ARD Algo (Ashley rose daug algorithm 😱)

"Optimizing Adaptive Bitrate Streaming in DASH Using BOLA and Lightweight GRU-Based Machine Learning Algorithms"

"Enhancing QoE in DASH Video Streaming with BOLA and GRU-Driven Adaptive Bitrate Selection"

"Efficient Adaptive Bitrate Streaming for Mobile Networks: Integrating BOLA and GRU-Based Approaches"

"Leveraging BOLA and Gated Recurrent Units for Adaptive Bitrate Streaming in DASH: A Solution for Network Fluctuations"

"Machine Learning-Enhanced DASH Streaming: Combining BOLA and GRU for Improved Video Quality and Stability"

"Enhancing Adaptive Bitrate Streaming Efficiency in DASH Using Combined BOLA and GRU Techniques"

"Advanced Techniques for Adaptive Bitrate Streaming in DASH: Merging BOLA Optimization with GRU-Based Predictions"

I. Introduction

The increasing demand for high-quality video streaming, coupled with the unpredictable nature of wireless networks, makes Adaptive Bitrate Streaming (ABS) a pivotal technique for enhancing users' Quality of Experience (QoE). ABS adjusts the video quality dynamically to match the varying network conditions, ensuring a smoother and more enjoyable viewing experience[1]. ABS algorithms are crucial for delivering high-quality video experiences over networks with differing speeds, maximizing the QoE [1,2]. One of the most widely adopted streaming technology standards for adaptive streaming is Dynamic Adaptive Streaming over HTTP (DASH). DASH is renowned for its compatibility across devices and efficient utilization of Hypertext Transfer Protocol (HTTP) [2]. It works by dividing multimedia content's information into small segments into a manifest file called the Media Presentation Description (MDP) and storing them on a server. Clients first download a file from the server that details the video, including quality options and segment sizes. The server and client then utilize an ABS algorithm to pick the best bitrate quality for each video segment based on the network conditions corresponding to a specific bitrate. Each segment is then downloaded individually and stored in a temporary buffer with limited space. Concurrently, the ABS algorithm matches the download speed (based on chosen bitrates) to the network speed. If the stream playback uses data faster than the download, the buffer empties, causing a pause or rebuffering in playback. Ideally, the ABS algorithm aims to deliver the highest quality video stream possible while minimizing rebuffering, and initial delays, and preventing frequent quality changes, taking into account playback status, such as buffer length and requested video quality, and the network's bandwidth [1, 2, 3, 4]. High bitrate, less rebuffering, minimal initial delays, and fewer quality changes serve as metrics of QoE (Quality of Experience) in this study. To achieve better QoE with time-varying network conditions on top of the adaptive streaming standard DASH, it is imperative to design an effective ABS algorithm for the DASH system.

-dash algos

-

BOLA (Buffer Occupancy-based Lyapunov Algorithm) **is** a notable bitrate adaptation algorithm that **gains** traction for its effectiveness. BOLA **leverages** Lyapunov optimization to derive an online bitrate adaptation algorithm that **achieves** utility close to the maximum possible in large video regimes. It **offers** a balance between video quality and reduced chances of rebuffering, giving video content providers a way to control this trade-off. Although frequent bitrate switches during video playback **can be** disruptive, BOLA **includes** techniques to mitigate this issue, ensuring a more stable viewing experience. Furthermore, BOLA's buffer-based approach **avoids** the complexities and overheads of bandwidth prediction, making it more resilient to bandwidth fluctuations. This resilience **makes** it a preferred choice over other existing algorithms, which **may not offer** theoretical guarantees on achieved utility. Despite these advantages, BOLA **has** some limitations. Its buffer-based approach, while avoiding the overheads of bandwidth prediction, **may not fully capitalize** on the potential benefits of bandwidth or throughput prediction. Predictive models **can offer** more proactive bitrate adjustments, potentially leading to improved QoE by better-anticipating network conditions. However, traditional bandwidth prediction approaches often **come** with their own set of complexities and computational demands. Moreover, mobile networks, especially prevalent in regions like the Philippines where mobile data penetration **exceeds** fixed internet access, **are** particularly susceptible to fluctuations in speed due to factors like signal interference and handovers. These fluctuations **pose** a significant obstacle to maintaining consistent video quality and uninterrupted streaming experiences for users. Traditionally, proposed ABS methods **rely** on heuristics, categorized as throughput-based, buffer-based, or a hybrid of both. However, these heuristic approaches, with their fixed-rule nature, **struggle** to adapt to the dynamic conditions of mobile networks and playback environments. Achieving optimal QoE with a fixed-control rule **is inherently** challenging. To address these limitations, machine learning (ML) techniques **offer** a promising alternative. While these techniques **offer** a powerful approach to ABS, it's important to acknowledge the computational demands they **place** on devices. Complex calculations and large model sizes **can strain** low-to-mid-range smartphones, potentially **leading** to slow performance and impacting the viewing experience. In this context, Gated Recurrent Units (GRUs), a type of Recurrent Neural Network (RNN), **offer** significant advantages. GRUs **are designed** to handle sequence prediction problems and **are** particularly effective in capturing temporal dependencies in data, making them suitable for predicting network bandwidth fluctuations. Their architecture, which **includes** gating mechanisms to control the flow of information, **allows** for efficient learning and reduced computational load compared to traditional RNNs. This **makes** GRUs a compelling choice for implementing bandwidth prediction in adaptive bitrate streaming, offering a balance between performance and computational efficiency. The scheme proposed in [17] **utilizes** GRUs for bandwidth prediction and adaptive bitrate selection, demonstrating their strengths in handling the dynamic conditions of heterogeneous mobile networks. This study **introduces** an innovative approach to enhance bitrate selection within the DASH framework. Taking into account the computational constraints of various devices, including mobile devices and streaming devices, our objective **is** to develop an adaptive bitrate selection algorithm that **utilizes** lightweight and efficient machine learning techniques, such as GRUs. This approach **aims** to provide a smooth and reliable streaming experience for users across all devices without placing undue strain on their processing capabilities. Our strategy **addresses** network fluctuations and **minimizes** buffering interruptions, enhancing the overall QoE for video viewers on all platforms.

II. Review of Related Literature

a. Traditional Heuristic-Based ABS Approaches Among the early ABS algorithms built on DASH, BBA (Buffer-Based Approach) **was** one of the pioneers that utilized a buffer occupancy-based approach [8]. It **adjusted** the bitrate according to the buffer fullness to maintain a stable buffer level, ensuring smooth playback. Although simple and efficient, BBA **did not always select** the optimal bitrate to maximize video quality or bandwidth utilization. Another notable ABS algorithm **was** MPC (Model Predictive Control), which **employed** a control-theoretic approach using predictive models to optimize bitrate selection [9]. MPC **predicted** future buffer and throughput conditions to minimize a cost function over a future time horizon, offering potentially better performance by considering future network dynamics. However, its implementation complexity and the need for accurate prediction models **were** challenging. BOLA **was** an ABS algorithm that **focused** on buffer occupancy but **used** a more linear approach compared to BBA [10]. It **maintained** a sliding window of recent throughput measurements to estimate available bandwidth and **selected** bitrate based on buffer and estimated bandwidth. BOLA's efficiency and responsiveness to changing network conditions often **resulted** in better video quality, although it **could be** sensitive to estimation errors leading to buffer oscillations. Panda, on the other hand, **employed** a hybrid approach by combining buffer-based and throughput-based strategies [11]. This ABR algorithm **considered** both buffer fullness and available throughput to make bitrate decisions, striking a balance between stability and video quality. While Panda **adapted** well to varying network conditions, its complexity in combining different approaches and the need for fine-tuning **were** challenging. Lastly, Festive (Fair and Efficient Streaming Technique for Internet Video Everywhere) **was designed** to provide fairness among users sharing the same network [12]. Beyond considering buffer and throughput, Festive **focused** on fairness in bandwidth allocation. This algorithm **aimed** to prevent congestion and improve user experience in shared networks by ensuring fair bandwidth allocation. However, Festive's complexity in fairness calculations **sometimes sacrificed** individual video quality for the sake of fairness.

b. Machine-Learning-Based Approaches Several machine-learning-based approaches for ABS **were proposed** with DASH. One approach **utilized** a deep neural network (DNN) trained on video streaming data to recommend the next video segment's bitrate based on the current network state and buffer occupancy [13]. Another **leveraged** client-side and server-side information for informed decision-making, including video content, network congestion, and client buffer status [14]. Additionally, an online Q-learning algorithm **was proposed** for real-time adaptation based on continuous learning from experience [15]. Beyond optimizing video quality, RL **contributed** to cost-effectiveness and scalability. One approach **combined** convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract video features and predict future network conditions, enabling cost-effective bitrate selection on the server side of the DASH system [16]. Similarly, gated recurrent units (GRUs) on HTTP Adaptive Streaming (HAS) systems **were used** to predict bandwidth availability in heterogeneous mobile networks, tailoring bitrate selection for optimal quality of experience (QoE) [17].

III. Methodology

a. Proposed System Model

The primary goal of this paper was to develop an efficient and lightweight model to enhance adaptive bitrate streaming while also maximizing the QoE. Drawing inspiration from the framework proposed in [17], we employed a variant of Recurrent Neural Network (RNN), known as the Gated Recurrent Unit (GRU), for bandwidth estimation. In essence, the GRU model's output served as input to a video segment selector, which determined the most appropriate quality level for the subsequent video segment to optimize the QoE metric.

B. Dataset and preprocessing

Building upon the framework proposed in [17], which initially considered three features for their GRU model, transferred data, elapsed time, and calculated bitrate from the two, we employed two distinct datasets:

1. Dataset 1: This dataset comprised 4G/LTE network traces with a granularity of 1 second. Bandwidth was originally measured in microseconds but has been converted to seconds to reduce the dataset's size while preserving its inherent trends and patterns.

- Dataset 2: This dataset included 3G network traces, also with a granularity of 1 second. Here, bandwidth was initially measured in milliseconds and similarly converted to seconds for consistency and ease of analysis.

To evaluate the performance of our model and determine the impact of the selected features on its predictive accuracy, the researchers utilized two commonly accepted metrics: mean squared error (MSE) and mean absolute error (MAE). These metrics are well-suited for measuring prediction errors, providing a comprehensive assessment of the model's predictive capabilities. By incorporating these features and using MSE and MAE as our evaluation criteria, we aim to enhance the model's accuracy and reliability.

c. GRU Model

A (GRU) is a type of recurrent neural network (RNN) that is designed to capture long-term dependencies in sequential data while mitigating some of the issues like vanishing gradients faced by traditional RNNs. GRUs have gating mechanisms that control the flow of information through the network, making them particularly effective for tasks like sequence modeling and time series prediction.

In the context of our study, the selected features serve as a flow of information in the network, allowing the GRU model to predict the hidden state (ht) at time t based on the previously known values. The update gate (Eq. 1) determines how much of the previous hidden state (Eq. 3) should be retained, while the reset gate (Eq. 2) decides how much of the previous state should be forgotten or reset. The candidate hidden state (Eq. 3) computes a new potential hidden state based on the current input (xt) and the reset gate (Eq. 2). Finally, the hidden state (Eq. 4) is updated using the update gate (Eq. 1) to combine the previous hidden state (Eq. 3) and the new candidate hidden state (Eq. 3), resulting in the updated hidden state.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z \quad (1)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad \mathbf{(2)}$$

$$\widehat{h}_t = \phi(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widehat{h}_t \quad (4)$$

By utilizing these formulas, the GRU model can effectively learn and capture the underlying patterns in the data, making it suitable for bandwidth estimation in our proposed system. The flow of information facilitated by the selected features allows the GRU model to adapt and predict bandwidth requirements based on various context factors, thereby enhancing the accuracy and robustness of our proposed system.

d. BOLA Model as Video Segment Selector

In the followed study, the framework **separated** throughput prediction from the video chunk selector. For our study, the researcher **chose** to incorporate BOLA as the video chunk selector algorithm. BOLA, a DASH-based ABS algorithm, dynamically **adjusted** video bitrate based on buffer occupancy and estimated network bandwidth. While effective in many scenarios, BOLA **had** certain limitations that **could** affect its performance.

One key weakness of BOLA **lay** in its reactivity. Because it **relied** on current buffer and network conditions, BOLA's bitrate adjustments **could** be reactive, potentially causing oscillations in bitrate selection. These fluctuations **could** result in inconsistent QoE for viewers, impacting their overall streaming experience. Additionally, BOLA **lacked** predictive capability, meaning it **could not anticipate** future network conditions or adapt proactively to changing dynamics. This limitation **could** lead to unsustainable or unsuitable bitrate selections for anticipated future conditions, further compromising QoE.

To address these weaknesses and enhance the video segment selection process, researchers **proposed** integrating the GRU model's bandwidth estimation capabilities with BOLA's bitrate selection mechanism. The GRU model, with its ability to estimate bandwidth based on multiple context features, **provided** a more comprehensive understanding of network conditions and viewer context. By incorporating the GRU model's enhanced context awareness, researchers **could include** additional contextual information into the bitrate selection process. This information **included** factors such as device type, screen resolution, and time-varying network conditions, which **could** be used alongside buffer occupancy and network bandwidth by BOLA for more informed bitrate selection.

e. Proposed video chunk selector algorithm

The GRU model's predictive capability allows for proactive adaptation to future network conditions. By anticipating changes in network bandwidth and stability, the GRU model can guide BOLA in making bitrate adjustments that are optimal for current conditions and sustainable for anticipated future conditions. Pseudo codes of the proposed algorithm are given as follows.

Algorithm 1 *ARD*

Input:

R^*: Bitrate of the previously downloaded chunk
R_i**:** Bitrate of the previously downloaded chunk
L: Length of the video chunk
$B_{current}$: The current buffer occupancy
B_{prev}: The buffer occupancy for the previous video chunk
C: The network bandwidth predicted by the GRU model

Output:

$qualityLevel$**:** The optimized quality level of the next video chunk

1: V_{min} : Minimum buffer occupancy threshold
2: V_{max} : Maximum buffer occupancy threshold
3: γ: Control parameter for BOLA
4: $QoE_{max} := 0$
5: $qualityLevel := 0$

$V := \frac{B_{current} - V_{min}}{V_{max} - V_{min}}$

For $i = 0$ **to** $qualityLevel_{max}$ **:**

$Utility(R_i) := \frac{R_i}{C} - \gamma \cdot (\frac{R_i \cdot L}{C} - B_{current})$

For $i = 0$ **to** $qualityLevel_{max}$ **:**

  If $(Utility(R_i) > QoE_{max})$ and $(R_i < C)$ **:**

    $QoE_{max} = Utility(R_i)$

    $qualityLevel = 0$

Return qualityLevel

The modified algorithm integrates BOLA's buffer occupancy-based approach to select the video chunk's bitrate. The algorithm starts by initializing necessary parameters for BOLA and calculating the buffer occupancy state V Then, it calculates the utility for each possible bitrate level based on the current buffer occupancy and the predicted network bandwidth using the GRU model. The utility calculation incorporates BOLA's principle by factoring in the buffer occupancy and bandwidth estimation. Finally, it selects and returns the quality level that maximizes the utility.

This integration aims to leverage BOLA's ability to adjust bitrate dynamically based on buffer occupancy while enhancing it with the GRU model's predictive bandwidth estimation capabilities to ensure more stable and proactive bitrate adjustments, thereby improving the overall Quality of Experience (QoE) for video streaming. In essence, by integrating the bandwidth estimation capabilities of the GRU model with BOLA's bitrate selection mechanism, researchers created a hybrid approach that combines the strengths of both methods while mitigating their respective weaknesses. This synergistic integration can lead to improve QoE for viewers by ensuring more informed, proactive, and sustainable bitrate selections.

IV. PERFORMANCE EVALUATION

a. EXPERIMENT SETUP

b. QoE FACTORS EVALUATION

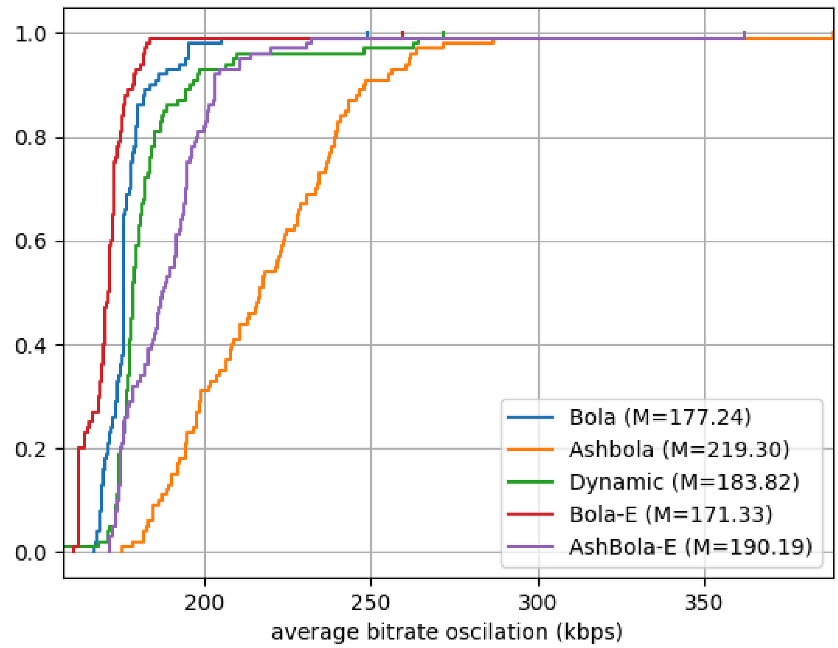

FIGURE 1.

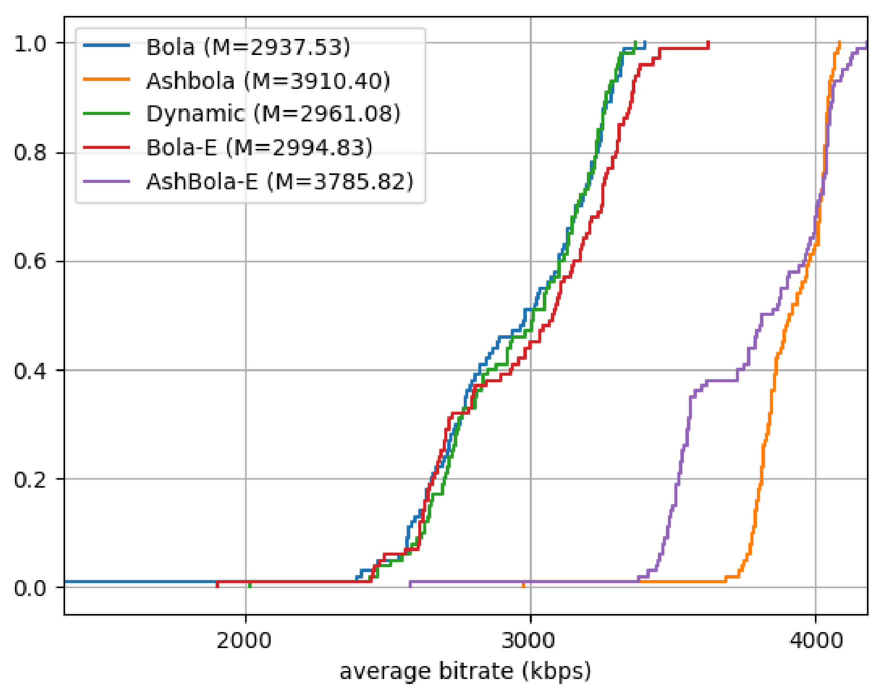

FIGURE 2.

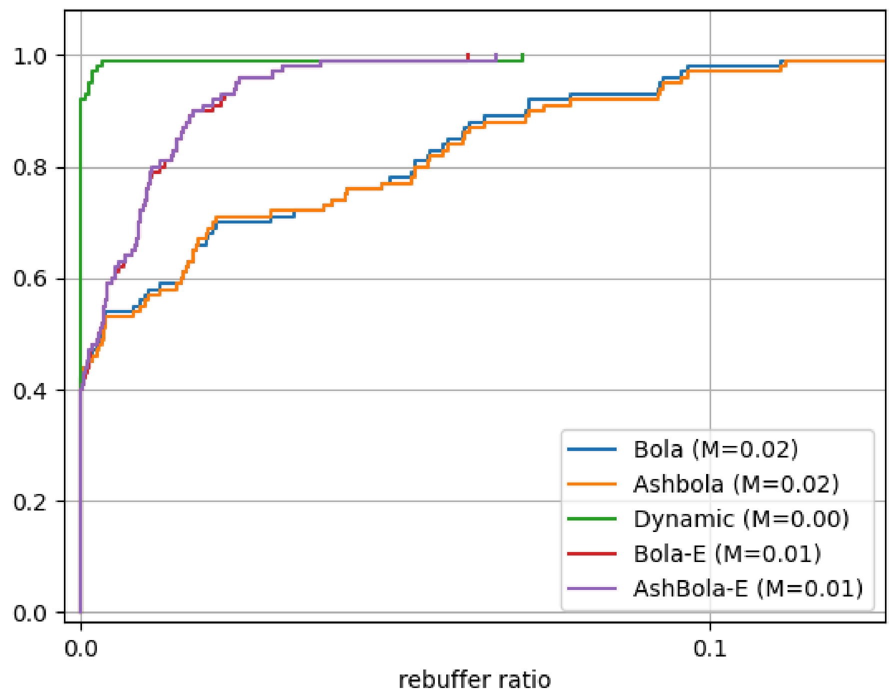

FIGURE 3.

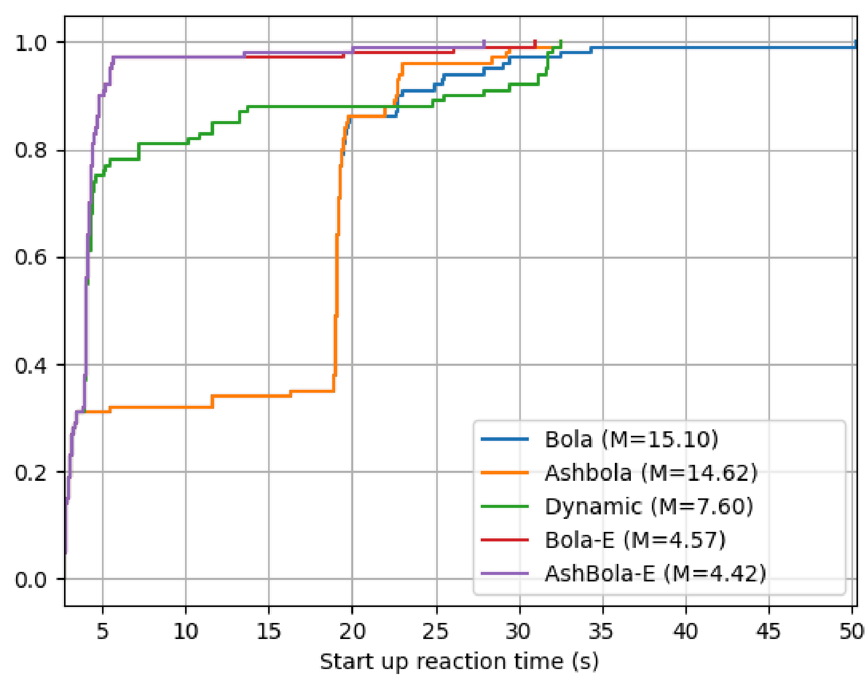

FIGURE 4.

Buffer occupation results, quality levels, bitrate oscilation and QoE of the four compared methods are shown in Figures 1, 2, 3 and 4. Buffers accumulate and change with bandwidth changes, as seen in Figures 2 and 4 which depict buffer occupation as well as QoE therefore the more accurate the estimation of bandwidth is done, the better a performance is obtained. It can be observed that in most cases adopting our developed bandwidth prediction outperforms other methods except at initial streaming stage. Quality levels' adaptation via adjusting bitrate and resolution of video chunks for improving QoE by BOLA algorithm can be seen in Figure 4 during video streaming process. In addition,

References:

1. Khan, K. (2024). Optimization Models for Adaptive Video Streaming: Balancing Quality, Buffering, and Bandwidth. International Journal of Latest Engineering Research and Applications (IJLERA), 9(1). http://ijlera.com/papers/v9-i1/1.202401614.pdf
2. Van, L., MA, Park, J., Nam, J., Ryu, H., Kim, J.: A Fuzzy-Based adaptive streaming algorithm for reducing entropy rate of DASH bitrate fluctuation to improve mobile quality of service. Entropy. 19, 477 (2017). https://doi.org/10.3390/e19090477.
3. Khan, K.: A taxonomy for deep learning in dynamic adaptive video streaming over HTTP. International Journal of Multidisciplinary Research and Analysis. 06, (2023). https://doi.org/10.47191/ijmra/v6-i5-60.
4. Petrangeli, S., Van Der Hooft, J., Wauters, T., De Turck, F.: Quality of Experience-Centric Management of Adaptive video streaming services. ACM Transactions on Multimedia Computing, Communications and Applications/ACM Transactions on Multimedia Computing Communications and Applications. 14, 1–29 (2018). https://doi.org/10.1145/3165266.
5. Kemp, S.: Digital in the Philippines: All the statistics you need in 2021 — DataReportal – Global Digital Insights, https://datareportal.com/reports/digital-2021-philippines.
6. Karagkioules, T., Paschos, G.S., Liakopoulos, N., Fiandrotti, A., Tsilimantos, D., Cagnazzo, M.: Online learning for adaptive video streaming in mobile networks. ACM Transactions on Multimedia Computing, Communications and Applications/ACM Transactions on Multimedia Computing Communications and Applications. 18, 1–22 (2022). https://doi.org/10.1145/3460819.

7. M. J. Khan, S. Harous and A. Bentaleb, "Client-driven Adaptive Bitrate Techniques for Media Streaming over HTTP: Initial Findings," 2020 IEEE International Conference on Electro Information Technology (EIT), Chicago, IL, USA, 2020, pp. 053-059, doi: 10.1109/EIT48999.2020.9208253.
8. Huang, T.-Y., Johari, R., McKeown, N., Trunnell, M., Watson, M.: A buffer-based approach to rate adaptation. Computer Communication Review. 44, 187–198 (2014). https://doi.org/10.1145/2740070.2626296.
9. Yin, X., Jindal, A., Sekar, V., & Sinopoli, B. (2015). A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP. ACM SIGCOMM Computer Communication Review, 45(4). https://doi.org/10.1145/2785956.2787486
10. BOLA: Near-optimal bitrate adaptation for online videos. IEEE Conference Publication | IEEE Xplore. (2016).
11. Li, Z., Zhu, X., Gahm, J., Pan, R., Hu, H., Begen, A.C., & Oran, D. (2013). Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale. IEEE Journal on Selected Areas in Communications, 32, 719-733.
12. Jiang, J., Sekar, V., Zhang, H.: Improving fairness, efficiency, and stability in HTTP-Based adaptive video streaming with Festive. IEEE/ACM Transactions on Networking. 22, 326–340 (2014). https://doi.org/10.1109/tnet.2013.2291681.
13. Souane, N., Bourenane, M., Douga, Y.: Deep Reinforcement Learning-Based Approach for Video Streaming: Dynamic Adaptive Video Streaming over HTTP. Applied Sciences. 13, 11697 (2023). https://doi.org/10.3390/app132111697.

14. Hafez, N.A., Hassan, M.S., Landolsi, T.: Reinforcement Learning-Based rate adaptation in dynamic video streaming. Research Square (Research Square). (2022). https://doi.org/10.21203/rs.3.rs-1616726/v1.

15. Kang, J., Chung, K.: HTTP Adaptive Streaming Framework with Online Reinforcement Learning. Applied Sciences. 12, 7423 (2022). https://doi.org/10.3390/app12157423.

16. Darwich, M., Bayoumi, M.: Video quality adaptation using CNN and RNN models for cost-effective and scalable video streaming Services. Cluster Computing. (2024). https://doi.org/10.1007/s10586-024-04315-8.

17. Huu, T.V., Pham, V.S., Huong, T.N.T., Le, H.-C.: QOE aware video streaming scheme utilizing GRU-based bandwidth prediction and adaptive bitrate selection for heterogeneous mobile networks. IEEE Access. 1 (2024). https://doi.org/10.1109/access.2024.3382155.

18. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. In *NIPS 2014 Workshop on Deep Learning, December 2014*