Online Caching with Optimal Switching Regret

Samrat Mukhopadhyay

Joint work with

Abhishek Sinha

Dept. of Electrical Engineering
IIT Madras

ISIT 2021

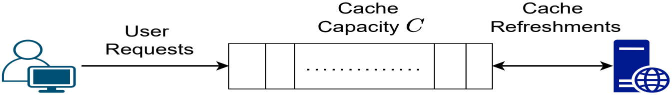
History and Related Literature

- The caching (*a.k.a.* paging, $k-$sets) problem has been studied for more than sixty years and is still a very active area of research
- Two distinct lines of work for single caches:
  - Adversarial requests: minimize the Competitive Ratio
  - Stochastic requests: maximize the hit-rate (*e.g.*, with Zipf's popularity distribution)
- With ever-changing content popularity, the stationarity assumption does not hold in practice
- This work: uncoded caching with adversarial requests to minimize the ~~competitive ratio~~ switching regret using tools from online learning theory

History and Related Literature

- The caching (*a.k.a.* paging, $k-$sets) problem has been studied for more than sixty years and is still a very active area of research
- Two distinct lines of work for single caches:
  - Adversarial requests: minimize the Competitive Ratio
  - Stochastic requests: maximize the hit-rate (*e.g.*, with Zipf's popularity distribution)
- With ever-changing content popularity, the stationarity assumption does not hold in practice
- This work: uncoded caching with adversarial requests to minimize the ~~competitive ratio~~ switching regret using tools from online learning theory

Problem Setup



The online caching setup

- A file server of N distinct files.
- A cache of capacity C.
- A user requests (probably adversarial requests) at most one file per slot
- Cache incurs a download cost to download new files during cache refreshment
- Goal is to design an online caching policy with high hit rates and low switching rates

Setup for Cache Refreshments

- At each time t, the caching algorithm predicts a cache configuration $\boldsymbol{y}_t$.
- The set of admissible caching configuration vectors is
$$\mathcal{Y} = \left\{ \boldsymbol{y} \in \{0,1\}^N : \sum_{f=1}^{N} y_f \leq C \right\}.$$
- The user request vector is $\boldsymbol{x}_t \in \{0,1\}^N$ such that $\sum_{f=1}^{N} x_{t,f} = 1$.
- At slot t, Hit rate is $\boldsymbol{x}_t \cdot \boldsymbol{y}_t$ and switching loss is $\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1$.
- Objective: Maximize the total reward accrued over a time interval of time T:

$$\sum_{t=1}^{T} \underbrace{\boldsymbol{x}_t \cdot \boldsymbol{y}_t}_{\text{Hit rate}} - \frac{D}{2} \sum_{t=2}^{T} \underbrace{\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1}_{\text{Switching loss}}. \quad (1)$$

Regret formulation

- Since the requests might be adversarial, we consider minimizing the switching regret
- The switching regret is defined as the difference between the reward obtained by the best static caching configuration and the reward of an online policy considering switching cost:

$$R_T = \max_{\boldsymbol{y} \in \mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \right) \cdot \boldsymbol{y} - \sum_{t=1}^{T} \boldsymbol{x}_t \cdot \boldsymbol{y}_t + \frac{D}{2} \sum_{t=2}^{T} \|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1$$

The Follow-The-Perturbed-LeaderFTPL based Caching Policy

Algorithm 1 The FTPL Caching Policy

Learning rate $\{\eta_t\}_{t\geq 1}$, switching cost $D \geq 0$, cache capacity C, initial cache-configuration $\boldsymbol{y}_0$
$\boldsymbol{X}_1 \leftarrow \mathbf{0}$
Sample: $\gamma \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.
for $t = 1$ to T **do**
    Cache the top C files corresponding to the perturbed cumulative count vector $\boldsymbol{X}_t + \eta_t \gamma$, *i.e.*,

$$\boldsymbol{y}_t \leftarrow \arg\max_{\boldsymbol{y} \in \mathcal{Y}} \langle \boldsymbol{y}, \boldsymbol{X}_t + \eta_t \boldsymbol{\gamma} \rangle.$$

    User requests a file corresponding to the request vector $\boldsymbol{x}_t$
    The policy receives a reward $q_t = \langle \boldsymbol{y}_t, \boldsymbol{x}_t \rangle - \frac{D}{2} \|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1$.
    Update $\boldsymbol{X}_{t+1} \leftarrow \boldsymbol{X}_t + \boldsymbol{x}_t$.
end for

Review of Online Caching without Cache Refreshments[1] ($D = 0$)

- Lower bound: The regret of any online caching policy is lower bounded as

$$R_T \geq \sqrt{\frac{CT}{2\pi}} - \Theta(\frac{1}{\sqrt{T}}).$$

- Upper-Bound: The FTPL-based caching policy without switching cost and fixed learning rate $\eta_t = \eta$ achieves

$$\mathbb{E}(R_T) \leq 1.51(\log(N/C))^{1/4}\sqrt{CT}.$$

[1] Rajarshi Bhattacharjee, Subhankar Banerjee, and Abhishek Sinha. "Fundamental Limits on the Regret of Online Network-Caching". In: *Abstracts of the 2020 SIGMETRICS/Performance Joint International Conference on Measurement and Modeling of Computer Systems*. 2020, pp. 15–16.

This Paper: Online Caching with Switching, Constant Learning Rate

Regret Upper Bound of FTPL based caching Policy for constant learning rate

With $\eta_t = \eta = \sqrt{T(D+1)/C}(4\pi \ln(N/C))^{-1/4}$, the expected regret of FTPL-based caching policy, including switching cost, is upper bounded as below,

$$\mathbb{E}(R_T) \leq 1.51\sqrt{C(D+1)}(\ln(N/C))^{1/4}\sqrt{T},$$

[a] Amit Daniely and Yishay Mansour. "Competitive ratio vs regret minimization: achieving the best of both worlds". In: *Algorithmic Learning Theory*. 2019, pp. 333–368.

This Paper: Online Caching with Switching, Constant Learning Rate

Regret Upper Bound of FTPL based caching Policy for constant learning rate

With $\eta_t = \eta = \sqrt{T(D+1)/C}(4\pi \ln(N/C))^{-1/4}$, the expected regret of FTPL-based caching policy, including switching cost, is upper bounded as below,

$$\mathbb{E}(R_T) \leq 1.51\sqrt{C(D+1)}(\ln(N/C))^{1/4}\sqrt{T},$$

[a] Daniely and Mansour, "Competitive ratio vs regret minimization: achieving the best of both worlds".

In the above, the expectation is taken w.r.t. the random perturbation γ added by the policy.

This Paper: Online Caching with Switching, Constant Learning Rate

Regret Upper Bound of FTPL based caching Policy for constant learning rate

With $\eta_t = \eta = \sqrt{T(D+1)/C}(4\pi \ln(N/C))^{-1/4}$, the expected regret of FTPL-based caching policy, including switching cost, is upper bounded as below,

$$\mathbb{E}(R_T) \leq 1.51\sqrt{C(D+1)}(\ln(N/C))^{1/4}\sqrt{T},$$

This improves the best known switching regret bound for caching problem by a factor of $\mathcal{O}(\sqrt{C})$![a]

[a] Daniely and Mansour, "Competitive ratio vs regret minimization: achieving the best of both worlds".

In the above, the expectation is taken w.r.t. the random perturbation γ added by the policy.

A Closer Look: Upper Bounding The Switching Loss, $\eta_t = \eta$

$$\mathbb{E}(R_T) = \max_{\boldsymbol{y}\in\mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t\right)\cdot \boldsymbol{y} - \sum_{t=1}^{T}\mathbb{E}(\boldsymbol{x}_t\cdot\boldsymbol{y}_t) + \underbrace{\frac{D}{2}\sum_{t=2}^{T}\mathbb{E}(\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1)}_{\text{Switching Loss}}$$

$$\leq \underbrace{C\eta\sqrt{2\ln(N/C)} + \frac{T}{\eta\sqrt{2\pi}}}_{\text{Bhattacharjee et al.}^2} + ?$$

[2] Bhattacharjee, Banerjee, and Sinha, "Fundamental Limits on the Regret of Online Network-Caching".

[2] Bhattacharjee, Banerjee, and Sinha, "Fundamental Limits on the Regret of Online Network-Caching".

Upper Bounding The Switching Loss, $\eta_t = \eta$

Some Notations:

- $S_t = \{i \in [N] : y_{t,i} = 1\}$, support of cache configuration at time t
- f_t, index of file requested at time t

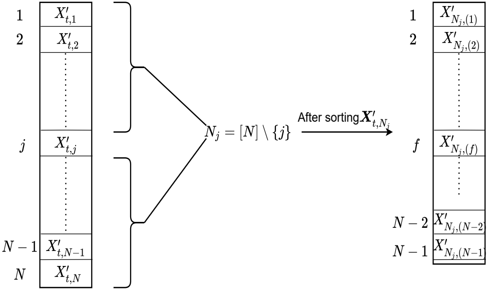
At most one eviction per slot

For FTPL with constant learning rate it is guaranteed that $|S_t \setminus S_{t-1}| \leq 1$.

Consequently,

$$\mathbb{E}\left(\|\boldsymbol{y}_{t+1} - \boldsymbol{y}_t\|_1\right) = 2\mathbb{P}\left(\boldsymbol{y}_{t+1} \neq \boldsymbol{y}_t\right) = 2\mathbb{P}\left(f_t \notin S_t, f_t \in S_{t+1}\right).$$

Upper Bounding $\mathbb{P}(f_t \notin S_t, f_t \in S_{t+1})$, $\eta_t = \eta$: Some Notations



- Perturbed cumulative count of file $f \in [N]$, $X'_{t,f} = X_{t,f} + \eta\gamma_f$,
- For any $j \in [N]$, $N_j = [N] \setminus \{j\}$,
- $X'_{N_j,(f)}$ denotes the f^{th} component of the sorted vector $\mathbf{X}'_t$, in decreasing order, ignoring the j^{th} file.

Upper Bounding $\mathbb{P}\left(f_t \notin S_t, f_t \in S_{t+1}\right), \ \eta_t = \eta$

By the FTPL selection criterion,

$$\mathbb{P}\Big(f_t \in S_{t+1}, f_t \notin S_t\Big)$$
$$= \mathbb{P}\Big(X'_{N_{f_t},(C)} - X_{t,f_t})/\eta \geq \gamma_{f_t} > (X'_{N_{f_t},(C)} - X_{t,f_t})/\eta) - 1/\eta\Big)$$

- The indices $\{f_t\}_{t\geq 1}$ are determined by an agnostic adversary, hence γ_{f_t} and $\gamma_j, \ j \in N_{f_t}$ are stochastically independent
- Therefore, one can derive, using properties of Gaussian distribution,

$$\mathbb{P}\Big(f_t \in S_{t+1}, f_t \notin S_t\Big) \leq \frac{1}{\eta\sqrt{2\pi}}.$$

Taking Everything Together

$$\mathbb{E}(R_T) = \max_{\boldsymbol{y}\in\mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t\right)\cdot \boldsymbol{y} - \sum_{t=1}^{T}\mathbb{E}(\boldsymbol{x}_t\cdot\boldsymbol{y}_t) + \underbrace{\frac{D}{2}\sum_{t=2}^{T}\mathbb{E}(\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1)}_{\text{Switching Loss}}$$

$$\leq \underbrace{C\eta\sqrt{2\ln(N/C)} + \frac{T}{\eta\sqrt{2\pi}}}_{\text{Bhattacharjee et al.}^3} + \underbrace{\frac{DT}{\eta\sqrt{2\pi}}}_{\text{This paper}}$$

$$= C\eta\sqrt{2\ln(N/C)} + \frac{T(D+1)}{\eta\sqrt{2\pi}}$$

Choosing optimal $\eta = \sqrt{T(D+1)/C}(4\pi\ln(N/C))^{-1/4}$ yields the desired result

[3] Bhattacharjee, Banerjee, and Sinha, "Fundamental Limits on the Regret of Online Network-Caching".

[3] Bhattacharjee, Banerjee, and Sinha, "Fundamental Limits on the Regret of Online Network-Caching".

Online Caching with Switching, anytime policy

Regret Upper Bound of FTPL based caching Policy for time varying learning rate

With $\eta_t = \alpha\sqrt{t}, t \geq 1, \alpha = \sqrt{(2+3D)/C}(4\pi \ln(Ne/C))^{-1/4}$,

$$\mathbb{E}(R_T) \leq c_1\sqrt{T} + c_2 \ln T + c_3,$$

where $c_1 = \mathcal{O}(\sqrt{CD}(\ln(Ne/C))^{1/4})$, and c_2, c_3 are small constants depending on N, C.

Online Caching with Switching, anytime policy

Regret Upper Bound of FTPL based caching Policy for time varying learning rate

With $\eta_t = \alpha\sqrt{t}, t \geq 1, \alpha = \sqrt{(2+3D)/C}(4\pi \ln(Ne/C))^{-1/4}$,

$$\mathbb{E}(R_T) \leq c_1\sqrt{T} + c_2 \ln T + c_3,$$

where $c_1 = \mathcal{O}(\sqrt{CD}(\ln(Ne/C))^{1/4})$, and c_2, c_3 are small constants depending on N, C.

In the above, the expectation is taken w.r.t. the random perturbation γ added by the policy.

Bounding the Switching Regret for anytime FTPL policy

$$\mathbb{E}(R_T) = \max_{\boldsymbol{y} \in \mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \right) \cdot \boldsymbol{y} - \sum_{t=1}^{T} \mathbb{E}(\boldsymbol{x}_t \cdot \boldsymbol{y}_t) + \underbrace{\frac{D}{2} \sum_{t=2}^{T} \mathbb{E}(\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1)}_{\text{Switching Loss}}$$

- The first term can be upper bounded by an adaption of the proof of Cohen et al.[4] for time-varying learning rate:

$$\max_{\boldsymbol{y} \in \mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t \right) \cdot \boldsymbol{y} - \sum_{t=1}^{T} \mathbb{E}(\boldsymbol{x}_t \cdot \boldsymbol{y}_t) \leq \eta_1 C \sqrt{2 \log(N/C)} +$$

$$+ \eta_T C \sqrt{2 \ln(Ne/C)} + \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{T} \frac{1}{\eta_t}.$$

[4] Alon Cohen and Tamir Hazan. "Following the perturbed leader for online structured learning". In: *International Conference on Machine Learning*. 2015, pp. 1034–1042.

Bounding the Switching Cost of anytime FTPL

- Bounding the switching cost is trickier: now multiple files can be fetched at a slot because of time varying learning rate with two possibilities at time $t+1$:
  - Fetching the requested file at time t: desirable
  - Fetching any other file: undesirable

Switching cost upper bound for anytime FTPL

For anytime FTPL caching policy with $\eta_t = \alpha\sqrt{t}$:

$$\sum_{t=2}^{T} \mathbb{E}\left(\|\boldsymbol{y}_{t+1} - \boldsymbol{y}_t\|_1\right) \leq \underbrace{\frac{3\sqrt{2}}{\alpha\sqrt{\pi}}\left(\sqrt{T}-1\right)}_{\text{Desirable Switches}}$$

$$+ \underbrace{(N-1)\frac{2+\sqrt{2e\ln(2N)}}{\sqrt{e}}\ln T + \frac{3(N-1)(2+\sqrt{2e\ln(2N)})}{\sqrt{2\pi e}\alpha}}_{\text{Undesirable Switches}}.$$

Taking everything together

$$\mathbb{E}\left(R_T\right) = \max_{\boldsymbol{y}\in\mathcal{Y}} \left(\sum_{t=1}^{T} \boldsymbol{x}_t\right) \cdot \boldsymbol{y} - \sum_{t=1}^{T} \mathbb{E}\left(\boldsymbol{x}_t \cdot \boldsymbol{y}_t\right) + \underbrace{\frac{D}{2} \sum_{t=2}^{T} \mathbb{E}\left(\|\boldsymbol{y}_t - \boldsymbol{y}_{t-1}\|_1\right)}_{\text{Switching Loss}}$$

$$\leq \eta_1 C \sqrt{2\log(N/C)} + \eta_T C \sqrt{2\ln(Ne/C)} + \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{T} \frac{1}{\eta_t}$$

$$+ \frac{3\sqrt{2}}{\alpha\sqrt{\pi}}\left(\sqrt{T} - 1\right) + (N-1)\frac{2 + \sqrt{2e\ln(2N)}}{\sqrt{e}} \ln T$$

$$+ \frac{3(N-1)(2 + \sqrt{2e\ln(2N)})}{\sqrt{2\pi e}\alpha}.$$

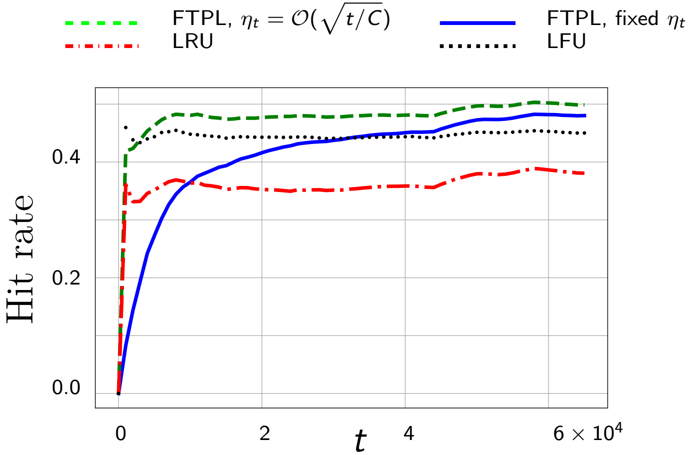
Using the expression of η_t, the inequality $\sum_{t=2}^{T} 1\sqrt{t} \leq 2(\sqrt{T} - 1)$ and choosing optimal $\alpha = \sqrt{(2+3D)/C}(4\pi\ln(Ne/C))^{-1/4}$ results in the desired bound.

Consequence of anytime FTPL: vanishing asymptotic download rate

- Define the Fetch Rate at slot t to be $\mathrm{FR}_t \triangleq \frac{\sum_{\tau=2}^{t} \|\boldsymbol{y}_\tau - \boldsymbol{y}_{\tau-1}\|_1}{t}$, that can be understood as the average download rate
- Then the expected regret upper bound for anytime FTPL, along with the Bounded Convergnece Theorem (BCT) implies that

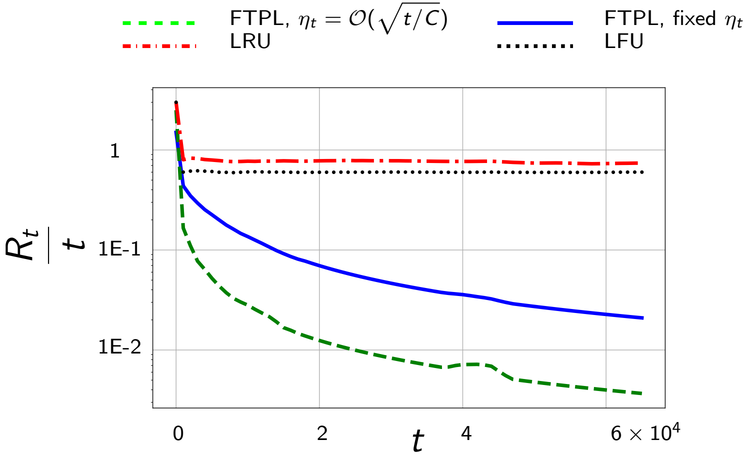
$$\limsup_t \mathrm{FR}_t = 0 \;\; a.s.$$

Numerical Simulations: Hit rate



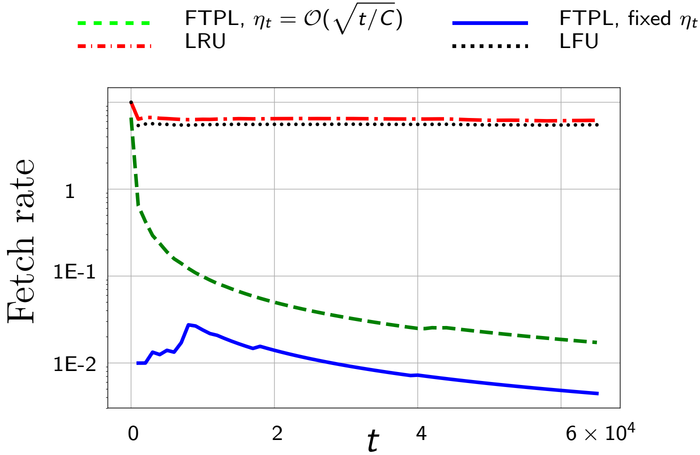
Hit rates comparison, C is 1% of N.

Numerical Simulations: Normalized Regret Rate



Normalized regret rates comparison, C is 1% of N.

Numerical Simulations: Fetch rate



Fetch rates comparison, C is 1% of N.

Conclusions

- We proved that FTPL based caching policy has order optimal switching regret,
- Our result improves best-known bound by a factor of $\mathcal{O}(\sqrt{C})$,
- We prove that the FTPL based anytime caching policy enjoys vanishing asymptotic download rate.

THANK YOU!