

On the Regret of Online Edge Service Hosting

R Sri Prakash, Nikhil Karamchandani, Sharayu Moharir

^aIndian Institute of Technology Bombay, Powai, 400076, Mumbai, India

Abstract

We consider the problem of service hosting where a service provider can dynamically rent edge resources via short term contracts to ensure better quality of service to its customers. The service can also be partially hosted at the edge, in which case, customers' requests can be partially served at the edge. The total cost incurred by the system is modeled as a combination of the rent cost, the service cost incurred due to latency in serving customers, and the fetch cost incurred as a result of the bandwidth used to fetch the code/databases of the service from the cloud servers to host the service at the edge. In this paper, we compare multiple hosting policies with regret as a metric, defined as the difference in the cost incurred by the policy and the optimal policy over some time horizon T . In particular we consider the Retro Renting (RR) and Follow The Perturbed Leader (FTPL) policies proposed in the literature and provide performance guarantees on the regret of these policies. We show that under i.i.d stochastic arrivals, RR policy has linear regret while FTPL policy has constant regret. Next, we propose a variant of FTPL, namely Wait then FTPL (W-FTPL), which also has constant regret while demonstrating much better dependence on the fetch cost. We also show that under adversarial arrivals, RR policy has linear regret while both FTPL and W-FTPL have regret $O(\sqrt{T})$ which is order-optimal.

1. Introduction

Software as a Service (SaaS) instances like online navigation platforms, Video-on-Demand services, etc., have stringent latency constraints in order

Email addresses: prakash.14191@gmail.com; sriprakash@ee.iitb.ac.in (R Sri Prakash), nikhilk@ee.iitb.ac.in (Nikhil Karamchandani), sharayum@ee.iitb.ac.in (Sharayu Moharir)

to provide a good quality of experience to their customers. While most SaaSs use cloud resources, low latency necessitates the use of storage/computational resources at the edge, i.e., close to the end-user. A service is said to be hosted at the edge if the code and databases needed to serve user queries are stored on the edge servers and requests can be served at the edge. In this work, the service can also be partially hosted at the edge, in which case, customers' requests can be partially served at the edge [1, 2]. For instance, consider a SaaS application which provides access to news articles on demand. A typical news article includes some text, some images and possibly some embedded videos. One way to partially host such a service is to store the text corresponding to the news articles at the edge and store the images/videos in the cloud. Each user request will then be served partially by the edge.

We consider the setting where third-party resources can be rented via short-term contracts to host the service, and the edge hosting status of the SaaS can be dynamically changed over time. If the service is not hosted at the edge, it can be fetched from the cloud servers by incurring a fetch cost. The performance of a hosting policy is a function of the rent cost, the fetch cost, and the quality of experience of the users. We refer to the algorithmic challenge of determining what fraction of the service to host at the edge over time as the *service hosting problem*. Note that we study the service hosting problem from the perspective of a specific SaaS provider which can rent third-party edge resources to improve their customers' quality of experience by reducing the latency of their service.

Novel online service hosting policies with provable performance guarantees have been proposed in [3, 4]. The metric of interest in [3, 4] is the *competitive ratio*, defined as the ratio of the cost incurred by an online policy to the cost incurred by an optimal offline policy for the same request arrival sequence. Since the competitive ratio is multiplicative by definition, the absolute value of the difference in the cost incurred by an online policy and the optimal policy can be large even though the competitive ratio of the online policy is close to one. This motivates studying the performance of candidate policies in terms of *regret*, defined as the difference in the cost incurred by the policy and the optimal policy. Regret is a widely used metric in online learning [5], including recently for the caching problem [6, 7], which is closely related to the service hosting problem. In this work, one of our goals is to design hosting policies with provable guarantees in terms of the regret. Another key dimension in the performance guarantees of online policies is the assumption made on the request arrival process. Commonly studied settings

include the *stochastic setting* and the *adversarial setting*, and we provide performance guarantees for both settings.

1.1. Our Contributions

We propose a variant of the FTPL policy called Wait then Follow the Perturbed Leader (W-FTPL). W-FTPL is a randomized policy that takes into account the fetch cost in its decision-making. More specifically, W-FTPL does not fetch the service for an initial wait period which depends on the request arrivals and is an increasing function of the fetch cost. Following the wait period, W-FTPL mimics the FTPL policy.

For i.i.d. stochastic request arrivals and the regret metric, we show that RR is sub-optimal while FTPL and W-FTPL are both order-optimal with respect to the time horizon. While the regret of FTPL can increase linearly with the fetch cost, the regret of W-FTPL increases at most polylogarithmically. The improved performance of W-FTPL over FTPL is a consequence of the fact that W-FTPL avoids most of the fetches made by FTPL in the initial time-slots and by the end of the wait period, its estimate of the arrival rate is accurate enough to avoid multiple fetches. Further, we characterize the competitive ratio of FTPL and W-FTPL for the setting where a finite number of partial hosting levels are allowed. We compare these results with the competitive ratio of RR when a single partial hosting level is allowed [2] to conclude that FTPL and W-FTPL outperform RR with respect to the competitive ratio.

For the adversarial setting, we first characterize a fundamental lower bound on the regret of any online hosting policy. In terms of regret, we then show that RR is strictly suboptimal, while FTPL and W-FTPL have order-optimal performance with respect to time. We show that the competitive ratios of FTPL and W-FTPL are upper bounded by a constant (with respect to time) for any finite number of partial hosting levels. This is a more general setting than the one considered in [2], where the competitive ratio of a variant of RR is characterized when only one partial hosting level is permitted.

1.2. Related work

Several emerging applications such as Augmented / Virtual Reality (AR/VR), online gaming, the Internet of Things (IoT), and autonomous driving have very stringent latency requirements. In addition, many of these applications are also compute-and-storage heavy. This necessitates the migration of some

of the storage and computation requirements of these services from remote cloud servers to the edge. To enable such offloading, several edge computing architectures have been proposed in the literature and their performance has been analyzed extensively, see for example [8, 9, 10] and references therein.

The service hosting problem has received a lot of attention recently, and various settings have been studied. One approach has been to pose the design problem of where to host which service as a one-shot large-scale optimization problem, see for example [11, 12, 13]. Our work differs from this line of work in that we consider an online setting where we adaptively decide when (and what fraction of) service to host at any time, depending on the varying number of requests seen thus far.

There have been recent works [4, 14] which consider the online service hosting problem, and design adaptive schemes whose performance is characterized in terms of the competitive ratio with respect to an oracle which known the entire request sequence apriori. Our work differs from these in a couple of ways. Firstly, the above mentioned works study the problem from the perspective of an edge resource provider who has to decide which amongst a collection of services to host at each time on a storage-constrained edge server. On the other hand, similar to some other works [15, 3], we study the problem from the perspective of a service provider which needs to decide when to host its application at the edge so as to minimize the overall cost. Secondly, most of the works mentioned above do not consider partial hosting of the service, which is allowed in our framework. While [2] does study partial hosting and proposed the Retro-Renting (RR) policy which is shown to have a constant competitive ratio, it only allows one level of partial hosting of the service. On the other hand, we consider a much more general setting where multiple partial service hosting levels are allowed and are able to show that both the FTPL and its variant W-FTPL are able to achieve a constant competitive ratio.

While the works mentioned above study competitive ratio as a performance metric, another popular measure of performance for online algorithms is the regret which considers the difference in the cost incurred by an online policy and the optimal (static) policy. In particular, for the closely related caching problem, several policies inspired by the recent advances in online convex optimization have been proposed, including Online Gradient Ascent [6], Online Mirror Descent [16], and Follow the Perturbed Leader (FTPL) [17, 7]. In particular, FTPL has been shown to have order-optimal regret for the caching problem in the adversarial setting [17, 7]. In this work, we

study regret for the RR and FTPL policies for the service hosting problem under both the stochastic and adversarial settings. Since RR is a deterministic policy, its regret performance in the adversarial setting is poor. While FTPL does achieve order-optimal regret (with respect to time), one limitation is that it makes hosting decisions agnostic of the fetch cost. As a result, in some cases, FTPL is prone to fetching and evicting the service multiple times in the initial time-slots when its estimate of the request arrival rate is noisy, thus leading to poor performance. This motivated our design and regret analysis of W-FTPL, a variant of FTPL which takes into account the fetch cost. One recent work which also considers regret as a performance metric is [18], which studied the problem of joint online service hosting and routing and proposed a two time-scale algorithm based on online gradient descent with provably order-optimal regret under adversarial requests.

Finally, there are several other approaches which have been used to address the online service hosting problem, including Markov Decision Processes [19, 20, 1], Reinforcement Learning [21], and Multi-Armed Bandits [22, 23].

2. System Setup

We consider a system consisting of a back-end server and an edge-server to serve customers of a service. The back-end server always hosts the service and can serve any requests that are routed to it. In this work, we allow the service to be partially hosted at the edge-server. When the service is partially hosted at the edge, requests are partially served at the edge and partially at the back-end server. Further, the fraction of service hosted at the edge can be changed over time. If the service is not hosted or partially hosted at the edge, parts of the service can be fetched from the back-end server to host at the edge.

Sequence of events in each time-slot: We consider a time-slotted system. In each time-slot, we first make the service hosting decision for that time-slot. Following this, requests may arrive and are served at the edge and/or the back-end server.

Request arrivals: We consider two types of arrival processes in this paper.

- *Stochastic arrivals:* In this case, request arrivals are i.i.d. across time-slots with mean μ .

- *Adversarial arrivals*: In this case, the request arrival sequence is generated by an oblivious adversary¹.

Let r_t denote the number of request arrivals in time-slot t . We consider bounded requests, specifically, $0 \leq r_t \leq \kappa$, and let $R_t = \sum_{i=1}^t r_i$ denote the cumulative requests observed by the end of time-slot t . In both cases, we assume that there are at most κ arrivals per slot.

Costs: We model three types of costs.

- *Rent cost* ($\mathcal{C}_{R,t}^{\mathcal{P}}$): The system incurs a cost of c units per time-slot to host the entire service at the edge. For hosting f fraction of service, the system incurs a cost of cf units per time-slot.
- *Service cost* ($\mathcal{C}_{S,t}^{\mathcal{P}}$): This is the cost incurred to use the back-end server to serve (parts of) requests. If f fraction of service is hosted at the edge, the system incurs a service cost of $g(f)$ units per request, where $g(f)$ is a decreasing function of f . We fix $g(0) = 1$, i.e., service cost is one unit when the service is not hosted at the edge.
- *Fetch cost* ($\mathcal{C}_{F,t}^{\mathcal{P}}$): The system incurs a cost of $M > 1$ units for each fetch of the entire service from the back-end server to host on the edge-server. For fetching f fraction of service, the system incurs a cost of Mf units.

Let $\rho_t^{\mathcal{P}}$ denote the edge hosting status of the service in time slot t under policy \mathcal{P} . We consider the setting where $\rho_t^{\mathcal{P}} \in \{\alpha_1, \alpha_2, \dots, \alpha_{K-1}, \alpha_K\}$ and $\rho_t^{\mathcal{P}} = \alpha_i$ implies that α_i fraction of the service is hosted at the edge-server in time-slot t . Note that $\alpha_i \in [0, 1]$ and in particular we consider $\alpha_1 = 0$, $\alpha_K = 1$ and $\alpha_i < \alpha_j$ for $i < j$. It follows that $\rho_t^{\mathcal{P}} = 1$ implies that the entire service is hosted at the edge in time-slot t , and $\rho_t^{\mathcal{P}} = 0$ implies that the service is not hosted at the edge in time-slot t .

We make some assumptions about the various system parameters.

Assumption 1. $c \leq \kappa$.

This assumption is motivated by the fact that if $\kappa < c$, it is strictly sub-optimal to host the entire service at the edge, irrespective of the number of arrivals in a time-slot.

¹The entire request sequence is assumed to be fixed apriori.

Assumption 2. For a fraction of service α_i , $1 \leq i \leq K$, $\alpha_i + g(\alpha_i) \leq 1$.

In [1, 2], it is shown that the benefits of partial hosting are limited to the setting where $\alpha_i + g(\alpha_i) \leq 1$. Motivated by this, we consider the case where $\alpha_i + g(\alpha_i) \leq 1$ for all candidate values of α_i .

Assumption 3. For any fraction of service α_i , $1 \leq i \leq K$, $c\alpha_i + g(\alpha_i)r_t \leq \kappa$

Assumption 3 follows from Assumptions 1 and 2

Assumption 4. For stochastic arrivals, $\mu > 0$ and for adversarial arrivals, $R_T \geq 1$.

This assumption eliminates degenerate cases where there are no request arrivals in the time-horizon of interest.

The total cost incurred in time-slot t by policy \mathcal{P} , denoted by $\mathcal{C}_t^{\mathcal{P}}(r_t)$, is the sum of the rent, service, and fetch costs. It follows that,

$$\mathcal{C}_t^{\mathcal{P}}(r_t) = c\rho_t^{\mathcal{P}} + g(\rho_t^{\mathcal{P}})r_t + M(\rho_t^{\mathcal{P}} - \rho_{t-1}^{\mathcal{P}})^+. \quad (1)$$

Let $r = \{r_t\}_{t \geq 1}$ denote the request arrival sequence and $\mathcal{C}^{\mathcal{P}}(T, r)$ denote the cumulative cost incurred by policy \mathcal{P} in time-slots 1 to T . It follows that,

$$\mathcal{C}^{\mathcal{P}}(T, r) = \sum_{t=1}^T \mathcal{C}_t^{\mathcal{P}}(r_t).$$

Performance metrics: We consider two performance metrics, namely regret and competitive ratio.

For i.i.d. stochastic arrivals, the regret of a policy \mathcal{P} , denoted by $\mathcal{R}_S^{\mathcal{P}}(T)$, is defined as the difference in the total expected cost incurred by the policy and the oracle static hosting policy. The oracle static hosting policy makes a hosting decision at $t = 1$ using the knowledge of the statistics of the request arrival process but not the entire sample-path. For ease of notation, we define $\mu = \mathbb{E}[r_t]$, $\mu_i = c\alpha_i + g(\alpha_i)\mu$. It follows that the expected cost incurred by the optimal static hosting policy in time-slots 1 to T is $\min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\}$, and therefore,

$$\begin{aligned} \mathcal{R}_S^{\mathcal{P}}(T) &= \mathbb{E}_{\mathcal{P}, r}[\mathcal{C}^{\mathcal{P}}(T, r)] - \min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\} \\ &= \mathbb{E}_{\mathcal{P}, r}[\mathcal{C}^{\mathcal{P}}(T, r)] - \min_i \{\mu_i T + M\alpha_i\}, \end{aligned}$$

$$\leq \mathbb{E}_{\mathcal{P},r}[\mathcal{C}^{\mathcal{P}}(T,r)] - \min_i \{\mu_i T\}, \quad (2)$$

where $\mathbb{E}_{\mathcal{P},r}$ represents the expectation over the randomization in policy \mathcal{P} and the request arrival sequences r .

For i.i.d. stochastic arrivals, the competitive ratio of a policy \mathcal{P} , denoted by $\sigma_S^{\mathcal{P}}(T)$, is defined as the ratio of the expected total cost incurred by the policy and expected total cost incurred by the optimal static hosting policy as discussed above. It follows that,

$$\sigma_S^{\mathcal{P}}(T) = \frac{\mathbb{E}_{\mathcal{P},r}[\mathcal{C}^{\mathcal{P}}(T,r)]}{\min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\}}.$$

For adversarial arrivals, the regret of a policy \mathcal{P} , denoted by $\mathcal{R}_A^{\mathcal{P}}(T)$, is defined as the worst case difference in the total expected cost incurred by the policy and the optimal static hosting policy. The expectation is taken over the randomization in policy \mathcal{P} . It follows that for any request sequence r , the total cost incurred by the optimal static hosting policy in time-slots 1 to T is $\min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\}$, and therefore,

$$\begin{aligned} \mathcal{R}_A^{\mathcal{P}}(T,r) &= \mathbb{E}_{\mathcal{P}}[\mathcal{C}^{\mathcal{P}}(T,r)] - \min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\}, \\ \mathcal{R}_A^{\mathcal{P}}(T) &= \sup_r (\mathcal{R}_A^{\mathcal{P}}(T,r)). \end{aligned}$$

We have the following relation between the regret induced by a policy under i.i.d stochastic arrivals and adversarial arrivals.

Lemma 1. *For any online policy \mathcal{P} and stochastic arrivals we have, $\mathcal{R}_S^{\mathcal{P}}(T) \leq \mathcal{R}_A^{\mathcal{P}}(T)$.*

The proof can be found in Section 7.

For adversarial arrivals, the competitive ratio of a policy \mathcal{P} , denoted by $\sigma_A^{\mathcal{P}}(T)$, is defined as the worst case ratio of the expected total cost incurred by the policy and the cost incurred by the offline optimal policy. The optimal offline policy can change the hosting status over time using the knowledge of the entire request process a priori. Let the cost incurred by the optimal offline hosting policy in time-slots 1 to T be denoted by $\mathcal{C}^{\text{OFF-OPT}}(T,r)$. It follows that,

$$\sigma_A^{\mathcal{P}}(T) = \sup_r \frac{\mathbb{E}_{\mathcal{P}}[\mathcal{C}^{\mathcal{P}}(T,r)]}{\mathcal{C}^{\text{OFF-OPT}}(T,r)}.$$

Goal: The goal is to design online hosting policies with provable performance guarantees with respect to regret and competitive ratio.

Notations: We define some notations which will be used in the rest of this paper. Let $\mathbf{s} = [0, \alpha_2, \alpha_3, \dots, 1]'$, $\mathbf{f} = [1, g(\alpha_2), g(\alpha_3), \dots, 0]'$, and $\boldsymbol{\rho}_t^{\mathcal{P}} \in \mathcal{X}$, where \mathcal{X} is the collection of all possible one hot vectors in $\{0, 1\}^K$ and thus $|\mathcal{X}| = K$. The position of 1 in the one hot vector $\boldsymbol{\rho}_t^{\mathcal{P}}$ represents the level of service hosted at the edge in slot t , that is $\rho_t^{\mathcal{P}} = \langle \boldsymbol{\rho}_t^{\mathcal{P}}, \mathbf{s} \rangle$. The total cost in slot t as given in (1) can be rewritten using the above notations as follows,

$$\begin{aligned} \mathcal{C}_t^{\mathcal{P}}(r_t) &= c\langle \boldsymbol{\rho}_t^{\mathcal{P}}, \mathbf{s} \rangle + r_t\langle \boldsymbol{\rho}_t^{\mathcal{P}}, \mathbf{f} \rangle + M(\rho_t^{\mathcal{P}} - \rho_{t-1}^{\mathcal{P}})^+, \\ &= \langle \boldsymbol{\rho}_t^{\mathcal{P}}, \boldsymbol{\theta}_t \rangle + M(\rho_t^{\mathcal{P}} - \rho_{t-1}^{\mathcal{P}})^+, \end{aligned}$$

where $\boldsymbol{\theta}_t = c\mathbf{s} + r_t\mathbf{f}$.

For ease of notation, for stochastic arrivals we define $\mu = \mathbb{E}[r_t]$, $\mu_i = c\alpha_i + g(\alpha_i)\mu$, $\Delta_{ij} = \mu_i - \mu_j$, $i^* = \arg \min_i \mu_i$, $\Delta_i = \mu_i - \mu_{i^*}$, $\Delta_{\min} = \min_i \Delta_i$, $\Delta_{\max} = \max_i \Delta_i$.

We summarize important notations used in this paper in Table 1.

3. Policies

Our Policy: Our policy called Wait then Follow the Perturbed Leader (W-FTPL), is a variant of the popular FTPL policy. A formal definition of FTPL is given in Algorithm 1. FTPL is a randomized policy and is known to perform well for the traditional caching problem, in fact achieving order-wise optimal regret for adversarial arrivals [7]. Under FTPL, in any time slot t , for each i , FTPL considers the cumulative cost that would have been incurred had the system used α_i hosting fraction in the entire duration $[1, t-1]$, and then perturbs this by an appropriately scaled version of a sample of a standard Gaussian random variable. FTPL then hosts the fraction of service for which the perturbed cost is minimum in that slot.

We propose a policy called Wait then Follow The Perturbed Leader (W-FTPL). The key idea behind the W-FTPL policy is to not host the service for an initial wait period. This is to reduce the number of fetches made initially when the estimate of the arrivals is noisy. The duration of this wait period is not fixed apriori and is a function of the arrival pattern seen till that time. Following the wait period, W-FTPL mimics the FTPL policy. Refer to Algorithm 2 for a formal definition of W-FTPL. Let T_s be the time slot, after which W-FTPL starts acting as FTPL. Formally we define $T_s = \min\{t : t < \frac{(\max_{i \neq j} (\Theta_{t+1,i} - \Theta_{t+1,j}))^2}{\kappa^2 \beta (\log M)^{1+\delta}}\}$, where $\beta > 1$, $\delta > 0$ are constants.

Symbol	Description
c	Rent cost for hosting complete service in a slot
M	Fetch cost
K	Number of hosting levels allowed
t	Time index
$\rho_t^{\mathcal{P}}$	Fraction of service hosted in time slot t under policy \mathcal{P}
κ	Maximum number of arrivals in a slot
α_i	Fraction of service
$g(f)$	Fraction of request forwarded if f fraction is hosted
r_t	Number of requests in time slot t
R_t	Cumulative number of requests up to time slot t
μ	$\mathbb{E}[r_t]$
T	Time horizon
$\mathcal{C}^{\mathcal{P}}(T, r)$	Cost incurred till T under policy \mathcal{P} for request sequence r
$\mathcal{R}_S^{\mathcal{P}}(T)$	Regret under stochastic arrival setting
$\sigma_S^{\mathcal{P}}(T)$	Competitive ratio under stochastic arrival setting
$\mathcal{R}_A^{\mathcal{P}}(T)$	Regret under adversarial arrival setting
$\sigma_A^{\mathcal{P}}(T)$	Competitive ratio under adversarial arrival setting
μ_i	$c\alpha_i + g(\alpha_i)\mu$
Δ_{ij}	$\mu_i - \mu_j$
i^*	$\arg \min_i \mu_i$
Δ_i	$\mu_i - \mu_{i^*}$
Δ_{\min}	$\min_i \Delta_i$
Δ_{\max}	$\max_i \Delta_i$
\mathbf{s}	$[0, \alpha_2, \alpha_3, \dots, 1]'$
\mathbf{f}	$[1, g(\alpha_2), g(\alpha_3), \dots, 0]'$
$\boldsymbol{\theta}_t$	$c\mathbf{s} + r_t\mathbf{f}$
$\boldsymbol{\Theta}_t$	$\sum_{i=1}^{t-1} \boldsymbol{\theta}_i$
η_t	Learning rate
α	Hyper parameter in FTPL, W-FTPL policies
β, δ	Hyperparameters in W-FTPL policy
$\boldsymbol{\gamma}$	IID Gaussian vector
\mathcal{X}	Collection of one hot vectors in $\{0, 1\}^K$
T_s	Waiting time for W-FTPL policy

Table 1: Notations

Algorithm 1: Follow The Perturbed Leader (FTPL)

Input: $c, \{\eta_t\}_{t \geq 1}, \{r_t\}_{t \geq 1}$
1 Set $\Theta_1 \leftarrow \mathbf{0}$
2 Sample $\gamma \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
3 **for** $t = 1$ **to** T **do**
4 host $\rho_t \in \arg \min_{\rho \in \mathcal{X}} \langle \rho, \Theta_t + \eta_t \gamma \rangle$
5 $\theta_t = c\mathbf{s} + r_t \mathbf{f}$
6 Update $\Theta_{t+1} = \Theta_t + \theta_t$

Algorithm 2: Wait then Follow The Perturbed Leader (W-FTPL)

Input: $c, M, \{\eta_t\}_{t \geq 1}, \{r_t\}_{t \geq 1}, \beta, \delta, \kappa$
1 Set $\Theta_1 \leftarrow \mathbf{0}$, wait=1
2 Sample $\gamma \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
3 **for** $t = 1$ **to** T **do**
4 **if** $wait = 1$ **then**
5 $\rho_t = \mathbf{0}$
6 **else**
7 host $\rho_t \in \arg \min_{\rho \in \mathcal{X}} \langle \rho, \Theta_t + \eta_t \gamma \rangle$
8 $\theta_t = c\mathbf{s} + r_t \mathbf{f}$
9 Update $\Theta_{t+1} = \Theta_t + \theta_t$
10 wait = $\min\{\text{wait}, \mathbb{1}_{t \geq \frac{(\max_{i \neq j} (\Theta_{t+1,i} - \Theta_{t+1,j}))^2}{\kappa^2 \beta (\log M)^{1+\delta}}}\}$

Retro-Renting (RR) [3]: The RR policy is a deterministic hosting policy, and it either hosts the complete service or hosts nothing in a slot. The key idea behind this policy is to use recent arrival patterns to make hosting decisions. We refer the reader to Algorithm 1 in [3] for a formal definition of the RR policy. The performance of RR with respect to the competitive ratio was analyzed in [3].

α -RR [2] is a variant of RR where one partial level of hosting the service is allowed. The formal definition of α -RR, borrowed from [2, Algorithm 1] is given in Algorithm 3. Similar to RR, the key idea behind α -RR policy is to check whether the current hosting status is optimal in the hindsight using recent arrival patterns to make hosting decisions. α -RR checks the optimal offline policy cost for the recent request arrivals and chooses the fraction of service with the minimum cost. For $\rho_t^{\mathcal{P}} \in \{0, 1\}$, α -RR behaves as RR [2]. Therefore we refer α -RR as RR in this paper.

4. Main Results and Discussion

In this section, we state and discuss our key results.

4.1. Regret Analysis

Our first result characterizes the regret performance of the policies discussed in Section 3 and the fundamental limit on the performance of any online policy for adversarial arrivals.

Theorem 1 (Adversarial Arrivals). *Let the arrivals be generated by an oblivious adversary under the constraint that at most κ requests arrive in each time-slot. Then,*

(a) $\mathcal{R}_A^{\mathcal{P}}(T) = \Omega(\sqrt{T})$ for any online policy \mathcal{P} .

(b) $\mathcal{R}_A^{RR}(T) = \Omega(T)$.

(c) For $\eta_t = \alpha\sqrt{t}$, $\alpha > 0$,

$$\mathcal{R}_A^{FTPL}(T) \leq \sqrt{2T \log K} \left(\alpha + \frac{4\kappa^2}{\alpha} \right) + \frac{K^2 M(c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T + 1}.$$

(d) For $\eta_t = \alpha\sqrt{t}$, $\alpha > 0$,

$$\mathcal{R}_A^{FTPL}(T) \geq M\alpha_2/K.$$

Algorithm 3: α -RetroRenting (α -RR)

```

1 Input: Fetch cost  $M$ , partial hosting level  $\alpha_2$ , latency cost under
   partial hosting  $g(\alpha_2)$ , rent cost  $c$ , request arrival sequence  $\{x_l\}_{l \geq 0}^t$ 
2 Output: service hosting strategy  $r_{t+1}$ ,  $t > 0$ 
3 Initialize:  $r_1 = t_{\text{recent}} = 0$ 
4 for each time-slot  $t$  do
5    $I_t = (M, g(\alpha_2), t, t_{\text{recent}}, c, \{x_l\}_{l \geq t_{\text{recent}}}^t)$ 
6    $R_0^{(\tau_0)} = [\underbrace{r_t, r_t, \dots, r_t}_{\tau_0 - t_{\text{recent}}}, \underbrace{0, 0, \dots, 0}_{t - \tau_0}]$ 
7    $R_\alpha^{(\tau_\alpha)} = [\underbrace{r_t, r_t, \dots, r_t}_{\tau_\alpha - t_{\text{recent}}}, \underbrace{\alpha_2, \alpha_2, \dots, \alpha_2}_{t - \tau_\alpha}]$ 
8    $R_1^{(\tau_1)} = [\underbrace{r_t, r_t, \dots, r_t}_{\tau_1 - t_{\text{recent}}}, \underbrace{1, 1, \dots, 1}_{t - \tau_1}]$ 
9    $\text{minCost}(0) = \min_{\tau_0 \in (t_{\text{recent}}, t)} \text{totalCost}(R_0^{(\tau_0)}, I_t)$ 
10   $\text{minCost}(\alpha_2) = \min_{\tau_\alpha \in (t_{\text{recent}}, t)} \text{totalCost}(R_\alpha^{(\tau_\alpha)}, I_t)$ 
11   $\text{minCost}(1) = \min_{\tau_1 \in (t_{\text{recent}}, t)} \text{totalCost}(R_1^{(\tau_1)}, I_t)$ 
12   $r_{t+1} = \arg \min_{i \in \{0, \alpha_2, 1\}} \text{minCost}(i)$ 
13  if  $r_{t+1} \neq r_t$  then
14     $t_{\text{recent}} = t$ 
15  end
16 end
17 Function  $\text{totalCost}(R, I_t)$ :
18    $g(0) = 1, g(1) = 0;$ 
19    $\text{cost} = R(1) \times c + x_1 \times g(R(1));$ 
20   for  $j \leftarrow 2$  to  $t - t_{\text{recent}}$  do
21      $\text{cost} = \text{cost} + R(j) \times c + x_j \times g(R(j))$ 
22      $\quad + M \times |R(j) - R(j-1)|;$ 
23   end
24   return  $\text{cost};$ 
25 end

```

(e) For $\eta_t = \alpha\sqrt{t}$, $\alpha > 0$, $\beta > 1$, $\delta > 0$,

$$\begin{aligned}\mathcal{R}_A^{W\text{-}FTPL}(T) &\leq \sqrt{\beta T (\log M)^{1+\delta}} + \sqrt{2T \log K} \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\ &\quad + \frac{K^2 M (c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T+1}.\end{aligned}$$

The key take-away from Theorem 1 is that RR has linear regret and FTPL, W-FTPL have order-optimal regret, i.e., $O(\sqrt{T})$ with respect to the time horizon under adversarial arrivals. The proof of Theorem 1 is in Section 6.

Our second result characterizes the regret performance of the policies discussed in Section 3 for i.i.d. stochastic arrivals.

Theorem 2 (Stochastic Arrivals). *Let the arrivals in each time-slot be i.i.d. stochastic with mean μ and $\Delta_{\min} \neq 0$. Then we have*

(a) $\mathcal{R}_S^{RR}(T) = \Omega(T)$.

(b) For $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$,

$$\begin{aligned}\mathcal{R}_S^{FTPL}(T) &\leq \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) + \frac{16\alpha^2 + 4\kappa^2}{\Delta_{\min}} \\ &\quad + (16\alpha^2 + 3\kappa^2) \frac{M}{\Delta_{\min}^2},\end{aligned}$$

where $h_1 = 4 \max\{8\alpha^2, \kappa^2\}$.

(c) For $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$, $\beta > 1$, $\delta > 0$, and M large enough we have,

$$\begin{aligned}\mathcal{R}_S^{W\text{-}FTPL}(T) &\leq 1 + \beta\kappa^2 (\log M)^{1+\delta} \left(\frac{4}{\Delta_{\min}^2} + \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2 \Delta_{\max}^2} \right) \\ &\quad + (16\alpha^2 + 4\kappa^2) \left(\frac{1}{\Delta_{\min}} + \frac{1}{\Delta_{\min}^2} \right) \\ &\quad + \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right),\end{aligned}$$

where $h_1 = 4 \max\{8\alpha^2, \kappa^2\}$.

Corollary 1 (Stochastic Arrivals). *Let the arrivals in each time-slot be i.i.d. stochastic with mean μ and $\Delta_{\min} \neq 0$. For the W-FTPL policy, if $\rho_t^P \in \{0, 1\}$, $\beta = \max\{(1 + 4\alpha/\kappa)^2, (1 + \sqrt{2}/\kappa)^2\}$, then the regret scales at most logarithmically with M .*

Remark 1. *While the result in 2 (c) is stated for $\delta > 0$, it can be shown that for $\delta = 0$ and β chosen suitably large, the regret of W-FTPL scales logarithmically with M .*

The key take-aways from Theorem 2 are that RR has linear regret with respect to time, and FTPL/W-FTPL has regret that does not scale with time. Also, the upper bound on the regret of W-FTPL scales as $(\log M)^{1+\delta}$ for $\delta \geq 0$ where as for FTPL, the upper bound on the regret scales linearly with M . W-FTPL thus outperforms FTPL w.r.t. to the dependence of regret on the fetch cost M . We validate these results via simulations in Section 5.

4.2. Competitive Ratio Analysis

Our next result characterizes the competitive ratio of the policies discussed in Section 3. The competitive ratio of RR is characterized in [3, 2], and it is shown that RR has a constant competitive ratio and the competitive ratio decreases with an increase in fetch cost M . We present the results for FTPL and W-FTPL policies below.

Theorem 3 (Adversarial Arrivals). *Let the arrivals be generated by an oblivious adversary under the constraint that at most κ requests arrive in each time-slot. Then we have*

(a) For $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$,

$$\begin{aligned} \sigma_A^{FTPL}(T) &\leq \frac{\kappa^2(3 + 2M/c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \\ &\quad + \frac{\kappa^2(M + c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \sum_{i=2}^K \frac{16\alpha^2}{c^2\alpha_i^2}. \end{aligned}$$

(b) For $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$, $\beta > 1$, $\delta \geq 0$,

$$\sigma_A^{W-FTPL}(T) \leq \frac{\kappa^2(3 + 2M/c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i}$$

$$\begin{aligned}
& + \frac{\kappa^2(M+c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \sum_{i=2}^K \frac{16\alpha_i^2}{c^2\alpha_i^2} \\
& + \frac{\kappa^2}{\min_i(c\alpha_i + g(\alpha_i)\kappa)}.
\end{aligned}$$

The key take-away from Theorem 3 is that the competitive ratios of both FTPL and W-FTPL are bounded by a constant. The competitive ratio of FTPL and W-FTPL deteriorates with M , while the competitive ratio of RR improves with M when only one partial hosting level is permitted [2]. Note that the competitive ratio of FTPL and W-FTPL is bounded by a constant for the setting where any finite number of partial hosting levels are allowed. Compared to this, for RR in [2], the competitive ratio result holds only for one partial hosting level.

Theorem 4 (Stochastic Arrivals). *Let the arrivals in each time-slot be i.i.d. stochastic with mean μ . Then we have*

- (a) *For T large enough, $\sigma_S^{RR}(T) > 1$.*
- (b) *For $\eta_t = \alpha\sqrt{t-1}$, $\sigma_S^{FTPL}(T) \leq 1 + O(1/T)$.*
- (c) *For $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$, $\beta > 1$, $\delta > 0$, $\sigma_S^{W-FTPL}(T) \leq 1 + O(1/T)$.*

The key take-away from Theorem 4 is that for T large enough, FTPL and W-FTPL outperform RR in terms of competitive ratio for stochastic arrivals.

In Section 5 we supplement the results stated in this section through simulations.

5. Simulations

In this section, we compare RR, FTPL and W-FTPL policies via simulations using synthetic request sequences, as well as real trace data from [24]. In all our simulations, we consider $\eta_t = 0.1\sqrt{t}$ for FTPL and W-FTPL, and $\beta = 6$ for W-FTPL.

5.1. Synthetic arrivals

In Figure 1, we plot the regret of different policies as a function of time horizon T . The parameters considered are $c = 0.1$, $M = 50$, $\kappa = 5$. The request arrival considered is similar to the one used in the proof of Theorem

3(b), i.e., we divide time into frames, and each frame starts with $\lceil \frac{M}{\kappa - c} \rceil$ slots of κ requests each followed by $\lceil \frac{M}{c} \rceil$ slots of zero requests. We consider 100 such frames for this experiment. The regret is averaged over 50 experiments. We observe that RR has linear regret in this case which agrees with Theorem 1.

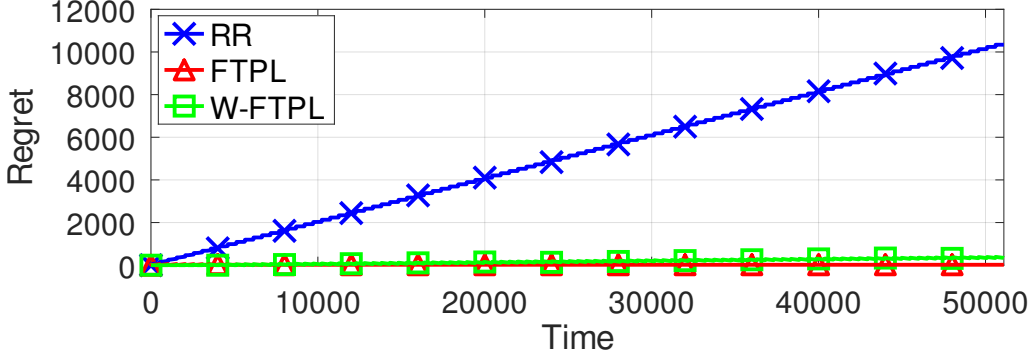


Figure 1: Regret as a function of Time horizon (T)

5.2. Stochastic arrivals

For the stochastic case, we consider arrivals in each slot to be Bernoulli with parameter μ . We will restrict attention to the case where there is only one partial level of hosting, i.e., $\rho_t^P \in \{0, \alpha_2, 1\}$. We consider $\alpha_2 = 0.5$ and $g(\alpha_2) = 0.45$. In Figure2 and Figure3, we consider $\mu = 0.4$, $c = 0.45$. All the results plotted are averaged over 50 independent experiments. In Figure2, we fix $M = 5$ and compare the regret performance of the policies as a function of time. We observe that W-FTPL performs better than other policies, and RR has linear regret, which agrees with Theorem 2. In Figure3, we compare the performance of the policies with respect to M for $T = 10^4$. We observe that W-FTPL performs better than other policies, and the main difference between the cost of policies is due to fetch cost. Also note that W-FTPL has better dependence on M as compared to FTPL, which agrees with Theorem 2. From Theorem 2 we observe that for large values of M the regret of FTPL is proportional to M/Δ_{\min}^2 whereas W-FTPL is proportional to $\log M/\Delta_{\min}^2$. Therefore in Figure4, we compare the performance of the policies by varying rent cost c for $M = 500$, $\mu = 0.4$, and $T = 10^4$. Again we note that W-FTPL outperforms RR and FTPL. As c gets closer to μ , the FTPL policy does multiple fetches, and as a result, the regret increases when c is close to μ .

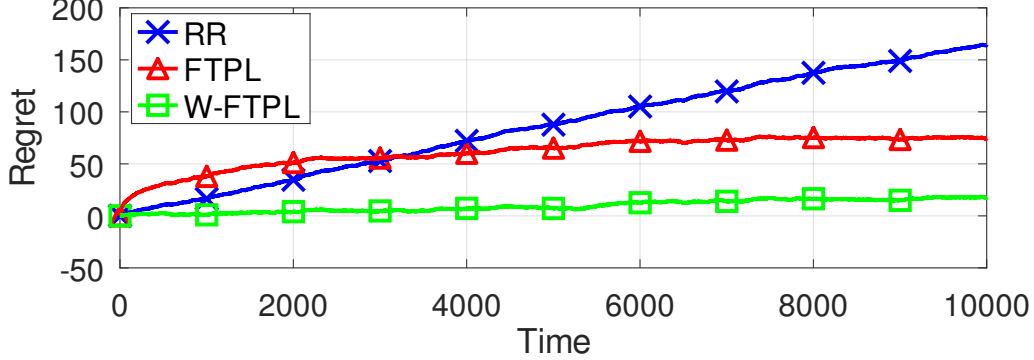


Figure 2: Regret as a function of time horizon (T)

5.3. Trace driven arrivals

For trace-driven simulations, we use the data collected by Grid Workloads Archive [24], which consists of requests arriving at computational servers. Among the traces mentioned in [24] we use DAS-2 traces, which were provided by the Advanced School for Computing and Imaging (ASCI). We consider slot duration as 1 hour and plot a snapshot of the trace in Figure 5. We consider $\kappa = 300$ for all the experiments in this subsection. All the results plotted for trace driven experiments are averaged over 50 independent experiments.

We consider partial service hosting with one partial level for our simulations. This dataset provides information on the number of processors required to serve each request, and this number lies between 1 and 128 for all requests. For this experiment, we consider hosting the entire service as equivalent to having 128 processors at the edge and hosting the partial service as having 16 processors at the edge. It follows that the hosting status $\rho_t^P \in \{0, 16/125 (= 0.125), 1\}$. To compute $g(0.125)$, we use the dataset to compute the fraction of requests that require more than 16 processors for service and obtain that $g(0.125) = 0.0328$.

In Figure 6, we consider $M = 500$, $c = 100$, $\kappa = 300$ and compare the regret performance of policies as a function of time. We observe that RR performs better than other policies for most values of time horizon. This is because RR attempts to mimic the optimal offline policy, which is not a static policy for this arrival sequence. Note that FTPL and W-FTPL attempt to mimic the optimal static policy and perform worse than RR.

In Figure 7, we compare the performance of different policies with respect

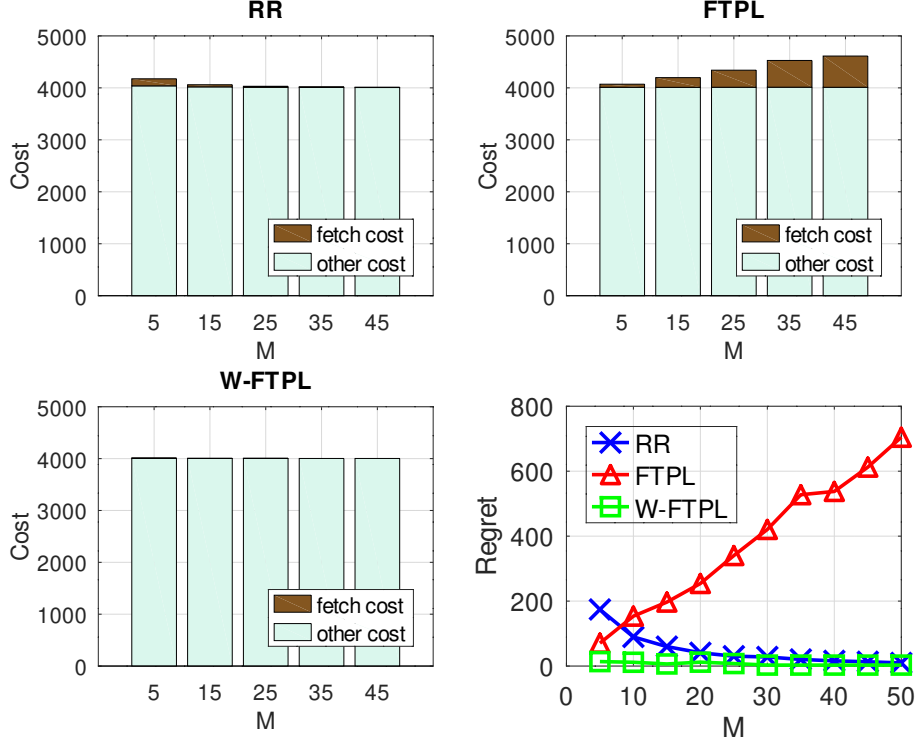


Figure 3: Cost, Regret as a function of fetch cost (M)

to M and fix $c = 100$, $\kappa = 300$. We observe that RR performs better than the other policies for some values of M because of the same reason mentioned before. In Figure 8, we compare the performance of different policies by varying rent cost c and fix $M = 500$, $\kappa = 300$. Again we note that RR outperforms W-FTPL and FTPL in some cases.

Our simulation results thus indicate that FTPL and W-FTPL outperform RR for i.i.d. stochastic arrivals. In the case where the arrival process is such that the offline optimal policy is not static, RR can outperform FTPL and W-FTPL. Additionally, RR can have linear regret with respect to time, while FTPL and W-FTPL have sub-linear regret with respect to time for all arrival processes considered.

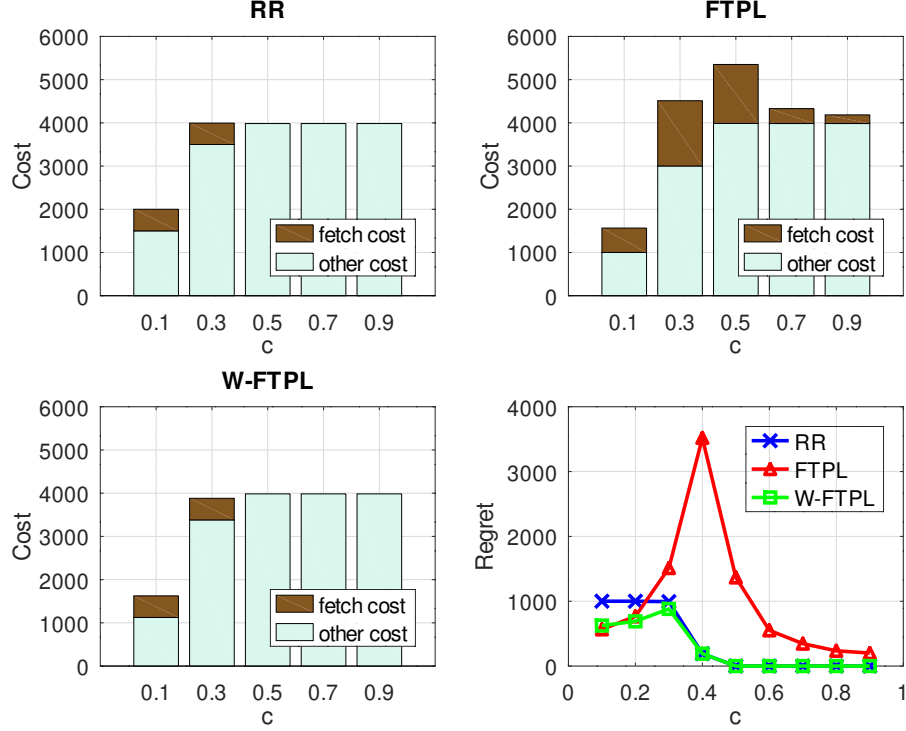


Figure 4: Cost, Regret as a function of rent cost (c)

6. Proof of theorems

6.1. Proof of Theorem 1(a)

Proof of Theorem 1(a). Recall that regret of any policy is given by

$$\begin{aligned} \mathcal{R}_A^{\mathcal{P}}(T, r) &= \sum_{t=1}^T c\rho_t + g(\rho_t)r_t + M(\rho_t - \rho_{t-1})^+ \\ &\quad - \min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\}. \end{aligned}$$

Note that $\sup_r \mathcal{R}_A^{\mathcal{P}}(T, r) \geq \mathbb{E}_r[\mathcal{R}_A^{\mathcal{P}}(T, r)]$ where the expectation is w.r.t some specified distribution over request sequences. We lower bound $\mathbb{E}_r[\mathcal{R}_A^{\mathcal{P}}(T, r)]$ to get a lower bound on $\mathcal{R}_A^{\mathcal{P}}(T)$.

$$\mathbb{E}_r[\mathcal{R}_A^{\mathcal{P}}(T, r)] = \sum_{t=1}^T c\rho_t + g(\rho_t)\mathbb{E}_r[r_t] + M(\rho_t - \rho_{t-1})^+$$

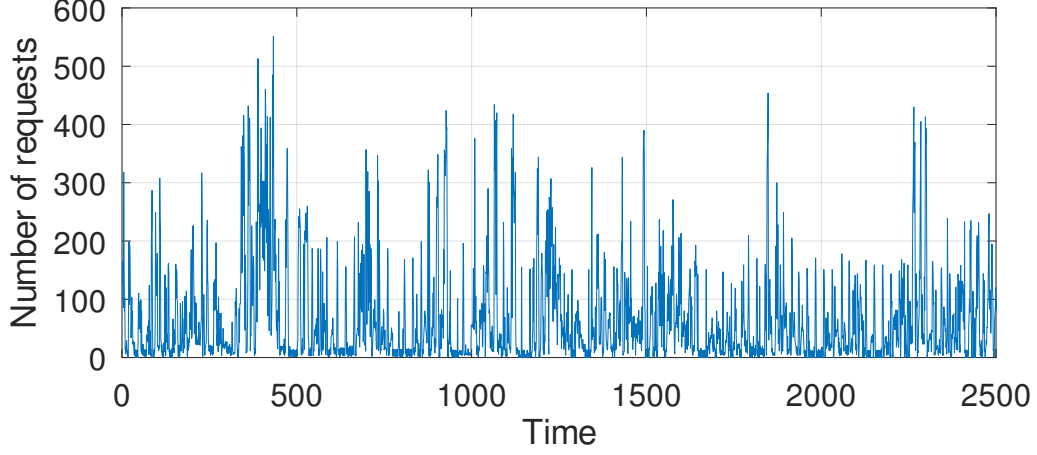


Figure 5: Arrivals in a slot for trace data considered

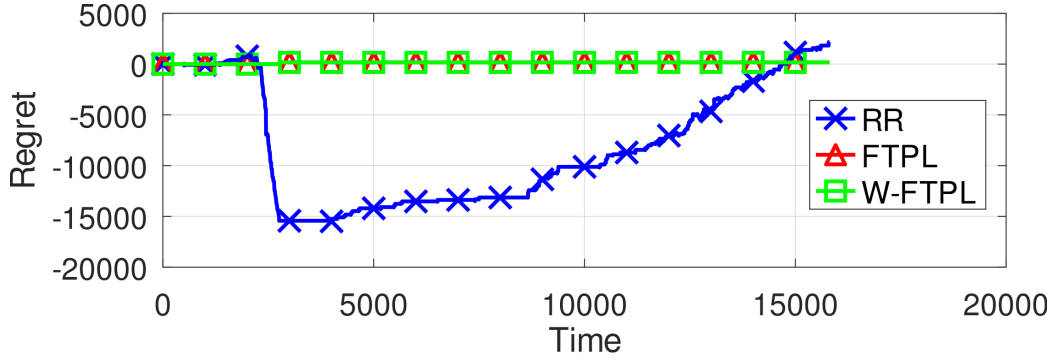


Figure 6: Regret as a function of time horizon

$$\begin{aligned}
& -\mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\}] \\
& \geq \sum_{t=1}^T c\rho_t + g(\rho_t)\mathbb{E}_r[r_t] \\
& \quad -\mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i)R_T\}] - M.
\end{aligned}$$

Let $\ell = \arg \min_{i \neq 1} \frac{\alpha_i}{1-g(\alpha_i)}$, $X \sim \text{Ber}(\frac{c\alpha_\ell}{\kappa(1-g(\alpha_\ell))})$, and $r_t = \kappa X$ and i.i.d over-time, therefore $\mathbb{E}_r[r_t] = \frac{c\alpha_\ell}{1-g(\alpha_\ell)}$. From Assumption 2 we have $\alpha_i + g(\alpha_i)r_t \leq \kappa$ for all $1 \leq i \leq K$ and $0 \leq r_t \leq \kappa$ therefore $\frac{c\alpha_\ell}{\kappa(1-g(\alpha_\ell))} \leq 1$ and X is a valid

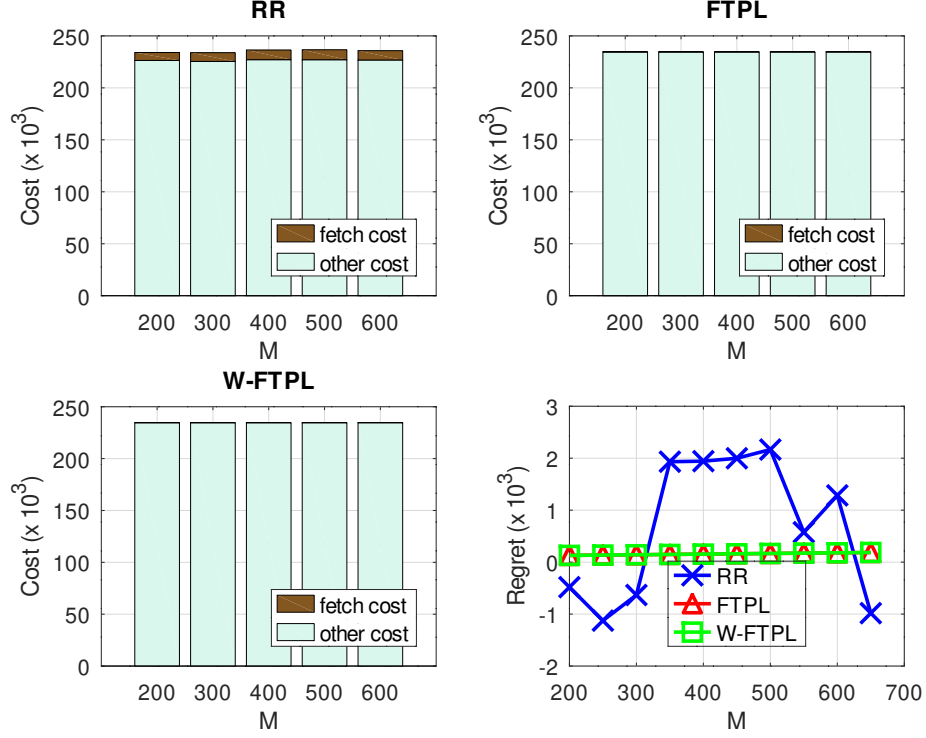


Figure 7: Cost, Regret as a function of fetch cost (M)

Bernoulli random variable.

$$\begin{aligned}
\mathbb{E}_r[\mathcal{R}_A^{\mathcal{P}}(T, r)] &\geq \sum_{t=1}^T c\rho_t + g(\rho_t) \frac{c\alpha_\ell}{1 - g(\alpha_\ell)} \\
&\quad - \mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i) R_T\}] - M \\
&= \sum_{t=1}^T (1 - g(\rho_t)) \frac{c\rho_t}{(1 - g(\rho_t))} \mathbb{1}_{\rho_t \neq 0} + g(\rho_t) \frac{c\alpha_\ell}{1 - g(\alpha_\ell)} \\
&\quad - \mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i) R_T\}] - M \\
&\stackrel{(a)}{\geq} \sum_{t=1}^T (1 - g(\rho_t)) \frac{c\alpha_\ell}{(1 - g(\alpha_\ell))} + g(\rho_t) \frac{c\alpha_\ell}{1 - g(\alpha_\ell)} \\
&\quad - \mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i) R_T\}] - M
\end{aligned}$$

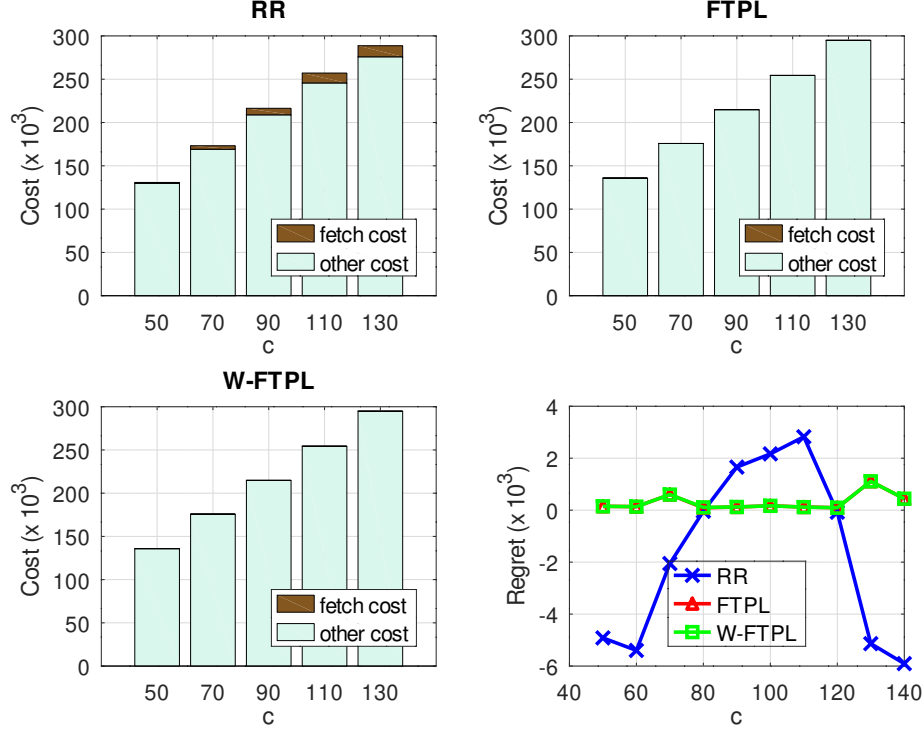


Figure 8: Cost, Regret as a function of rent cost (c)

$$\begin{aligned}
&\stackrel{(b)}{\geq} \frac{cT\alpha_\ell}{1 - g(\alpha_\ell)} - \mathbb{E}_r[\min\{R_T, cT\alpha_\ell + g(\alpha_\ell)R_T\}] - M \\
&\stackrel{(c)}{=} \frac{1}{2} \mathbb{E}_r[|R_T - cT\alpha_\ell - g(\alpha_\ell)R_T|] - M \\
&= \frac{\kappa(1 - g(\alpha_\ell))}{2} \mathbb{E}_r \left[\left| \frac{R_T}{\kappa} - \frac{Tc\alpha_\ell}{\kappa(1 - g(\alpha_\ell))} \right| \right] - M \\
&\stackrel{(d)}{=} \Omega(\sqrt{T}),
\end{aligned}$$

where (a) follows from definition of ℓ , (b) follows because of $\min_{1 \leq k \leq K} a_k \leq \min\{a_i, a_j\}$ for $1 \leq i, j \leq K$, (c) follows from Lemma 15 in Section 7, and (d) follows from Lemma 16 in Section 7. \square

6.2. Proof of Theorem 1(b)

Note that the RR algorithm in [3] can have only one intermediate level of partial hosting i.e., $\rho_t^{\text{RR}} \in \{0, \alpha_2, 1\}$. So throughout the is subsection we

consider $\rho_t^{\text{RR}} \in \{0, \alpha_2, 1\}$. We use the following lemmas to prove Theorem 1(b).

Lemma 2. *RR does not fetch any non zero fraction of service for at least $\left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$ slots after eviction of complete service, where $\alpha_{i'} = \arg \min_{\alpha_i \neq 0} \frac{M\alpha_i}{\kappa - c\alpha_i - \kappa g(\alpha_i)}$.*

Proof. Without loss of generality we consider $t = 0$ is when RR last evicted the service. If RR policy fetches $\alpha_i \neq 0$ fraction of service in slot $t + 1$, then $c\alpha_i t + g(\alpha_i)R_t + M\alpha_i < R_t$.

We now prove the lemma by contradiction. Let us assume that RR fetches α_i fraction of the service in time slot $t_f + 1$ and $t_f < \left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$. Then by using the definition of $\alpha_{i'}$, we get $t_f < \frac{M\alpha_i}{\kappa - c\alpha_i - g(\alpha_i)\kappa}$ and $c\alpha_i t_f + g(\alpha_i)\kappa t_f + M\alpha_i > \kappa t_f \geq R_{t_f}$ which is a contradiction. \square

Lemma 3. *Once RR fetches any non zero fraction of service, it hosts (some positive fraction of the service) for at least $\lceil \frac{M}{c} \rceil$ slots before it evicts the completely.*

Proof. Without loss of generality we consider $t = 0$ is when RR last fetched the service and say it hosted $\alpha_i \neq 0$ fraction of the service. For the RR policy to evict the service in slot $t + 1$, we need $c\alpha_i t + g(\alpha_i)R_t > M\alpha_i + R_t$, or equivalently $R_t < \frac{c\alpha_i t - M\alpha_i}{1 - g(\alpha_i)}$.

We now prove the lemma by contradiction. Let us assume that RR evicts the service in time slot $t_e + 1$ and $t_e < \lceil \frac{M}{c} \rceil$. Then from the assumption above, we must have $R_{t_e} < \frac{c\alpha_i t_e - M\alpha_i}{1 - g(\alpha_i)} < 0$ which is a contradiction. \square

Proof of Theorem 1(b). We split the entire time duration T into frames of size $t_f + t_e$ and consider the arrival sequence in each frame as $r_t = \kappa$ for $1 \leq t \leq t_f$ and $r_t = 0$ for $t_f + 1 \leq t \leq t_f + t_e$, where $\alpha_{i'} = \arg \min_{\alpha_i \neq 0} \frac{M\alpha_i}{\kappa - c\alpha_i - \kappa g(\alpha_i)}$, $t_f = \left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$, and $t_e = \lceil \frac{M}{c} \rceil$.

From Lemma 2 we know that RR does not fetch any fraction of service till time t_f . At $t = t_f + 1$ RR fetches some positive fraction of the service because $c\alpha_{i'} t_f + g(\alpha_{i'})R_{t_f} + M\alpha_{i'} < R_{t_f}$ holds by definition of t_f .

From Lemma 3 we know that RR does not evict the service for $t \in [t_f + 1, t_f + t_e]$. At $t = t_f + t_e + 1$ RR evicts the service because for $\alpha_i \neq 0$, $c\alpha_i t_e + g(\alpha_i) \sum_{t=t_f+1}^{t_f+t_e} r_t < M\alpha_i + \sum_{t=t_f+1}^{t_f+t_e} r_t$ holds by definition of t_e .

So amongst the first $t_f + t_e$ time-slots, the RR policy does not host for the first t_f slots and then hosts some positive fraction of service for the following t_e slots. Let $i^* = \arg \min_i c\alpha_i T + g(\alpha_i R_T + M\alpha_i)$ be the fraction of service to be hosted by optimal static policy. If the optimal static policy is to host $\alpha_{i^*} \neq 0$ fraction of service then we have a regret of at least $(\kappa - c\alpha_{i^*} - g(\alpha_{i^*})\kappa)t_f - M\alpha_{i^*}$ for initial frame and $(\kappa - c\alpha_{i^*} - g(\alpha_{i^*})\kappa)t_f = (\kappa - c\alpha_{i^*} - g(\alpha_{i^*})\kappa)\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \rceil > 0$ from second frame onwards because of forwarding, which is a constant and does not depend on T . If the optimal static policy is not to host the service, then we have a regret of $c\alpha_i t_e > 0$ till $t_f + t_e$ for some $\alpha_i \neq 0$, which is a constant and does not depend on T . So in either case, a regret larger than some positive constant (say d) is occurred in a frame of size $t_f + t_e$. Since we split the entire time duration T into frames of size $t_f + t_e = \lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \rceil + \lceil \frac{M}{c} \rceil$ and the request sequence is repeated in each frame, we get linear regret.

$$\begin{aligned}
\mathcal{R}_A^{\text{RR}}(T) &\geq \sum_{f=1}^{\lfloor T/(t_f+t_e) \rfloor} d - M\alpha_{i^*} \\
&\geq \lfloor T/(t_f+t_e) \rfloor d - M\alpha_{i^*} \\
&\geq d \left(\frac{T}{\frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} + \frac{M}{c} + 2} - 1 \right) - M\alpha_{i^*} \\
&= \Omega(T).
\end{aligned}$$

□

6.3. Proof of Theorem 1(c)

The optimal static policy only fetches the optimal fraction of service to be hosted once. Therefore,

$$\begin{aligned}
\mathcal{R}_A^{\text{FTPL}}(T, r) &= \mathbb{E}[\mathcal{C}^{\text{FTPL}}(T, r)] \\
&\quad - \min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\} \\
&\leq \mathbb{E}[\mathcal{C}^{\text{FTPL}}(T, r)] - \min_i \{c\alpha_i T + g(\alpha_i)R_T\}. \tag{3}
\end{aligned}$$

Since FTPL policy does not consider fetch cost M while making decisions we can decouple the fetch cost and the non fetch cost incurred by the FTPL policy. Therefore we can also decouple the expected regret incurred by FTPL

into expected regret without fetch cost ($M = 0$) and the fetch cost incurred by FTPL to bound (3). We first bound the expected regret of the FTPL policy with fetch cost $M = 0$ and then add it's expected fetch cost to get the final regret bound.

Lemma 4. *The regret for FTPL policy with non decreasing learning rate $\{\eta_t\}_{t=1}^T$ and $M = 0$ is given by*

$$\mathcal{R}_A^{\text{FTPL}}(T) \leq \sqrt{2 \log K} \left(\eta_T + \kappa^2 \sum_{t=1}^T \frac{1}{\eta_t} \right).$$

Proof. The proof of this theorem follows along the same lines as Theorem 1 of [25] and, Proposition 4.1 of [17]. Recall that $\rho_t^{\mathcal{P}}$ represents the fraction of service hosted in slot t by policy \mathcal{P} and $\rho_t^{\mathcal{P}}$ is a one hot vector where position of one represents the level of service hosted at the edge. In slot t FTPL hosts fraction of service $\rho_t^{\text{FTPL}} = \langle \rho_t^{\text{FTPL}}, \mathbf{s} \rangle$, where $\rho_t^{\text{FTPL}} = \arg \min_{\rho \in \mathcal{X}} \langle \rho, \Theta_t + \eta_t \gamma \rangle$. For ease of notation we suppress the policy in super script and denote $\rho_t^{\mathcal{P}}$ as ρ_t and $\rho_t^{\mathcal{P}}$ as ρ_t in further discussion. Define a time varying potential function

$$\Phi_t(\theta) = \mathbb{E}_{\gamma} \left[\min_{\rho \in \mathcal{X}} \langle \rho, \theta + \eta_t \gamma \rangle \right], \quad (4)$$

and observe that the gradient of Φ_t at Θ_t is $\mathbb{E}_{\gamma}[\rho_t]$. So $\langle \nabla \Phi_t(\Theta_t), \theta_t \rangle = \mathbb{E}_{\gamma}[\langle \rho_t, \theta_t \rangle]$. Consequently

$$\begin{aligned} \mathbb{E}_{\gamma}[\langle \rho_t, \theta_t \rangle] &= \langle \nabla \Phi_t(\Theta_t), \Theta_{t+1} - \Theta_t \rangle, \\ &= \Phi_t(\Theta_{t+1}) - \Phi_t(\Theta_t) - \frac{1}{2} \langle \theta_t, \nabla^2 \Phi_t(\tilde{\theta}_t) \theta_t \rangle, \end{aligned}$$

where $\tilde{\theta}_t$ is the line segment joining Θ_{t+1} and Θ_t which follows from Taylor's expansion of $\Phi_t(\Theta_{t+1})$.

$$\begin{aligned} &\sum_{t=1}^T \mathbb{E}_{\gamma} [\langle \rho_t, \theta_t \rangle] \\ &= \sum_{t=1}^T \Phi_t(\Theta_{t+1}) - \Phi_t(\Theta_t) - \frac{1}{2} \langle \theta_t, \nabla^2 \Phi_t(\tilde{\theta}_t) \theta_t \rangle, \end{aligned}$$

$$\begin{aligned}
&= \Phi_T(\Theta_{T+1}) - \Phi_1(\Theta_1) + \sum_{t=2}^T \Phi_{t-1}(\Theta_t) - \Phi_t(\Theta_t) \\
&\quad - \sum_{t=1}^T \frac{1}{2} \langle \theta_t, \nabla^2 \Phi_t(\tilde{\theta}_t) \theta_t \rangle.
\end{aligned} \tag{5}$$

By Jensen's inequality

$$\begin{aligned}
\Phi_T(\Theta_{T+1}) &= \mathbb{E} \left[\min_{\rho \in \mathcal{X}} \langle \rho, \Theta_{T+1} + \eta_t \gamma \rangle \right] \\
&\leq \min_{\rho \in \mathcal{X}} \mathbb{E} [\langle \rho, \Theta_{T+1} + \eta_t \gamma \rangle] \\
&\stackrel{(a)}{=} \min_{\rho \in \mathcal{X}} \langle \rho, \Theta_{T+1} \rangle,
\end{aligned}$$

where (a) follows because γ_i 's are standard Gaussian random variables. Recall that $\min_{\rho \in \mathcal{X}} \langle \rho, \Theta_{T+1} \rangle$ is the loss incurred by the offline static policy. Therefore by rearranging (5) and using Jensen's inequality we get

$$\begin{aligned}
\mathcal{R}_A^{\text{FTPL}}(T) &\leq -\Phi_1(\Theta_1) + \sum_{t=2}^T \Phi_{t-1}(\Theta_t) - \Phi_t(\Theta_t) \\
&\quad - \sum_{t=1}^T \frac{1}{2} \langle \theta_t, \nabla^2 \Phi_t(\tilde{\theta}_t) \theta_t \rangle.
\end{aligned} \tag{6}$$

We bound each term in the RHS separately to get the final result.

The first term is $-\Phi_1(\Theta_1) = -\Phi_1(0) = -\eta_1 \mathbb{E}_\gamma [\min_{\rho} \langle \rho, \gamma \rangle]$
 $= \eta_1 \mathbb{E}_\gamma [\max_{\rho \in \mathcal{X}} \langle \rho, \gamma \rangle]$ since Gaussian random variables are symmetric. By using the result of Lemma 9 in [25], we get

$$-\Phi_1(0) \leq \eta_1 \sqrt{2 \log K}. \tag{7}$$

The second term can be analyzed by considering

$$\begin{aligned}
&\Phi_{t-1}(\Theta_t) - \Phi_t(\Theta_t) \\
&= \mathbb{E}_\gamma [\min_{\rho \in \mathcal{X}} \langle \rho, \theta_t + \eta_{t-1} \gamma \rangle - \min_{\rho \in \mathcal{X}} \langle \rho, \theta_t + \eta_t \gamma \rangle] \\
&\stackrel{(a)}{\leq} \mathbb{E}_\gamma [\max_{\rho \in \mathcal{X}} \langle \rho, (\eta_{t-1} - \eta_t) \gamma \rangle] \\
&\leq |\eta_t - \eta_{t-1}| \sqrt{2 \log K},
\end{aligned} \tag{8}$$

where (a) follows because $\max_{\mathbf{y} \in \mathcal{Y}} (f(\mathbf{y}) - g(\mathbf{y})) \geq \max_{\mathbf{y} \in \mathcal{Y}} f(\mathbf{y}) - \max_{\mathbf{y} \in \mathcal{Y}} g(\mathbf{y})$ for any arbitrary functions f, g . We obtain (8) by using Lemma 9 in [25].

Now we bound the third term in (6). Let $H = \nabla^2 \Phi_t(\tilde{\boldsymbol{\theta}}_t)$. By using Lemma 7 of [26] we have,

$$H_{i,j} = \frac{1}{\eta_t} \mathbb{E}[\hat{x}(\tilde{\boldsymbol{\theta}}_t + \eta_t \boldsymbol{\gamma})_i \gamma_j].$$

If we replace η_t with η_t/κ and θ_t with θ_t/κ and apply Lemma 2 of [25] we get

$$\begin{aligned} -\left\langle \frac{\boldsymbol{\theta}_t}{\kappa}, \kappa H \frac{\boldsymbol{\theta}_t}{\kappa} \right\rangle &\leq \frac{2\kappa}{\eta_t} \sqrt{2 \log K} \\ \implies -\langle \boldsymbol{\theta}_t, \nabla^2 \Phi_t(\tilde{\boldsymbol{\theta}}_t) \boldsymbol{\theta}_t \rangle &\leq \frac{2\kappa^2}{\eta_t} \sqrt{2 \log K}. \end{aligned} \quad (9)$$

By using (7), (8) and (9), we bound (6),

$$\begin{aligned} \mathcal{R}_A^{\text{FTPL}}(T) &\leq \eta_1 \sqrt{2 \log K} + \sum_{t=2}^T |\eta_t - \eta_{t-1}| \sqrt{2 \log K} \\ &\quad + \sum_{t=1}^T \frac{2\kappa^2}{\eta_t} \sqrt{2 \log K} \\ &\leq \eta_T \sqrt{2 \log K} + \sum_{t=1}^T \frac{2\kappa^2}{\eta_t} \sqrt{2 \log K} \\ &\leq \eta_T \sqrt{2 \log K} + 2\kappa^2 \sqrt{2 \log K} \sum_{t=1}^T \frac{1}{\eta_t}. \end{aligned}$$

□

Now we bound the expected fetch cost incurred by FTPL policy.

Lemma 5. *The expected fetch cost under FTPL policy upto time T denoted by $\mathbb{E}[\mathcal{C}_F^{\text{FTPL}}(T)]$ with learning rate $\eta_t = \alpha\sqrt{t}$ is bounded as follows*

$$\mathbb{E}[\mathcal{C}_F^{\text{FTPL}}(T)] \leq \frac{MK^2(c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T + 1}.$$

Proof. The proof of this theorem follows similar a line of arguments as in [17] for the case with $N = 2$ and cache size 1 case. The fetch cost is incurred only

when the service is fetched. The overall cost of fetching is $\sum_{t=1}^T M(\rho_{t+1} - \rho_t)^+$. So, the expected fetch cost is given by

$$\begin{aligned}\mathbb{E}[\mathcal{C}_F^{\text{FTPL}}(T)] &= \mathbb{E} \left[\sum_{t=1}^T M(\rho_{t+1} - \rho_t) \mathbb{1}_{\{\rho_{t+1} > \rho_t\}} \right] \\ &\leq M \sum_{t=1}^T \mathbb{P}(\rho_{t+1} > \rho_t) \\ &\leq M \sum_{t=1}^T \sum_{j>i} \mathbb{P}(\rho_{t+1} = \alpha_j, \rho_t = \alpha_i).\end{aligned}$$

Let $\mathcal{E}_{i,j}^{(t)}$ be the event that $\rho_t = \alpha_i$, $\rho_{t+1} = \alpha_j$. $\mathcal{E}_{i,j}^{(t)}$ occurs only if $\Theta_{t,i} + \eta_t \gamma_i < \Theta_{t,j} + \eta_t \gamma_j$, $\Theta_{t+1,i} + \eta_{t+1} \gamma_i > \Theta_{t+1,j} + \eta_{t+1} \gamma_j$. Let $\delta_\alpha = \alpha_j - \alpha_i$, $\delta_g = g(\alpha_i) - g(\alpha_j)$, note that $0 < \delta_\alpha \leq 1$, $0 < \delta_g \leq 1$.

$$\begin{aligned}\mathbb{P}(\mathcal{E}_{i,j}^{(t)}) &\leq \mathbb{P}(\Theta_{t,i} + \eta_t \gamma_i < \Theta_{t,j} + \eta_t \gamma_j, \\ &\quad \Theta_{t+1,i} + \eta_{t+1} \gamma_i > \Theta_{t+1,j} + \eta_{t+1} \gamma_j) \\ &= \mathbb{P} \left(\frac{\Theta_{t+1,j} - \Theta_{t+1,i}}{\sqrt{2}\eta_{t+1}} < \frac{(\gamma_i - \gamma_j)}{\sqrt{2}} < \frac{\Theta_{t,j} - \Theta_{t,i}}{\sqrt{2}\eta_t} \right) \\ &\stackrel{(a)}{\leq} \frac{\Theta_{t,j} - \Theta_{t,i}}{2\sqrt{\pi}\eta_t} - \frac{\Theta_{t+1,j} - \Theta_{t+1,i}}{2\sqrt{\pi}\eta_{t+1}} \\ &= \frac{1}{2\sqrt{\pi}} \left((c(t-1)\delta_\alpha - \delta_g R_{t-1}) \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t+1}} \right) \right. \\ &\quad \left. - \frac{1}{2\sqrt{\pi}} \left(\frac{c\delta_\alpha - \delta_g r_t}{\eta_{t+1}} \right) \right) \\ &\leq \frac{1}{2\sqrt{\pi}} \left(ct \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t+1}} \right) + \frac{\delta_g r_t - c\delta_\alpha}{\eta_{t+1}} \right) \\ &\stackrel{(b)}{\leq} \frac{1}{2\sqrt{\pi}} \left(ct \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t+1}} \right) + \frac{\kappa}{\eta_{t+1}} \right) \\ &= \frac{1}{2\alpha\sqrt{\pi}} \left(ct \left(\frac{1}{\sqrt{t}} - \frac{1}{\sqrt{t+1}} \right) + \frac{\kappa}{\sqrt{t+1}} \right) \\ &= \frac{1}{2\alpha\sqrt{\pi}} \left(c\sqrt{\frac{t}{t+1}} \frac{1}{\sqrt{t} + \sqrt{t+1}} + \frac{\kappa}{\sqrt{t+1}} \right)\end{aligned} \tag{10}$$

$$\stackrel{(c)}{\leq} \frac{(c + 2\kappa)}{4\alpha\sqrt{\pi(t+1)}}.$$

Here (a) follows from the fact that $\mathbb{P}(a < Z < b) \leq \frac{b-a}{\sqrt{2\pi}}$, where $Z \sim \mathcal{N}(0, 1)$, (b) follows because $\delta_g \leq 1$ and $r_t \leq \kappa$ and, (c) follows because $\frac{\sqrt{t}}{\sqrt{t}+\sqrt{t+1}} \leq \frac{1}{2}$. Therefore,

$$\begin{aligned} \mathbb{E}[\mathcal{C}_f^{\text{FTPL}}(T)] &\leq \frac{M(c + 2\kappa)}{4\alpha\sqrt{\pi}} \sum_{t=1}^T \sum_{j>i} \frac{1}{\sqrt{t+1}} \\ &\leq \frac{MK^2(c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T+1}. \end{aligned}$$

□

Proof of Theorem 1(c). By using Lemma 4 and 5 we get,

$$\begin{aligned} \mathcal{R}_A^{\text{FTPL}}(T) &\leq \sqrt{2 \log K} \left(\alpha\sqrt{T} + 2\kappa^2 \sum_{t=1}^T \frac{1}{\alpha\sqrt{t}} \right) \\ &\quad + \frac{K^2 M(c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T+1} \\ &\leq \sqrt{2T \log K} \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\ &\quad + \frac{K^2 M(c + 2\kappa)}{2\alpha\sqrt{\pi}} \sqrt{T+1}. \end{aligned}$$

□

6.4. Proof of Theorem 1(d)

Proof of Theorem 1(d). We characterize the fetch cost under FTPL to analyze the impact of M on the total cost incurred under FTPL. Fetch cost is incurred if $\rho_{t-1} = 0$, $\rho_t = 1$. Consider a request sequence $r_1 = 1$ and $r_t = 0$ for all $t \geq 2$. The optimal static policy is to not host any fraction of service i.e., to forward all the requests to back-end server.

The probability of hosting α_i , $1 \leq i \leq K$ fraction of service in first slot is $1/K$. Hosting any non zero fraction of service will incur at least $M\alpha_2$ fetch cost. Therefore the expected fetch cost of FTPL policy will be at least $M\alpha_2/K$ which is a lower bound on the expected cost of FTPL policy for

this request sequence. For FTPL we can decouple the fetch cost and the non fetch cost and FTPL does not consider M while making decisions. Therefore $M\alpha_2/K$ is also a lower bound for regret of FTPL. This completes the proof. \square

6.5. Proof of Theorem 1(e)

Proof of Theorem 1(e). Recall that $T_s = \min\{t : t < \frac{(\max_{i \neq j}(\Theta_{t+1,i} - \Theta_{t+1,j}))^2}{\kappa^2 \beta (\log M)^{1+\delta}}\}$. For arrival sequences with requests arriving in time-slots 1 to T , we have two possible cases, namely, $T_s \geq T$ and $T_s < T$. We consider each case to bound the regret of W-FTPL policy.

Case I ($T_s \geq T$): Let $r^{(1)}$ be a request sequence chosen by adversary such that $T_s \geq T$, then

$$\begin{aligned} & |\max_{i \neq j}(\Theta_{T+1,j} - \Theta_{T+1,i})| < \kappa \sqrt{\beta T (\log M)^{1+\delta}} \\ \implies & |\max_i(\Theta_{T+1,1} - \Theta_{T+1,i})| < \kappa \sqrt{\beta T (\log M)^{1+\delta}} \\ & \xrightarrow{(a)} \mathcal{R}_A^{\text{W-FTPL}}(T) < \kappa \sqrt{\beta T (\log M)^{1+\delta}}, \end{aligned}$$

(a) follows because $\rho_t = 0$ for $t < T_s$ for W-FTPL.

Case II ($T_s < T$): Let $r^{(2)}$ be a request sequence chosen by adversary such that $T_s < T$, then by using case I we can bound

$$\begin{aligned} & |\max_{i \neq j} \Theta_{T_s,i} - \Theta_{T_s,j}| < \kappa \sqrt{\beta (T_s - 1) (\log M)^{1+\delta}} \\ \implies & \mathcal{R}_A^{\text{W-FTPL}}(T_s - 1) < \kappa \sqrt{\beta T (\log M)^{1+\delta}}. \end{aligned}$$

Thus combining both cases we get upper bound on regret in wait phase as

$$\mathcal{R}_A^{\text{W-FTPL}}(T_s - 1) < \kappa \sqrt{\beta T (\log M)^{1+\delta}}. \quad (11)$$

Note that W-FTPL policy follows FTPL policy after its waiting time i.e., from time T_s . Therefore W-FTPL and FTPL take same decisions and have same cost from time T_s to T . Thus the regret in a slot is also same for W-FTPL and FTPL from time T_s to T . If $T_s \geq T$ then only wait phase will be there and regret in FTPL phase is considered to be zero. We denote regret from time t_1 to t_2 as $\mathcal{R}_A^{\mathcal{P}}(t_1 : t_2)$.

$$\mathcal{R}_A^{\text{W-FTPL}}(T) = \mathcal{R}_A^{\text{W-FTPL}}(T_s - 1) + \mathcal{R}_A^{\text{W-FTPL}}(T_s : T)$$

$$\begin{aligned}
&= \mathcal{R}_A^{\text{W-FTPL}}(T_s - 1) + \mathcal{R}_A^{\text{FTPL}}(T_s : T) \\
&\leq \mathcal{R}_A^{\text{W-FTPL}}(T_s - 1) + \mathcal{R}_A^{\text{FTPL}}(1 : T).
\end{aligned} \tag{12}$$

By (11), (12) and Theorem 1(c) we get

$$\begin{aligned}
\mathcal{R}_A^{\text{W-FTPL}}(T) &\leq \kappa \sqrt{\beta T (\log M)^{1+\delta}} \\
&\quad + \sqrt{2T \log K} \left(\alpha + \frac{4\kappa^2}{\alpha^2} \right) \\
&\quad + \frac{K^2 M (c + 2\kappa)}{2\alpha \sqrt{\pi}} \sqrt{T + 1}.
\end{aligned}$$

□

6.6. Proof of Theorem 2(a)

Proof of Theorem 2(a). From Lemma 3 we know that once RR fetches any non-zero fraction of service, it hosts (some positive fraction of the service) for at least $\lceil \frac{M}{c} \rceil$ slots before it evicts completely. From Lemma 2, we know that, once RR evicts the complete service from the edge, then it does not fetch any non zero fraction of the service for at least $\left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$ slots. We divide the entire time T into frames of size $f_s = \lceil \frac{M}{c} \rceil + \left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$ slots. Let us define an event \mathcal{F} as the frame that starts with $\left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil$ slots with κ arrivals in each slot and followed by $\lceil \frac{M}{c} \rceil$ slots of zeros arrivals. Let the probability of event \mathcal{F} occurring be $\mathbb{P}(\mathcal{F})$ (independent of T). Conditioned on event \mathcal{F} , if optimal static policy is to host non zero fraction of service then we have regret at least $(\kappa - c\alpha_i + g(\alpha_i)\kappa) \left\lceil \frac{M\alpha_{i'}}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} \right\rceil > 0$ in a frame where $\alpha_i \neq 0$. If optimal is not to host any fraction of service then we have regret of $c\alpha_i \lceil \frac{M}{c} \rceil > 0$ in a frame where $\alpha_i \neq 0$. Therefore conditioned on event \mathcal{F} RR always has a finite nonzero regret say $d(> 0)$ which is independent of T . Therefore,

$$\begin{aligned}
\mathcal{R}_S^{\text{RR}}(T) &= \mathbb{E}_r \left[\sum_{f=1}^{\lfloor T/f_s \rfloor} d \mathbb{1}(\text{event } \mathcal{F} \text{ occurred}) \right] \\
&\geq \left(\frac{T}{f_s} - 1 \right) \mathbb{P}(\mathcal{F}) d
\end{aligned}$$

$$\begin{aligned}
&\geq \left(\frac{T}{\frac{M}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} + \frac{M}{c} + 2} - 1 \right) \mathbb{P}(\mathcal{F})d \\
&= \Omega(T).
\end{aligned}$$

So even in the stochastic case RR observes linear regret. \square

6.7. Proof of Theorem 2(b)

Lemma 6. *Probability of hosting a sub-optimal fraction of service $\alpha_j \neq \alpha_{i^*}$ under the FTPL policy in time slot $t + 1$ is bounded as follows*

$$\mathbb{P}(\rho_{t+1} = \alpha_j) \leq \exp\left(-\frac{t^2 \Delta_j^2}{16\eta_{t+1}^2}\right) + \exp\left(\frac{-\Delta_j^2 t}{2\kappa^2}\right).$$

Proof. FTPL hosts α_j fraction of service if $c\alpha_j t + g(\alpha_j)R_t + \eta_{t+1}\gamma_j > c\alpha_j t + g(\alpha_j)R_t + \eta_{t+1}\gamma_j$ in time slot $t + 1$, for all $i \neq j$. Let $p_{t,j} = \mathbb{P}(\rho_{t+1} = \alpha_j)$.

$$\begin{aligned}
p_{t,j} &\leq \mathbb{P}(c\alpha_j t + g(\alpha_j)R_t + \eta_{t+1}\gamma_j \\
&\quad < c\alpha_{i^*} t + g(\alpha_{i^*})R_t + \eta_{t+1}\gamma_{i^*}) \\
&= \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \\
&\quad \geq c(\alpha_j - \alpha_{i^*})t + (g(\alpha_j) - g(\alpha_{i^*}))R_t).
\end{aligned}$$

There can be two possibilities one is $\alpha_j > \alpha_{i^*}$ or $\alpha_j < \alpha_{i^*}$ we bound the probability $p_{t,j}$ by considering each case separately.

Case 1 ($\alpha_j > \alpha_{i^*}$): Since $g(\cdot)$ is a decreasing function $g(\alpha_j) < g(\alpha_{i^*})$. Let \mathcal{E}_j be the event that $R_t < \mu t + t\Delta_j/2$.

$$\begin{aligned}
p_{t,j} &\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq c(\alpha_j - \alpha_{i^*})t \\
&\quad + (g(\alpha_j) - g(\alpha_{i^*}))R_t, \mathcal{E}_j) + \mathbb{P}(\mathcal{E}_j^c) \\
&\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq t\Delta_j - (g(\alpha_{i^*}) - g(\alpha_j))t\Delta_j/2) \\
&\quad + \mathbb{P}(\mathcal{E}_j^c) \\
&\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq t\Delta_j/2) + \mathbb{P}(\mathcal{E}_j^c) \\
&\stackrel{(a)}{\leq} \exp\left(-\frac{t^2 \Delta_j^2}{16\eta_{t+1}^2}\right) + \exp\left(\frac{-\Delta_j^2 t}{2\kappa^2}\right),
\end{aligned}$$

where (a) is obtained by using the fact that complementary CDF of standard Gaussian $Q(x) \leq e^{-x^2/2}$ for $x > 0$ and Hoeffding's inequality [27].

Case 2 ($\alpha_j < \alpha_{i^*}$): Let \mathcal{E}_j be an event $R_t > \mu t - t\Delta_j/2$.

$$\begin{aligned}
p_{t,j} &\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq c(\alpha_j - \alpha_{i^*})t \\
&\quad + (g(\alpha_j) - g(\alpha_{i^*}))R_t, \mathcal{E}_j) + \mathbb{P}(\mathcal{E}_j^c) \\
&\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq t\Delta_j - (g(\alpha_j) - g(\alpha_{i^*}))t\Delta_j/2) \\
&\quad + \mathbb{P}(\mathcal{E}_j^c) \\
&\leq \mathbb{P}(\eta_{t+1}(\gamma_{i^*} - \gamma_j) \geq t\Delta_j/2) + \mathbb{P}(\mathcal{E}_j^c) \\
&\stackrel{(a)}{\leq} \exp\left(-\frac{t^2\Delta_j^2}{16\eta_{t+1}^2}\right) + \exp\left(\frac{-\Delta_j^2 t}{2\kappa^2}\right),
\end{aligned}$$

where (a) is obtained by using the fact that $Q(x) \leq e^{-x^2/2}$ for $x > 0$ and Hoeffding's inequality [27].

Therefore by combining both cases we get

$$\mathbb{P}(\rho_{t+1} = \alpha_i) \leq \exp\left(-\frac{t^2\Delta_j^2}{16\eta_{t+1}^2}\right) + \exp\left(\frac{-\Delta_j^2 t}{2\kappa^2}\right). \quad (13)$$

□

Lemma 7. For $M = 0$,

$$\begin{aligned}
\mathcal{R}_S^{FTPL}(T) &\leq \left(\sqrt{2\log K} + \frac{2\sqrt{2h_1}\log K}{\Delta_{\min}}\right) \left(\alpha + \frac{4\kappa^2}{\alpha}\right) \\
&\quad + \frac{16\alpha^2 + 4\kappa^2}{\Delta_{\min}},
\end{aligned}$$

where $h_1 = 2\max\{16\alpha^2, 2\kappa^2\}$.

Proof. Let $t_0 = \lceil h_1 \frac{\log K}{\Delta_{\min}^2} \rceil$ where $h_1 = 2\max\{16\alpha^2, 2\kappa^2\}$. From (2) we have

$$\begin{aligned}
\mathcal{R}_S^{FTPL}(T) &\leq \left(\sum_{t=1}^T c\rho_t + g(\rho_t)\mu - \mu_{i^*}\right) \\
&= \sum_{t=1}^T \left(\sum_i \mu_i \mathbb{P}(\rho_t = \alpha_i)\right) - \mu_{i^*} \\
&= \sum_{t=1}^T \sum_{i \neq i^*} \Delta_i \mathbb{P}(\rho_t = \alpha_i)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^{t_0} \sum_{i \neq i^*} \Delta_i \mathbb{P}(\rho_t = \alpha_i) \\
&\quad + \sum_{t=t_0+1}^T \sum_{i \neq i^*} \Delta_i \mathbb{P}(\rho_t = \alpha_i) \\
&\stackrel{(a)}{\leq} \sqrt{2t_0 \log K} \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\
&\quad + \sum_{t=t_0+1}^T \sum_{i \neq i^*} \Delta_i \mathbb{P}(\rho_t = \alpha_i),
\end{aligned}$$

where (a) is obtained by using adversarial regret bound when $M = 0$. Now we consider each term in RHS separately and bound them.

$$\sqrt{2t_0 \log K} \leq \sqrt{2 \log K} \left(1 + \frac{\sqrt{h_1 \log K}}{\Delta_{\min}} \right).$$

By using Lemma 6 we have

$$\begin{aligned}
&\Delta_j \mathbb{P}(\rho_{t+1} = \alpha_j) \\
&\leq \Delta_j \left[\exp \left(-\frac{t^2 \Delta_j^2}{16\eta_{t+1}^2} \right) + \exp \left(\frac{-\Delta_j^2 t}{2\kappa^2} \right) \right] \\
&= \Delta_j \left[\exp \left(-\frac{t \Delta_j^2}{16\alpha^2} \right) + \exp \left(\frac{-\Delta_j^2 t}{2\kappa^2} \right) \right] \\
&\stackrel{(b)}{\leq} \Delta_{\min} \left[\exp \left(-\frac{t \Delta_{\min}^2}{16\alpha^2} \right) + \exp \left(\frac{-\Delta_{\min}^2 t}{2\kappa^2} \right) \right].
\end{aligned}$$

Here (b) is obtained using the fact that $f(u) = ue^{-u^2}$ is a decreasing function over $u \in [1, \infty)$, and for $t > t_0$, $\frac{t \Delta_j^2}{16\alpha^2} > \frac{h_1 \log K \Delta_j^2}{16\alpha^2 \Delta_{\min}^2} \geq 1$ and $\frac{t \Delta_j^2}{2\kappa^2} > \frac{h_1 \log K \Delta_j^2}{2\kappa^2 \Delta_{\min}^2} \geq 1$. Therefore,

$$\begin{aligned}
&\mathcal{R}_S^{\text{FTPL}}(T) \\
&\leq \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1 \log K}}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\
&\quad + \sum_{t=t_0}^T K \Delta_{\min} \left[\exp \left(-\frac{t \Delta_{\min}^2}{16\alpha^2} \right) + \exp \left(\frac{-\Delta_{\min}^2 t}{2\kappa^2} \right) \right]
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(c)}{\leq} \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\
& \quad + \frac{16\alpha^2 + 2\kappa^2}{\Delta_{\min}} + 2\Delta_{\min} \\
& \leq \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\
& \quad + \frac{16\alpha^2 + 4\kappa^2}{\Delta_{\min}}.
\end{aligned}$$

Here (c) follows because for $t > t_0$ we have $K \exp(-t\Delta_{\min}^2/16\alpha^2) < 1$, $K \exp(-t\Delta_{\min}^2/2\kappa^2) < 1$ and $\sum_{t=1}^{\infty} e^{-at} \leq 1/a$ for $a > 0$. \square

Lemma 8. *Fetch cost under FTPL policy with learning rate $\eta_t = \alpha\sqrt{t-1}$ is bounded as follows*

$$\mathbb{E}[\mathcal{C}_f^{FTPL}(T)] \leq M \frac{16\alpha^2 + 2\kappa^2}{\Delta_{\min}^2}.$$

Proof. The fetch cost is incurred when we fetch extra fraction of service. Similar to the proof of Lemma 5 we get

$$\begin{aligned}
\mathbb{E}[\mathcal{C}_f^{FTPL}(T)] & \leq M \sum_{t=1}^T \mathbb{P}(\rho_t < \rho_{t+1}) \\
& = M \sum_{t=1}^T \sum_{j>i} \mathbb{P}(\rho_t = \alpha_i, \rho_{t+1} = \alpha_j).
\end{aligned}$$

Let $\mathcal{E}_{i,j}^{(t)}$ be the event that $\rho_t = \alpha_i$, $\rho_{t+1} = \alpha_j$. $\mathcal{E}_{i,j}^{(t)}$ occurs if $\Theta_{t,i} + \eta_t \gamma_i < \Theta_{t,j} + \eta_t \gamma_j$, $\Theta_{t+1,i} + \eta_{t+1} \gamma_i > \Theta_{t+1,j} + \eta_{t+1} \gamma_j$.

$$\mathbb{P}(\mathcal{E}_{i,j}^{(t)}) = \mathbb{P} \left(\frac{\Theta_{t+1,j} - \Theta_{t+1,i}}{\sqrt{2}\eta_{t+1}} \leq \frac{\gamma_i - \gamma_j}{\sqrt{2}} \leq \frac{\Theta_{t,j} - \Theta_{t,i}}{\sqrt{2}\eta_t} \right).$$

We consider two cases $\mu_i > \mu_j$ and $\mu_i < \mu_j$ and bound the probability.

Case 1 ($\mu_i > \mu_j$): Let \mathcal{E}_1 be the event that $|R_{t-1} - \mu(t-1)| \leq |\Delta_{ij}|(t-1)/2$,

$$\begin{aligned}
& \mathbb{P}(\mathcal{E}_{i,j}^{(t)}) \\
& \leq \mathbb{P} \left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \leq \frac{\Theta_{t,j} - \Theta_{t,i}}{\sqrt{2}\eta_t} \right)
\end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{P}\left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \leq \frac{\Theta_{t,j} - \Theta_{t,i}}{\sqrt{2}\eta_t}, \mathcal{E}_1\right) + \mathbb{P}(\mathcal{E}_1^c) \\
&= \mathbb{P}\left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \leq \frac{c(\alpha_j - \alpha_i)(t-1)}{\sqrt{2}\eta_t} \right. \\
&\quad \left. - \frac{(g(\alpha_i) - g(\alpha_j))R_{t-1}}{\sqrt{2}\eta_t}, \mathcal{E}_1\right) + \mathbb{P}(\mathcal{E}_1^c) \\
&\leq \mathbb{P}\left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \leq \frac{-|\Delta_{ij}|(t-1)}{2\sqrt{2}\eta_t}\right) + \mathbb{P}(\mathcal{E}_1^c) \\
&\stackrel{(a)}{\leq} \exp\left(-\frac{(t-1)^2\Delta_{ij}^2}{16\eta_t^2}\right) + \exp\left(\frac{-\Delta_{ij}^2(t-1)}{2\kappa^2}\right) \\
&\leq \exp\left(-\frac{(t-1)^2\Delta_{\min}^2}{16\eta_t^2}\right) + \exp\left(\frac{-\Delta_{\min}^2(t-1)}{2\kappa^2}\right) \\
&\leq \exp\left(-\frac{(t-1)\Delta_{\min}^2}{16\alpha^2}\right) + \exp\left(\frac{-\Delta_{\min}^2(t-1)}{2\kappa^2}\right),
\end{aligned}$$

where (a) is obtained by using the fact that $Q(x) \leq e^{-x^2/2}$ for $x > 0$ and Hoeffding's inequality [27].

Case 2 ($\mu_i < \mu_j$): Let \mathcal{E}_2 be the event that $|R_t - \mu t| \leq |\Delta_{ij}|t/2$,

$$\begin{aligned}
&\mathbb{P}(\mathcal{E}_{i,j}^{(t)}) \\
&\leq \mathbb{P}\left(\frac{\Theta_{t+1,j} - \Theta_{t+1,i}}{\sqrt{2}\eta_{t+1}} \leq \frac{\gamma_i - \gamma_j}{\sqrt{2}}\right) \\
&\leq \mathbb{P}\left(\frac{c(\alpha_j - \alpha_i)t - (g(\alpha_i) - g(\alpha_j))R_t}{\sqrt{2}\eta_t} \leq \frac{\gamma_i - \gamma_j}{\sqrt{2}}\right) \\
&\leq \mathbb{P}\left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \geq \frac{c(\alpha_j - \alpha_i)t - (g(\alpha_i) - g(\alpha_j))R_t}{\sqrt{2}\eta_t}, \mathcal{E}_2\right) \\
&\quad + \mathbb{P}(\mathcal{E}_2^c) \\
&\leq \mathbb{P}\left(\frac{\gamma_i - \gamma_j}{\sqrt{2}} \geq \frac{|\Delta_{ij}|t}{2\sqrt{2}\eta_t}\right) + \mathbb{P}(\mathcal{E}_2^c) \\
&\stackrel{(b)}{\leq} \exp\left(-\frac{t^2\Delta_{ij}^2}{16\eta_t^2}\right) + \exp\left(\frac{-\Delta_{ij}^2 t}{2\kappa^2}\right) \\
&\leq \exp\left(-\frac{t^2\Delta_{\min}^2}{16\eta_t^2}\right) + \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right) \\
&\leq \exp\left(-\frac{t\Delta_{\min}^2}{16\alpha^2}\right) + \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right)
\end{aligned}$$

$$\leq \exp\left(-\frac{(t-1)\Delta_{\min}^2}{16\alpha^2}\right) + \exp\left(\frac{-\Delta_{\min}^2(t-1)}{2\kappa^2}\right),$$

where (b) is obtained by using the fact that $Q(x) \leq e^{-x^2/2}$ for $x > 0$ and Hoeffding's inequality [27].

From combining both cases we get

$$\mathbb{P}(\mathcal{E}_{i,j}^{(t)}) \leq \exp\left(-\frac{(t-1)\Delta_{\min}^2}{16\alpha^2}\right) + \exp\left(\frac{-\Delta_{\min}^2(t-1)}{2\kappa^2}\right). \quad (14)$$

Therefore,

$$\begin{aligned} \mathbb{E}[\mathcal{C}_F^{\text{FTPL}}] &\leq M \sum_{t=1}^T \sum_{j>i} \mathbb{P}(\mathcal{E}_{i,j}^{(t)}) \\ &\leq MK^2 \sum_{t=1}^{T-1} \exp\left(-\frac{t\Delta_{\min}^2}{16\alpha^2}\right) + \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right) \\ &\quad + MK^2 \\ &\leq MK^2 \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2}. \end{aligned}$$

□

Proof of Theorem 2(b). Note that FTPL policy does not consider fetch cost M while taking the decisions. By using Lemma 7, 8 we get the result stated.

□

6.8. Proof of Theorem 2(c)

Lemma 9. *Under the W-FTPL policy, $T_s > t'$ if and only if there exist $t < t'$ such that the following condition holds:*

$$(g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t > -c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}},$$

or

$$(g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t < -c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}},$$

where $(\tilde{i}_t, \tilde{j}_t) = \arg \max_{i \neq j} \Theta_{t+1,i} - \Theta_{t+1,j}$, $\hat{\mu}_t = R_t/t$.

Proof. Let $(\tilde{i}_t, \tilde{j}_t) = \arg \max_{i \neq j} \Theta_{t+1,i} - \Theta_{t+1,j}$, $\hat{\mu}_t = R_t/t$. $T_s > t'$ if and only if there exist $t < t'$ such that

$$\begin{aligned}
& \frac{|\max_{i \neq j} (\Theta_{t+1,i} - \Theta_{t+1,j})|}{t} < \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}} \\
& \iff \frac{|\Theta_{t+1,\tilde{j}_t} - \Theta_{t+1,\tilde{i}_t}|}{t} > \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}} \\
& \iff \\
& |c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) + (g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t| > \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}} \\
& \iff \\
& (g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t > -c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) + \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}} \\
& \text{or} \\
& (g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t < -c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) - \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}.
\end{aligned}$$

This completes the proof. \square

Lemma 10. *Under the W-FTPL policy, $T_s > \frac{(\sqrt{\beta}-1)^2(\log M)^{1+\delta}}{\Delta_{\max}^2}$ with probability at least $1 - \frac{(\sqrt{\beta}-1)^2(\log M)^{1+\delta}}{M^{2(\log M)^\delta} \Delta_{\max}^2}$.*

Proof. Let \mathcal{E}_t be the event that

$$|\mu - \hat{\mu}_t| \leq \kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}},$$

where $\hat{\mu}_t = R_t/t$. By using Hoeffding's inequality [27] we get

$$\mathbb{P}(\mathcal{E}_t) \geq 1 - \frac{1}{M^{2(\log M)^\delta}}.$$

Let $\mathcal{E} = \cap_{t=1}^{T_0} \mathcal{E}_t$, where $T_0 = \frac{(\sqrt{\beta}-1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}$. Using the union bound, we get:

$$\mathbb{P}(\mathcal{E}) \geq 1 - \frac{T_0}{M^{2(\log M)^\delta}}.$$

Let $\epsilon_t = \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}$. We prove the lemma by contradiction. Let us consider

$$T_s < \frac{(\sqrt{\beta}-1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2},$$

and \mathcal{E} holds, then by Lemma 9, $\exists t < \frac{(\sqrt{\beta}-1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}$ such that

$$(g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-))\hat{\mu}_t > -c(\alpha_{j_t}^- - \alpha_{i_t}^-) + \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}},$$

or

$$(g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-))\hat{\mu}_t < -c(\alpha_{j_t}^- - \alpha_{i_t}^-) - \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}},$$

holds. There are two possibilities $\alpha_{j_t}^- < \alpha_{i_t}^-$, $\alpha_{j_t}^- > \alpha_{i_t}^-$ we consider them separately to prove the final result.

Case 1 ($\alpha_{j_t}^- < \alpha_{i_t}^-$): It follows that

$$\begin{aligned} & (g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-))\hat{\mu}_t > -c(\alpha_{j_t}^- - \alpha_{i_t}^-) + \epsilon_t \\ \implies & \hat{\mu}_t - \mu > \frac{c(\alpha_{i_t}^- - \alpha_{j_t}^-) + \epsilon_t}{g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-)} - \mu \\ \implies & \kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} > \frac{-\Delta_{\max} + \epsilon_t}{g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-)} \\ \implies & (\sqrt{\beta} - g(\alpha_{j_t}^-) + g(\alpha_{i_t}^-))\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} < \Delta_{\max} \\ \implies & (\sqrt{\beta} - 1)\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} < \Delta_{\max} \\ \implies & t > \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}, \end{aligned}$$

which is a contradiction. Alternatively,

$$\begin{aligned} & (g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-))\hat{\mu}_t < -c(\alpha_{j_t}^- - \alpha_{i_t}^-) - \epsilon_t \\ \implies & \hat{\mu}_t - \mu < \frac{c(\alpha_{i_t}^- - \alpha_{j_t}^-) - \epsilon_t}{g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-)} - \mu \\ \implies & -\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} < \frac{\Delta_{\max} - \epsilon_t}{g(\alpha_{j_t}^-) - g(\alpha_{i_t}^-)} \\ \implies & (\sqrt{\beta} - g(\alpha_{j_t}^-) + g(\alpha_{i_t}^-))\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} < \Delta_{\max} \\ \implies & (\sqrt{\beta} - 1)\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} < \Delta_{\max} \end{aligned}$$

$$\implies t > \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2},$$

which is also a contradiction.

Case 2 ($\alpha_{j_t}^{\sim} > \alpha_{i_t}^{\sim}$): It follows that:

$$\begin{aligned} (g(\alpha_{j_t}^{\sim}) - g(\alpha_{i_t}^{\sim}))\hat{\mu}_t &> -c(\alpha_{j_t}^{\sim} - \alpha_{i_t}^{\sim}) + \epsilon_t \\ \implies \hat{\mu}_t - \mu &< \frac{c(\alpha_{j_t}^{\sim} - \alpha_{i_t}^{\sim}) - \epsilon_t}{g(\alpha_{i_t}^{\sim}) - g(\alpha_{j_t}^{\sim})} - \mu \\ \implies -\kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} &< \frac{\Delta_{\max} - \epsilon_t}{g(\alpha_{i_t}^{\sim}) - g(\alpha_{j_t}^{\sim})} \\ \implies t &> \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}, \end{aligned}$$

which is a contradiction. Alternatively,

$$\begin{aligned} (g(\alpha_{j_t}^{\sim}) - g(\alpha_{i_t}^{\sim}))\hat{\mu}_t &< -c(\alpha_{j_t}^{\sim} - \alpha_{i_t}^{\sim}) - \epsilon_t \\ \implies \hat{\mu}_t - \mu &> \frac{c(\alpha_{j_t}^{\sim} - \alpha_{i_t}^{\sim}) + \epsilon_t}{g(\alpha_{i_t}^{\sim}) - g(\alpha_{j_t}^{\sim})} - \mu \\ \implies \kappa \sqrt{\frac{(\log M)^{1+\delta}}{t}} &> \frac{-\Delta_{\max} + \epsilon_t}{g(\alpha_{i_t}^{\sim}) - g(\alpha_{j_t}^{\sim})} \\ \implies t &> \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}, \end{aligned}$$

which is also a contradiction.

Therefore $T_s > \frac{(\sqrt{\beta}-1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2}$, if \mathcal{E} occurs, which happens with probability at least $1 - \frac{(\sqrt{\beta}-1)^2 \kappa^2 (\log M)^{1+\delta}}{M^2 (\log M)^\delta \Delta_{\max}^2}$. \square

Lemma 11. *Under the W-FTPL policy, $\mathbb{P}(T_s > t) \leq \exp(-\frac{\Delta_{\min}^2 t}{2\kappa^2})$ for $t > \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2}$.*

Proof. For $t > \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2}$, we have $\frac{\Delta_{\min}}{2} > \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}$. For $T_s > t$, we have

$$(g(\alpha_{j_t}^{\sim}) - g(\alpha_{i_t}^{\sim}))\hat{\mu}_t < -c(\alpha_{j_t}^{\sim} - \alpha_{i_t}^{\sim}) + \kappa \sqrt{\frac{\beta(\log M)^{1+\delta}}{t}},$$

and

$$(g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))\hat{\mu}_t > -c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}.$$

Let us consider the following cases

Case 1 ($\alpha_{\tilde{j}_t} < \alpha_{\tilde{i}_t}$, $\Delta_{\tilde{i}_t\tilde{j}_t} > 0$):

$$\begin{aligned} \mathbb{P}(T_s > t) &\leq \mathbb{P}\left(\hat{\mu}_t > \frac{-c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{(g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{\Delta_{\tilde{i}_t\tilde{j}_t} - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t})}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{\Delta_{\min} - \Delta_{\min}/2}{g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t})}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{\Delta_{\min}}{2}\right) \\ &\stackrel{(a)}{\leq} \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right), \end{aligned}$$

where (a) is obtained by using Hoeffding's inequality [27].

Case 2 ($\alpha_{\tilde{j}_t} < \alpha_{\tilde{i}_t}$, $\Delta_{\tilde{i}_t\tilde{j}_t} < 0$):

$$\begin{aligned} \mathbb{P}(T_s > t) &\leq \mathbb{P}\left(\hat{\mu}_t < \frac{-c(\alpha_{\tilde{j}_t} - \alpha_{\tilde{i}_t}) + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{(g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t}))}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{\Delta_{\tilde{i}_t\tilde{j}_t} + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t})}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{-\Delta_{\min} + \Delta_{\min}/2}{g(\alpha_{\tilde{j}_t}) - g(\alpha_{\tilde{i}_t})}\right) \\ &\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{-\Delta_{\min}}{2}\right) \\ &\stackrel{(a)}{\leq} \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right). \end{aligned}$$

where (a) is obtained by using Hoeffding's inequality [27].

Case 3 ($\alpha_{j_t}^{\tilde{\cdot}} > \alpha_{i_t}^{\tilde{\cdot}}$, $\Delta_{i_t j_t}^{\tilde{\cdot}} > 0$):

$$\begin{aligned}
\mathbb{P}(T_s > t) &\leq \mathbb{P}\left(\hat{\mu}_t < \frac{-c(\alpha_{j_t}^{\tilde{\cdot}} - \alpha_{i_t}^{\tilde{\cdot}}) - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{(g(\alpha_{j_t}^{\tilde{\cdot}}) - g(\alpha_{i_t}^{\tilde{\cdot}}))}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{-\Delta_{i_t j_t}^{\tilde{\cdot}} + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{g(\alpha_{i_t}^{\tilde{\cdot}}) - g(\alpha_{j_t}^{\tilde{\cdot}})}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{-\Delta_{\min} + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{g(\alpha_{i_t}^{\tilde{\cdot}}) - g(\alpha_{j_t}^{\tilde{\cdot}})}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu < \frac{-\Delta_{\min}/2}{g(\alpha_{i_t}^{\tilde{\cdot}}) - g(\alpha_{j_t}^{\tilde{\cdot}})}\right) \\
&\leq \mathbb{P}(\hat{\mu}_t - \mu < -\Delta_{\min}/2) \\
&\stackrel{(a)}{\leq} \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right).
\end{aligned}$$

where (a) is obtained by using Hoeffding's inequality [27].

Case 4 ($\alpha_{j_t}^{\tilde{\cdot}} > \alpha_{i_t}^{\tilde{\cdot}}$, $\Delta_{i_t j_t}^{\tilde{\cdot}} < 0$):

$$\begin{aligned}
\mathbb{P}(T_s > t) &\leq \mathbb{P}\left(\hat{\mu}_t > \frac{-c(\alpha_{j_t}^{\tilde{\cdot}} - \alpha_{i_t}^{\tilde{\cdot}}) + \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{(g(\alpha_{j_t}^{\tilde{\cdot}}) - g(\alpha_{i_t}^{\tilde{\cdot}}))}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{-\Delta_{i_t j_t}^{\tilde{\cdot}} - \kappa\sqrt{\frac{\beta(\log M)^{1+\delta}}{t}}}{g(\alpha_{i_t}^{\tilde{\cdot}}) - g(\alpha_{j_t}^{\tilde{\cdot}})}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{\Delta_{\min} - \Delta_{\min}/2}{g(\alpha_{i_t}^{\tilde{\cdot}}) - g(\alpha_{j_t}^{\tilde{\cdot}})}\right) \\
&\leq \mathbb{P}\left(\hat{\mu}_t - \mu > \frac{\Delta_{\min}}{2}\right) \\
&\stackrel{(a)}{\leq} \exp\left(\frac{-\Delta_{\min}^2 t}{2\kappa^2}\right),
\end{aligned}$$

where (a) is obtained by using Hoeffding's inequality [27]. The result follows by combining these cases. \square

Lemma 12. *Under the W-FTPL policy*

$$\mathbb{E}[T_s] \leq 1 + \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2} + \frac{1}{M^{2\beta}} \frac{2\kappa^2}{\Delta_{\min}^2}.$$

Proof. Let $T_1 = \lceil \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2} \rceil$. It follows that

$$\begin{aligned} \mathbb{E}[T_s] &= \sum_{t=1}^T \mathbb{P}(T_s > t) \\ &= \sum_{t=1}^{T_1} \mathbb{P}(T_s > t) + \sum_{t=T_1+1}^T \mathbb{P}(T_s > t) \\ &\stackrel{(a)}{\leq} T_1 + \sum_{t=T_1}^T \exp\left(-\frac{\Delta_{\min}^2 t}{2\kappa^2}\right) \\ &\leq 1 + \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2} + \frac{1}{M^{2\beta(\log M)^\delta}} \frac{2\kappa^2}{\Delta_{\min}^2}, \end{aligned}$$

where (a) is obtained by using Lemma 11. \square

Proof of Theorem 2(c). Under W-FTPL,

$$\begin{aligned} \mathbb{E}_r[\mathcal{C}^{\text{W-FTPL}}(T, r)] &= \mathbb{E}_r[T_s] + \mathbb{E}_r\left[\sum_{t=T_s+1}^T (c\rho_t + g(\rho_t)r_t)\right] \\ &\quad + M\mathbb{E}_r\left[\sum_{t=T_s+1}^T \mathbb{P}(\rho_t > \rho_{t-1})\right]. \end{aligned}$$

By the using definition of regret we get,

$$\begin{aligned} \mathcal{R}_S^{\text{W-FTPL}}(T) &\leq \mathbb{E}[T_s] + \sum_{t=1}^T (c\rho_t + \mu g(\rho_t)) - \mu_{i^*} \\ &\quad + M\mathbb{E}_r\left[\sum_{t=T_s+1}^T \mathbb{P}(\rho_t > \rho_{t-1})\right]. \end{aligned}$$

By Lemma 7, we get

$$\begin{aligned}\mathcal{R}_S^{\text{W-FTPL}}(T) &\leq \mathbb{E}[T_s] + \frac{16\alpha^2 + 4\kappa^2}{\Delta_{\min}} \\ &\quad + \left(\sqrt{2\log K} + \frac{2\sqrt{2h_1}\log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\ &\quad + M\mathbb{E}_r \left[\sum_{t=T_s+1}^T \mathbb{P}(\rho_t > \rho_{t-1}) \right].\end{aligned}$$

Let $\mathbb{E}_r[\mathcal{C}_f] = M\mathbb{E}_r \left[\sum_{t=T_s+1}^T \mathbb{P}(\rho_t > \rho_{t-1}) \right]$, then

$$\begin{aligned}\mathbb{E}_r[\mathcal{C}_f] &= \mathbb{E}_r[\mathcal{C}_f|T_s \leq T_0]\mathbb{P}(T_s \leq T_0) \\ &\quad + \mathbb{E}_r[\mathcal{C}_f|T_s > T_0]\mathbb{P}(T_s > T_0) \\ &\leq \mathbb{E}[\mathcal{C}_f|T_s = 1]\mathbb{P}(T_s \leq T_0) + \mathbb{E}[\mathcal{C}_f|T_s = \lceil T_0 \rceil] \\ &\stackrel{(a)}{\leq} M\mathbb{P}(T_s \leq T_0)\mathbb{E}[\mathcal{C}_F^{\text{FTPL}}] \\ &\quad + M \sum_{t=\lceil T_0 \rceil}^T \left(\exp\left(-\frac{\Delta_{\min}^2 t}{16\alpha^2}\right) + e^{-\Delta_{\min}^2 t/2\kappa^2} \right) \\ &\leq M\mathbb{P}(T_s \leq T_0) \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2} \\ &\quad + \frac{16\alpha^2}{\Delta_{\min}^2} \exp\left(-\frac{\Delta_{\min}^2 T_0}{16\alpha^2}\right) \\ &\quad + \frac{2\kappa^2}{\Delta_{\min}^2} \exp\left(-\frac{\Delta_{\min}^2 T_0}{2\kappa^2}\right).\end{aligned}$$

Here, (a) is obtained by using inequality (14). By considering $T_0 = \frac{(\sqrt{\beta}-1)^2\kappa^2(\log M)^{1+\delta}}{\Delta_{\max}^2}$ and using Lemma 10, 12 we get,

$$\begin{aligned}\mathcal{R}_S^{\text{W-FTPL}}(T) &\leq 1 + \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2} + \frac{1}{M^{2\beta(\log M)^\delta}} \frac{2\kappa^2}{\Delta_{\min}^2} \\ &\quad + \frac{(16\alpha^2 + 4\kappa^2)}{\Delta_{\min}} \\ &\quad + \left(\sqrt{2\log K} + \frac{2\sqrt{2h_1}\log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right)\end{aligned}$$

$$\begin{aligned}
& + M \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2 M^{2(\log M)^\delta}} \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2} \\
& + \left(\frac{16\alpha^2 M}{M \frac{(\sqrt{\beta}-1)^2 \kappa^2 \Delta_{\min}^2 (\log M)^\delta}{16\alpha^2 \Delta_{\max}^2}} + \frac{2M\kappa^2}{M \frac{(\sqrt{\beta}-1)^2 \Delta_{\min}^2 (\log M)^\delta}{2\Delta_{\max}^2}} \right) \\
& \times \frac{1}{\Delta_{\min}^2}.
\end{aligned}$$

For large values of M we have

$$\begin{aligned}
& \mathcal{R}_S^{\text{W-FTPL}}(T) \\
& \leq 1 + \frac{4\beta\kappa^2(\log M)^{1+\delta}}{\Delta_{\min}^2} + \frac{2\kappa^2}{\Delta_{\min}^2} \\
& \quad + \frac{(16\alpha^2 + 4\kappa^2)}{\Delta_{\min}} \\
& \quad + \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right) \\
& \quad + \frac{(\sqrt{\beta} - 1)^2 \kappa^2 (\log M)^{1+\delta}}{\Delta_{\max}^2} \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2} \\
& \quad + \frac{16\alpha^2 + 2\kappa^2}{\Delta_{\min}^2} \\
& \leq 1 + \beta\kappa^2(\log M)^{1+\delta} \left(\frac{4}{\Delta_{\min}^2} + \frac{16\alpha^2 + 3\kappa^2}{\Delta_{\min}^2 \Delta_{\max}^2} \right) \\
& \quad + (16\alpha^2 + 4\kappa^2) \left(\frac{1}{\Delta_{\min}} + \frac{1}{\Delta_{\min}^2} \right) \\
& \quad + \left(\sqrt{2 \log K} + \frac{2\sqrt{2h_1} \log K}{\Delta_{\min}} \right) \left(\alpha + \frac{4\kappa^2}{\alpha} \right).
\end{aligned}$$

□

6.9. Proof of Theorem 3(a)

Lemma 13. *The cost of the offline optimal policy is lower bounded as follows*

$$\mathcal{C}^{\text{OPT-OFF}}(T, r) \geq \frac{R_T}{\kappa^2} (\min_i c\alpha_i + g(\alpha_i)\kappa).$$

Proof. The cost incurred by the optimal offline policy with $M > 0$ is lower bounded by the cost incurred by the optimal offline policy when $M = 0$. For $M > 0$, there is an additional fetching cost. Therefore,

$$\begin{aligned}
\mathcal{C}^{\text{OPT-OFF}}(T, r) &\geq \sum_{t=1}^T \min_i (c\alpha_i + g(\alpha_i)r_t) \\
&\geq \sum_{t=1}^T \min_i (c\alpha_i/\kappa + g(\alpha_i)r_t) \\
&\geq \sum_{t=1}^T \min_i (c\alpha_i/\kappa + g(\alpha_i)) \mathbf{1}_{\{r_t \geq 1\}} \\
&\geq \sum_{t=1}^T \min_i (c\alpha_i/\kappa + g(\alpha_i)) r_t / \kappa \\
&\geq \frac{R_T}{\kappa^2} (\min_i c\alpha_i + g(\alpha_i)\kappa).
\end{aligned}$$

□

Lemma 14. *The expected cost of FTPL policy for any request sequence r for $\eta_t = \alpha\sqrt{t-1}$, $\alpha > 0$, can be upper bounded as follows*

$$\begin{aligned}
\mathbb{E}[\mathcal{C}^{\text{FTPL}}(T, r)] &\leq \left((3 + 2M/c) \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \right) R_T \\
&\quad + (c + M) \sum_{i=2}^K \frac{16\alpha^2}{c^2\alpha_i^2}.
\end{aligned}$$

Proof. FTPL does not host any fraction the service for time $t + 1$ if $c\alpha_i t + g(\alpha_i)R_t + \eta_{t+1}\gamma_i > R_t + \eta_{t+1}\gamma_1$, for all $i \neq 1$. The probability of hosting any non zero fraction of service ($p_{h,t}$) in time slot $t + 1$ is given as

$$\begin{aligned}
p_{h,t} &\leq \sum_{i=2}^K \mathbb{P}(c\alpha_i t + g(\alpha_i)R_t + \eta_{t+1}\gamma_i < R_t + \eta_{t+1}\gamma_1) \\
&= \sum_{i=2}^K \mathbb{P}\left(\frac{\gamma_1 - \gamma_i}{\sqrt{2}} > \frac{c\alpha_i t + g(\alpha_i)R_t - R_t}{\sqrt{2}\eta_{t+1}}\right)
\end{aligned}$$

$$= \sum_{i=2}^K \mathbb{P} \left(\frac{\gamma_1 - \gamma_i}{\sqrt{2}} > \frac{c\alpha_i t - (1 - g(\alpha_i))R_t}{\sqrt{2}\eta_{t+1}} \right).$$

Let $T' = \frac{2R_T}{c} \max_{\alpha_i \neq 0} \frac{1-g(\alpha_i)}{\alpha_i}$, if $t \geq T'$ then $R_t \leq R_T \leq \frac{ct}{2} \min_{\alpha_i \neq 0} \frac{\alpha_i}{1-g(\alpha_i)}$. Therefore for $t \geq T'$, $R_t \leq \frac{c\alpha_i t}{2(1-g(\alpha_i))}$ for all $\alpha_i \neq 0$ and we get

$$\begin{aligned} p_{h,t} &\leq \sum_{i=2}^K \mathbb{P} \left(\frac{\gamma_1 - \gamma_i}{\sqrt{2}} > \frac{c\alpha_i t - c\alpha_i t/2}{\sqrt{2}\eta_{t+1}} \right) \\ &\leq \sum_{i=2}^K \exp \left(-\frac{c^2 \alpha_i^2 t^2}{16\eta_{t+1}^2} \right) \\ &= \sum_{i=2}^K \exp \left(-\frac{c^2 \alpha_i^2 t}{16\alpha^2} \right). \end{aligned} \tag{15}$$

Therefore, the expected cost under FTPL can be bounded as

$$\begin{aligned} &\mathbb{E}[\mathcal{C}^{\text{FTPL}}(T, r)] \\ &\leq \sum_{t=1}^T (c + M)p_{h,t} + r_t \\ &= R_T + \sum_{t=1}^T (c + M)p_{h,t} \\ &\leq R_T + (c + M)T' + (c + M) \sum_{t=\lceil T' \rceil}^T p_{h,t} \\ &\stackrel{(a)}{\leq} R_T + (c + M)T' \\ &\quad + (c + M) \sum_{t=\lceil T' \rceil}^T \sum_{i=2}^K \exp \left(-\frac{c^2 \alpha_i^2 t}{16\alpha^2} \right) \\ &\leq \left((3 + 2M/c) \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \right) R_T \\ &\quad + (c + M) \sum_{i=2}^K \frac{16\alpha^2}{c^2 \alpha_i^2}. \end{aligned}$$

Here (a) follows from (15). If $T < T'$, then the bound will only have the first term and even in that case the above bound holds. Therefore the above bound holds for any T . \square

Proof of Theorem 3(a). By using Lemma 13, we have

$$\begin{aligned}
\sigma_A^{\text{FTPL}} &\leq \sup_r \frac{\mathbb{E}[\mathcal{C}^{\text{FTPL}}(T, r)]}{(\min_i c\alpha_i + g(\alpha_i)\kappa)R_T/\kappa^2} \\
&\stackrel{(a)}{\leq} \frac{\kappa^2(3 + 2M/c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \\
&\quad + \sup_r \frac{\kappa^2(M + c)}{R_T \min_i(c\alpha_i + g(\alpha_i)\kappa)} \sum_{i=2}^K \frac{16\alpha^2}{c^2\alpha_i^2} \\
&\leq \frac{\kappa^2(3 + 2M/c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \\
&\quad + \frac{\kappa^2(M + c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \sum_{i=2}^K \frac{16\alpha^2}{c^2\alpha_i^2}.
\end{aligned}$$

Here (a) is obtained by using Lemma 14. \square

6.10. Proof of Theorem 3(b)

Proof of Theorem 3(b). By using Lemma 13, we have

$$\begin{aligned}
\sigma_A^{\text{W-FTPL}} &\leq \sup_r \frac{R_{T_s} + \mathbb{E}[\mathcal{C}_A^{\text{FTPL}}(T, r)]}{R_T(\min_i \{c\alpha_i + g(\alpha_i)\kappa\})/\kappa^2} \\
&\stackrel{(a)}{\leq} \frac{\kappa^2(3 + 2M/c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \max_{\alpha_i \neq 0} \frac{1 - g(\alpha_i)}{\alpha_i} \\
&\quad + \frac{\kappa^2(M + c)}{\min_i(c\alpha_i + g(\alpha_i)\kappa)} \sum_{i=2}^K \frac{16\alpha^2}{c^2\alpha_i^2} \\
&\quad + \frac{\kappa^2}{\min_i(c\alpha_i + g(\alpha_i)\kappa)}.
\end{aligned}$$

Here (a) is obtained by using Theorem 3(a) and the fact that $R_{T_s} \leq R_T$. \square

6.11. Proof of Theorem 4(a)

Proof of Theorem 4(a).

$$\begin{aligned}\sigma_S^{\text{RR}}(T) &= \frac{\mathbb{E}[\mathcal{C}^{\text{RR}}(T, r)]}{\min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\}} \\ &= \frac{\mathcal{R}_S^{\text{RR}}(T) + \min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\}}{\min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\}} \\ &\leq 1 + \frac{\mathcal{R}_S^{\text{RR}}(T)}{\mu_{i^*} T}.\end{aligned}$$

By using Theorem 2 we have,

$$\sigma_S^{\text{RR}}(T) \geq 1 + \frac{M}{\mu_{i^*}} \left(\frac{1}{\frac{M}{\kappa - c\alpha_{i'} - g(\alpha_{i'})\kappa} + \frac{M}{c} + 2} - \frac{1}{T} \right) pd.$$

For large T we get,

$$\sigma_S^{\text{RR}}(T) > 1.$$

□

6.12. Proof of Theorem 4(b)

Proof of Theorem 4(b). By using Theorem 2 we have

$$\begin{aligned}\sigma_S^{\text{FTPL}}(T) &\leq 1 + \frac{1}{\mu_{i^*} T} (16\alpha^2 + 1) \left(\frac{M}{\Delta_{\min}^2} + \sum_{i \neq i^*} \frac{1}{\Delta_i} \right) \\ &= 1 + O(1/T).\end{aligned}$$

□

6.13. Proof of Theorem 4(c)

Proof of Theorem 4(c). By using Theorem 2 we have $\mathcal{R}_S^{\text{W-FTPL}}(T)$ as a constant and

$$\sigma_S^{\text{W-FTPL}}(T) \leq 1 + \frac{\mathcal{R}_S^{\text{W-FTPL}}(T)}{T\mu_{i^*}}.$$

Therefore we get,

$$\sigma_S^{\text{W-FTPL}}(T) \leq 1 + O(1/T).$$

□

References

- [1] R. S. Prakash, N. Karamchandani, V. Kavitha, S. Moharir, Partial service caching at the edge, in: 2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT), IEEE, 2020, pp. 1–8.
- [2] V. S. C. L. Narayana, M. Agarwala, R. S. Prakash, N. Karamchandani, S. Moharir, Online partial service hosting at the edge, CoRR abs/2103.00555 (2021). [arXiv:2103.00555](https://arxiv.org/abs/2103.00555).
URL <https://arxiv.org/abs/2103.00555>
- [3] V. C. L. Narayana, S. Moharir, N. Karamchandani, On renting edge resources for service hosting, ACM Transactions on Modeling and Performance Evaluation of Computing Systems 6 (2) (2021) 1–30.
- [4] T. Zhao, I.-H. Hou, S. Wang, K. Chan, Red/led: An asymptotically optimal and scalable online algorithm for service caching at the edge, IEEE Journal on Selected Areas in Communications 36 (8) (2018) 1857–1870.
- [5] T. L. Lai, H. Robbins, Asymptotically efficient adaptive allocation rules, Advances in applied mathematics 6 (1) (1985) 4–22.
- [6] G. S. Paschos, A. Destounis, L. Vigneri, G. Iosifidis, Learning to cache with no regrets, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, IEEE, 2019, pp. 235–243.
- [7] R. Bhattacharjee, S. Banerjee, A. Sinha, Fundamental limits on the regret of online network-caching, Proceedings of the ACM on Measurement and Analysis of Computing Systems 4 (2) (2020) 1–31.
- [8] C. Puliafito, E. Mingozzi, F. Longo, A. Puliafito, O. Rana, Fog computing for the internet of things: A survey, ACM Trans. Internet Technol. 19 (2) (2019) 18:1–18:41. doi:10.1145/3301443.
URL <http://doi.acm.org/10.1145/3301443>
- [9] Y. Mao, C. You, J. Zhang, K. Huang, K. B. Letaief, A survey on mobile edge computing: The communication perspective, IEEE Communications Surveys & Tutorials 19 (4) (2017) 2322–2358.

- [10] P. Mach, Z. Becvar, Mobile edge computing: A survey on architecture and computation offloading, *IEEE Communications Surveys & Tutorials* 19 (3) (2017) 1628–1656.
- [11] S. Pasteris, S. Wang, M. Herbster, T. He, Service placement with provable guarantees in heterogeneous edge computing systems, in: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, IEEE, 2019, pp. 514–522.
- [12] S. Bi, L. Huang, Y. A. Zhang, Joint optimization of service caching placement and computation offloading in mobile edge computing systems, *IEEE Transactions on Wireless Communications* 19 (7) (2020) 4947–4963. doi:10.1109/TWC.2020.2988386.
- [13] O. Ascigil, A. Tasiopoulos, T. K. Phan, V. Sourlas, I. Psaras, G. Pavlou, Resource provisioning and allocation in function-as-a-service edge-clouds, *IEEE Transactions on Services Computing* (2021) 1–1doi:10.1109/TSC.2021.3052139.
- [14] H. Tan, S. H.-C. Jiang, Z. Han, M. Li, Asymptotically optimal online caching on multiple caches with relaying and bypassing, *IEEE/ACM Transactions on Networking* (2021).
- [15] X. Xia, F. Chen, Q. He, J. Grundy, M. Abdelrazek, H. Jin, Online collaborative data caching in edge computing, *IEEE Transactions on Parallel and Distributed Systems* 32 (2) (2020) 281–294.
- [16] T. S. Salem, G. Neglia, S. Ioannidis, No-regret caching via online mirror descent, in: *ICC 2021-IEEE International Conference on Communications*, IEEE, 2021, pp. 1–6.
- [17] S. Mukhopadhyay, A. Sinha, Online caching with optimal switching regret, in: *2021 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2021, pp. 1546–1551.
- [18] S. Fan, I.-H. Hou, L. B. Van Sy Mai, Online service caching and routing at the edge with unknown arrivals (2022).
- [19] S. Wang, R. Uргаonkar, M. Zafer, T. He, K. Chan, K. K. Leung, Dynamic service migration in mobile edge-clouds, in: *2015 IFIP Networking Conference (IFIP Networking)*, IEEE, 2015, pp. 1–9.

- [20] G. Xiong, R. Singh, J. Li, Learning augmented index policy for optimal service placement at the network edge, arXiv preprint arXiv:2101.03641.
- [21] D. Zeng, L. Gu, S. Pan, J. Cai, S. Guo, Resource management at the network edge: A deep reinforcement learning approach, *IEEE Network* 33 (3) (2019) 26–33.
- [22] L. Chen, J. Xu, Budget-constrained edge service provisioning with demand estimation via bandit learning, arXiv preprint arXiv:1903.09080 (2019).
- [23] W. He, D. He, Y. Huang, Y. Zhang, Y. Xu, G. Yun-feng, W. Zhang, Bandit learning-based service placement and resource allocation for mobile edge computing, in: 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, IEEE, pp. 1–6.
- [24] A. Iosup, H. Li, M. Jan, S. Anoep, C. Dumitrescu, L. Wolters, D. H. Epema, The grid workloads archive, *Future Generation Computer Systems* 24 (7) (2008) 672–686.
- [25] A. Cohen, T. Hazan, Following the perturbed leader for online structured learning, in: *International Conference on Machine Learning*, PMLR, 2015, pp. 1034–1042.
- [26] J. Abernethy, C. Lee, A. Sinha, A. Tewari, Online linear optimization via smoothing, in: *Conference on Learning Theory*, PMLR, 2014, pp. 807–823.
- [27] W. Hoeffding, Probability inequalities for sums of bounded random variables, in: *The collected works of Wassily Hoeffding*, Springer, 1994, pp. 409–426.
- [28] D. Berend, A. Kontorovich, A sharp estimate of the binomial mean absolute deviation with applications, *Statistics & Probability Letters* 83 (4) (2013) 1254–1259.
- [29] H. Robbins, A remark on stirling’s formula, *The American mathematical monthly* 62 (1) (1955) 26–29.

7. Appendix

Proof of Lemma 1. For i.i.d. stochastic arrivals, we have

$$\begin{aligned}
\mathcal{R}_S^{\mathcal{P}}(T) &= \mathbb{E}_{\mathcal{P},r}[\mathcal{C}^{\mathcal{P}}(T,r)] - \min_i \{c\alpha_i T + g(\alpha_i)\mu T + M\alpha_i\} \\
&= \mathbb{E}_{\mathcal{P},r}[\mathcal{C}^{\mathcal{P}}(T,r)] - \min_i \{\mathbb{E}_r[c\alpha_i T + g(\alpha_i)R_T + M\alpha_i]\} \\
&\stackrel{(a)}{\leq} \mathbb{E}_{\mathcal{P},r}[\mathcal{C}^{\mathcal{P}}(T,r)] - \mathbb{E}_r[\min_i \{c\alpha_i T + g(\alpha_i)R_T + M\alpha_i\}] \\
&= \mathbb{E}_r[\mathcal{R}_A(T,r)] \leq \sup_r (\mathcal{R}_A^{\mathcal{P}}(T,r)) = \mathcal{R}_A^{\mathcal{P}}(T),
\end{aligned}$$

where (a) is obtained by Jensen's inequality. \square

Lemma 15. *If X is a random variable and $\mathbb{E}[f_1(X)] = \mathbb{E}[f_2(X)] = m$, then*

$$m - \mathbb{E}[\min\{f_1(X), f_2(X)\}] = \frac{1}{2}\mathbb{E}[|f_1(X) - f_2(X)|].$$

Proof.

$$\begin{aligned}
\min\{X, Y\} &= \frac{1}{2}(X + Y) - \frac{1}{2}(|X - Y|) \\
\implies m - \min\{f_1(X), f_2(X)\} &= m - \frac{1}{2}(f_1(X) + f_2(X)) \\
&\quad + \frac{1}{2}(|f_1(X) - f_2(X)|) \\
\implies m - \mathbb{E}[\min\{f_1(X), f_2(X)\}] &= \frac{1}{2}\mathbb{E}[|f_1(X) - f_2(X)|].
\end{aligned}$$

\square

Lemma 16. *If $X \sim \text{Bin}(T, p)$ then*

$$\begin{aligned}
\frac{1}{2}\mathbb{E}[|X - Tp|] &\geq \frac{\sqrt{Tp(1-p)}}{e\sqrt{2\pi}} \left[1 + \frac{1}{12T+1} - \frac{1}{12\lfloor Tp \rfloor} \right. \\
&\quad \left. - \frac{1}{12(T - \lfloor Tp \rfloor - 1)} \right].
\end{aligned}$$

Proof. By using equation (1) from [28] we get

$$\mathbb{E}[|X - Tp|]$$

$$\begin{aligned}
&= 2(1-p)^{T-\lfloor Tp \rfloor} p^{\lfloor Tp \rfloor + 1} (\lfloor Tp \rfloor + 1) \binom{T}{\lfloor Tp \rfloor + 1} \\
&= 2(1-p)^{T-\lfloor Tp \rfloor} p^{\lfloor Tp \rfloor + 1} \frac{T!}{\lfloor Tp \rfloor! (T - \lfloor Tp \rfloor - 1)!} \\
&\stackrel{(a)}{\geq} 2(1-p)^{T-\lfloor Tp \rfloor} p^{\lfloor Tp \rfloor + 1} \frac{1}{e\sqrt{2\pi}} \frac{T^{T+\frac{1}{2}} e^{\frac{1}{12T+1}}}{\lfloor Tp \rfloor^{T+\frac{1}{2}} e^{\frac{1}{12\lfloor Tp \rfloor}}} \\
&\quad \times \frac{1}{(T - \lfloor Tp \rfloor - 1)^{T-\lfloor Tp \rfloor-1+\frac{1}{2}} e^{\frac{1}{12(T-\lfloor Tp \rfloor-1)}}} \\
&\stackrel{(b)}{\geq} \frac{2}{e\sqrt{2\pi}} (1-p)^{T-\lfloor Tp \rfloor} p^{\lfloor Tp \rfloor + 1} \frac{T^{T+\frac{1}{2}} e^{\frac{1}{12T+1}}}{(Tp)^{\lfloor Tp \rfloor + \frac{1}{2}} e^{\frac{1}{12\lfloor Tp \rfloor}}} \\
&\quad \times \frac{1}{(T - Tp)^{T-\lfloor Tp \rfloor - \frac{1}{2}} e^{\frac{1}{12(T-\lfloor Tp \rfloor-1)}}} \\
&= \frac{2\sqrt{Tp(1-p)}}{e\sqrt{2\pi}} e^{\frac{1}{12T+1} - \frac{1}{12\lfloor Tp \rfloor} - \frac{1}{12(T-\lfloor Tp \rfloor-1)}} \\
&\stackrel{(c)}{\geq} \frac{2\sqrt{Tp(1-p)}}{e\sqrt{2\pi}} \left[1 + \frac{1}{12T+1} - \frac{1}{12\lfloor Tp \rfloor} \right. \\
&\quad \left. - \frac{1}{12(T-\lfloor Tp \rfloor-1)} \right],
\end{aligned}$$

where (a) follows by using the result in [29] which is $\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n}}$. (b) follows by using the fact that $x-1 \leq \lfloor x \rfloor \leq x$. (c) follows from the fact $e^x \geq 1+x$. \square