A Production Dataset for Studying Mobile Device Heterogeneity in Hybrid CDN–P2P VoD Streaming

Yi Lu and Soufiane Rouibia
Easybroadcast
Nantes, France
lu.yi, rouibia@easybroadcast.fr

Guillaume Urvoy-Keller and Ramon Aparicio-Pardo
Université Côte d'Azur, CNRS, I3S
Sophia-Antipolis, France
guillaume.urvoy-keller, ramon.aparicio-pardo@univ-cotedazur.fr

***Abstract*—Hybrid CDN–P2P delivery can complement CDN-based VoD streaming with peer upload capacity, but its behavior in mobile environments remains difficult to study due to heterogeneous end-user devices. This paper presents an open production dataset from a mobile hybrid CDN–P2P VoD platform deployed in a North African market. The dataset contains approximately 700,000 viewer sessions, aggregated from 30-second client-side reports, and combines device, network, P2P QoS, playback QoE, and peer-link transfer measurements. We characterize device heterogeneity using a lightweight score based on SoC, RAM, and OS version, and show that higher-capability devices achieve higher P2P upload throughput without contributing proportionally more traffic under the current network- and locality-driven peer selection algorithm. These observations motivate future device-aware peer selection and chunk scheduling.**

***Index Terms*—VoD, Hybrid CDN-P2P, Mobile hardware, Peer selection, Chunk scheduling**

I. Introduction

Mobile devices generate a large and growing share of video streaming traffic [1], yet they form a highly heterogeneous execution environment [2]. At the same time, hybrid CDN–P2P delivery has been deployed in some commercial streaming platforms to complement CDN delivery by exploiting peer upload capacity [3], [4]. In such systems, peer selection is typically driven by network- and content-level information, such as ISP locality and geographic proximity [5]–[7]. However, in mobile environments, differences in SoC capability, memory capacity, operating system version, and battery constraints may also affect peer-assisted delivery.

In this paper, we present a production dataset for studying mobile device heterogeneity in a hybrid CDN–P2P VoD system. The dataset is collected from a North African deployment and contains approximately 700,000 viewer sessions. Each session combines device information, network context, P2P transfer metrics, playback QoE indicators, and directional peer-link measurements.

To characterize device heterogeneity at scale, we build a lightweight operational device score based on SoC, RAM, and OS version. We observe that higher-capability devices achieve substantially higher upload throughput, but their total upload contribution grows more modestly under the current network- and locality-driven peer selection policy, motivating future work on device-aware peer selection and chunk scheduling.

978-3-903176-83-6 ©2026 IFIP

II. Production System

The production system consists of a React Native video player, an embedded client-side P2P library, a signaling tracker for peer discovery, and a data collection pipeline, as shown in Fig. 1. Instead of retrieving every video segment directly from the CDN, the player delegates segment requests to the P2P library. The library attempts to fetch chunks from neighboring peers through WebRTC data channels and falls back to the CDN when peer delivery is unavailable or too slow. The client-side library also periodically reports measurements to our data collection pipeline at fixed 30-second intervals.

Peer discovery is coordinated by a remote tracker. The tracker groups viewers consuming the same content at the same quality level and returns up to 10 candidate peers. In the current production system, peer selection is mainly driven by ISP affinity, geographic proximity, and VoD playback position.

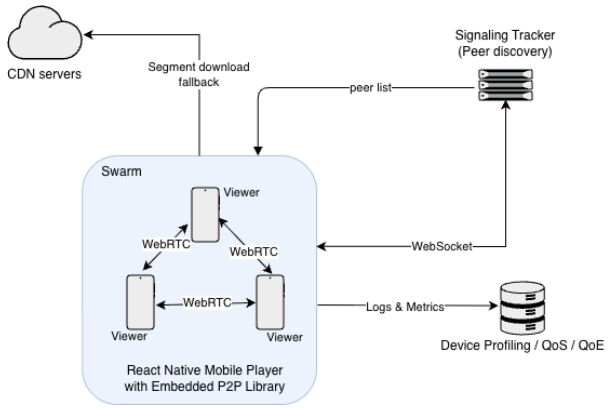


Fig. 1: System overview of the hybrid CDN–P2P mobile VoD platform.

III. Dataset

The raw dataset covers the period from October 2025 to March 2026. Since VoD audience density varies strongly across contents and time periods, we retain measurements only from peak viewing hours, defined as 20:00–01:00 Paris time, and select the most popular content item in each observation window. We further keep only valid sessions lasting between 3 and 60 minutes, excluding very short sessions that do not allow

P2P swarms to stabilize and unusually long sessions likely caused by background playback or abnormal client behavior.

After filtering and aggregation, the final dataset contains approximately 700,000 viewer sessions. Each session-level sample represents an aggregated viewer session and combines four families of measurements: (i) *device information*, (ii) *network context*, (iii) *P2P QoS*, and (iv) *playback QoE*. Table I summarizes the main fields of the dataset.

Device information. It describes the mobile client participating in the swarm, including the device brand, model identifier, RAM capacity, used RAM, OS version, and battery level.

Network context. It captures the external connectivity conditions experienced by the session. It includes the average CDN download throughput, and WebRTC connection counters including successful peer connections, connection timeouts, and ping-pong timeouts. The latter two capture, respectively, failures to establish a WebRTC data channel and disconnections triggered by unhealthy keep-alive responses.

P2P QoS metrics. They describe the efficiency and reliability of peer-assisted delivery. They include the P2P download ratio, P2P upload throughput, P2P download throughput, average peer delay, and Chunk Loss Rate (CLR). The P2P download ratio measures the fraction of downloaded content served by peers rather than by the CDN. Upload and download throughput are computed from the transferred volume and the corresponding transfer duration. Finally, the directional Chunk Loss Rate, for each seeder–leecher pair (A, B), is defined as:

$$CLR_{A\to B} = \frac{S_{A\to B} - R_{B\leftarrow A}}{S_{A\to B}}, \tag{1}$$

where $S_{A\to B}$ is the volume reported as uploaded by the seeder and $R_{B\leftarrow A}$ is the corresponding volume reported as received by the leecher. For each session, the average CLR is obtained by aggregating the directional CLR values over all observed P2P links.

Playback QoE metrics. They capture the user-facing quality of the session. They include the number of rebuffering events, total rebuffering duration, and average buffer length. We also include device-side operational indicators such as memory usage and battery drain.

Dataset availability. An anonymized and aggregated version of the dataset is publicly available on Zenodo [8]. The release includes the session-level dataset, the link-level dataset, the dataset schema, and a README file describing the data source, time range, filtering process, and citation information.

TABLE I: Overview of the session-level dataset schema.

Field	Type	Description
Device		
brand	STRING	Device brand
device_name_raw	STRING	Raw device model identifiers
total_mem_mb	FLOAT	Total available memory
os_full	STRING	Device OS version
avg_battery	FLOAT	Avg. battery level
Network		
cdn_thrpt	FLOAT	Avg. CDN throughput
webrtc_established	INTEGER	Established WebRTC conns.
webrtc_timeout	INTEGER	Failed WebRTC attempts
pingpong_timeout	INTEGER	Health-check disconnects
avg_peers	FLOAT	Avg. connected peers
QoS		
p2p_ratio	FLOAT	Fraction of traffic delivered via P2P
avg_p2p_throughput_Mbps	FLOAT	Avg. P2P download throughput
avg_p2p_upload_throughput_Mbps	FLOAT	Avg. P2P upload throughput
avg_peer_delay_ms	FLOAT	Avg. peer delay with connected peers
avg_chunk_loss_rate	FLOAT	Avg. Chunk loss rate
QoE		
avg_used_memory_mb	FLOAT	Avg. memory usage
battery_drain	FLOAT	Battery consumption
startup_delay_ms	FLOAT	Startup delay
avg_bufferlength	FLOAT	Avg. buffer length
total_rebuffer_time	FLOAT	Total rebuffering time

IV. Device Profiling

To study hardware heterogeneity, we construct a lightweight operational device score based on three dimensions: System-on-Chip (SoC) capability, RAM capacity, and OS version.

A practical challenge is that device identifiers are often ambiguous and may contain internal codenames rather than commercial model names (e.g., *Redmi earth*, *Samsung pineapple*). We therefore maintain a manually curated hardware-reference table that maps frequent device keywords observed in our production traffic to broad SoC performance classes, using public benchmark aggregations such as NanoReview [9] as external guidance.

The device score is computed as:

$$Score = 0.6\times SoC_{\text{Base}} + 0.3\times RAM_{\text{Score}} + 0.1\times OS_{\text{Score}}. \tag{2}$$

The SoC component receives the largest weight, while RAM and OS are included to capture memory constraints and platform recency, and to complement the SoC mapping when device identifiers are imprecise. RAM and OS values are grouped into four discrete score levels (20, 50, 80, and 100), as the resulting score is intended only for coarse device-tier classification in the measurement analysis.

V. Dataset Insights

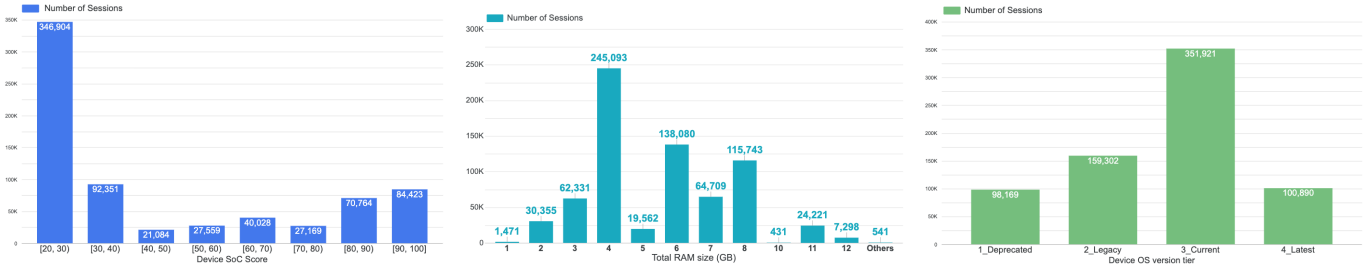
A. Raw Device Characteristics

We first characterize the mobile device population through the three individual dimensions used in our device profiling method: SoC capability, RAM capacity, and OS version tier. Fig. 2 reports their distributions before they are combined into the aggregate device score.

The three dimensions reveal heterogeneous hardware and software profiles in the production traffic. The SoC score distribution is uneven: a large fraction of sessions comes from low-score devices, while high-score devices are also represented. RAM capacity varies across common smartphone configurations such as 4 GB, 6 GB, and 8 GB, with smaller groups below and above these values. The OS tier distribution shows that most sessions run on current OS versions, while legacy and latest versions are also present.

B. From Individual Dimensions to Device Capability Score

We then aggregate the three dimensions into the operational device capability score defined in Section IV. The purpose of this score is simply to group production clients into broad capability tiers that can be compared at scale.



(a) SoC score. (b) RAM size. (c) OS tier.

Fig. 2: Distribution of the three individual dimensions used for device profiling.

Since Samsung, Apple, and Redmi are the three dominant brands in our dataset, we further examine how their sessions are distributed across the resulting device capability tiers. Fig. 3 shows that Samsung and Redmi sessions appear across both low- and higher-capability tiers, while Apple sessions are more concentrated in the upper tiers. This indicates that brand-level concentration does not imply a homogeneous device population: even within a traffic mix dominated by a few vendors, the observed clients cover a wide range of hardware capabilities.

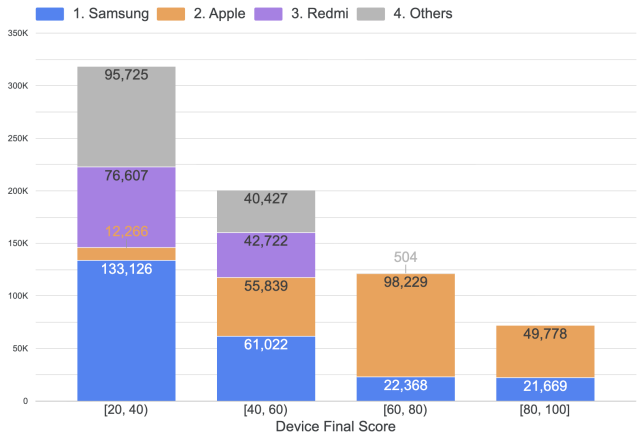


Fig. 3: Distribution of dominant device brands across device capability tiers.

C. Initial Relation with P2P Upload Behavior

Finally, we provide an initial example of how the dataset can be used to relate device profiles to P2P performance. Fig. 4 compares average P2P upload throughput and total P2P upload traffic across device capability tiers.

The average upload throughput increases strongly with device capability, from 10.8 Mbps in the lowest tier to 51.7 Mbps in the highest tier. This suggests that the device score captures a meaningful performance-related dimension for upload behavior. However, the total amount of uploaded traffic grows much more moderately, from 139.4 MB to 247.7 MB across the same tiers. In other words, high-capability devices provide substantially higher upload capacity, but their actual contribution to the P2P overlay does not increase proportionally under the current network- and locality-driven peer selection policy.

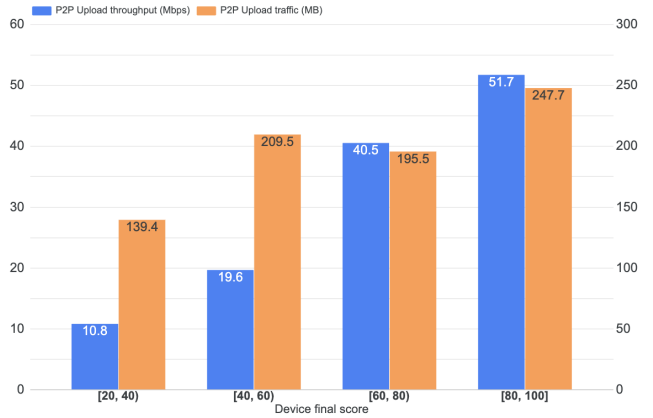


Fig. 4: Average P2P upload throughput and total P2P upload data across device capability tiers.

VI. Conclusion and Research Opportunities

This paper presents a production dataset for studying mobile device heterogeneity in a hybrid CDN–P2P VoD system. The dataset combines device characteristics, network context, P2P transfer metrics, playback QoE indicators, and directional peer-link measurements. Our initial observations show that mobile clients exhibit heterogeneous hardware and software profiles, and that device capability is related to P2P upload behavior.

A natural next step is to study more precisely how individual device-side dimensions affect peer-assisted delivery. Instead of relying only on an aggregate device score, future work can analyze the separate roles of SoC capability, RAM capacity, OS version, battery level, and memory usage, or investigate how these dimensions should be weighted to better predict the P2P upload capacity of a joining device.

These measurements may eventually inform device-aware peer-assisted streaming mechanisms. For instance, device-aware chunk scheduling could route requests toward peers with higher expected upload capacity, while upload-side admission control could help peers decide whether to serve requests based on device score, battery level, or current upload load.

REFERENCES

[1] Ericsson Report, "Mobile network traffic q4 2025," https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-update, 2025.

[2] E. C. Paim *et al.*, "Hidden impact of hardware technologies on throughput: a case study on a brazilian mobile web network," in *WWW*, 2025.

[3] Z. Ma *et al.*, "Neighbor selection strategies in the wild for cdn/v2v webrtc live streaming: Can we learn what a good neighbor is?" *Computer Networks*, 2026.

[4] R. Farahani *et al.*, "Hybrid p2p-cdn architecture for live video streaming: An online learning approach," in *GLOBECOM*, 2022.

[5] H. Xie *et al.*, "P4p: provider portal for applications," *SIGCOMM Comput. Commun. Rev.*, 2008.

[6] M. El Dick *et al.*, "Building a Peer-to-Peer Content Distribution Network with High Performance, Scalability and Robustness," *Information Systems*, 2011.

[7] Z. Ma *et al.*, "When locality is not enough: Boosting peer selection of hybrid cdn-p2p live streaming systems using machine learning," in *IFIP TMA*, 2021.

[8] Y. Lu *et al.*, "Production mobile hybrid cdn–p2p vod dataset for studying mobile device heterogeneity," 2026. [Online]. Available: https://doi.org/10.5281/zenodo.20374907

[9] "Nanoreview smartphone soc rating list," https://nanoreview.net/en/soc-list/rating, accessed: 2026-01-01.