Longitudinal Analysis of Inter-City Network Delays

Selim Ozcan
SimulaMet & University of Oslo
selim@simula.no

Ioana Livadariu
SimulaMet
ioana@simula.no

Georgios Smaragdakis
Delft University of Technology
g.smaragdakis@tudelft.nl

Carsten Griwodz
University of Oslo
griff@ifi.uio.no

***Abstract*—During the last decades, public and private investments contributed to building the Internet infrastructure, including undersea cables, long-distance fiber links, broadband networks, and satellite constellations to reduce end-to-end delay. However, a systematic longitudinal analysis of end-to-end network delays between major city pairs is still missing. In this study, we measure the inter-city delays over the last six years, considering 17 major metropolitan areas around the globe. Our analysis shows that the delay for 88% of city pairs end-to-end delay has decreased. Moreover, we study delay changes for regional and long-haul (intercontinental) pairs. Our analysis shows that end-to-end delay has decreased for 80% and 55% of city pairs in Europe and North America, respectively. Our study also shows that despite the overall decrease in inter-city delays, global phenomena, e.g., the COVID-19 pandemic, profoundly impact many inter-city connections simultaneously while not affecting others.**

***Index Terms*—Internet measurement, RTT delay, longitudinal analysis, big network data analysis.**

I. Introduction

The Internet connects people and devices worldwide, and numerous applications and services rely on the Internet infrastructure to function, including but not limited to communication, health services, and cloud gaming. Latency (delay) is an important metric for assessing network performance as it directly impacts end-user experience. In networking by latency is measured in milliseconds (ms) and includes the total propagation and processing time as the time it takes for a data packet to travel from a source to a destination, including queuing time.

End-to-end Internet connections might span wide areas on different continents via undersea cables or satellites. Improving hardware and software systems such as cables, wireless connections, routing devices, and algorithms is the primary instrument to decrease the connection delay. Although deploying cutting-edge solutions from individuals or organizations can significantly reduce latency, not all cities have the means to do so. Furthermore, millions of people live in metropolitan areas, putting a strain on the cities' Internet connections during special events, festivals, elections, and pandemics. For example, in 2017, a popular augmented reality mobile game organized a festival in a city's central park in a densely populated metropolitan area [1]. Due to the high concentration of connection requests, the mobile networks in the area were overloaded and collapsed. The game was playable again after actions had been taken to ease the congestion on the network.

978-3-903176-58-4 ©2023 IFIP

In the past decade, much research and development focused on decreasing the latency by relying on content delivery networks (CDN), edge computing, and cloud centers to convert data transfers from the global to the local level. However, a variety of scenarios remain that call for the global and intercity transfer of data [2]–[4]. Situations where long distances data transfer cannot be avoided include interpersonal communication like social networks, teleconferences, distributed games and business transaction.

Despite several studies on the changing Internet latency between cities [5]–[9], their extent is limited, they also have yet to focus on historical evolution, and further research is required. This paper comprehensively analyzes network latency evolution between major cities around the globe. To this end, we leverage delay data collected from Réseaux IP Européens Network Coordination Centre (RIPE) Atlas [10] vantage points (probes) that are scattered around the world. Using a platform developed for the study, we report on pairs between major cities where the network latency increased or decreased over the last six years. RIPE Atlas probes can be hosted in different locations like end-user homes and data centers. Hence, our latency analysis between cities is not limited to only end-users in the cities. In summary, this study presents the following contributions:

- A contemporary and historical examination of Internet measurements spanning long distances using statistical methods.
- We comprehensively analyze and compare the evolution and trends of inter-city delays between major metropolitan areas.
- We conducted tests to identify the limitations and biases of RIPE Atlas vantage point features on the Internet latency performance.
- We present our approaches for detecting anomalies in Internet delay evolution using COVID-19 impact as a test case, and empirically assess the impact of the COVID-19 pandemic and lockdowns on the inter-city delays.

The remainder of this paper is organized as follows. Section II describes the background of technologies, sources of the datasets, and the previous related research. Section III introduces the datasets, pre-processing steps, and implementation details. An overview of the analysis and overall results are presented in Section IV. Further, Section V provides city and continent-level delay evolution results. A case study for evaluating the COVID-19 pandemic impact on delay evolution is given in Section VI. Finally, Section VII discusses the results and provides conclusions.

II. Related Work

Retrospective research on Internet measurements plays an important part in understanding the development of the Internet, discovering the critical points, and understanding the potential evolution of delay in the future [11]. Previous studies investigated end-to-end delays between Autonomous Systems (AS) [12]–[14] and server-to-server delays [15] using traceroute data. The study published by Chowdhury et al. [14] relied on the Ark [16] and iPlane [17] to analyze delay from 2008 to 2013 between different countries. Conversely, we focus on the evolution of delay between cities using ping measurements. Festivals, pandemics, infrastructure upgrades, and the influence of high population on delay can all be included in the city-level delay study. Data-based measurement-driven studies have emerged that analyze the overall Internet capacity, capabilities, malfunction, irregularity, network robustness, and other aspects [18]–[20].

A massive amount of network measurement data is provided by public sources and organizations such as RIPE Atlas, which collects different networking raw data, including ping, traceroute, NTP, and DNS. RIPE Atlas probe devices (vantage points) perform measurements between two points, from probe devices as a source to any destination. Since its establishment in 2010, until this study performed, RIPE Atlas has composed probe devices with different hardware versions, including system-v1, system-v2, system-v3, system-v4, system-software, system-anchor, and system-virtual. Our classification of the probes revealed that system-v3 is the dominant hardware version for the city-matched probes. On the other hand, probes with system-v1 and system-v2 account for less than 15% of the probes in 2021. Recent studies have shown that congestion and heavy load on probes (mostly on hardware version system-v1 and system-v2) cause interference on delay measurements [21], [22].

RIPE Atlas has been extensively employed to infer Internet measurements [21]–[24]. Unlike this study's focus on using wider date-ranged data produced by RIPE Atlas probes, Davisson et al. [9] assess the Internet latency in a limited longitudinal way using the data for the first two weeks of every year between 2016 and 2021 produced by 124 RIPE anchors. Dönni et al. [25] used RIPE Atlas probes in the Schengen Area for traceroute measurements to infer whether the routes stay in the Schengen area. Fontugne et al. [26] extensively investigated network conditions and use cases using RIPE Atlas traceroute measurements. Outside of Europe, Fanou et al. [27] analyzed the topology of interdomain routing using data from 214 probes spread across 32 African countries and more than one million traceroute measurements. Fiadino et al. [28] used the RIPE Atlas platform to investigate the IP addresses of an instant multimedia messaging platform and assess latency to the servers as well as to map the discovered IP addresses to cities. Cicalese et al. [29] used data from the RIPE Atlas infrastructure for IP anycast enumeration and geolocation.

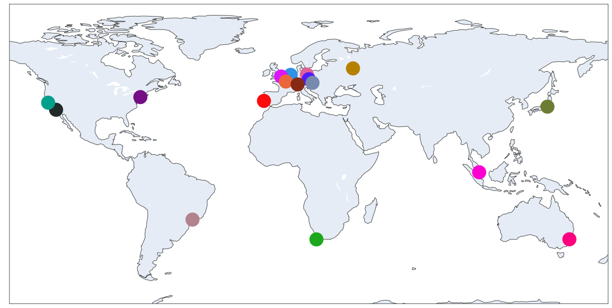


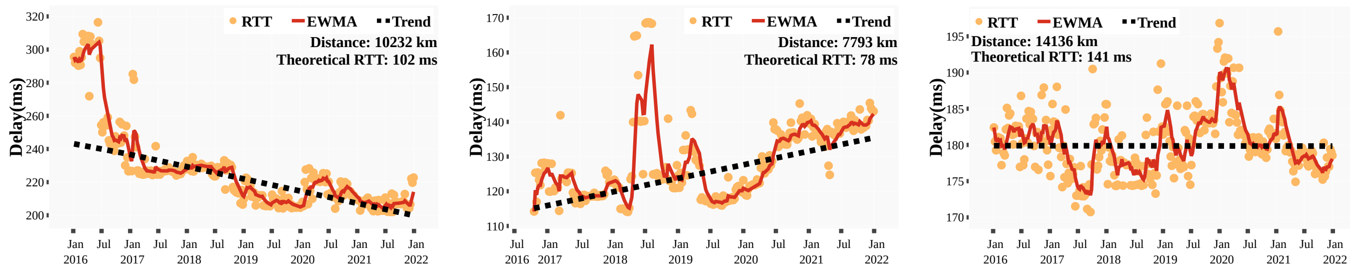
Fig. 1. Selected cities on the World map

In recent years, we have experienced an unusual phenomenon known as the novel Coronavirus Disease (COVID-19), which began spreading in the latter months of 2019 and was declared a pandemic in the early months of 2020. Beginning in the spring of 2020, numerous countries imposed a series of restrictions (stay-at-home orders, curfews, quarantines) as a result of the pandemic [30], [31]. These restrictions caused most of the population to stay at home for long periods of time, replacing the physical interactions with digital (video meetings, virtual classrooms, remote working), thus affecting the Internet performance [8], [32]–[34]. The impact of the COVID-19 pandemic has been investigated in several reports, news, and blogs analyzing the extensive growth in traffic [35]–[38]. During the pandemic, studies for Internet performance were conducted from the perspective of a single or a few operators/IXPs to a broader multi-perspective view, identifying performance changes and an increase in outages [7], [39]–[41]. Network delays are monitored at the Internet Health Report (IHR) platform [13] to study congestion during COVID-19 lockdowns using RIPE Atlas traceroute measurements.

III. Dataset & Implementation Details

In this section, we briefly describe the datasets used in this work as well as our processing approach. This study focuses on IPv4 ping measurements from RIPE Atlas probes every Wednesday between 2016 and 2022 for the 17 major cities (see Fig. 1). Three additional meta datasets are utilized in addition to the ping measurements: (i) the RIPE Atlas probe dataset, (ii) the geolocation dataset, and (iii) RIPE Atlas measurement metadata.

Ping measurements. The ping measurement dataset contains information on Round Trip Time (RTT), source probe, destination IP address, measurement time, and other results. The selected subset of the RIPE Atlas that we use in this study yielded more than *9.6* billion ping measurements, which is



(a) São Paulo → Berlin: delay decreasing trend (b) Tokyo → Sydney: delay increasing trend (c) Singapore → Los Angeles: no trend in delay

Fig. 2. Intercity delay over time: decreasing (a), increasing (b) and no trend (c) in the delay evolution.

approximately *4.5 TB* in size. We use the RIPE Atlas REST API [42] to download the ping measurements.

Probe dataset. The downloaded probe dataset includes probes that were active during our six-year measurement period, i.e., from 2016 to 2022. It includes probe location coordinates, identification numbers, IP addresses, country information, and hardware version, among other details. RIPE Atlas probes are hosted in both end-user homes and data centers. For some probes, this information can be extracted by using the user or system tags in the probe metadata, and thus it does not consistently appear across all probes. Hence, we do not break down the probes into different classes using their hosting environment. We utilize probe and geolocation data to find the source and destination cities of the ping measurements.

Geolocation dataset. The geolocation dataset comprises information on the country name, city name, population, and coordinates for *2594* cities in *247* countries [43]. The geolocation dataset is used to map probe coordinates with a densely populated city within a 25-kilometer radius of the city center.

Measurement metadata. To find probe-to-probe IPv4 ping measurements for the selected dates and cities, we downloaded more than *122* million metadata records about measurements. The measurement metadata includes meta information for all measurements that run on the RIPE Atlas platform. After filtering the metadata to have probe-to-probe IPv4 ping measurements for the selected dates and cities, we download ping measurements. Furthermore, probes are mapped to ping measurements as source and destination cities to find the vantage points for our delay study.

The metadata for every pair of probes was augmented with the geodesic distance between the two probes. It determines the physical distance for computing the theoretical minimum latency in the speed of light in fiber cables between the cities and constitutes to allow us to compute a definite lower latency bound for the probe pair.

Processing method. We set up a platform that utilizes distributed big data processing and virtualization technologies because raw network data can easily exceed terabytes in size. A heterogeneous high-performance computing cluster provides the required processing power and data storage units. We set up an Apache Spark (version 3.2.0) [44] big data processing tool in standalone cluster mode on top of the virtualized Docker [45] containers. On top of the platform, we develop four delay evolution analysis modules that look into the delay trends over time and the probe impact of delay. The *Delay RTT Evolution* module focuses on delay changes between pairwise cities over time. *Probe Analysis* examines the probe metadata and matches probes to cities. The *Stable Probe Analysis* module reports information on probes that are online over a period of time. The *Probe Hardware Version Analysis* module investigates the probe hardware versions and their differences. Additionally, we tested the platform's run time performance by running modules on 1, 2, 3, and 4 nodes four times and calculating the average run duration. We measure 9 hours 52 seconds to finish using only one node. Using two, three, and four nodes results in 7'38", 6'44", and 6'03", respectively. Comparing the performances of four nodes and one node shows that using a four nodes cluster increases the performance by 39%. In other words, using only one server takes 61% more time to finish.

We handle the ping measurements using probe features to investigate the delay-affecting causes in three ways. The first approach is using ping measurements between all probes without any filter. In the second method, we present results by using ping data that is extracted for being only between probes with specific hardware versions of system-v3, system-v4, system-anchor, system-virtual, and system-software. The ping measurements use 4,494 probes for the 17 cities and with selected hardware versions. In the last method, we investigate the delay evolution utilizing ping measurements between different sets of stable probes by their minimum period of online status. We classify probe stability as being online for more than any 52, 104, 156, and 208 Wednesdays (one to four years) across six years.

IV. Analysis

In this section, we present the results of our analysis.

A. Overview

We rely on our platform to study the evolution of intercity delays over six years. We compute the median of the measurements for each day and city pair to be considered the delay performance. An exponentially weighted moving average (EWMA) is calculated to enhance the delay evolution analysis using the RTT median values. We compute EWMA over four months sliding time window. The EWMA smooths the time series delay values to highlight the spikes and drops.

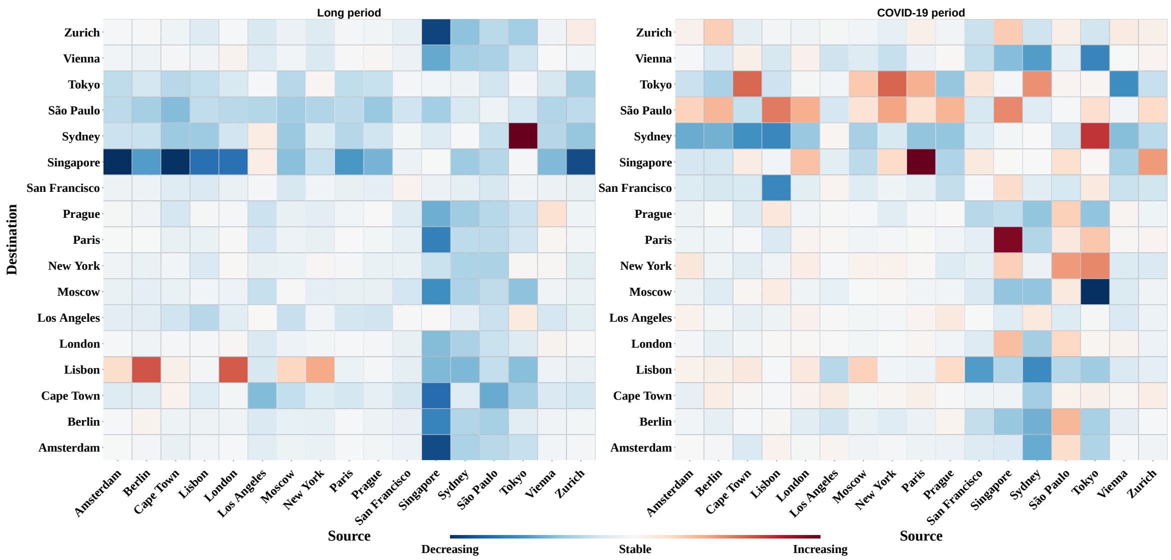


Fig. 3. City level regional delay trend slope results

Our next step is to classify the delay evolution for each city pair by applying a trend analysis method to the delay data. Using the non-parametric Mann-Kendall (MK) trend test [46], [47], we determine whether the delay has increased, decreased, or remained unchanged over time. The MK test is a standard statistical approach to inferring the trends in time series, and it does not require that the data follows a specific distribution [48]. Intuitively, the test does not compare the delay values but rather the relative magnitude of the delay. Thus, it is not sensitive to outliers. However, the test does not offer information on whether the increase is linear and does not report any local fluctuations in the data. We employ the 'pyMannKendall' Python module for the MK trend test. We show in Fig. 2 the delay evolution over time and the fitted trend lines for three pairs of cities. Note that we select these city pairs as examples to illustrate the three classes of delay trends that we encounter in our analysis. The delays from São Paulo to Berlin and from Tokyo to Sydney has decreasing and increasing trends, respectively. However, the delay from Singapore to Los Angeles varies, and we cannot detect any trend in the extracted values.

This study's decreasing/increasing trend implies that the overall delay decrease/increase is significant over time, even though there may be sudden fluctuations. The trend test shows a no-trend when the p-value does not pass the significance level. In this study, *N/A* denotes the number of city pairs with no time series data to evaluate. Our results indicate a decreasing trend in the delay for 91% of the 289 city pairs. However, for 11 city pairs, the delay results in an increasing trend. A closer look at these pairs reveals that the increase occurs both in the same city and between different cities.

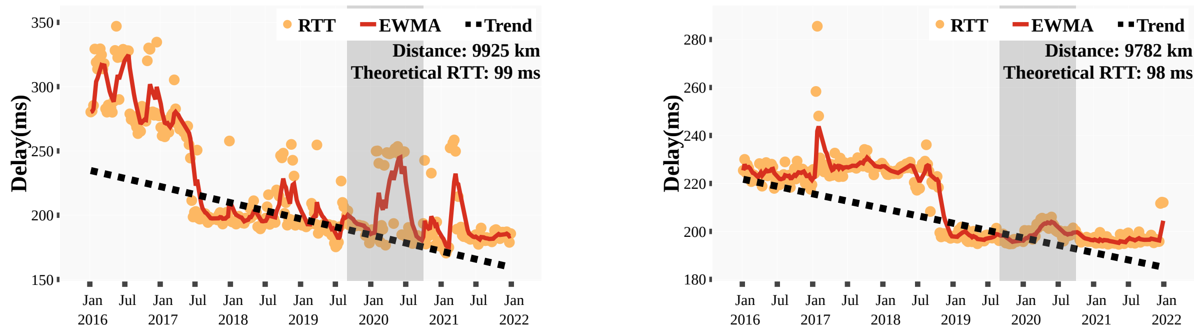
B. Local and temporal effect on the delay

The raw delay data we employ in our analysis depends on the available probes. Thus, we further analyze whether the inferred trend decrease is sensitive to a few probes which are stable over time. To this end, we consider probes that are stable over a minimum period of four years. We filter out 73 city pairs from our initial set as we do not find stable probes between these cities for at least four years. Our trend analysis reveals decreasing trends for 71% of the 216 remaining pairwise cities.

TABLE I
DELAY TREND COUNTS (PERCENTAGE) OF CITY PAIRS USING ONLY PROBES WITH HARDWARE VERSION 3-4-ANCHOR-VIRTUAL-SOFTWARE

Minimum Weeks(Years)	Trends			
	Decreasing	Increasing	No Trend	N/A
0 (0)	254 (88%)	16 (5%)	19 (7%)	0 (0)
52 (1)	249 (86%)	24 (8%)	16 (6%)	0 (0)
104 (2)	247 (86%)	24 (8%)	17 (6%)	1 (0%)
156 (3)	210 (73%)	36 (12%)	17 (6%)	26 (9%)
208 (4)	105 (36%)	28 (10%)	24 (8%)	132 (46%)

We also investigate the local effects, like vantage point features of RIPE Atlas probes, on the delay values. Using the probe hardware version filtered ping measurements, we analyze the delay trends. Table I summarizes these results. Recall that the collection of these measurements excludes probes with hardware versions system-v1 and system-v2. As expected, we find that the delay has a decreasing trend for a significant percentage of the city pairs. Moreover, considering only stable probes for at least one, two, and three years yields the same conclusion. Specifically, we find that the delay has a decreasing trend for approximately 73%, 86%, and 86% of city pairs when we consider stable probes for at least three, two, and one year, respectively. The percentage of city pairs with no data to process climbs to 46% when the stability period is extended to four years, which lowers the percentage of pairs with delay-decreasing trends to 36%. Our analysis shows that a significant number of locations worldwide experience an improvement in performance over time regardless of the local and temporal effects.



(a) Berlin → Singapore has decreasing delay trend

(b) São Paulo → Amsterdam has decreasing delay trend

Fig. 4. Comparison of COVID-19 periods over six years: (a) spike exists, but no trend (b) no spike exists, but increasing trend

V. Regional Delay Evolution

Having observed an overall performance improvement, we are interested in analyzing whether this improvement is limited to one geographical area or can be observed across the world. We further focus on how the delay between different regions evolved over time. Specifically, we study and present the city-level and continent-level delay evolution. For the remainder of the paper, the results are from using ping measurement data filtered for specific probe hardware versions to eliminate any bias from the local effect. Note that we include all such probes regardless of stability period as this covers all the initial data, i.e., probes and city pairs.

A. City level delay evolution

We study the delay evolution between different cities worldwide and compute the trend per city pair. Thus, for each pair, we extract the delay trend results and the rate of the increase/decrease and show these results in the left part of Fig. 3. Note that we include in Appendix A a similar figure that highlights only the trend results per city pair. The dark red/blue indicates a high increase/decrease in a delay, while white corresponds to city pairs for which the trend is most likely stable over time. The gradual shift from blue/red to white reduces the decreasing/increasing delay rate.

The results show that Tokyo to Sydney and Amsterdam to Singapore have the most increasing and decreasing delay change amount over time, respectively, and Prague to Prague is the most stable. The delay increases from around 110-120 ms to 140-150 ms from Tokyo to Sydney, decreasing from around 300-340 ms to 160-180 ms from Amsterdam to Singapore. The city pairs sourcing from and destined to Singapore have the highest delay decreasing amounts. Lisbon is a destination city that has the highest delay in increasing quantities. Turning our attention to intra-city level, the city pairs have mostly a stable delay over a long period.

B. Continent level delay evolution

Our next step is to analyze performance within and between different regions. To this end, we group the cities and generate the delay trends per region. Specifically, we group nine cities in Europe and three in North America. Note that we exclude Africa, Australia, South America, and Asia from this analysis as our data comprised only one or two cities within these regions.

The delay has a decreasing trend for approximately 80% and 55% of city pairs within Europe and North America, respectively. However, within the same two continents, we also find city pairs for which the delay increases over time. Approximately 12% and 11% of city pairs for intra-Europe and intra-North America have an increasing delay trend, respectively. The remaining 8% and 33% of city pairs have no trend for delay evolution within Europe and North America.

When examining the intercontinental trend, the city pairs from Europe to North America and vice versa have a decreasing delay trend of approximately 93% and 96%, respectively. At the same time, we find that the remaining part of city pairs from North America towards Europe has an increasing delay trend. We also find no delay trend between the remaining city pairs from Europe to North America.

These findings suggest that performance improvements are not limited to one region but are a global phenomenon likely resulting from many advancements, such as public and private infrastructure investments made worldwide. Our subsequent study will investigate the causes of delay changes, taking into account more cities from every possible region of the world.

VI. COVID-19 Impact on Delay

Our proposed method infers long-term changes in the delay between different locations. However, we are further interested in whether we can detect relatively short-term changes caused by external phenomena. Moreover, we investigate whether such phenomena have a long-term effect on the delay.

At the beginning of 2020, the coronavirus (COVID-19) spread worldwide, forcing strict lockdowns in many locations [31]. Several studies have shown that these periods clearly impacted Internet performance [8], [32]–[34]. Thus, we consider the first lockdown period our external phenomenon and analyze the delayed evolution between our selected cities. Since we are interested in capturing the delay trend changes, we consider fourteen months (from September 2019 to October 2020) as our measurement period.

Fig. 5 shows an example of the delay trend values between New York and Tokyo during our chosen measurement

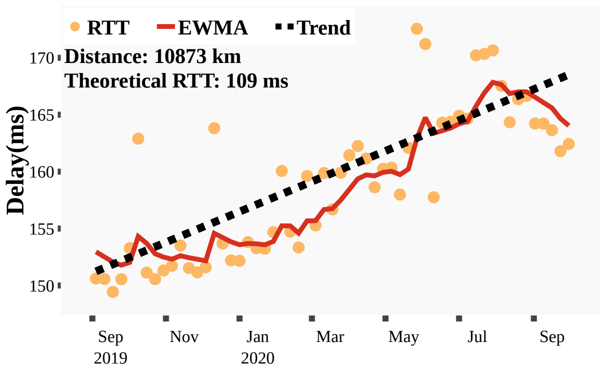


Fig. 5. New York → Tokyo delay evolution during COVID-19 period

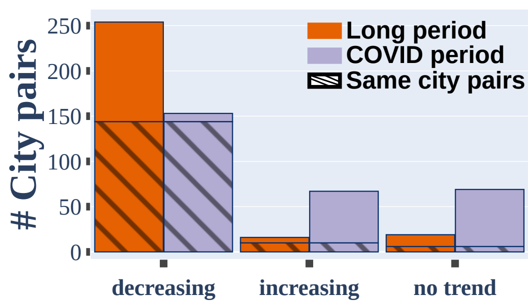


Fig. 6. Trend counts comparison

period. The delays, as expected, remain roughly the same value (150 ms) during the first months. However, we observe an apparent increase of 20 ms between February and July 2020 and a slight decline during the last few months. In this case, the delay values do not return to their initial value. As a result, our analysis concludes that there is an increasing trend during the COVID-19 period. The majority of this period corresponds to the lockdown period and most likely captures the effect on performance.

Fig. 3 details the delay trends per city pair. Recall that the left side of the figure showed the corresponding trend slope results for the six-year periods. Similarly, we show the same results filtered for our chosen COVID-19 period on the right side of the figure. We observe a clear difference when comparing the delay trend results during the two periods. During the pandemic, we find that the delay increases between 23% (67) of the pairs, while the delay values decrease for 53% (153) of the pairs. Our analysis does not infer any trend for the remaining 24% (69) city pairs. Fig. 6 shows the summary of this comparison. The striped area highlights the number of city pairs with the same trend classification during both periods. During the pandemic period, we notice a significant decrease in the number of cities that experience performance improvement. However, for another set of city pairs, we do not infer any trend hinting that the delay value most likely follows a diurnal pattern during this period.

To further quantify the scale of the lockdown impact on delay, we compute the increase in delay during this period as follows. First, we calculate the minimum and maximum delay values collected during our COVID-19 measurement periods.

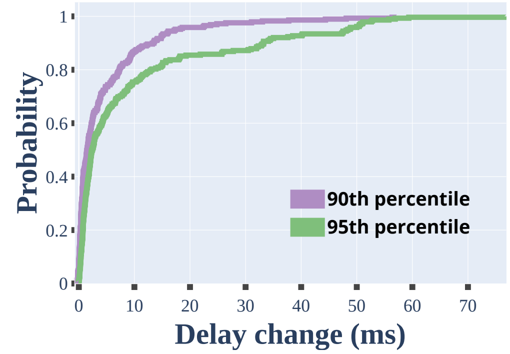


Fig. 7. COVID-19 period delay change amounts comparison

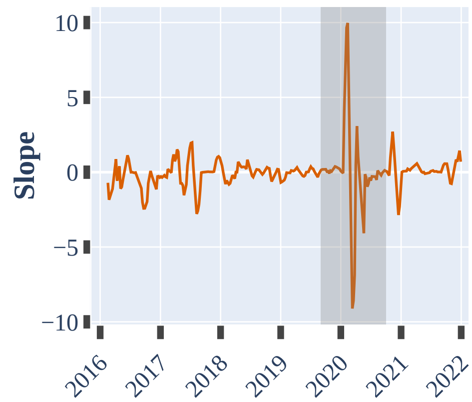


Fig. 8. Cape Town to London derivative changes of delay evolution

We consider the minimum delay value equal to the mean delay value over the first three months, i.e., from September to November 2019. As a maximum delay, we consider the 90$^{\text{th}}$ and 95$^{\text{th}}$ percentile values of all the collected delay values. Next, we compute the difference between the minimum and the maximum values. Fig. 7 shows the distribution of these differences. Our results show that 87% and 75% of city pairs have approximately less than 10 milliseconds of delay change for the 90$^{\text{th}}$ and 95$^{\text{th}}$ percentile threshold, respectively. At the same time, we find city pairs for which the delay changes by more than 50 ms.

The study also investigates the regional impact of COVID-19 on delays considering only Europe and North America as in Section V-B. We extract the percentage values of the trend test results for city pairs between continents. Trend test shows approximately 26%, and 11% of city pairs have an increasing trend at the intra-level of Europe and North America, respectively. Decreasing trend percentages of city pairs are 60% and 33% within Europe and North America, respectively. The trend test finds a no-trend result for the remaining city pairs on the delay evolution. There are no city pairs with an increasing trend from North America to Europe at the inter-level. Still, approximately 15% of city pairs have an increasing delay trend from Europe toward North America. We find that the delay shows a decreasing trend for 67% and 63% of city pairs from North America toward Europe and vice versa, respectively. There is no trend in the delay for the remaining 33% of city pairs from North America to Europe and 22% of the pairs in the opposite direction.

Having seen that the pandemic has an increasing effect on the delay, we are further interested in whether the pandemic has a long-term impact. Fig. 4 plots the delay evolution for two pairs of cities, i.e., from Berlin to Singapore and from São Paulo to Amsterdam. We highlight in the gray area our chosen COVID-19 measurement period and observe a clear delay increase of different scales during this period. Specifically, we notice a jump of approximately 70 ms in delay between Berlin and Singapore. In contrast, the impact on the delay between São Paulo and Amsterdam is less than 5 ms. In both cases, however, we find these increases to correspond only to the pandemic period, as the long-term delay trends appear not to be affected. Moreover, the comparison from Fig. 3 confirms that this observation is valid across most of our studied pairs. On the other hand, even when a pair of cities shows both a small overall change and similar behavior globally and in the COVID-19 period, this does not preclude strong temporary change. To this end, we compute derivative of delay changes (slope) for the traffic between our selected set of cities. This metric is useful to identify the time where significant changes occur in the series. We provide one example of values for this metric for the traffic from Cape Town to London in Fig. 8. We note a clear signal during the pandemic period and plan to explore in the future how we can leverage this metric for event detection.

Our analysis shows that the pandemic periods impacted Internet performance. Similar results were reported by other studies that focused on COVID-19 short-term effects [7], [39]. Surprisingly, our results do not infer any long-term effect of the pandemic on the delay. Thus, when applied over shorter periods of time, our proposed trend approach appears to infer the impact of external phenomena correctly.

VII. Conclusion

We have provided a comprehensive analysis to reveal the historical and longitudinal Internet delay variations between 17 cities every Wednesday over the six years from 2016 until 2022. In the following, we discuss our findings and their implications and limitations.

In the analysis results, the delay for some city pairs experiences sudden increases or decreases for a period of time. However, a gradual decrease, increase, or stable historical delay evolution is observable. The findings of this study clearly show that there is a decreasing trend in the majority for both the regions of city level and continent level. The delay trend results were also in the same direction when we analyzed the delay evolution excluding ping measurements with the early RIPE Atlas probes with hardware versions of system-v1 and system-v2 to prevent interference. Furthermore, we obtain similar results when we impose different stability periods for the vantage points. However, we acknowledge that for a four-year stability period, there is a significant decrease in the number of available probes. In our next study, we plan to extend our understanding of the local impact on the delay. Specifically, we plan to analyze other features related to the probe.

By analyzing the delay change amounts, we find that the most decreased amount is much more changed than the most increased amount. The findings show that the majority of the delay change amounts are toward decreasing or stable over the long period. Although the delay decreasing amount is not limited to specific regions, long-haul connections from/to Singapore still have the most reducing values, while inter-city connections show stability.

Our findings clearly show that the COVID-19 pandemic increasingly impacts the delay in trend results and change amounts. We observe a shift from the decreasing and no-trend delay trends during our six-year period to increasing delay trends during the COVID-19 measurement period. The delay increase amount exceeds 70 ms for some city pairs during lockdowns due to the COVID-19 pandemic. However, we observe that for most of the city pairs the pandemic does not affect the delay trend over the long period.

This paper presents several key findings about the delayed evolution and COVID-19 impact but with limitations on city and date coverages. Even though we rely on samples taken only one day per week, we can still demonstrate how the COVID-19 pandemic has caused the delay to increase. There may be more than one factor to consider when identifying stable probes over time. Moreover, this study was limited to only one city for some continents.

Our study comes with some limitations. We acknowledge that in our study we focus on a subset of cities, i.e., 17 large cities. In the future, we plan to extend our analysis to include other cities. We are well aware that it only partially represents the continent's delay, and we want to cover the continents more extensively in the future as part of our study goal. This study uses only ping measurements to study the long-haul connection between cities. We also investigated the usability of different probe features like probe tags, time-to-live (TTL), and stability to further understand the delay evolution. However, most of our results were inconclusive. Thus, our future steps include incorporating in our analysis paths information from RIPE Atlas traceroute data [42]. Moreover, we plan to develop an event detection algorithm that identifies the significant events similar that have an impact similar to pandemic period. As part of our future agenda we will study both long-haul and intra-city connections.

In this paper, we assess the evolution of intercity Internet latencies due to numerous investments. Our analysis relies on end-to-end measurements collected between probes at major metropolitan areas around the globe. The findings reveal that the delay shows a decreasing trend for almost all city pairs in general over the long period. Our study also assesses the impact of global or local events such as calendar cycles and a once-in-a-generation global event: the COVID-19 outbreak. We acknowledge that the observed delay evolution can be a result of many aspects/factors of network operations and practices. In future work, we intend to expand on this study by examining the effects of other possible factors, such as infrastructure upgrades, congestion along the path, packet loss rate, or changes in the routing policies on the evolution of intercity delays. We also acknowledge that delay variations may due to social and political events, military conflicts, and financial sanctions. Moreover, we plan to include more cities and expand our analysis to include detailed geographical analyzes and connections between end-user and data centers.

Acknowledgment

The research presented in this paper has benefited from the Experimental Infrastructure for Exploration of Exascale Computing (eX3), which is financially supported by the Research Council of Norway under contract 270053. Ioana Livadariu is internally funded by SimulaMet. Georgios Smaragdakis was supported by the European Research Council (ERC) Starting Grant ResolutioNet (ERC-StG-679158).

References

[1] The Guardian, "Pokémon Go fans enraged as first festival ends in connectivity disaster," Jul. 2017, visited on March 2023. [Online]. Available: https://www.theguardian.com/technology/2017/jul/24/pokemon-go-fest-fans-enraged-connectivity-disaster

[2] T. Leighton, "Improving Performance on the Internet," *Communications of the ACM*, vol. 52, no. 2, p. 44–51, 2009.

[3] T. Flach, N. Dukkipati, A. Terzis, B. Raghavan, N. Cardwell, Y. Cheng, A. Jain, S. Hao, E. Katz-Bassett, and R. Govindan, "Reducing Web Latency: the Virtue of Gentle Aggression," in *In Proceedings of SIGCOMM*, 2013.

[4] O. Victor Babasanmi and J. Chavula, "Measuring cloud latency in africa," in *CloudNet*, 2022.

[5] R. Noordally, X. Nicolay, P. Anelli, R. Lorion, and P. U. Tournoux, "Analysis of Internet Latency: The Reunion Island Case," in *AINTEC*, 2016.

[6] P. Sevcik, J. Alan, and R. Wetzel, "2020 Internet Latency Benchmark Report," NetForecast, techreport, Apr. 2021.

[7] M. Candela, V. Luconi, and A. Vecchio, "Impact of the COVID-19 pandemic on the internet latency: A large-scale study," *Computer Networks*, vol. 182, p. 107495, dec 2020.

[8] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis, "The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic," in *IMC*, 2020.

[9] L. Davisson, J. Jakovleski, N. Ngo, C. T. M. Pham, and J. Sommers, "Reassessing the Constancy of End-to-End Internet Latency," in *TMA*, 2021.

[10] E. Aben, "Measuring More Internet with RIPE Atlas," 2016. [Online]. Available: https://labs.ripe.net/author/emileaben/measuring-more-internet-with-ripe-atlas/

[11] C. Williamson, "Internet Traffic Measurement," *IEEE Internet Computing*, vol. 5, no. 6, pp. 70–74, 2001.

[12] F. Mazzola, P. Marcos, I. Castro, M. Luckie, and M. Barcellos, "On the Latency Impact of Remote Peering," in *PAM*, 2022.

[13] Internet Health Report, visited on March 2023. [Online]. Available: https://ihr.iijlab.net/ihr/en-us

[14] M. Chowdhury, R. Agarwal, V. Sekar, and I. Stoica, "A Longitudinal and Cross-Dataset Study of Internet Latency and Path Stability," 2014.

[15] B. Chandrasekaran, G. Smaragdakis, A. Berger, M. Luckie, and K.-C. Ng, "A Server-to-Server View of the Internet," in *Proceedings of ACM CoNEXT 2015*, 2015.

[16] Center for Applied Internet Data Analysis (CAIDA), "Archipelago (Ark) Measurement Infrastructure," https://www.caida.org/projects/ark/, [Online; accessed 1-June-2023].

[17] H. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane: An information plane for distributed services." OSDI 06, Nov. 2006.

[18] W. Willinger and M. Roughan, "Internet Topology Research Redux," *Recent Advances in Networking ACM SIGCOMM eBook*, Aug. 2013.

[19] J. Liu, F. Liu, and N. Ansari, "Monitoring and analyzing big traffic data of a large-scale cellular network with Hadoop," *IEEE Network*, 2014.

[20] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, "InterTubes: A Study of the US Long-Haul Fiber-Optic Infrastructure," *SIGCOMM*, 2015.

[21] V. Bajpai, S. J. Eravuchira, and J. Schönwälder, "Lessons Learned From Using the RIPE Atlas Platform for Measurement Research," *SIGCOMM*, 2015.

[22] T. Holterbach, C. Pelsser, R. Bush, and L. Vanbever, "Quantifying Interference between Measurements on the RIPE Atlas Platform," in *IMC*, 2015.

[23] P. Gigis, V. Kotronis, E. Aben, S. D. Strowes, and X. Dimitropoulos, "Characterizing User-to-User Connectivity with RIPE Atlas," in *ANRW'17*.

[24] Y. Li and L. Iannone, "Performance evaluation of locator/identifier separation protocol through ripe atlas," in *SIGCOMM*, 2016.

[25] D. Dönni, G. S. Machado, C. Tsiaras, and B. Stiller, "Schengen Routing: A Compliance Analysis," in *IFIP AIMS*, 2015.

[26] R. Fontugne, E. Aben, C. Pelsser, and R. Bush, "Pinpointing Delay and Forwarding Anomalies Using Large-Scale Traceroute Measurements," 2017.

[27] R. Fanou, P. Francois, and E. Aben, "On the Diversity of Interdomain Routing in Africa," in *PAM*, 2015.

[28] P. Fiadino, M. Schiavone, and P. Casas, "Vivisecting WhatsApp in Cellular Networks: Servers, Flows, and Quality of Experience," in *Traffic Monitoring and Analysis*, 2015.

[29] D. Cicalese, D. Joumblatt, D. Rossi, M. Buob, J. Augé, and T. Friedman, "A fistful of pings: Accurate and lightweight anycast enumeration and geolocation," in *INFOCOM*, Apr. 2015.

[30] DW, "How covid-19 is affecting internet performance," 2020, visit: 03.2023. [Online]. Available: https://p.dw.com/p/3Zz2f

[31] Wikipedia, "Covid-19 lockdowns," visited on March 2023. [Online]. Available: https://en.wikipedia.org/wiki/COVID-19_lockdowns

[32] C. Labovitz, "Pandemic Impact on Global Internet Traffic," 2020.

[33] T. Böttger, G. Ibrahim, and B. Vallis, "How the internet reacted to covid-19: A perspective from facebook's edge network," ser. IMC '20, 2020.

[34] A. Feldmann, O. Gasser, F. Lichtblau, E. Pujol, I. Poese, C. Dietzel, D. Wagner, M. Wichtlhuber, J. Tapiador, N. Vallina-Rodriguez, O. Hohlfeld, and G. Smaragdakis, "A Year in Lockdown: How the Waves of COVID-19 Impact Internet Traffic," *Commun. ACM*, July 2021.

[35] A. Bergman and J. Iyengar, "Coronavirus: What do the lockdowns mean in europe?" Apr. 2020, visited on March 2023. [Online]. Available: https://www.fastly.com/blog/how-covid-19-is-affecting-internet-performance

[36] D. Howdle and M. Ashton, "How global broadband speeds changed during covid-19 lockdown periods," 2020, visited on March 2023. [Online]. Available: https://www.cable.co.uk/broadband/speed/broadband-speeds-covid-19-lockdown/

[37] U. Hölzle, "Keeping our network infrastructure strong amid COVID-19," 2020, visited on March 2023. [Online]. Available: https://www.blog.google/inside-google/infrastructure/keeping-our-network-infrastructure-strong-amid-covid-19/

[38] A. Medina, "State of Internet Health During COVID-19," Mar. 2020, visited on March 2023. [Online]. Available: https://www.thousandeyes.com/blog/internet-health-during-covid-19

[39] A. Elmokashfi, A. Arouna, I. Livadariu, M.-R. Fida, A. Kvalbein, A. Al-Selwi, T. Dreibholz, and H. Bryhni, "A multi-perspective study of internet performance during the covid-19 outbreak," 2021.

[40] S. Liu, P. Schmitt, F. Bronzino, and N. Feamster, "Characterizing Service Provider Response to the COVID-19 Pandemic in the United States," in *PAM*, 2021.

[41] M. Rajiullah, A. S. Khatouni, C. Midoglu, O. Alay, A. Brunstrom, and C. Griwodz, "Mobile Network Performance during the COVID-19 Outbreak from a Testbed Perspective," in *WiNTECH*, 2020.

[42] RIPE Atlas REST API, visited on March 2023. [Online]. Available: https://atlas.ripe.net/api/v2/

[43] GeoNames, "Large cities," https://www.geonames.org/countries/, 2021, [Online; accessed 1-June-2021].

[44] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica, "Spark: Cluster Computing with Working Sets," ser. HotCloud, 2010.

[45] D. Merkel, "Docker: lightweight linux containers for consistent development and deployment," *Linux journal*, 2014.

[46] M. Kendall, *Rank correlation measures*. Charles Griffin, London, 1975.

[47] H. B. Mann, "Nonparametric tests against trend," *Econometrica*, 1945.

[48] M. Hussain and I. Mahmud, "pyMannKendall: a python package for non parametric Mann Kendall family of trend tests." 7 2019.

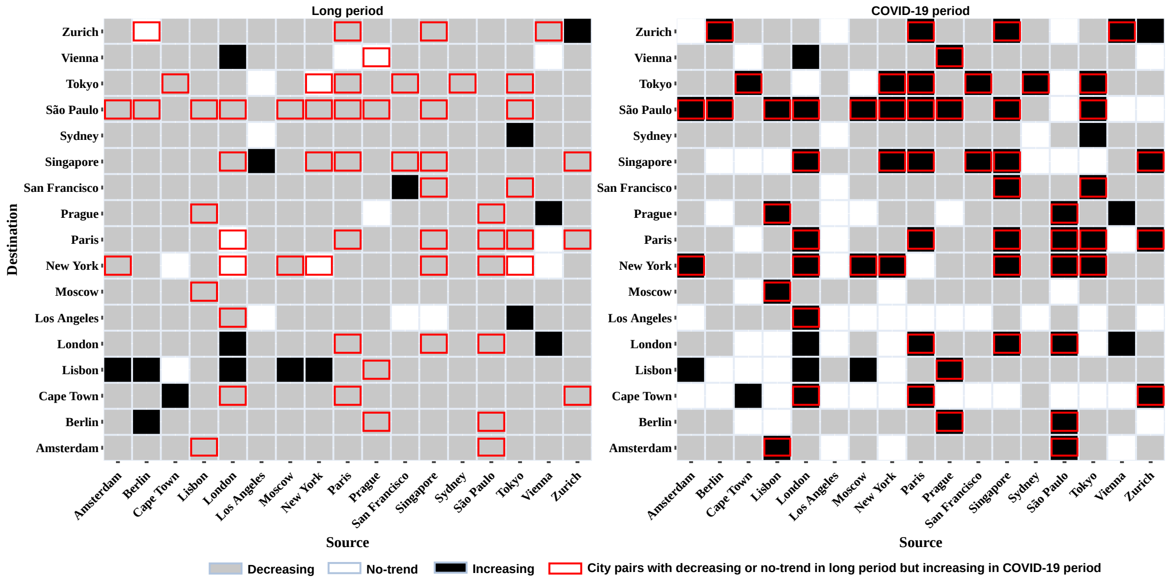


Fig. 9. City level delay trend results for 6 years long and COVID-19 measurement periods

Appendix A
City Level Delay Trend Results

We analyze and compare the delay trend test results on the city level for both long and COVID-19 measurement periods. The left side of the Fig. 9 shows the delay trend results for a long period from source to destination cities in a heatmap. Refer to Table I for the overall counts of trend results. The black color indicates an increasing trend, white is for a no-trend, and gray shows the decreasing trend result of city pairs. The majority of city pairs have a decreasing trend over the six years, considering the figure. However, this is different for some city pairs, which continue to have an increasing delay trend at both intra- and inter-levels. Intra-level means source and destination cities are the same, and inter-level means different source and destination cities. Cities of Berlin, Cape Town, London, San Francisco, and Zurich have an increasing delay trend at the intra-level. New York to Lisbon, Tokyo to Los Angeles, Tokyo to Sydney, and Los Angeles to Singapore are the long-distance inter-level city pairs with an increasing trend result. Furthermore, we can infer more by treating cities as sources or destinations. Lisbon, for example, as a destination city, has the most increasing delay trend with a source city both within Europe and from North America. However, when viewed as a source city, Lisbon exhibits decreasing delay trends to all cities, including intra-level. We find that the delay has a decreasing trend from Amsterdam towards 16 cities and from 17 cities towards the same location. Moscow as a destination city and Sydney as a source city have exclusively decreasing trend results. São Paulo has all decreasing trend results as a bidirectional city. The red boxes in the Fig. 9 represent the 57 city pairs with a long-term trend of ‘decreasing’ or ‘no-trend’ and an ‘increasing’ delay trend during the COVID-19 era. The long-period results in the Fig. 9 indicate a decreasing trend in the delay for the majority of city pairs. Using the city-level view, this finding indicates that the performance improvement is not limited to one region but rather a global feature.

We also offer trend results for each city pair to demonstrate the effect of the pandemic on delay in greater depth and to compare the results from the extended period. The right side of the Fig. 9 shows the delay trend results for the COVID period from source to destination cities in a heatmap. The exceeded number of city pairs with an increasing trend or no-trend results in the figure can infer the effect of the pandemic on the delay. There is no increasing delay trend when we take Los Angeles as a source city. São Paulo is one of the cities with the most increasing trends when considering both source and destination. The inter-city delay trends are also affected by the pandemic. In view of this, many city pairs have changed trends from decreasing trends. As destination and source cities, respectively, Moscow and Sydney having all decreasing trends in the long period, are also affected by the pandemic and experiencing changed trends. The performance-decreasing effect of the pandemic indicates that the effect is not restricted to one region or any distance with a pattern.