Learning Video Transmission via Requirements

Tianchi Huang
Dept. of Computer Science and Technology, Tsinghua University, Beijing, China
htc19@mails.tsinghua.edu.cn

Abstract

Off-the-shelf learning-based video transmission techniques optimize itself towards the linear-combination of several weighted metrics with mutual restriction rather than deterministic requirements, which might finally generalize a strategy that violates the original demand. To eliminate this concern, *Zwei* aims to utilize the fundamental requirement to update the policy. Considering the given requirement often fails to directly provide any gradients, we propose a novel deep learning-based method which can effectively update the network via a binarized signal, symboling which one is closer to the assigned requirement, between two trajectories sampled from the same environment. To build Zwei, we have to develop video transmission simulation environments, design adequate neural network models, and invent training methods for dealing with different requirements on various video transmission scenarios. As expected, trace-driven analysis over three representative tasks shows that Zwei optimizes itself according to the assigned requirement faithfully, and rivals or outperforms the best existing scheme.

1 Introduction

Thanks to the dynamic growth of video encoding technologies and basic Internet services [8], currently we are living with the great help of video transmission services. For example, users often watch interesting videos or live streaming from video content providers, such as Youtube and Kuaishou. Also, they prefer communicating with each other via real-time video streaming rather than a phone call. Technically, users require video streaming with high bitrate and less rebuffering time, while high bitrate may increase the probabilities of the rebuffering event. Moreover, video content providers aim to provide less stalling ratio video streaming services with lower costs, where it's also necessary to trade off the stalling ratio against the cost. The observations above are being left on the horns of a classic dilemma: both qualities of experience and service are evaluated with contradicted metrics. Hence, how to bridge the gap between the positive and negative factors?

Unfortunately, as much as the fundamental problem has already been published about *two decades* [26], current approaches, either heuristics or learning-based methods, fall short of achieving this goal. On the one hand, heuristic-based schemes often use existing models or specific domain knowledge as the designing principle [15, 17, 22]. However, though most solutions are quite simple and effective, they sometimes require careful tuning and will backfire under the circumstance that violated with presumptions. To this end, heuristics like MPC [35] fail to perform well under all considered scenarios [21]. On the other hand, learning-based methods [13, 18] leverage deep reinforcement learing (DRL) for training a neural network (NN) by interacting with the environments towards a reward function without any presumptions, where the function is often defined as a linear-based equation with the combination of weighted sum metrics. Unsurprisingly, the learned policy outperforms heuristics, with the improvements on the overall performance of more than 18% [5]. Nevertheless, we empirically illustrate the weakness of existing learning-based approaches: an inaccurate reward function may mislead the algorithm to generalize the bad strategies. But considering the diversity

Preprint. Under review.

of real-world network environments, we can hardly present an accurate reward function that fits *all* considered network conditions.

Taking a look from another perspective, we observe that the aforementioned problem can be written as a deterministic goal or requirement [12]. E.g., *the goal of the adaptive streaming algorithm is to achieve lower rebuffering time first, and next, reaching higher bitrate [19].* It is pretty straightforward that it can be easily understood and refined by others. To this end, we attempt to train the NN based on the assigned requirement without reward engineering. Unfortunately, off-the-shelf learning-based algorithms are unable to optimize the policy like this since it lacks the abilities to provide any gradients information to guide the algorithm towards a better performance directly. *Hence, we ask if machine learning can help taming the complexity of video transmission services with the actual requirement and especially, without reward engineering.*

Inspired by this opportunity, we envision a learning-based approach called Zwei [1], which can be viewed as a solution for tackling the video transmission dilemma. Unlike previously proposed methods, Zwei trains the NN by answering one question: given two strategies collected from the same environment, which one is closer to the actual demand? Following the answer to this question, Zwei then marks the closer one as *win* and the other one as *loss*. Finally, Zwei will update the NN via increasing the probabilities of the winning sample and reducing the possibilities of the failure sample. Based on this thought, we further add NN-based baseline into Zwei to enhance its overall performance.

In the rest of the paper, we attempt to evaluate the potential of Zwei using trace-driven analysis of various representative video transmission scenarios. To achieve this, we build several faithful video transmission simulators which can accurately replicate the environment via real-world network dataset. Specifically, we validate Zwei on three different tasks, including client-to-server, server-to-client, and client-to-client service. Note that each of them has individual requirements and challenges. As expected, experimental results demonstrate the superiority of Zwei against existing state-of-the-art approaches on all tasks. In particular, we find that Zwei perfectly follows the guidance of the given requirement since it demonstrates two distinct stages during the training process.

Contributions: This paper makes three key contributions.

- We point out the shortcoming of learning-based schemes in video transmission tasks and present the idea that update networks without reward engineering.
- We then propose a novel learning-based method Zwei to make the idea practical.
- Results illustrate that Zwei works perfectly well on all considered video transmission scenarios.

References

[1] Dash industry forum | catalyzing the adoption of mpeg-dash, 2019.

[2] Http live streaming. `https://developer.apple.com/streaming/`, 2019.

[3] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning. In *OSDI*, volume 16, 2016.

[4] V. K. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-L. Zhang. Unreeling netflix: Understanding and improving multi-cdn movie delivery. In *INFOCOM'12*, 2012.

[5] Z. Akhtar, P. Adria, M. Mirza, et al. Oboe: auto-tuning video abr algorithms to network conditions. In *SIGCOMM 2018*, 2018.

[6] A. Bentaleb, B. Taani, A. C. Begen, C. Timmerer, and R. Zimmermann. A survey on bitrate adaptation schemes for streaming media over http. *IEEE Communications Surveys & Tutorials*, 2018.

[7] G. Carlucci, L. De Cicco, S. Holmer, and S. Mascolo. Analysis and design of the google congestion control for web real-time communication (webrtc). In *MMSys'16*, 2016.

[1] Zwei (German: *two*)

[8] Cisco. Cisco visual networking index: Forecast and methodology, 2016-2021, 2017.

[9] A. Dethise, M. Canini, and S. Kandula. Cracking open the black box: What observations can tell us about reinforcement learning agents. In *Proceedings of the 2019 Workshop on Network Meets AI & ML*, 2019.

[10] A. Elo. *The rating of chessplayers, past and present*. Arco Pub., 1978.

[11] I. J. Goodfellow, M. Mirza, D. Xiao, A. Courville, and Y. Bengio. An empirical investigation of catastrophic forgetting in gradient-based neural networks. *arXiv preprint arXiv:1312.6211*, 2013.

[12] T. Huang, X. Yao, C. Wu, R.-X. Zhang, and L. Sun. Tiyuntsong: A self-play reinforcement learning approach for abr video streaming. *arXiv preprint arXiv:1811.06166*, 2018.

[13] T. Huang, R.-X. Zhang, C. Zhou, and L. Sun. Qarc: Video quality aware rate control for real-time video streaming based on deep reinforcement learning. In *Multimedia'18*, 2018.

[14] T. Huang, C. Zhou, R.-X. Zhang, C. Wu, X. Yao, and L. Sun. Comyco: Quality-aware adaptive video streaming via imitation learning. *arXiv preprint arXiv:1908.02270*, 2019.

[15] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson. A buffer-based approach to rate adaptation: Evidence from a large video streaming service. *SIGCOMM'14*, 44(4), 2015.

[16] J. Jiang, V. Sekar, and H. Zhang. Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive. *TON*, 22(1), 2014.

[17] J. Jiang, S. Sun, V. Sekar, and H. Zhang. Pytheas: Enabling data-driven quality of experience optimization using group-based exploration-exploitation. In *NSDI*, volume 1, 2017.

[18] H. Mao, P. Adria, M. Mirza, et al. Neural adaptive video streaming with pensieve. In *SIGCOMM'17*, 2017.

[19] H. Mao, S. Chen, D. Dimmery, S. Singh, D. Blaisdell, Y. Tian, M. Alizadeh, and E. Bakshy. Real-world video adaptation with reinforcement learning. 2019.

[20] H. Mao, P. Negi, A. Narayan, H. Wang, J. Yang, H. Wang, R. Marcus, R. Addanki, M. Khani, S. He, et al. Park: An open platform for learning augmented computer systems. 2019.

[21] H. Mao, S. B. Venkatakrishnan, M. Schwarzkopf, and M. Alizadeh. Variance reduction for reinforcement learning in input-driven environments. *ICLR'19*, 2019.

[22] J. Mo, R. J. La, V. Anantharam, and J. Walrand. Analysis and comparison of tcp reno and vegas. In *INFOCOM'99*, volume 3, 1999.

[23] R. Netravali, A. Sivaraman, S. Das, A. Goyal, K. Winstein, J. Mickens, and H. Balakrishnan. Mahimahi: Accurate record-and-replay for {HTTP}. In *ATC'15)*, 2015.

[24] A. Panse, T. Madheshia, A. Sriraman, and S. Karande. Imitation learning on atari using non-expert human annotations. 2018.

[25] P. G. Pereira, A. Schmidt, and T. Herfet. Cross-layer effects on training neural algorithms for video streaming. In *Proceedings of the 28th ACM SIGMM Workshop on Network and Operating Systems Support for Digital Audio and Video*, 2018.

[26] R. Rejaie, M. Handley, D. Estrin, et al. Quality adaptation for congestion controlled video playback over the internet. In *SIGCOMM'99*, 1999.

[27] M. F. B. Report. Raw data measuring broadband america 2016. https://www.fcc.gov/reports-research/reports/measuring-broadband-america/raw-data-measuring-broadband-america-2016, 2016. [Online; accessed 19-July-2016].

[28] H. Riiser, P. Vigmostad, C. Griwodz, and P. Halvorsen. Commute path bandwidth traces from 3g networks: analysis and applications. In *Proceedings of the 4th ACM Multimedia Systems Conference*, 2013.

[29] D. Rossi et al. Ledbat: The new bittorrent congestion control protocol. In *ICCCN*, 2010.

[30] A. Singh, L. Yang, K. Hartikainen, C. Finn, and S. Levine. End-to-end robotic reinforcement learning without reward engineering. *arXiv preprint arXiv:1904.07854*, 2019.

[31] K. Spiteri et al. Bola: Near-optimal bitrate adaptation for online videos. In *INFOCOM'16*, 2016.

[32] K. Spiteri, R. Sitaraman, and D. Sparacio. From theory to practice: improving bitrate adaptation in the dash reference player. In *Proceedings of the 9th MMSys*, 2018.

[33] Y. Tang. Tf. learn: Tensorflow's high-level module for distributed machine learning. *arXiv preprint arXiv:1612.04251*, 2016.

[34] K. Winstein, A. Sivaraman, and H. Balakrishnan. Stochastic forecasts achieve high throughput and low delay over cellular networks. 2013.

[35] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli. A control-theoretic approach for dynamic adaptive video streaming over http. In *SIGCOMM'15*, 2015.

[36] R.-X. Zhang, T. Huang, M. Ma, H. Pang, X. Yao, C. Wu, and L. Sun. Enhancing the crowd-sourced live streaming: a deep reinforcement learning approach. In *NOSSDAV'19*, 2019.