Bandwidth Prediction in Low-Latency Media Transport

Abdelhak Bentaleb★, Mehmet N. Akcay§, May Lim+, Ali C. Begen§ and Roger Zimmermann+
★Concordia University, §Ozyegin University, +National University of Singapore
★Canada, §Turkiye, +Singapore

ABSTRACT

Designing a robust bandwidth prediction algorithm for low-latency media transport that can quickly adapt to varying network conditions is challenging. In this paper, we present the working principles of a hybrid bandwidth predictor (termed BoB, Bang-on-Bandwidth) we developed recently for real-time communications and discuss its use with the new Media-over-QUIC (MOQ) protocol proposals.

CCS CONCEPTS

• **Networks** → **Application layer protocols** .

KEYWORDS

RTC, MOQ, bandwidth prediction, QoE, latency, DRL.

ACM Reference Format:

Abdelhak Bentaleb★, Mehmet N. Akcay§, May Lim+, Ali C. Begen§ and Roger Zimmermann+. 2023. Bandwidth Prediction in Low-Latency Media Transport. In *Mile-High Video Conference (MHV '23), May 7–10, 2023, Denver, CO, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3588444.3591017

1 INTRODUCTION AND THE BOB DESIGN

Today, real-time communications (RTC) is used in a range of applications such as videoconferencing, cloud gaming, e-learning and immersive media. There have been many studies to date (*e.g.*, congestion control optimization at the transport layer [6], bitrate decision optimization for video codecs [13], and mixed techniques [9]) to improve the quality of experience (QoE) in RTC systems. Despite the advantages of these solutions, accurate bandwidth prediction during dynamic network conditions still needs to be solved as it plays a critical role in maintaining a good QoE. The past studies are mainly based on the traditional heuristic-based bandwidth prediction for RTC systems, following the design principle of the Google Congestion Control (GCC) algorithm [1]. GCC shows some problems [8, 10] in the bandwidth prediction and, therefore, in adapting video bitrate and latency in networks with diverse bandwidth and latency characteristics.

Recent research has turned to deep reinforcement learning (DRL) [11] to solve these problems as an alternative to hand-tuned heuristics. Leveraging a DRL approach that can dynamically adapt to heterogeneous environments, we recently designed BoB, a receiver-side hybrid bandwidth prediction solution for RTC, which combines a heuristic-based controller (inspired by the GCC algorithm) with a DRL controller [5].

The BoB design is depicted in Figure 1, which consists of DRL and heuristic controllers. It uses the heuristic-based controller only at the beginning of an RTC session when input data is scarce. Then it switches to the DRL controller to perform bandwidth prediction. As a result, BoB is able to adapt quickly to any change in network conditions while considering video quality and packet delays/loss. The DRL controller implements actor-critic neural networks (NN) for model training. Later, at each time interval (*e.g.*, 200 ms), the NN takes receiving rate, packet delay, packet loss ratio and the eight most recent predicted bandwidth samples as inputs.

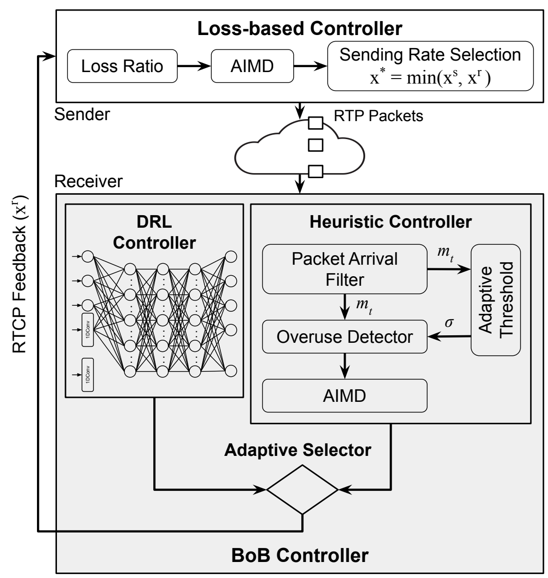

Figure 1: Receiver-side BoB controller design (shown in gray).

We implemented BoB on Microsoft's OpenNetLab platform termed AlphaRTC [4] and validated its performance. Experimental results showed > 80% accuracy in bandwidth prediction across several network traces, outperforming the traditional GCC-based bandwidth predictors. Further detailed results are presented in [5], and the BoB code is publicly available at [2].

RTC traffic is typically transported using the Real-time Transport Protocol (RTP, RFC 3550) running over UDP, and we assumed the usage of RTP when designing BoB. However, the IETF has recently established a new working group (Media over QUIC [3]) to develop a new low-latency transport protocol for media applications. While a few proposals are already under consideration (*e.g.*, [7, 12]), bandwidth prediction is still an open issue. Our current work involves investigating the feasibility of BoB in the MOQ protocol proposals and testing its performance.

This work is licensed under a Creative Commons Attribution International 4.0 License.

MHV '23, May 7–10, 2023, Denver, CO, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0160-3/23/05.
https://doi.org/10.1145/3588444.3591017

ACKNOWLEDGMENTS

This research is supported by the Singapore Ministry of Education Academic Research Fund Tier 2 under MOE's official grant number T2EP20221-0023.

REFERENCES

[1] A Google Congestion Control Algorithm for Real-Time Communication. [Online] Available: https://datatracker.ietf.org/doc/html/draft-ietf-rmcat-gcc-02. Accessed on Dec. 10, 2022.
[2] BoB Code. [Online] Available: https://github.com/NUStreaming/BoB. Accessed on Dec. 10, 2022.
[3] Charter for Working Group - Media Over QUIC. [Online] Available: https://datatracker.ietf.org/doc/charter-ietf-moq/. Accessed on Dec. 10, 2022.
[4] OpenNetLab AlphaRTC. [Online] Available: https://github.com/OpenNetLab/AlphaRTC. Accessed on Dec. 10, 2022.
[5] A. Bentaleb, M. N. Akcay, M. Lim, A. C. Begen, and R. Zimmermann. BoB: Bandwidth prediction for real-time communications using heuristic and reinforcement learning. *IEEE Trans. Multimedia*, to appear (DOI: 10.1109/TMM.2022.3216456).
[6] G. Carlucci, L. De Cicco, S. Holmer, and S. Mascolo. Congestion control for web real-time communication. *IEEE/ACM Trans. Networking*, 25(5):2629–2642, 2017.
[7] L. Curley, K. Pugin, and S. Nandakumar. Warp - Segmented Live Media Transport. [Online] Available: https://datatracker.ietf.org/doc/draft-lcurley-warp/. Accessed on Dec. 10, 2022.
[8] J. Fang, M. Ellis, B. Li, S. Liu, Y. Hosseinkashi, M. Revow, A. Sadovnikov, Z. Liu, P. Cheng, S. Ashok, et al. Reinforcement learning for bandwidth estimation and congestion control in real-time communications. *arXiv preprint arXiv:1912.02222*, 2019.
[9] S. Fouladi, J. Emmons, E. Orbay, C. Wu, R. S. Wahby, and K. Winstein. Salsify: Low-latency network video through tighter integration between a video codec and a transport protocol. In *USENIX NSDI*, 2018.
[10] M. Guerrero Viveros. Performance analysis of Google congestion control algorithm for webrtc. *Master Thesis*, 2019.
[11] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim. Applications of deep reinforcement learning in communications and networking: a survey. *IEEE Communications Surveys & Tutorials*, 21(4):3133–3174, 2019.
[12] Pugin, Kirill and Frindell, Alan and Cenzano, Jordi and Weissman, Jake. RUSH - Reliable (unreliable) streaming protocol. [Online] Available: https://datatracker.ietf.org/doc/draft-kpugin-rush/. Accessed on Dec. 10, 2022.
[13] Y. Zhang, S. Kwong, and S. Wang. Machine learning based video coding optimizations: a survey. *Information Sciences*, 506:395–423, 2020.