additionally, the number of domains to be studied should also be extended, as well as the operating systems or web browsers from where the measurements are done. Moreover, domains could be classified in different groups according to their application purpose (social network, video streaming, file storage, etc.) When doing the experiments, they should consider clearing the browser cache, in addition to the DNS, in order to mitigate the effects of persistent connections. Finally, existing studies on web page load times [28] (ITU-G1030) should be updated to determine and harmonize thresholds to current habits.

References

[1] E. Gavaletz, D. Hamon, and J. Kaur, "In-browser network performance measurement," *W3C Workshop on Web Performance 8*, 2012.

[2] T. Janevski, *QoS for Fixed and Mobile Ultra-Broadband.* Wiley - IEEE, Hoboken, NJ: Wiley-Blackwell, Apr. 2019.

[3] J. L. García-Dorado, J. A. Hernández, J. Aracil, J. E. López de Vergara, F. J. Monserrat, E. Robles, and T. P. de Miguel, "On the duration and spatial characteristics of internet traffic measurement experiments," *IEEE Communications Magazine*, vol. 46, no. 11, pp. 148–155, 2008.

[4] A. Morichetta and M. Mellia, "Lenta: Longitudinal exploration for network traffic analysis," in *2018 30th International Teletraffic Congress (ITC 30)*, vol. 01, pp. 176–184, 2018.

[5] ITU, "ITU-T Recomendation E.806: Measurement campaigns, monitoring systems and sampling methodologies to monitor the quality of service in mobile networks," *ITU-T Serie E*, 2019.

[6] E. Bocchi, L. De Cicco, and D. Rossi, "Measuring the quality of experience of web users," *ACM SIGCOMM Computer Communication Review*, vol. 46, no. 4, pp. 8–13, 2016.

[7] Yoav Einav, "Amazon found every 100ms of latency cost them 1% in sales." Available: https://www.gigaspaces.com/blog/amazon-found-every-100ms-of-latency-cost-them-1-in-sales. Last accessed: 24 de November 2022, 2022.

[8] W. Zhiheng, "W3c recommendation "navigation timing". 17 december 2012."

[9] MDN,Web,Docs, "Navigation and resource timings." Available: https://developer.mozilla.org/en-US/docs/Web/Performance/"" Navigation˙and˙resource˙timings. Last accessed: 13 de November 2022, 2022.

[10] M. Trevisan, D. Giordano, I. Drago, M. M. Munafò, and M. Mellia, "Five years at the edge: Watching internet from the isp network," *IEEE/ACM Tran. Networking*, vol. 28, no. 2, pp. 561–574, 2020.

[11] D. Perdices, G. Perna, M. Trevisan, D. Giordano, and M. Mellia, "When Satellite is All You Have: Watching the Internet from 550 ms," in *Proceedings of the 22nd ACM Internet Measurement Conference*, IMC '22, pp. 137–150, 2022.

[12] W3Techs, " A Usage statistics of Default protocol https for websites." Available: https://w3techs.com/technologies/details/ce-httpsdefault. Last accessed: 24 de November 2022, 2022.

[13] Page Speed, "Page speed monitoring." Available: https://www.pingdom.com/product/page-speed/. Last accessed: 24 de November 2022, 2022.

[14] A. Bouchefra, *Performance tools*. SitePoint Pty, 2017.

[15] Apache, "Apache JMeter." Available: https://jmeter.apache.org/. Last accessed: 26 de November 2022, 2022.

[16] Sitespeed.io, "Introduction to sitespeed.io and web performance testing." Available: https://www.sitespeed.io/documentation/sitespeed.io/introduction/. Last accessed: 13 November 2022, 2022.

[17] I. Taibi, Y. Hadjadj-Aoul, and C. Barakat, "Data driven network performance inference from within the browser," in *2020 IEEE Symposium on Computers and Communications (ISCC)*, pp. 1–6, IEEE, 2020.

[18] K. Ruth, D. Kumar, B. Wang, L. Valenta, and Z. Durumeric, "Toppling top lists: Evaluating the accuracy of popular website lists," in *Proc. 22nd ACM Internet Measurement Conference*, pp. 374–387, 2022.

[19] N. Abdelwahed, A. B. Letaifa, and S. E. Asmi, "Monitoring web qoe based on analysis of client-side measures and user behavior," *Multimedia Tools and Applications*, pp. 1–27, 2022.

[20] S. K. Shivakumar, *Modern web performance optimization*. Berlin, Germany: APress, 1 ed., Nov. 2020.

[21] J. Kline, A. Aelony, B. Carpenter, and P. Barford, "Triangulated rank-ordering of web domains," in *2020 32nd International Teletraffic Congress (ITC 32)*, pp. 1–9, 2020.

[22] N.-A. Le-Khac and K.-K. R. Choo, *A Practical Hands-on Approach to Database Forensics*, vol. 116. Springer Nature, 2022.

[23] S. R. Departmnet, "Global market share held by operating systems for desktop pc's from january 2013 to june 2022," *statista*, 2022.

[24] L. de Pedro, M. Martínez Redondo, C. Mancha, and J. E. López de Vergara, "Estimating server load based on its correlation with TCP SYN response time," in *2020 IFIP Networking Conference (Networking)*, pp. 379–385, 2020.

[25] J. Schmuller, *Statistical Analysis with R For Dummies*. John Wiley & Sons, 2017.

[26] J. L. Myers and A. D. Well, *Research Design & Statistical Analysis*. New York, NY: Routledge Member of the Taylor and Francis Group, 2 ed., Dec. 2002.

[27] K. Ruth, A. Fass, J. Azose, M. Pearson, E. Thomas, C. Sadowski, and Z. Durumeric, "A world wide view of browsing the world wide web," in *Proc. 22nd ACM Internet Measurement Conference*, pp. 317–336, 2022.

[28] ITU, "ITU-T Recomendation G.1030: Performance in IP Networks for Data Applications," *ITU-T Recommendation G Series*, 2005.