WebRTC-based measurement tool for peer-to-peer applications and preliminary findings with real users

Kosuke Nakagawa
nakagawa-kosuke973@g.ecc.u-tokyo.ac.jp
The University of Tokyo
Bunkyo, Tokyo, Japan

Keiichi Shima
keiichi@iijlab.net
Internet Initiative Japan
Chiyoda, Tokyo, Japan

Manabu Tsukada
mtsukada@g.ecc.u-tokyo.ac.jp
The University of Tokyo
Bunkyo, Tokyo, Japan

Hiroshi Esaki
hiroshi@wide.ad.jp
The University of Tokyo
Bunkyo, Tokyo, Japan

1 INTRODUCTION

With the new coronavirus pandemic sweeping the world, many people are now working remotely, thus increasing the opportunities for remote collaboration. For example, we currently hold remote face-to-face meetings more frequently using web conferencing applications such as Zoom. However, for communications that require accurate synchronization, such as musical ensembles and interactive gaming, applications that use server-client communication have significant delays, which negatively affect the user experience.

To address this problem, applications that reduce latency by using peer-to-peer (P2P) to connect home networks have emerged. For example, Yamaha’s Syncroom1 is an application that allows up to five people to play music together remotely by connecting their home networks via a P2P mesh topology connection. However, depending on the P2P network performance, this application has yet not been widely adopted because it can cause loss of synchronization and voice disruption. It is essential to measure the actual performance of the network between homes to make the application widely used.

However, few previous studies have measured the performance of inter-home networks, and no standard measurement method has been established thus far. This study proposes a method to measure the P2P performance between home environments using Web Real-Time Communication (WebRTC) that only requires a personal computer (PC) and a web browser. WebRTC is an open-standard technology that provides real-time communication to browsers and mobile applications through a simple application programming interface (API).

This study aims to develop a tool to measure the performance between home networks using WebRTC without using dedicated hardware or software, and reports the results of such measurement in the case of real users.

The rest of this paper is organized as follows. Section 2 provides an overview of related works. Section 3 analyses requirements for the proposed method. Section 4 provides the detailed description of proposed methods and implementations. Section 5 validates our implementation. Section 6 describes the measurement experiment of the proposed methods. Section 7 discusses the results. Section 8 concludes this paper.

---

1 https://syncroom.yamaha.com/
2 RELATED WORKS

The overview of related works is shown in Table 1. In this research, we aim to measure the P2P QoS of the Internet on a web browser by using WebRTC.

There are two types of approaches for the measurement of home network environments; one relies on dedicated hardware and the other relies on dedicated software. As for those that rely on dedicated hardware, Agarwal et al. proposed a P2P latency prediction system [4] based on geographic information to achieve low latency matching in Halo 3, a P2P game available for Xbox 360. Youngki et al. measured the P2P communication quality for 5.6 million IP addresses playing Halo 3 on Xbox 360 and determined the round trip time (RTT) and throughput for 120 million probes in [13]. The authors also used the MaxMind GeoIP City Database to obtain the geographic information of the acquired IP addresses. RIPE Atlas [5] is a system that obtains and visualizes the latency and network outage of Internet connections at various parts of the world by deploying more than 11000 monitoring nodes worldwide. Fontugne et al. proposed a method to detect network failure points in a wide area by analyzing the measurement results by the traceroute command collected by RIPE Atlas in [7]. Sundaresan et al. used the measurement infrastructure of the US Federal Communications Administration to obtain TCP dump information on OpenWrt (Linux distribution) routers deployed in 2652 home networks in [18]. The authors created a classifier to separate upstream network congestion from home network congestion based on packet arrival and RTT variation patterns.

As for measurements that depend on dedicated software, SIN-DAN [1] aims to establish a method to accurately determine the network status observed from the actual environment of a user to solve claims that are often ambiguous, such as “my network is broken,” so that operators can find the root cause of a network issue. Home Area Latency Measurement2 aimed to provide a service that estimates the latency of a home network by calculating the difference between the RTT from a home node to a measurement server and the measurement server to the exit point of the home network. Fontugne et al. revealed congestion and communication bottlenecks in the last mile by analyzing the traceroute data from RIPE Atlas in [8].

There are also web-browser-based measurement approaches. The iNonius project3 aims to provide a more accurate and rapid Internet measurement environment by operating an independent Internet measurement site for each organization that demands such services as telework. WebDINO VideoMark [2] is a browser extension that measures the quality of experience (QoE) of video delivery services, such as YouTube, and is used to improve the service quality of telecommunications and video delivery service providers.

In recent years, P2P media streaming using WebRTC has become increasingly popular, and there are many investigations on quality of service (QoS) and QoE of WebRTC applications. Moulay et al. obtained WebRTC performance from mobile nodes on MONROE, a mobile measurement platform in the European Union and evaluated the QoS and QoE in [15]. Garcia et al. proposed a framework that provides video quality and end-to-end latency measurement capabilities for WebRTC-based real-time application development in [11]. In [9], Garcia et al. proposed a tool to measure the QoS and QoE of WebRTC-based applications and evaluates the tool on Kurento [10], an open-source media server compliant with the WebRTC standard. Barik et al. verified that the QoS requirements of WebRTC set by applications using DiffServ Code Point work as expected in [5]. Flohr et al. proposed a method for minimizing latency by resolving the inconsistency between the delay minimization function of real-time transport protocol and the throughput maximization function of the Stream Control Transmission Protocol (SCTP) when streaming over WebRTC in [6]. Taheri et al. proposed a benchmarking method to measure the connection overhead and response latency of a WebRTC protocol stack implementation itself and compares the measured values against Google Chrome and Firefox in [19]. Tanskanen proposed a tool to explore the latency factors of WebRTC-based remote control systems in [20], implying that there is a great need for WebRTC quality measurement even in use cases of remote control.

Research has also been conducted on WebRTC-based network measurement. In [14], McClellan proposed a tool to measure the network in a LAN using WebRTC. The above-mentioned studies focus on measuring the on-premises network environment, QoE of WebRTC applications, and WebRTC-based networks in LANs. However, there has been no research on P2P RTT measurement between home environments using WebRTC.

3 REQUIREMENTS

The functional requirements of the method proposed in this study are as follows.

Ease of measurement

The potential solution must be able to measure network performance without installing dedicated hardware or software. Measurements using RIPE Atlas and Xbox 360 depend on dedicated hardware, making it difficult to perform measurements in various environments. Although the requirements for measurements using dedicated software are less strict than those using dedicated hardware, there is still a certain amount of difficulty owing to the installation process required. Without the need for dedicated hardware or software, the solution can solve the problems mentioned above, and anyone can easily participate in the measurement.

Comprehensive analysis of P2P network

It is necessary to make measurements in an environment where a mesh topology network connects multiple nodes to make measurements close to the actual usage environment. Real applications such as Yamaha’s Syncroom enable multi-person collaboration by connecting multiple home environments through a mesh topology network. In addition, since a typical home environment uses a network address translation (NAT) box to connect to the Internet, it is necessary to support measurement in a NAT environment.

Intuitive visualization

In this study, we aim to create a visualization that would allow us to intuitively grasp the trend of the RTT of P2P

---

1https://iham.ijlab.net/latency/
2https://inonius.net/
3https://iNonius.net/
connections throughout the mesh topology networks. It is difficult to evaluate the statistics of a multi-node full-mesh topology connection intuitively in a raw data form. Data visualization is essential to achieve the original purpose of helping to solve P2P communication problems.

**Constrained node support**

This research aims to create a lightweight measurement tool that can be used in as many different environments as possible and run on a resource-constrained PC. By creating a tool that can run under the limitations of an old laptop computer, we can increase its accessibility and deploy our tool in more diverse home environments. Also, by creating a tool that can run on a single-board computer, such as a Raspberry Pi, it will be possible to use clients that are always connected and participating in the measurement. These nodes will increase the number of nodes participating in the measurement and keep the scale of the measurement at a certain level.

**Granularity of measurement**

In this research, we aim to measure an RTT of less than 60 ms. The goal of this study is to evaluate the performance of P2P connections for applications such as ensembles. Since [17] states that the network latency that enables an ensemble to perform without problems is approximately 30 ms, we need to adopt a measurement method that can detect the fatal case of approximately 60 ms of RTT (30 ms one-way latency) to achieve the above goal.

By fulfilling these functional requirements, we measure RTT in more diverse P2P mesh topology connection environments on a user-centered basis without requiring dedicated hardware or software.

### 4 PROPOSED METHOD

#### 4.1 Overview

This study proposes a method to measure the RTT of P2P connections between home environments using WebRTC. WebRTC supports P2P communication over NAT, and we can obtain statistics such as RTT by using the `getStats()` WebRTC API method. We can use the API with a browser such as Google Chrome, which is already deployed on many PCs. Therefore, our WebRTC-based measurement tool fulfills the requirements described in the **ease of measurement** and **comprehensive analysis of P2P networks**.

#### 4.2 Design

Fig. 1 depicts the design architecture of this study. A user in each home network accesses the measurement website using Google Chrome, and the JavaScript measurement program runs on the browser. The operation of the measurement program can be divided into the following three stages:

- **Signaling**
- **P2P Connection**
- **Data Collection and Visualization**

![Figure 1: System Design](image-url)
The signaling stage begins with the exchange of the Session Description Protocol (SDP) between clients. SDP is required to connect a local client and other clients connected to the measurement server over NAT boxes. After the signaling stage is complete, the P2P connection stage is performed, which initiates the P2P connection over NAT using WebRTC and sends data at the pre-configured size and interval. While the P2P connection stage is in progress, the data collection and visualization stage works in parallel. In the data collection and visualization stage, WebRTC regularly acquires the RTT of the P2P connection. The retrieved RTT is visualized in real-time in the front-end web browser. The data in the database can be visualized using a heatmap after the measurement. In this way, we aim to achieve intuitive visualization, which is one of the requirements of this research.

4.3 Implementation

Fig. 2 shows the implementation of the proposal. The measurement program for this method is a JavaScript file executed on the client’s browser. When the client accesses the measurement URL, the signaling stage begins. After the signaling stage is complete, a user is required to click on the “Connect” button shown on the web page to establish P2P connections with other clients over NAT using WebRTC. Once P2P connections are established, the local client starts sending data to other clients with pre-configured data sizes and transmission frequencies. While the P2P connection stage is in progress, the data collection and visualization stage is performed, which obtains the RTT, current UnixTime, and data transmission conditions every second. This statistical information is visualized in real-time on the client browser using Chart.js\(^4\) and sent to the data collection server in JSON format. The data contain the following information.

- **Date** The result of the `Date()` call on the client web browser [ms]
- **your ISP** Client ISP
- **your IP** Client global IP address
- **candidatePair_RTT** RTT of the P2P connection [ms]
- **yourID** An ID to identify the client.
- **peerID** An ID to identify the client.

\(^4\)https://www.chartjs.org/
The server consists of an API server using Flask⁵ and a database using SQLite⁶, and the collected data are registered in the database through the API server. The statistical information in the server database is analyzed and visualized on the server after the measurement process is completed. The statistical information that one client can obtain is limited to the P2P connection information between itself and other clients. In contrast, the server can perform data analysis in a full-mesh topology network by utilizing collected data from every client. For visualization on the server, we use techniques such as heatmaps to make the statistical information of the mesh topology networks intuitive to the user.

When sending statistics to the server, we also send the global IP address, ISP obtained by the ipinfo.io⁷ service at the beginning of the measurement in addition to the RTT of the WebRTC connection. This information is necessary for understanding the differences of RTT values between different IP addresses and ISPs.

5 VALIDATION OF OUR IMPLEMENTATION

In this section, we verify that the RTT of our method is negligibly different from the RTT of Ping, a well-known RTT evaluation method. The communication protocol of WebRTC is specified by RFC 8831 [12] as “SCTP over DTLS over UDP,” which is significantly different from conventional RTT measuring methods, such as the ping command. Ping is an RTT measurement tool that uses the Internet Control Message Protocol (ICMP) [16]. In this section, we compare the RTT measurements of WebRTC with those of the ping command to verify the measurement result of our tool.

We measured the RTT from an Ubuntu 18.04 PC to a Raspberry Pi 4 (2GB memory model), which were directly connected by Ethernet. Chrome and Chromium, respectively, were used as the browsers running WebRTC. Two types of measurements, WebRTC and ping, were performed once per second for 500 measurements. Table 2 presents the results obtained when constant latencies of 0, 10, and 100 ms are enforced with the tc command. Regardless of the order of the latency generated by the tc command, the results of the WebRTC cases are 1–3 ms longer on average.

Fig. 3 shows the boxplots of the measurement results with a constant latency of 30 ms and random jitter within 30 ms using the tc command. We can see 0.3–3 ms additional delay for WebRTC in all the quartiles compared to the ping results when we impose jitter. We assume that these differences are attributed to the processing overhead of WebRTC.

From these results, we can conclude there is no difference between the results of the WebRTC-based measurement and the conventional ping-based measurement except a negligible overhead seen in the WebRTC cases. We argue that the granularity of measurement described in Section 3 is satisfied by the measurement using WebRTC.

6 EXPERIMENT

Using this tool, we conducted a measurement experiment to connect real home networks. We asked the experiment participants to access the measurement site using Google Chrome from their laptops in the LAN environment of their homes and organizations to perform the measurements. The number of measurement nodes used in the Kanto and Chubu regions in Japan was 10. Seven nodes were located in Tokyo, one in Kanagawa, one in Chiba, and one in Gifu, as shown in Fig. 4. During four hours from 18:00 to 22:00 (with participation and exits during this period), each client sent 100 bytes of data to the other clients and measured the RTT every second. All clients sent the measured data to the collecting server. Fig. 5 shows the distribution of the group of RTT values obtained from the entire mesh topology network. It shows that 91.7% of the measurement results are between 0–60 ms, 7.5% are between 60–100 ms, and only approximately 0.78% measurement results are over 100 ms. Most of the measurement results meet a comfortable communication RTT of less than 60 ms, but a few high RTTs may spoil the overall experience.

The heatmaps in Fig. 6 and Fig. 7 illustrate the average and the 99th percentile points of the group of RTT values obtained for each P2P combination, respectively. We adopted the 99% point to visualize the phenomenon where the RTT value increases temporarily. Even if the average latency is maintained at a low value, the QoE of the real-time communication is severely impaired when a sizeable temporary delay occurs. The label numbers in these heatmaps correspond to each measurement node and the geographical relationship in Fig. 4. All the measurement results over 60 ms are indicated in red. The number of such combinations is 12 in Fig. 6 and 26 in Fig. 7. There are three important points we can observe from the averages in Fig. 6. The first is that node 3 is outstandingly bad; the second is that nodes 1, 4, 5, 7, 8, 9, and 10 have low RTT, with less than 20 ms for more than half of the peers; the third is that nodes 2 and 6 are in the middle, with approximately 30 ms RTT for most of the peers, suggesting a bottleneck in the home area.

Comparing the heatmaps in Fig. 6 and Fig. 7 shows that nodes 1 and 2 had low RTT on average, but they had high RTT at the 99th percentile point. This result suggests that, on average, they meet the requirements for real-time applications but that sudden delays
Table 2: Comparison of WebRTC and ping measured round trip times (RTT)

| tc command latency | 0 ms       | 10 ms      | 100 ms     |
|--------------------|------------|------------|------------|
| method             | WebRTC     | Ping       | WebRTC     | Ping       | WebRTC     | Ping       |
| Average RTT [ms]   | 2.5        | 0.284      | 11.686     | 10.329     | 101.806    | 100.343    |

Figure 4: Geographical location of the experiment participants

Figure 5: Distribution of the group of RTT values obtained from the entire mesh topology network

The number of peer connections with a RTT of less than 60 ms was 51 out of 77, indicating that the current Internet environment in Japan is not sufficient to allow many people to perform activities that require high synchronization, such as remote ensembles. Additionally, although the measurement results achieve less than 60 ms of RTT, this is a minimum requirement. In [17], a delay of 10–20 ms one way is more desirable, which is an even stricter requirement. We expect that the environment in which we can perform satisfactory activities is even more limited.

In this study, we proposed and implemented a browser-based P2P measurement tool using WebRTC as a method to measure the RTT of P2P connections between home environments under various network environments. Our tool is easy to use and platform-independent. We also used this tool to obtain and visualize the RTT of P2P connections in a mesh topology with 10 nodes. As a result, we found that the number of peers capable of performing activities are occurring at a high frequency. If we use real-time applications in such a communication environment, the QoE is poor.

7 DISCUSSION

In Fig. 7, node 1, which is located in Tokyo, has more than 60 ms of RTT with nodes 5 and 10, which are also located in Tokyo. In contrast, node 6, located in Gifu, is connected to nodes 2, 4, 5, 7, 8, and 10 in Tokyo with less than 60 ms of RTT. This result suggests that the P2P RTT between home environments is more likely to depend on home environment issues rather than geographic location. This result is understandable considering the fact that the time required to transit 270 km between Tokyo and Gifu is approximately 1 ms at the speed of light. In a country such as Japan, which is small enough not to consider the effects of the speed of light, geographic distance does not significantly affect the RTT between two home networks.

The number of peer connections with a RTT of less than 60 ms was 51 out of 77, indicating that the current Internet environment in Japan is not sufficient to allow many people to perform activities that require high synchronization, such as remote ensembles. Additionally, although the measurement results achieve less than 60 ms of RTT, this is a minimum requirement. In [17], a delay of 10–20 ms one way is more desirable, which is an even stricter requirement. We expect that the environment in which we can perform satisfactory activities is even more limited.

In this study, we aimed to establish a full-mesh topology connection to emulate a realistic P2P service network. However, in reality, we could only collect RTT measurement results from 77 out of 90 peer connections in the full-mesh topology network as we could not establish a P2P connection using WebRTC in some cases.

8 CONCLUSION

In this study, we proposed and implemented a browser-based P2P measurement tool using WebRTC as a method to measure the RTT of P2P connections between home environments under various network environments. Our tool is easy to use and platform-independent. We also used this tool to obtain and visualize the RTT of P2P connections in a mesh topology with 10 nodes. As a result, we found that the number of peers capable of performing activities...
that require high synchronization, such as ensemble music, is very limited.

The broader implications of this research are twofold. The first is an urgent need to develop technology to measure P2P communications’ performance objectively. Remote ensemble applications using P2P connections are already widespread, but most of them operate with unknown network performance. We have presented an idea for a tool that allows end-users to measure the performance quickly. The second is that it is significant to conduct extensive research on P2P performance in real networks. P2P technology allows us to communicate with other terminals in the shortest path without going through a central server. This property will become increasingly important to achieve higher bandwidth and lower latency communications in areas such as edge computing, which has attracted much attention in recent years. During the experiments of this study, we were not able to confirm the measurement in a full mesh, and some peers did not get any results. For some peers, only one direction was successfully measured, and we are currently investigating the cause. Furthermore, studying the effect of geographic distance on RTT with globally distributed peer connections is recommended.

ACKNOWLEDGMENTS
This work was partly supported by JSPS KAKENHI (grant number: 19H04091).

REFERENCES
[1] [n.d.]. SINDAN Project | SINDAN Project. https://www.sindan-net.com/. (Accessed on 08/16/2021).
[2] [n.d.]. Web VideoMark. https://vm.webdino.org/about. (Accessed on 08/16/2021).
[3] [n.d.]. What is RIPE Atlas? | RIPE Atlas. https://atlas.ripe.net/about/. (Accessed on 08/16/2021).
[4] Sharad Agarwal and Jacob R Lorch. 2009. Matchmaking for online games and other latency-sensitive P2P systems. In Proceedings of the ACM SIGCOMM 2009 conference on Data communication (Barcelona, Spain) (SIGCOMM '09). Association

for Computing Machinery, New York, NY, USA, 315–326.
[5] Runa Barik, Michael Welzl, Ahmed Mustafa Elmokashfi, Thomas Dreihholz, and Stein Gjessing. 2018. Can WebRTC QoS work? a DSCP measurement study. In 2018 30th International Teletraffic Congress (ITC 30), Vol. 1. IEEE, 167–175.
[6] Julius Flurb, Ekaterina Vologina, and Erwin P Rathgeb. 2018. FSE-NG for managing real-time media flows and SCTP data channel in WebRTC. (2018). 315–318.
[7] Romain Fontugne, Cristel Pelser, Emile Aben, and Randy Bush. 2017. Pinpointing delay and forwarding anomalies using large-scale traceroute measurements. In Proceedings of the 2017 Internet Measurement Conference (London, United Kingdom) (IMC '17). Association for Computing Machinery, New York, NY, USA, 15–28.
[8] Romain Fontugne, Anant Shah, and Kenjiro Cho. 2020. Persistent Last-mile Congestion: Not so Uncommon. In Proceedings of the ACM Internet Measurement Conference. 420–427.
[9] Boni Garcia, Francisco Gortazar, Luis Lopez-Fernandez, Micael Gallego, and Miguel Paris. 2017. WebRTC Testing: Challenges and Practical Solutions. IEEE Communications Standards Magazine 1, 2 (2017), 36–42.
[10] Boni Garcia, Luis Lopez-Fernandez, Micael Gallego, and Francisco Gortazar. 2017. Kurento: The Swiss Army Knife of WebRTC Media Servers. IEEE Communications Standards Magazine 1, 2 (2017), 44–51.
[11] Boni Garcia, Luis Lopez-Fernandez, Francisco Gortazar, and Micael Gallego. 2016. Analysis of video quality and end-to-end latency in WebRTC. In 2016 IEEE Globecom Workshops (GC Wkshps). IEEE, 1–6.
[12] R. Jesup, S. Loretto, and M. Tüxen. 2021. WebRTC Data Channels. RFC 8831. https://doi.org/10.17487/RFC8831
[13] Youngju Lee, Sharad Agarwal, Chris Butcher, and Jitu Padhye. 2008. Measurement and Estimation of Network QoS Among Peer Xbox 360 Game Players. In Passive and Active Network Measurement. Springer Berlin Heidelberg, 41–50.
[14] Miranda McClellan. 2019. WebRTC based network performance measurements. Ph.D. Dissertation. Massachusetts Institute of Technology.
[15] Mohamed Moulay and Vincenzo Mancuso. 2018. Experimental performance evaluation of WebRTC video services over mobile networks. In IEEE INFOCOM 2018 – IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). ieeexplore.ieee.org, 541–546.
[16] Jon Postel. 1981. Internet Control Message Protocol. RFC 792. https://doi.org/10.17487/RFC0792
[17] Nathan Schuett. 2002. The effects of latency on ensemble music performance. Bachelor Thesis, CCRMA Department of Music, Stanford University (2002).
[18] Srikanth Sundaresan, Nick Feamster, and Renata Teixeira. 2016. Home network or access link? locating last-mile downstream throughput bottlenecks. In International Conference on Passive and Active Network Measurement. Springer, 111–123.
[19] Sajjad Taheri, Laleh Aghababaei Beni, Alexander V Veidenbaum, Alexandru Nicolau, Rosario Cammarota, Jianlin Qiu, Qiang Lu, and Mohammad H Haghighat. 2015. WebRTCbench: a benchmark for performance assessment of webRTC
implementations. In 2015 13th IEEE Symposium on Embedded Systems For Real-time Multimedia (ESTIMedia). 1–7.

[20] Sakari Tanskanen. 2021. Latency contributors in WebRTC-based remote control system. Master’s thesis. Aalto University.