Proof of APL

Haodan Zhang[1], Yixuan Ban[1], Xinggong Zhang[1,*], Zongming Guo[1]
Zhimin Xu[2], Shengbin Meng[2], Junlin Li[2], Yue Wang[2]
[1]Wangxuan Institute of Computer Technology, Peking University, Beijing, P.R. China
[2]Beijing Bytedance Technology Co., Ltd., Beijing, P.R. China
{pkuzhd, banyixuan, zhangxg, guozongming}@pku.edu.cn

I. Proof

Our objective is to solve:

Problem 1.

$$\arg\max_{\{a_{k,l}\}} \frac{\lambda I_k - \sum_{l=l_k}^{l_k+N} w_{l-l_k} \cdot Q_{k,l} \cdot a_{k,l} \cdot p}{T_k}. \tag{1}$$

The proof of **Theorem 1** need **Lemma 1**:

Lemma 1. *The one-slot Lyapunov drift is upper bounded:*

$$\begin{aligned}\Delta \boldsymbol{Q}_k \leq & \frac{p^2+\Psi}{2}\sum_{n=0}^{N} w_n \\ & - \sum_{l=l_k}^{l_k+N} w_{l-l_k} Q_{k,l} (T_{k,l}^{play} - a_{k,l}\cdot p)\end{aligned} \tag{2}$$

where Ψ is the upper bound of $\mathbb{E}\{T_k^2\}$ under any control algorithm.

Proof. We can derive the expression of $\Delta \boldsymbol{Q}_k$ as follows:

$$\begin{aligned}\Delta \boldsymbol{Q}_k = & \frac{1}{2}\sum_{l=l_k}^{l_k+N} (w_{l-l_{k+1}} Q_{k+1,l}^2 - w_{l-l_k} Q_{k,l}^2) \\ = & \frac{1}{2}\sum_{l=l_k}^{l_k+N} [(w_{l-l_{k+1}} - w_{l-l_k}) Q_{k+1,l}^2 \\ & + w_{l-l_k}(Q_{k+1,l}^2 - Q_{k,l}^2)] \\ \leq & \frac{1}{2}\sum_{l=l_k}^{l_k+N} w_{l-l_k}(Q_{k+1,l}^2 - Q_{k,l}^2)\end{aligned} \tag{3}$$

For $Q_{k+1,l}^2 - Q_{k,l}^2$, we consider two cases: $Q_{k,l} \geq T_{k,l}^{play}$ and $Q_{k,l} < T_{k,l}^{play}$. In the first case we have

$$\begin{aligned} & Q_{k+1,l}^2 - Q_{k,l}^2 \\ = & (Q_{k,l} - T_{k,l}^{play} + a_{k,l}\cdot p)^2 - Q_{k,l}^2 \\ = & Q_{k,l}^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p) + (T_{k,l}^{play} - a_{k,l}\cdot p)^2 - Q_{k,l}^2 \\ = & (T_{k,l}^{play} - a_{k,l}\cdot p)^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p) \\ = & (T_{k,l}^{play})^2 - 2T_{k,l}^{play}\cdot a_{k,l}\cdot p + a_{k,l}^2\cdot p^2 \\ & - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p) \\ \leq & (T_{k,l}^{play})^2 + a_{k,l}^2\cdot p^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p) \\ \leq & \Psi + p^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p)\end{aligned} \tag{4}$$

In the second case we have

$$\begin{aligned} & Q_{k+1,l}^2 - Q_{k,l}^2 \\ = & a_{k,l}^2\cdot p^2 - Q_{k,l}^2 \\ \leq & a_{k,l}^2\cdot p^2 - Q_{k,l}^2 + (Q_{k,l} - T_{k,l}^{play})^2 \\ \leq & a_{k,l}^2\cdot p^2 - 2Q_{k,l}\cdot T_{k,l}^{play} + (T_{k,l}^{play})^2 \\ \leq & (T_{k,l}^{play})^2 + a_{k,l}^2\cdot p^2 - 2Q_{k,l}\cdot T_{k,l}^{play} + 2Q_{k,l}\cdot a_{k,l}\cdot p \\ \leq & \Psi + p^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p)\end{aligned} \tag{5}$$

In both cases, $Q_{k+1,l}^2 - Q_{k,l}^2$ is bounded by

$$\begin{aligned} & Q_{k+1,l}^2 - Q_{k,l}^2 \\ & \leq \Psi + p^2 - 2Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p)\end{aligned} \tag{6}$$

where Ψ is the upper bound of $\mathbb{E}\{T_k^2\}$ under any control algorithm and is assumed to be finite. Sum on $l \in \{l_k, l_k+1, ..., l_k+N\}$, we get

$$\begin{aligned}\Delta \boldsymbol{Q}_k \leq & \frac{1}{2}\sum_{l=l_k}^{l_k+N} w_{l-l_k}(Q_{k+1,l}^2 - Q_{k,l}^2) \\ \leq & \frac{p^2+\Psi}{2}\sum_{n=0}^{N} w_n \\ & - \sum_{l=l_k}^{l_k+N} w_{l-l_k} Q_{k,l}(T_{k,l}^{play} - a_{k,l}\cdot p)\end{aligned} \tag{7}$$

□

Using **Lemma 1**, we can minimize $\Delta \boldsymbol{Q}_k$ by minimizing $\sum_{l=l_k}^{l_k+N} w_{l-l_k}\cdot Q_{k,l}\cdot a_{k,l}$.

Theorem 1. *The playback smoothness in our algorithm satisfies:*

$$S^{\dagger} \geq S^{*} - \frac{p^2+\Psi}{2\lambda}\sum_{n=0}^{N} w_n \tag{8}$$

where $S^{\dagger}$ is the playback smoothness of our algorithm, S^{*} is the optimal playback smoothness in theory.

Proof. Subtract $\lambda \times I_k^{\dagger}$ from both sides of (2) to get

$$\begin{aligned} &\Delta \boldsymbol{Q}_k - \lambda I_k^{\dagger} \\ \leq &\frac{p^2 + \Psi}{2} \sum_{n=0}^{N} w_n - \lambda I_k^{\dagger} \\ &- \sum_{l=l_k}^{l_k+N} w_{l-l_k} Q_{k,l} (T_{k,l}^{play} - a_{k,l}^{\dagger} \cdot p) \end{aligned} \tag{9}$$

We greedily maximize (1), the solution of which we denote by $I_k^{\dagger}$, $a_{k,l}^{\dagger}$ and $T_k^{\dagger}$. For any other solution I_k, $a_{k,l}$ and T_k, we get

$$\begin{aligned} &\frac{\lambda I_k - \sum_{l=l_k}^{l_k+N} w_{l-l_k} \cdot Q_{k,l} \cdot a_{k,l} \cdot p}{T_k} \\ \leq &\frac{\lambda I_k^{\dagger} - \sum_{l=l_k}^{l_k+N} w_{l-l_k} \cdot Q_{k,l} \cdot a_{k,l}^{\dagger} \cdot p}{T_k^{\dagger}} \end{aligned} \tag{10}$$

Using (10), we can rewrite (9) as

$$\begin{aligned} &\Delta \boldsymbol{Q}_k - \lambda I_k^{\dagger} \\ \leq &\frac{p^2 + \Psi}{2} \sum_{n=0}^{N} w_n - \lambda T_k^{\dagger} \frac{I_k}{T_k} \\ &- \sum_{l=l_k}^{l_k+N} w_{l-l_k} \cdot Q_{k,l} T_k^{\dagger} (\frac{T_{k,l}^{play}}{T_k^{\dagger}} - \frac{a_{k,l} \cdot p}{T_k}) \end{aligned} \tag{11}$$

where $\frac{a_{k,l} \cdot p}{T_k}$ denotes the expected arrival rate and can't exceed $\frac{T_{k,l}^{play}}{T_k^{\dagger}}$ which is the video service rate because the buffer is stable. Thus we have

$$\Delta \boldsymbol{Q}_k - \lambda I_k^{\dagger} \leq \frac{p^2 + \Psi}{2} \sum_{n=0}^{N} w_n - \lambda T_k^{\dagger} \frac{I_k}{T_k} \tag{12}$$

According to Theorem 4.5 in [1], for any $\delta > 0$, there is a policy that satisfies

$$\mathbb{E}\{\frac{I_k}{T_k}\} \geq S^* - \delta \tag{13}$$

Using (13), we can rewrite (12) as

$$\Delta \boldsymbol{Q}_k - \lambda I_k^{\dagger} \leq \frac{p^2 + \Psi}{2} \sum_{n=0}^{N} w_n - \lambda T_k^{\dagger} S^* \tag{14}$$

Summing over all $k \in \{1, 2, ..., K\}$, we get

$$\boldsymbol{Q}_{K+1} - \boldsymbol{Q}_1 - \lambda \sum_{k=1}^{K} I_k^{\dagger} \leq \frac{p^2 + \Psi}{2} \sum_{n=0}^{N} w_n K - \lambda S^* \sum_{k=1}^{K} T_k^{\dagger} \tag{15}$$

Dividing both sides by $\lambda \sum_{k=1}^{K} T_k^{\dagger}$ and taking the limit as $K \to \infty$ yields the bound in (8). □

References

[1] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.