Improving Upon the generalized c-mu rule: a Whittle approach

Zhouzi Li* Keerthana Gurushankar Mor Harchol-Balter† Alan Scheller-Wolf

ABSTRACT

Scheduling a stream of jobs whose holding cost changes over time is a classic and practical problem. Specifically, each job is associated with a holding cost (penalty), where a job's instantaneous holding cost is some increasing function of its current age (the time it has spent in the system since its arrival) and its class. The goal is to schedule the jobs to minimize the time-average total holding cost across all jobs.

The seminal paper on this problem, by Van Mieghem in 1995 [10], introduced the generalized $c\mu$ rule for scheduling jobs. Since then, this problem has attracted significant interest but remains challenging due to the absence of a finite-dimensional state space formulation.

This paper translates the holding cost minimization problem to a novel Restless Multi-Armed Bandit (R-MAB) problem with a *finite* number of arms. Based on our R-MAB, we derive a novel Whittle Index policy, which is both elegant and intuitive. Our heuristic empirically improves upon the generalized $c\mu$ rule and all existing heuristics.

1. INTRODUCTION

Since the seminal paper by Van Mieghem in 1995 [10], the problem of scheduling jobs with Time-Varying Holding Cost (the TVHC problem) has been an important topic in the operations literature: In a single-server multi-class system with k classes of jobs, each job incurs a (time-varying) holding cost for every unit of time it remains in the system, and the goal of the scheduling policy is to minimize the time-average total holding cost across all jobs. Specifically, define the *age* of a job to be the time it has spent in the system since it arrived. For a job of class i, let $c_i(t)$ be the instantaneous holding cost when the job's age is t. We allow different classes to have different holding-cost functions (see Figure 1), but assume these functions are non-decreasing. Note that non-decreasing *instantaneous holding cost* is equivalent to convex *accumulated holding cost*, which is assumed in [10] and all its follow-on works. Also, throughout this paper, we assume that job sizes are *exponentially* distributed unless otherwise specified, and we assume that the job arrival process is Poisson. Let λ_i and μ_i, respectively, denote the arrival rate and the completion rate of class i jobs.

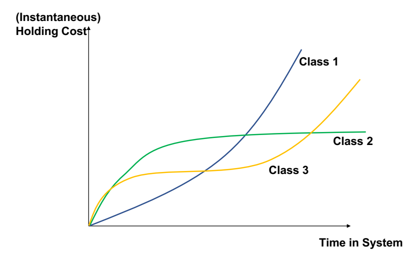

Figure 1: Classes with different holding costs.

In the special case when the holding cost of each class is *a constant function* (where $c_i(t) = c_i$), the optimal policy is the famous $c\mu$ rule, which always (preemptively) runs the job with the highest product $c_i \cdot \mu_i$ [5]. However, in the general case, this problem is much more complicated and the optimal policy is not known. In [10], an analogue of the $c\mu$ rule is proposed, which is now commonly referred to as the *generalized $c\mu$ rule*: The priority of a job is given by the product of its instantaneous holding cost, $c_i(t)$, and its instantaneous failure rate $\mu_i(t)$. While the generalized $c\mu$ rule is defined for general job sizes in [10], if the job sizes are exponentially distributed, the failure rate for class i is just μ_i. We discuss the generalized $c\mu$ rule in Section 1.1.

Note that both the generalized $c\mu$ rule and the $c\mu$ rule are *index policies*: Each job has an index (in both cases with value $c_i(t)\mu_i(t)$) which is a function of only the job's state, and the policy always serves that job with the highest index. Index policies are both simple and powerful. This paper aims to solve the TVHC problem within the class of preemptive *index policies*, yielding a heuristic solution to the problem in general. We now review the literature, and motivate our approach.

1.1 Generalized $c\mu$ rule is suboptimal

The generalized $c\mu$-rule has been shown to be asymptotically optimal for M/G/1 queues in the diffusion limit regime, where both the arrival and service rates go to infinity (and the total load goes to 1) [10]. However, outside of the asymptotic regime, it is known that the generalized $c\mu$-rule can perform poorly. Here we give an example to illustrate this.

Consider a system with two classes of jobs: one with deadlines and one without. For deadline-based jobs, the holding cost is zero before their deadline but becomes significantly higher once the deadline is passed, while jobs without dead-

*{zhouzil,kgurusha,harchol,awolf}@andrew.cmu.edu
†All authors are at Carnegie Mellon University. This work is supported by NSF-CIF-2403194, NSF-III-2322973, and NSF-CMMI-2307008.

lines have a constant holding cost. Assume that both classes have job sizes following the same $\text{Exp}(\mu)$ distribution. Under the generalized $c\mu$ rule, the system grants priority to deadline-based jobs only *after* they have missed their deadlines. Intuitively, however, by prioritizing these jobs *before* their deadline is reached, the total cost can be hugely reduced.

1.2 Prior attempts to improve upon the generalized $c\mu$ rule

One of the difficulties in solving the TVHC problem lies in the fact that a job's holding cost depends on its age. Hence, we need to track the ages of all jobs in the system, which requires an *infinite-dimensional* state space. The existing literature has considered more tractable versions of the TVHC problem that have a *finite-dimensional* state space. Such a finite-dimensional state space allows the authors to represent their problem as a finite-dimensional Markov Decision Problem (MDP), or, more specifically, a Restless Multi-Armed Bandit problem (R-MAB) with a finite number of arms.

One way to create a finite-dimensional state space for the TVHC problem is to assume a static setting, where all n jobs are present at time 0 and there are no new arrivals. This is the approach taken in [3, 2]. By limiting the number of jobs, the problem can now be translated to an n-arm R-MAB. From this R-MAB the authors then derive a Whittle index policy. The drawback of the static version setting is that the arrival rate (and consequently the load) of each class cannot be incorporated in the policy. This is unfortunate because, for example, in the example raised in Section 1.1, it is reasonable to expect that if load is higher, we might want to start working on the deadline-oriented jobs sooner than we would under lower load.

Another way to create a finite-dimensional state space is to consider a complementary setting, where the total holding cost for class i is a function of the *number* of class i jobs present. With this change, one only needs to track the number of jobs within each class, enabling the problem to be translated into a k-arm R-MAB. Several papers, [4, 7, 6], use the k-arm R-MAB to derive a Whittle index policy. This queue-length holding cost setting is complementary to our age-based holding cost setting, as the policy derived for one setting cannot be translated into a policy in the other setting.

1.3 Our approach and contributions

Our approach for solving the TVHC problem is shown in Figure 2. First, Theorem 1 (see Section 2) translates the TVHC problem to a novel *finite-arm* R-MAB problem. Second, Theorem 2 (see Section 3) proves indexability and derives the Whittle Index policy, which is known to be a good heuristic for R-MAB problems (see [9] for a discussion).

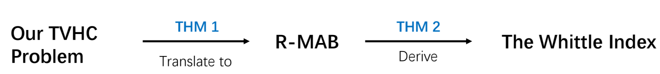

Figure 2: Using Theorem 1 and Theorem 2, we follow this road-map to derive the Whittle Index policy.

Our final result is a Whittle-based policy which always preemptively runs the job with the highest Whittle Index. The Whittle Index has an elegant and intuitive formulation given in Theorem 2 (see Section 3) and repeated here: Our Whittle Index for a class i job with age t, $W_i(t)$, is given by

$$W_i(t) = \mu_i \cdot \mathbf{E}\left[c_i(t+X)\right], \quad \text{where } X \sim \text{Exp}(\mu_i - \lambda_i). \quad (1)$$

Intuitively, while the generalized $c\mu$ rule focuses on the current holding cost $c_i(t)$ and schedules according to the index $\mu_i c_i(t)$, our Whittle index looks a little further into the future to age $t + X$. Note that if the load is high (λ_i is close to μ_i), then we look further into the future. This aligns with the motivating example in Section 1.1, where it is preferable to prioritize a job before its deadline, particularly when there are likely to be many jobs in the system.

It is also interesting to contrast our policy with the heuristic given in [2], which is the Whittle index in the static version. Under exponential job sizes, their index has the form $\mu_i \cdot \mathbf{E}\left[c_i(t+Y)\right]$, where $Y \sim \text{Exp}(\mu_i)$. Note that our index given in (1) matches theirs when $\lambda_i = 0$, where our setting degenerates to the static setting.

2. TRANSLATION TO AN R-MAB

In this section, we present the first part of the road-map of Figure 2: Translating the TVHC problem to an R-MAB problem. We first specify the class of policies we consider in Section 2.1, then give the R-MAB formulation in Section 2.2.

2.1 Policies considered

In this paper, we focus only on index policies. For our setting, an index policy is specified by a set of functions $\{V_i(\cdot)\}_{i=1}^k$, where $V_i(t)$ is the index of a job of type i with age t.

Moreover, since the holding cost is increasing and the job sizes are exponentially distributed, to minimize mean holding cost, we should schedule jobs within each type in FCFS (First Come First Serve) order, since older jobs have higher cost. Mathematically, we have the following lemma:

Lemma 1. *The optimal policy must serve jobs within each type in FCFS order.*

Motivated by Lemma 1, throughout this paper, we focus on index policies that enforce FCFS order within each class.

2.2 The R-MAB problem

Now we translate the TVHC problem into an R-MAB problem. Intuitively, our key idea is to let each arm of the R-MAB problem track the age of the oldest job within each class in the TVHC problem. At first this seems insufficent, because we're not capturing the state of all the other jobs within each class. However we will show how the FCFS ordering within each class ensures that the distribution of the ages of younger jobs can be effectively captured through the stochastic behavior of the Poisson arrival process. Our key theorem is as follows:

Theorem 1 (Translation to an R-MAB). *There exists an R-MAB problem, such that for any set of non-decreasing index functions $\{V_i(\cdot)\}_{i=1}^k$, the corresponding index policies (breaking ties by FCFS) in the TVHC problem and the R-MAB problem incur the same cost.*

Proof Sketch. We give the key idea to construct the continuous R-MAB as follows.

- The R-MAB has k arms, each representing a class. There are two actions for each arm (active or passive), where arm i is active means the oldest class i job is served, and passive means the job served at this moment is of some other type.

- *Arm state* $T_i(t)$: The ith arm's state is $T_i(t) \in \mathbb{R}$. The arm state can be interpreted in the TVHC problem as the age of the oldest type i job at time t. If there is no type i job in the system, $T_i(t)$ is negative, which means the next type i arrival happens $-T_i(t)$ time later. Note that for arm i, the action active is only allowed at time t when $T_i(t)$ is positive.

- *Transition Probability:* If the action for arm i is passive, the ith arm state grows with rate 1. Otherwise if arm i is active, the ith arm state may drop an $\text{Exp}(\lambda_i)$ amount according to a Poisson process (when completions happen).

  This transition function can be interpreted in the TVHC problem as follows: A passive action means that the oldest class i job is not in service and its age grows with rate 1. If the action is active, then in the next δ time period, the oldest class i job is served. With probability $\mu_i\delta$, the job is completed and leaves the system, in which case the oldest class i job in the system becomes the previously second-oldest class i job. Since the inter-arrival time follows the distribution $\text{Exp}(\lambda_i)$, the age of the oldest job drops by an $\text{Exp}(\lambda_i)$ amount.

- *Cost Function* $r_i(s)$*:* Arm i incurs a cost of $r_i(s)$ at state s, where $r_i(s)$ is defined to be

$$r_i(s) := c_i(s) + \mathbf{E}\left[\sum_{j=1}^{N_i} c_i(Y_j)\right], \quad (2)$$

  where $Y_j = \sum_{m=1}^{j} X_m$, $X_m \sim \text{Exp}(\lambda_i)$, and N_i is the random variable such that $Y_{N_i} < s$ and $Y_{N_i+1} \geq s$. For $s < 0$, define $r_i(s) = 0$.

  To interpret (2), observe that $r_i(s)$ represents the expected total instantaneous holding cost of all class i jobs given that the age of the oldest class i job is s: Since the index policy serves class i jobs in FCFS order, no "young" class i jobs (class i jobs younger than the current oldest one) have been completed. Since the inter-arrival times are distributed as $\text{Exp}(\lambda_i)$, their ages are distributed as

$$s - X_1,\ s - X_1 - X_2,\ \ldots,\ s - \sum_{m=1}^{N_i} X_m.$$

  Thus the expected total instantaneous cost of all the young class i jobs is $\mathbf{E}\left[\sum_{j=1}^{N_i} c_i(s - Y_j)\right]$, which is equal to $\mathbf{E}\left[\sum_{j=1}^{N_i} c_i(Y_j)\right]$ as the N_i arrivals are distributed as uniform order statistics in $(0, s)$.

- *Objective:* The objective is to minimize the long-run expected cost. Mathematically,

$$\text{Cost} = \mathbf{E}\left[\lim_{x\to\infty} \frac{1}{x} \int_0^x \sum_{i=1}^{k} r_i(T_i(t))dt \;\middle|\; T_i(0) = 0\right].$$

The full proof uses a coupling argument over sample paths to build equivalence between the TVHC and the R-MAB problems. □

3. DERIVATION OF THE WHITTLE INDEX

Given the R-MAB problem, we now perform the second part of the road-map of Figure 2: the derivation of the Whittle Index. In particular, following [8], we actually compute the Whittle index for the R-MAB problem in Theorem 1 with discount factor, and then take the discount factor to a limit of 0 to obtain an index policy of the original problem. The main result is to prove the indexability and to give the form of the Whittle Index as follows:

THEOREM 2. *The Whittle Index for the scheduling problem is*

$$W_i(t_0) = \mu\mathbf{E}\left[c_i(t_0 + X)\right],$$

where $X \sim \text{Exp}(\mu_i - \lambda_i)$.

4. CONCLUSION

This paper studies the classical TVHC problem: Jobs of different classes arrive over time, where each class of jobs is associated with a holding cost that increases with the job's age. The objective is to schedule the jobs so as to minimize the expected time-average total holding cost.

Our work takes a principled approach to the TVHC problem: (i) We derive the first representation of our problem as an R-MAB with a *finite* number of arms, and (ii) We derive a novel Whittle index policy for the resulting R-MAB. While the analysis is involved, the resulting policy is extremely simple and elegant (see Theorem 2). A longer version of this paper can be found in [1].

There are many directions for future work. First, our construction of the R-MAB problem may apply to more general settings (e.g., general job size distributions with increasing hazard rate). Second, there is much more work needed to find an *optimal* policy.

5. REFERENCES

[1] Zhouzi Li's webpage. https://li-zhouzi.github.io.

[2] S. Aalto. Whittle index approach to the multi-class queueing systems with convex holding costs and IHR service times. *Mathematical Methods of Operations Research*, 100(3):603–634, 2024.

[3] A. Anand and G. de Veciana. A Whittle's index based approach for QoE optimization in wireless networks. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(1):1–39, 2018.

[4] P. Ansell, K. D. Glazebrook, J. Nino-Mora, and M. O'Keeffe. Whittle's index policy for a multi-class queueing system with convex holding costs. *Mathematical Methods of Operations Research*, 57:21–39, 2003.

[5] C. Buyukkoc, P. Varaiya, and J. Walrand. The c-mu rule revisited. *Advances in Applied Probability*, 17(1):237–238, 1985.

[6] K. D. Glazebrook, R. Lumley, and P. Ansell. Index heuristics for multiclass M/G/1 systems with nonpreemptive service and convex holding costs. *Queueing Systems*, 45:81–111, 2003.

[7] M. Larrañaga, U. Ayesta, and I. M. Verloop. Index policies for a multi-class queue with convex holding cost and abandonments. In *The 2014 ACM international conference on Measurement and modeling of computer systems*, pages 125–137, 2014.

[8] A. Mandelbaum and A. L. Stolyar. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized cμ-rule. *Operations Research*, 52(6):836–855, 2004.

[9] J. Niño-Mora. Dynamic priority allocation via restless bandit marginal productivity indices. *Top*, 15:161–198, 2007.

[10] J. A. Van Mieghem. Dynamic scheduling with convex delay costs: The generalized c— mu rule. *The Annals of Applied Probability*, pages 809–833, 1995.