# Abstract

Mobile video streaming constitutes a significant and rapidly growing portion of internet traffic, yet delivering a consistently high Quality of Experience (QoE) remains challenging due to the inherent variability of mobile networks. Latency, in its various manifestations including startup delays, buffering events, and particularly the playback progression discrepancy between concurrent viewers, severely impacts user satisfaction. While much research has focused on bandwidth adaptation, the distinct and critical role of network Round-Trip Time (RTT) as a primary contributor to these latency issues is often underexplored.

This research presents a comprehensive investigation into the impact of RTT on adaptive video streaming latency and proposes client-side Adaptive Bitrate (ABR) strategies to mitigate these effects. A simulation framework was developed, driven by real-world mobile HSDPA bandwidth traces and utilizing synthetically generated, controlled RTT profiles to emulate diverse network path delays for multiple concurrent clients. The performance of a novel RTT-aware ABR algorithm, `LatencyAwareAbr`, was evaluated against these conditions, focusing on its ability to manage playback latency and viewer discrepancy.

The empirical results demonstrate that increased RTT significantly degrades streaming performance, leading to longer startup times (average increase of 0.23s for high-RTT client), lower average selected bitrates (approximately 20% reduction for high-RTT client), and substantial playback latency discrepancies between users, which can exceed 18 seconds in challenging network scenarios. The `LatencyAwareAbr` algorithm showed adaptive behavior by making more conservative bitrate choices for higher-RTT clients, partially mitigating stalls in some conditions but still resulting in noticeable playback lag and reduced overall QoE compared to low-RTT clients. The findings quantify the trade-offs involved and highlight the limitations of simple predictive heuristics within ABRs when faced with high RTT and volatile bandwidth.

This study contributes by providing a focused, quantitative analysis of RTT's impact on viewer playback discrepancy, establishing a performance baseline for an RTT-aware ABR, and underscoring the critical need for mobile-centric ABR designs that robustly incorporate RTT management and advanced predictive capabilities. The insights derived inform the development of more sophisticated latency optimization techniques crucial for enhancing user experience in the evolving mobile video streaming landscape, particularly for latency-sensitive live and interactive applications.


# Chapter 1: Analyzing and Optimizing Latency in Video Streaming

## 1.1 Background of the Study

Video streaming has become a cornerstone of modern digital communication and entertainment. Its pervasive presence spans various applications, from on-demand content platforms like Netflix and YouTube to real-time services such as video conferencing and interactive gaming. The escalating demand for high-quality video streaming, particularly on mobile devices, necessitates continuous innovation to address the inherent challenges of delivering seamless and low-latency experiences.

The proliferation of mobile devices and the exponential growth of mobile video traffic have transformed the landscape of content consumption. This surge in demand underscores the critical need for efficient video delivery mechanisms that can cope with the constraints of mobile networks. However, mobile networks are inherently variable, characterized by fluctuating bandwidth, variable Round-Trip Times (RTTs), and signal strength, posing significant obstacles to achieving consistent QoE (Oyman & Singh, 2012).

Latency, the delay between content creation and playback, emerges as a key bottleneck in delivering satisfactory video streaming experiences. This includes not only client-side perceived latency like startup delay and stalling, but also, critically for live or co-viewing scenarios, the discrepancy in playback progression between different users due to varying network conditions. Even minor delays can significantly degrade QoE, leading to user frustration, reduced engagement, and potential revenue loss for content providers (Balachandran et al., 2013; Dobrian et al., 2011). Studies have revealed that a substantial portion of online video viewers encounter buffering, startup delays, and low resolution—all manifestations of latency—highlighting the urgent need for effective latency optimization strategies (Dobrian et al., 2011).

Adaptive HTTP streaming technologies, such as MPEG-DASH, HLS, and Smooth Streaming, have emerged as industry standards to address the challenges of video streaming over variable networks (Sodagar, 2011). These protocols employ techniques like adaptive bitrate selection, segmented delivery, and client-side buffering to mitigate the impact of network fluctuations and optimize QoE. This study will investigate these dynamics within a simulated Adaptive Bitrate (ABR) streaming environment, allowing for controlled manipulation of factors like RTT to evaluate novel client-side ABR logic. However, existing adaptive streaming protocols often prioritize high bitrates and video quality over minimizing latency, particularly in mobile environments (Yin et al., 2015). This prioritization can lead to suboptimal QoE in delay-sensitive applications and scenarios characterized by fluctuating bandwidth and high network delays (Petrangeli et al., 2017). Furthermore, research in video streaming optimization has predominantly focused on bandwidth as the primary limiting factor, often overlooking the critical role of network delay, specifically RTT, as a major contributor to latency (Oyman & Singh, 2012). While bandwidth is undoubtedly important, network delay directly impacts interactivity, responsiveness, and the efficiency of ABR decision-making, especially in real-time applications (Balachandran et al., 2013). In mobile environments, network delay can be particularly pronounced and unpredictable, further exacerbating latency issues and hindering QoE (da Costa Filho et al., 2016).

## 1.2 Statement of the Problem

The central problem this project addresses is the persistent challenge of high latency and playback discrepancies in adaptive video streaming, especially within the context of mobile networks. Despite significant advancements in video streaming technologies and protocols, latency remains a major impediment to delivering truly seamless and high-quality user experiences. Specifically, the problem statement can be articulated as follows:

High latency (including significant network RTT) in adaptive video streaming, particularly in mobile network environments, continues to degrade Quality of Experience (QoE) for end-users by increasing startup times, causing stalls, forcing suboptimal bitrate choices, and leading to significant playback progression discrepancies between concurrent viewers, hindering user engagement, and impacting the viability of video streaming services.

This overarching problem encompasses several specific issues:

1. **Suboptimal QoE in Mobile Networks:** Current adaptive streaming protocols often fail to deliver optimal QoE in mobile networks due to their tendency to prioritize high bitrates over low latency. This is particularly problematic in mobile environments characterized by fluctuating bandwidth and high network delays, where latency (including RTT) becomes a dominant factor in user perception. Existing protocols like DASH and HLS, while effective in adapting to varying bandwidth conditions, may not adequately address the latency-sensitive nature of real-time applications and user expectations in mobile scenarios (Petrangeli et al., 2017). Simulations in this research will specifically demonstrate this, particularly for clients experiencing higher RTTs.

2. **Insufficient Focus on Network Delay (RTT):** Existing research and optimization efforts in video streaming often prioritize bandwidth optimization, neglecting the crucial role of network delay (specifically RTT) as a primary latency contributor. This bandwidth-centric approach overlooks the fact that network delay, even with ample bandwidth, can significantly impact latency and QoE, especially in mobile environments where delay is often unpredictable and highly variable. This research directly addresses this by modeling and varying RTT in a controlled simulation environment to quantify its specific impact on ABR performance and viewer discrepancy, distinct from bandwidth. The custom `LatencyAwareAbr` evaluated herein attempts to factor in RTT. The complex interplay between bandwidth and delay and their combined impact on user experience require further investigation and tailored optimization strategies (Oyman & Singh, 2012).

3. **Lack of Adaptation to Heterogeneity (Specifically Network Path RTT):** The increasing heterogeneity of mobile devices and network conditions presents a significant challenge for latency optimization. While various forms of heterogeneity exist, this study's initial empirical work focuses on quantifying the impact of network path RTT heterogeneity between clients using the same ABR logic. Existing algorithms often assume homogeneous environments and fail to account for the diverse network

characteristics of real-world mobile devices and networks. This lack of adaptability results in suboptimal performance and QoE for a significant portion of users, particularly those operating in challenging network conditions with high RTT. Optimization strategies must be tailored to address the unique characteristics of individual devices and network environments to ensure consistent and high-quality QoE across diverse user populations (Petrangeli et al., 2015).

4. **Fundamental Latency Challenges for Live and Immersive Applications:** The advent of VR and 360-degree video streaming introduces new and amplified latency challenges. These immersive experiences demand ultra-low latency to avoid motion sickness and ensure a sense of presence and realism (Akhshabi et al., 2013). While specific viewport prediction and tile-based streaming techniques (Fan et al., 2017) are promising for bandwidth optimization in VR, the fundamental issues of RTT impact on ABR responsiveness and playback discrepancy are critical for all streaming, and understanding them is a prerequisite for tackling more complex scenarios. The current implemented work focuses on these foundational aspects rather than VR-specific ABR optimizations.

## 1.3 Aim and Objectives

**Aim:**

The overarching aim of this project is to significantly reduce latency and playback discrepancy in adaptive video streaming, particularly in mobile networks, to enhance Quality of Experience (QoE) for end-users and improve the overall performance and efficiency of video streaming services.

**Objectives:**

To achieve this aim, the project will pursue the following specific objectives:

1. **Analyze and quantify latency factors in client-side adaptive video streaming:** This objective involves a comprehensive investigation through simulation of the various factors that contribute to latency in video streaming, including network conditions (real-world bandwidth traces and controlled round-trip time profiles), and client-side factors (ABR algorithm logic, buffer management strategies). The analysis will focus on understanding the relative importance of each factor and their interplay in determining overall latency and QoE, with a key aspect being the quantification of playback latency discrepancy between users experiencing different RTTs.

2. **Evaluate the impact of different client-side adaptive bitrate (ABR) strategies on latency and QoE:** This objective aims to assess the latency performance of ABR strategies, particularly a novel `LatencyAwareAbr` designed to be RTT-sensitive, under

varying mobile network conditions (bandwidth and RTT). The evaluation will identify the strengths and weaknesses of these ABR approaches in terms of latency optimization and inform the development of more effective solutions.

3. **Propose and evaluate an optimized ABR strategy that reduces latency by at least 20% in key metrics:** This objective focuses on the design and implementation (within the simulation environment) of a novel or enhanced ABR strategy that incorporates advanced latency optimization techniques, informed by the findings from Objectives 1 and 2. The target latency reduction of at least 20% in metrics such as startup delay, buffering ratio, or playback latency discrepancy will serve as a benchmark for evaluating the success of the proposed strategy.

4. **Test and validate the proposed ABR strategies using simulations driven by real-world mobile bandwidth traces and controlled, synthetic RTT profiles on different network conditions:** This objective entails rigorous testing and validation of the proposed ABR strategies using experimental setups within the simulation environment that mimic diverse network conditions. The validation process will employ both objective metrics, such as latency, buffering ratio, quality switch frequency, and playback latency discrepancy, and a composite QoE score to comprehensively evaluate the ABR's performance and effectiveness in emulated real-world scenarios. Subjective QoE assessments are acknowledged as a valuable future validation step.

## 1.4 Justification of the Study

This research is justified by the ever-increasing importance of video streaming in today's digital world and the critical need to address the persistent challenge of latency, particularly in mobile networks. Several factors underscore the significance and timeliness of this study:

1. **Industry Relevance and User Demand:** The explosive growth of mobile video streaming and the increasing user demand for seamless, high-quality experiences make latency optimization a critical concern for the video streaming industry. Minimizing latency (including startup times, stalls, and playback discrepancies between viewers) directly addresses user expectations, enhances QoE, and improves user engagement, which translates to increased revenue and subscriber growth for content providers (Dobrian et al., 2011). Furthermore, latency optimization aligns with the industry's need for efficient utilization of mobile network resources, ensuring the sustainability of mobile data networks (Dahlman et al., 2014).

2. **Addressing Research Gaps:** This project directly addresses significant research gaps in the existing literature on video streaming latency optimization. By focusing on network RTT as a primary and distinct latency factor, systematically evaluating its impact through controlled simulation, and quantifying the resultant playback latency discrepancy between users, this research aims to bridge critical gaps in current knowledge and

contribute to the advancement of the field. The project also explores how understanding these fundamental RTT impacts is foundational for addressing latency in more complex applications like VR/360 video and live streaming.

3. **Potential for Innovation and Impact:** This project leverages a controlled simulation framework to enable the development and evaluation of a novel RTT-aware ABR algorithm (`LatencyAwareAbr`). The insights gained and any subsequent algorithm enhancements to explore new possibilities for latency reduction and develop more intelligent and data-driven optimization strategies (Mao et al., 2017). The development of a novel ABR strategy with demonstrably reduced latency and enhanced QoE, particularly in managing viewer discrepancy, has the potential to significantly impact the video streaming industry and improve the user experience for millions of mobile video consumers. The simulation framework itself is a contribution, enabling this controlled study.

## 1.5 Scope of the Study

The scope of this project is focused on analyzing and optimizing client-side perceived latency and playback discrepancy in adaptive video streaming, particularly in mobile network environments, through simulation-based evaluation. While the research will draw upon insights from studies using real-world data and experimental setups, the primary focus will be on:

1. **Mobile Networks:** The study will primarily focus on video streaming optimization for simulated mobile networks (e.g., utilizing HSDPA bandwidth traces representative of 4G/LTE-like conditions) and their inherent challenges of fluctuating bandwidth and controlled high delays (RTT) (Dahlman et al., 2014). The optimization strategies developed will be specifically tailored to address the characteristics and constraints of mobile network environments. While the findings may be relevant to other network types, extending the findings to fixed networks or other broadband networks is outside the primary scope of this project and may require further investigation.

2. **Latency as the Primary Metric Set:** The project will prioritize latency as the primary metric set for optimization and QoE enhancement. This includes startup delay, buffering events (stalls), and critically, the relative playback progression between clients (viewer discrepancy). While other QoE factors such as video quality (average bitrate) and quality switch frequency will be considered as important balancing factors and outcomes, the central focus will remain on minimizing latency and improving responsiveness, especially for delay-sensitive applications.

3. **Adaptive Bitrate Streaming Logic:** The research will concentrate on client-side adaptive bitrate streaming techniques, specifically the logic within HTTP-based adaptive streaming protocols like DASH and HLS (though protocol specifics are abstracted in the simulation). While other streaming techniques exist, adaptive HTTP streaming is the

dominant paradigm for video delivery over the internet, making the ABR client logic the most relevant and impactful area of study for this work.

4. **Foundational Latency Optimization for General and Immersive Applications:** The project will explore fundamental latency optimization techniques applicable to a broader range of video streaming applications. However, the primary focus will remain on general latency optimization techniques applicable to a broader range of video streaming applications, with the understanding of RTT impact being foundational for more specialized areas like VR and 360-degree video serving as a specific use case motivation, though specific tile-based or VR-centric ABR optimizations are outside the immediate scope of the implemented work.

5. **Algorithm Development and Evaluation through Simulation:** The project will encompass the development of a novel latency-aware adaptive bitrate algorithm (`LatencyAwareAbr`) and its rigorous evaluation and validation through simulations. However, the project will not involve the deployment of a fully functional, production-ready video streaming platform. The focus will be on developing and validating the core algorithms and optimization strategies within the simulation testbed, leaving the implementation and deployment at scale for future work.

By focusing on these specific areas, the project aims to provide a deep and impactful investigation into latency optimization for video streaming, addressing critical challenges and contributing valuable knowledge to the field.

**References**


i. Akhshabi, S., Begen, A. C., & Dovrolis, C. (2011). An experimental evaluation of rate-adaptation algorithms in adaptive streaming over HTTP. *Proceedings of the 2nd Annual ACM Multimedia Systems Conference (MMSys '11)*, 157–168.

ii. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., & Zhang, H. (2013). Developing a predictive model of quality of experience for Internet video. *Proceedings of the ACM SIGCOMM Conference*, 339–350.

iii. da Costa Filho, R. I. T., Lautenschlager, W., Kagami, N., Roesler, V., & Gaspary, L. P. (2016). Network fortune cookie: Using network measurements to predict video streaming performance and QoE. *2016 IEEE Global Communications Conference (GLOBECOM)*, 1–6.

iv. Dahlman, E., Mildh, G., Parkvall, S., Peisa, J., Sachs, J., Selén, Y., & Sköld, J. (2014). 5G wireless access: Requirements and realization. *IEEE Communications Magazine*, *52*(12), 42–47.

v. Dobrian, F., Sekar, V., Awan, A., Stoica, I., Chuang, D., Ganjam, A., ... Zhang, H. (2011). Understanding the impact of video quality on user engagement. *Proceedings of the ACM SIGCOMM Conference (SIGCOMM '11)*, 362–373.

vi. Fan, C.-L., Lee, J., Lo, W.-C., Huang, C.-Y., Chen, K.-T., & Hsu, C.-H. (2017). Fixation prediction for 360 video streaming in head-mounted virtual reality. *Proceedings of the*

*27th Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV '17)*, 67–72.

vii. Mao, H., Netravali, R., & Alizadeh, M. (2017). Neural adaptive video streaming with pensieve. *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM '17)*, 197-210.

viii. Oyman, O., & Singh, S. (2012). Quality of experience for HTTP adaptive streaming services. *IEEE Communications Magazine*, *50*(4), 20–27.

ix. Petrangeli, S., Swaminathan, V., Hosseini, M., & De Turck, F. (2017). An HTTP/2-based adaptive streaming framework for 360 virtual reality videos. *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*, 306–314.

x. Petrangeli, S., Famaey, J., Claeys, M., Latré, S., & De Turck, F. (2015). QoE-driven rate adaptation heuristic for fair adaptive video streaming. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *12*(2), Article 28.

xi. Sodagar, I. (2011). The MPEG-DASH standard for multimedia streaming over the Internet. *IEEE Multimedia*, *18*(4), 62–67.

xii. Yin, X., Jindal, A., Sekar, V., & Sinopoli, B. (2015). A control-theoretic approach for dynamic adaptive video streaming over HTTP. *SIGCOMM Computer Communication Review*, *45*(4), 325–338.