Controlling G-AIMD by Index Policy

Konstantin Avrachenkov, Vivek Borkar, Sarath Pattathil

Controlling G-AIMD by Index Policy

Konstantin E. Avrachenkov, Vivek S. Borkar and Sarath Pattathil*

Abstract— **We consider the Generalized Additive Increase Multiplicative Decrease (G-AIMD) dynamics for resource allocation with alpha fairness utility function. This dynamics has a number of important applications such as internet congestion control, charging electric vehicles, and smart grids. We prove indexability for the special case of MIMD model and provide an efficient scheme to compute the index. The use of index policy allows us to avoid the curse of dimensionality. We also demonstrate through simulations for another special case, AIMD, that the index policy is close to optimal and significantly outperforms a natural heuristic which penalizes the strongest user.**

I. Introduction

The method of alternating increase and decrease dynamics was proposed in [12] and implemented in TCP [20], [2] as a means to control congestion in the Internet. Since then, many modifications have been studied to improve further the performance of the Internet. We refer an interested reader to [1], [9], [11], [13], [23], [31].

It is a natural idea to apply control theoretic methods for analysis and improvement of congestion control and more generally, to resource allocation, see e.g., [8], [17], [18], [22], [24], [31] and references therein. Given the extent of randomness in the Internet traffic, it is surprising that not many works applied adaptive / learning control methods to congestion control, barring a few notable exceptions such as [26], [25], [19], [7]. Among these, only [19], [7] develop effective heuristic control policies based on the celebrated Whittle index [36]. We believe that adaptive / learning control methods based on multi-arm bandits and index policies are especially well suited for resource allocation problems, as they help avoid the curse of dimensionality and are naturally adapted to varying number of users.

Here we advance this line of research. Specifically, in [7], the index policy was proposed for the discrete state space setting without proof of indexability. The present model, G-AIMD, is much more general than that of [7] and includes as particular cases both AIMD and MIMD mechanisms. Furthermore, we consider here a general network topology. For a related continuous-space continuous-time model, it has been shown in [8] that the optimal policies have threshold structure. However, no Whittle index or primal-dual decomposition method has been proposed. Here, for the case of MIMD in the continuous space discrete-time model, we prove indexability (in fact, we consider a generalization of the Whittle index and discounted criterion) and propose an efficient procedure to compute the index. For AIMD, we argue that the heuristic reducing the rate of the strongest user [5], [28] coincides with Whittle-type index policy in the symmetric case. In the non-symmetric case, numerical experiments show that the former heuristic is significantly worse than the proposed index policy.

AIMD dynamics and its generalizations have been applied to smart grids, smart cities [14], [16], [21], [27] and charging of electric vehicles [6], [13], [15], [32]. This gives additional motivation to revisit such adaptive / learning control approaches to resource allocation.

The paper is structured as follows: in the next section we formally define the problem in a very general setting with discounted α-fairness as the optimization criterion. In Section III we prove indexability of the problem for MIMD and outline the computational procedure for the index. Then in Section IV, we show by numerical experiments for AIMD that the index policy significantly outperforms the natural heuristic that upon congestion reduces the allocation to the strongest user. Finally, we conclude in Section V with a roadmap for future research.

We conclude this section with a brief introduction to the Whittle index [36] adapted for discounted reward control problems. Let $X^i(t), t \geq 0, 1 \leq i \leq N$, be N Markov chains, each (say ith) with two modes of operations: active and passive, with associated transition kernels $p_1(\cdot|\cdot), p_0(\cdot|\cdot)$ resp. Let $r_1(X^i(t)), r_0(X^i(t))$ be instantaneous rewards in the respective modes with $r_1(\cdot) \geq r_0(\cdot)$. The objective is to schedule active/passive modes so as to maximize the total expected discounted reward $\sum_t \sum_j \beta^t E[r_{\nu^j(t)}(X^j(t))]$ where $\nu^j(t) = 1$ if jth process is active at time t and 0 if not, under the constraint $\sum_j \nu^j(t) \leq M \ \forall t$, i.e., at most M processes are active. This problem is provably hard [29], so one relaxes the constraint to $\sum_t \sum_j \beta^t E[\nu^j(t)] \leq M$. This makes it a problem with separable cost and constraints which, given the Lagrange multiplier λ, decouples into N individual problems with reward for passivity changed to $\lambda + r_0(\cdot)$. The problem is Whittle indexable if under optimal policy, the set of passive states for each of these problems increases monotonically from empty set to full state space. If so, the Whittle index for a given state x can be defined as the value of $\lambda = \lambda(x)$ for which both modes are equally desirable. The index policy is then to compute these for the current state profile, sort them in decreasing order, and render

This is the author-edited copy of the article accepted to IEEE CDC 2017.

*Authors are listed alphabetically. This work was partially supported by Indo-French CEFIPRA Collaboration Grants No.5100-IT1 “Monte Carlo and Learning Schemes for Network Analytics” and No. IFC/DST-Inria-2016-01/448 “Machine Learning for Network Analytics”.

K.E. Avrachenkov is with INRIA Sophia Antipolis, 2004 Route des Lucioles, 06902, Sophia Antipolis, France k.avrachenkov@inria.fr

V.S. Borkar is with the Department of Electrical Engineering, IIT Bombay, Powai, Mumbai 400076, India borkar.vs@gmail.com

S. Pattathil is with the Department of Electrical Engineering, IIT Bombay, Powai, Mumbai 400076, India sarathpattathil@iitb.ac.in

active the top M processes, the rest passive. The decoupling implies $O(N)$ growth of state space as opposed to the original problem for which it is exponential in N. Further, the processes are coupled only through a simple index policy. The latter is known to be asymptotically optimal as $N \uparrow \infty$ [35], [34]. However, no convenient general analytic bound on optimality gap seems available.

II. Problem formulation

Let us consider a Generalized AIMD (G-AIMD) dynamics with N users in discrete time. In absence of a control signal, the allocation to user k (e.g., transmission rate in Internet congestion control) increases according to the equation:

$$x_k(t+1) = x_k(t) + a_k x_k^{\gamma_k}(t), \tag{1}$$

with $\gamma_k \in [0,1]$ and $a_k > 0$. However, when the control signal is sent to user k, the resource allocation to user k abruptly decreases according to

$$x_k(t+1) = b_k x_k(t), \tag{2}$$

with $b_k \in (0,1)$. The above dynamics is fairly general and covers at least two important cases: if $\gamma_k = 0$ we retrieve the classical Additive Increase Multiplicative Decrease (AIMD) mechanism, and if $\gamma_k = 1$ we retrieve the Multiplicative Increase Multiplicative Decrease (MIMD) mechanism. MIMD is a very aggressive dynamics and in contrast AIMD is quite gentle, slowly probing the network capacity. The parametrized family of dynamics thus captures the extent of trade-off between the two as γ_k vary.

Denote by $x(t) = [x_1(t) \ \cdots \ x_N(t)]^T$ the vector of resource allocations at time t. The total resource is composed of R subresources (e.g., links in case of congestion control). The incidence matrix D indicates which subresources are used by which users. Namely,

$$d_{ij} = \begin{cases} 1, & \text{if subresource } i \text{ is used by user } j, \\ 0, & \text{otherwise.} \end{cases}$$

At each time step we operate the system under the constraint $Dx \leq \kappa$ where $\kappa = [\kappa_1 \ \cdots \ \kappa_R]^T$ is the vector of capacities of the subresources. This constraint is a generalization of the Whittle formulation and as therein, renders the problem hard. Following Whittle, we relax the above constraint to

$$(1-\beta)\sum_{t=0}^{\infty} \beta^t D x(t) \leq \kappa. \tag{3}$$

In contrast to [36], this is a discounted rather than average cost formulation. Our objective function is the discounted weighted α-fairness function

$$V^{\beta}(x_0) = (1-\beta)\sum_{t=0}^{\infty} \beta^t \sum_{k=1}^{N} w_k \frac{x_k^{1-\alpha}(t) - 1}{1-\alpha}. \tag{4}$$

The cases $\alpha = 0$, $\alpha \to 1$ and $\alpha \to \infty$ correspond to throughput maximization, proportional fairness and max-min fairness, resp. We shall consider the range $\alpha \in [0,1]$. As in [3], we have modified the α-fairness to remove discontinuity without affecting the optimal solution. Note also that optimization of the discounted sum of the α-fairness results in better instantaneous fairness [4]. Thus time-discounting is well suited to avoid imbalances on shorter time scales.

III. Index policy

The Whittle-type relaxation above makes it a separable control problem with separable constraints [36]. Thus we can focus on the individual dynamics of the processes, with that of the k^{th} process rewritten as:

$$x_k(t+1) = (1-u_k(t))(b_k x_k(t)) + u_k(t)(x_k(t) + a_k x_k^{\gamma_k}(t)).$$

where $u_k(t) = 1$ denotes that the rate is increased (this is the active state) at time t and $u_k(t) = 0$ indicates that the rate is decreased (this is the passive state). The relaxation of constraint (3) gives the following running cost

$$c_k(x_k) := w_k \frac{x_k^{1-\alpha} - 1}{1-\alpha} - \sum_j \mu_j (d_{jk} x_k - \kappa_j), \tag{5}$$

where μ_j's are the Lagrange multipliers. Alternatively, we can view them as a priori specified penalties for constraint violation. If μ_j are known, the optimization problem gets decoupled to individual optimization problems for single users. Therefore we consider the optimization problem for a single user. The dynamic programming equation for the β-discounted infinite horizon problem for user k is given by

$$\begin{aligned} V_{\lambda,k}^{\beta}(x_k) &= c_k(x_k) + \max(\lambda + \beta V_{\lambda,k}^{\beta}(b_k x_k), \\ &\quad \beta V_{\lambda,k}^{\beta}(x_k + a_k x_k^{\gamma_k})). \end{aligned} \tag{6}$$

Here λ can be interpreted as the subsidy for passivity.

The legitimacy of the Whittle-type index policy depends upon verification of a certain '*indexability*' property. We prove it below in Theorem 3.1 for the cases of $\gamma = 1$ (MIMD) alone. We do not have a complete proof for the case of $\gamma \in [0,1)$ at present, we leave it as a conjecture. Thus in what follows, $\gamma = 1$.

Henceforth we drop the subscript k from $V(.)$ for ease of notation and rewrite (6) as:

$$\begin{aligned} V_{\lambda}^{\beta}(x) &= c(x) + \max(\lambda + \beta V_{\lambda}^{\beta}(bx), \ \beta V_{\lambda}^{\beta}(x + a x^{\gamma})) \\ &= c(x) + \max_{u \in \{0,1\}} [(1-u)(\lambda + \beta V_{\lambda}^{\beta}(bx)) + \\ &\qquad u \beta V_{\lambda}^{\beta}(x + a x^{\gamma})]. \end{aligned} \tag{7}$$

Also, we assume that there is an upper bound on the permissible rate R, so that the dynamics of increase will be modified to $x(t+1) = ((1+a)x(t)) \wedge R$. In fact, TCP has limitation on the size of the congestion window, which limits the sending rate [2].

Lemma 3.1: $V_{\lambda}^{\beta}(\cdot)$ is a concave function in $[0,R]$ with

$$R \leq \left(\frac{w((1+a)^{1-\alpha} - b^{1-\alpha})}{\sum \mu_j d_j (1+a-b)} \right)^{\frac{1}{\alpha}}. \tag{8}$$

Proof: Define:

$$f(x) = w \frac{x^{1-\alpha} - 1}{1-\alpha} - \sum_j \mu_j (d_j x - \kappa_j), \tag{9}$$

where d_j is 1 if the user uses link j and 0 otherwise. Note that $f(.)$ is a concave function. Let the system start at rate $x(0) = x_0$. Then the evolution of the system is deterministic and the utility function will be of the form:

$$V_\lambda^\beta(x_0) = \max_{\{m_j\},\{k_j\}} \left(\sum_{i=0}^{\infty} \beta^i [f(m_i x_0) + k_i \lambda] \right) \quad (10)$$

for m_i, k_i defined as follows. Let $u(i) = \mathbb{1}_{\text{active at time i}}$. Then, if $u(i) = 1$, then $m_i = (1+a)m_{i-1}$ and $k_i = 0$. If $u(i) = 0$, then $m_i = bm_{i-1}$ and $k_i = 1$. We set $m_{-1} = 1$. Let

$$g(x_0) = \sum_{i=0}^{\infty} \beta^i [f(m_i x_0) + k_i \lambda]. \quad (11)$$

Now recall that the dynamics is of the form

$$x(t+1) = r_t x(t),$$

where $r_t \in \{b, 1+a\}$. We allow for randomized actions[1], thus picking in ith time slot $r_i = (1+a), k_i = 0$ with probability $p_i \in [0,1]$ and $r_i = b, k_i = 1$ with probability $1 - p_i$, then replacing the r.h.s. of (11) by its expectation. Note that $g(x_0)$ then is affine separately in each p_i when the rest are held constant. We have:

$$g'(x_0) = \mathbb{E}\left[\sum_{i=0}^{\infty} \beta^i m_i f'(m_i x_0) \right], \quad (12)$$

which is a decreasing function of x_0 since $f(\cdot)$ is concave. This yields:

$$\left(V_\lambda^\beta(x_0) \right)' = \left(\max_{\{p_j\}} g(x_0) \right)' \quad (13)$$

$$= g'(x_0) \Big|_{\underset{\{p_j\}}{\operatorname{argmax}}\, g(x_0)} \quad (14)$$

where we deduce equation (14) from equation (13) using Danskin's theorem (see Appendix B in [10]). To prove monotone decreasing property of a function with respect to several parameters, it suffices to prove it in each one keeping the rest fixed. To do so with respect to the value of p_i above in a fixed time slot, we may keep the values of $p_j, j \neq i$, fixed. A direct calculation then shows that under (8),

$$\frac{\partial^2 f(\bar{p}, x)}{\partial x \partial p_i} \geq 0,$$

for all i, where we have exhibited f as a function of $\bar{p} = \{p_i\}$ and x separately by abuse of notation. Using Theorem 10.4 and Theorem 10.7 of [33], we have that under (8),

$$x_0 \mapsto \underset{\{p_i\}}{\operatorname{argmax}}\ g(x_0)$$

is an increasing map, and therefore, the derivative of $V_\lambda^\beta(x_0)$ is decreasing. This shows that $V_\lambda^\beta(x_0)$ is concave in x_0. ■

[1] This does not affect the optimum which follows by a standard dynamic programming argument.

Concavity, as proven in Lemma 3.1, implies that $V_\lambda^\beta(.)$ has the property of decreasing differences, i.e.,

$$z > 0,\ x > y,\ \implies$$
$$V_\lambda^\beta(x+z) - V_\lambda^\beta(x) \quad \leq \quad V_\lambda^\beta(y+z) - V_\lambda^\beta(y). \quad (15)$$

The following results in the paper is stated for $\gamma \in [0,1]$. Note, however, that this is based on the conjecture that Lemma 3.1 holds for these γ. We have rigorous proof as above only for the case when $\gamma = 1$ under the rate constraint (8).

Lemma 3.2: The optimal policy is a threshold policy.

Proof: By (15), the difference

$$V_\lambda^\beta(x + ax^\gamma) - V_\lambda^\beta(bx) \quad (16)$$

is monotone decreasing in x. From the structure of the dynamic programming equation (7), it follows that the set of passive states is those x for which

$$\lambda \leq \beta(V_\lambda^\beta(x + ax^\gamma) - V_\lambda^\beta(bx)).$$

It then follows that the optimal policy must be a threshold policy such that for some threshold $x^*(\lambda)$, it is optimal to remain active below the threshold and passive above. ■

Theorem 3.1: This problem is 'Whittle-type indexable', i.e., as λ goes from $-\infty$ to ∞, the set of passive states goes from the empty set to the entire space monotonically.

Proof: Consider the measure ν defined as follows:

$$\int \phi d\nu = \sum_i \beta^i \phi(x(i))$$

for bounded continuous functions ϕ. Let $\Lambda = \sum_j \mu_j a_j$, then we can write the optimal reward (which depends on the controls used) in terms of λ as follows (we have used $\zeta(\lambda) = V_\lambda^\beta(x_0)$ for fixed β, x_0):

$$\zeta(\lambda) = \max_\nu \Bigg(\int w \frac{x^{1-\alpha} - 1}{1-\alpha} d\nu - \Lambda \int x d\nu + \Lambda L + \int \widetilde{\lambda}(\cdot) d\nu \Bigg) \quad (17)$$

where L is a constant which can be deduced from equation (9) and $\widetilde{\lambda}(x)$ is either λ or 0 depending on whether the state is passive or active, respectively. As a maximum of functions linear in λ, $\zeta(\lambda)$ is convex in λ. Let the optimal threshold be $x^*(\lambda)$, known to exist by Lemma 3.2. From equation (17) and Danskin's theorem, the right derivative (which, along with left derivative, exists by convexity) is:

$$\zeta'(\lambda) = \left(\frac{d}{d\lambda} \int \widetilde{\lambda}(\cdot) d\nu \right), \quad (18)$$

evaluated at $\nu^* :=$ the maximizing value of ν. Since by Lemma 3.2 there is an optimal threshold policy, we can maximize over the ν's corresponding to threshold policies. Thus, we may replace $\int \lambda(x) d\nu$ by $\lambda \sum_m \beta^m \mathbb{1}_{\{x_m \geq x_\nu\}}$ where x_ν is the threshold corresponding to the control policy ν and x_m is the rate sequence which evolves according to this control policy. Define $\{\widetilde{x}_m(\lambda)\}_{m\geq 0}$ to be the optimal rate sequence. Then by Danskin's Theorem,

$$\zeta'(\lambda) = \sum_m \beta^m \mathbb{1}_{\{\widetilde{x}_m(\lambda)\geq x^*(\lambda)\}}. \tag{19}$$

This function should be increasing in λ (since $\zeta(\lambda)$ is convex). However, $\widetilde{x}_m(\lambda)$ is decreasing in λ (This is because the utility in staying passive is linearly increasing in λ, when the other parameters are fixed. Therefore, as λ increases, there is more utility in staying passive which leads to a decrease in the rate). These two statements imply that $x^*(\lambda)$ has to decrease with λ. Therefore, as λ goes from $-\infty$ to ∞, the set of passive states goes from the empty set to the entire space *monotonically*. This proves the Whittle-type indexability of the problem. ■

In particular, $\lambda \mapsto \zeta(\lambda)$ is Lipschitz.

Proposed Index Policy: The strategy that we propose for resource allocation is the following: Whenever at least one subresource (link) is saturated, we compare the index of all the users. The one with the lowest index on the congested subresource has to reduce its transmission rate. Thereafter the system continues to evolve as before.

Note that the Whittle-type index can be precomputed offline by a numerical procedure described below and each user can store the index values.

Computation of the 'Whittle-type Index': The 'Whittle-type Index' for rate x is obtained from the following set of iterations

$$V^\beta_{\lambda_n}(y) = c(y) + \lambda_n + \beta V^\beta_{\lambda_n}(by) \quad \text{if } y > x, \tag{20}$$

$$V^\beta_{\lambda_n}(y) = c(y) + \beta V^\beta_{\lambda_n}(y + ay^\gamma) \quad \text{if } y \leq x, \tag{21}$$

$$\lambda_{n+1} = \lambda_n + \eta(\beta V^\beta_{\lambda_n}(x + ax^\gamma) - \beta V^\beta_{\lambda_n}(bx) - \lambda_n), \tag{22}$$

where η is a small stepsize (e.g., taken as 0.01).

Theorem 3.2: The sequence λ_n converges to a small (to be precise, $O(\eta)$) neighborhood of $\lambda(x)$.

Proof: As the problem is indexable, the Whittle-type index $\lambda(x) = \beta V^\beta_\lambda(x+ax^\gamma) - \beta V^\beta_\lambda(bx)$. At each iteration n, (20)-(21) constitute the linear system for evaluating the value function V_n for the prescribed threshold policy with threshold x, and the current subsidy for passivity, λ_n. Therefore if λ_n is more than the index, the 'error' term in the RHS of equation (22) is < 0 and if λ_n is less than the index, the 'error' term in the RHS of equation (22) is > 0. Given this fact the dynamics of equation (22) is such that λ_n converges to an $O(\eta)$ neighborhood of the index $\lambda(x)$. ■

Note that we are interested only in the ordinal comparison of the indices, so small errors do not matter. A bigger problem is the fact that this computation is only for a fixed x and x takes values in a continuum. So one has to use a further approximation, e.g., evaluate $\lambda(x)$ for a judiciously chosen finite collection of x's and interpolate.

IV. Simulations

For simulations, we first consider a simple network in which there are two users sharing a single resource (e.g., link). We simulate the AIMD model (i.e. $\gamma = 0$) in which the increase factor is $a = 0.5$ and the decrease factor is $b = 0.5$ for both the users. We take the capacity of the link to be 10. We have taken $\beta = 0.7$ (it can be any number less than 1).

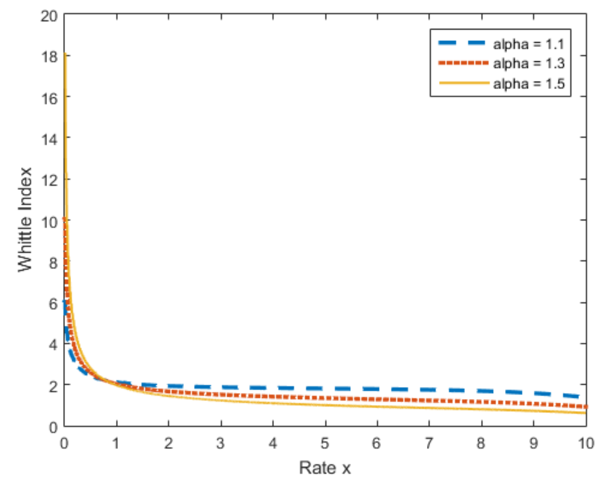

Fig. 1. Plot of Whittle Index on varying alpha (weight = 1).

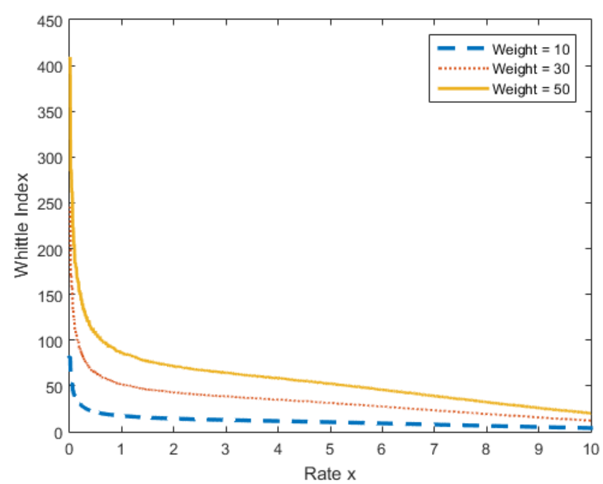

Fig. 2. Plot of Whittle Index varying with weight (alpha = 1.3).

Figures 1 and 2 show how the proposed 'Whittle-type Index' varies with the rate of the users for different values of α and weight w. From the plots, we see that the users who are too close to capacity are penalized by having a smaller value of the index and therefore their rates are reduced by the proposed scheme. Also, rates very near zero have a higher index which indicates that this policy tries to avoid users having near zero rates, irrespective of their weights. As the figures show, the 'Whittle-type Index' is a decreasing function of the rate. This implies that if all the weights, increase and decrease factors are the same, the 'Whittle-type index' will reduce to the scheme [5], [28], where we cut the rate of the user with the highest rate any time when there is congestion on the link. So now, in the completely symmetric case, the 'reduce maximum rate' heuristic [5], [28] has justification as a Whittle-type index policy. However, as will be demonstrated below, the heuristic 'reduce maximum rate' can drastically underperform in the non-symmetric case.

In Figure 3, we keep weights of the two users fixed at 1 with different increase factors. The index policy is seen to perform better than the scheme which penalizes the user with maximum rate and is closer to the optimal policy computed by value iteration [30]. Note that because of the curse of dimensionality, the optimal policy cannot be easily computed for more than 2-3 users.

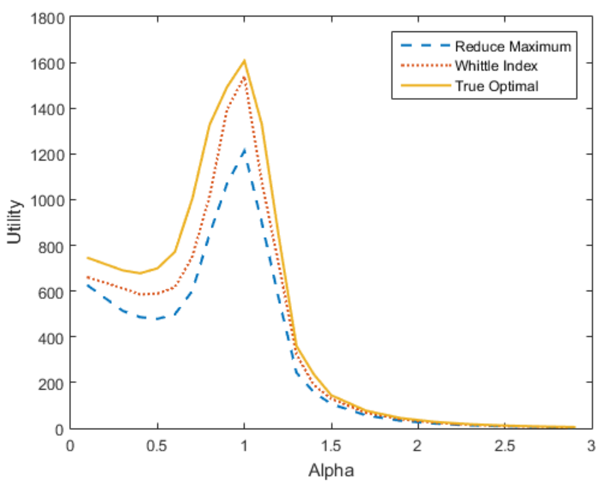

Fig. 3. Example with 2 users.

In Figures 4 and 5, we have extended the above setup, to 10 users (Capacity 100) and 100 (Capacity 1000) users, respectively. The weight for each user is still fixed at 1. The increase factor a_k is different for each user, chosen uniformly from the interval [0,1]. We observe that in the non-symmetric case with a large number of users, there is a big difference in the performance between the Whittle-type index and the 'reduce maximum rate' heuristic.

In Figure 6, we plot the results from the simulation of a system in which there are 3 users sharing 2 subresources (e.g., links). The capacity of both the links are kept the same at 10. One of the users uses both links, the other two use only one each. The incidence matrix in this case is given by

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The weights are taken to be 1 for all three users and the increase $a_k = 0.25$ for the user using both links and $a_k = 0.5$ for the other two users. This choice is motivated by the fact that TCP connections along longer routes take more time to

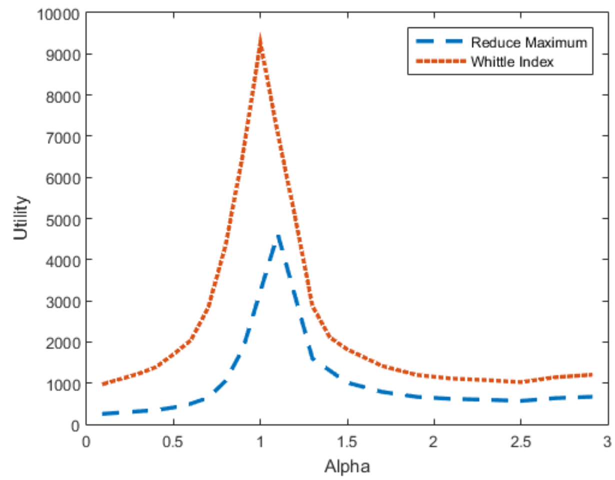

Fig. 4. Example with 10 users.

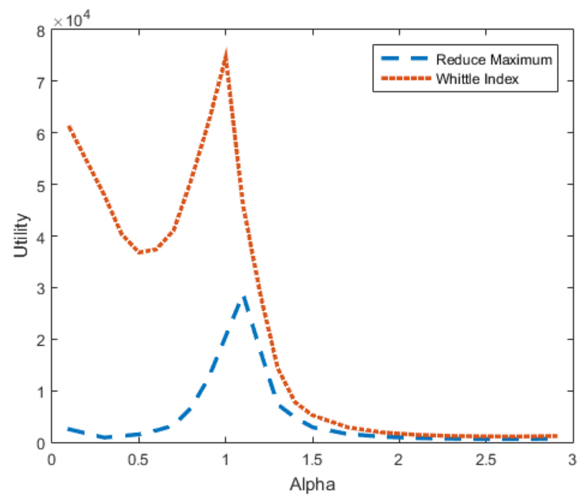

Fig. 5. Example with 100 users.

increase their transmission rates. The index policy performs significantly better than the 'reduce maximum rate' heuristic.

V. Conclusion and Future Research

We considered the G-AIMD dynamics (of which AIMD and MIMD are particular cases) for resource allocation with time-discounted α-fairness utility function. Time-discounting avoids imbalances on shorter time scales. We prove Whittle-type indexability of MIMD model and provide an efficient scheme to compute the index. We conjecture that G-AIMD dynamics is also indexable. The index policy allows us to avoid the curse of dimensionality. For AIMD, the index policy coincides with the 'reduce maximum rate' heuristic in the symmetric case. In the non-symmetric case, numerical experiments show that the index policy can significantly outperform this heuristic.

The G-AIMD dynamics has important applications, e.g., to internet congestion control, charging electric vehicles and

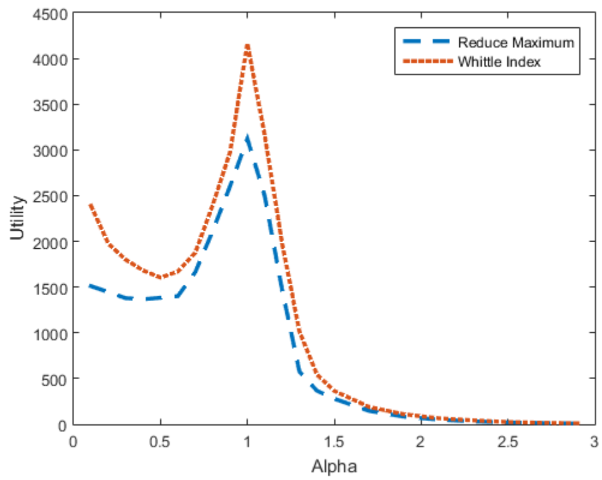

Fig. 6. Example with 3 users and 2 links.

smart grids. Thus one research and development agenda is to adapt the proposed general scheme to various application cases. Another interesting direction is to extend the index based approach to the case of varying number of users. One more direction of research is to incorporate adaptation with respect to other parameters such as the increase factor.

References

[1] Abed, G. A., Ismail, M., and Jumari, K., "A survey on performance of congestion control mechanisms for standard TCP versions." *Australian Journal of Basic and Applied Sciences*, 5(12), 1345-1352, 2011.

[2] Allman, M., Paxson, V., and Blanton, E., *TCP congestion control.* No. RFC 5681. 2009.

[3] Altman, E., Avrachenkov, K., and Garnaev, A., "Generalized α-fair resource allocation in wireless networks." *47th IEEE Conference on Decision and Control*, 2008.

[4] Altman, E., Avrachenkov, K. and Ramanath, S., "Multiscale fairness and its application to resource allocation in wireless networks." *Computer Communications*, 35(7), 820-828, 2012.

[5] Altman, E., El Azouzi, R., Ros, D., and Tuffin, B., "Loss strategies for competing AIMD flows." *Computer Networks*, 50(11), 1799-1815, 2006.

[6] Ardakanian, O., Rosenberg, C. and Keshav, S., "Real time distributed congestion control for electrical vehicle charging." *ACM SIGMETRICS Performance Evaluation Review*, 40(3), 38-42, 2012.

[7] Avrachenkov, K., Ayesta, U., Doncel, J. and Jacko, P., "Congestion control of TCP flows in Internet routers by means of index policy." *Computer Networks*, 57(17), 3463-3478, 2013.

[8] Avrachenkov, K., Habachi, O., Piunovskiy, A. and Zhang, Y., "Infinite horizon optimal impulsive control with applications to Internet congestion control." *International Journal of Control*, v 88(4), 703-716, 2015.

[9] Barakat, C., Altman, E. and Dabbous, W., "On TCP performance in a heterogeneous network: A survey." *IEEE Communications Magazine*, v. 38(1), 40-46, 2000.

[10] Basar, T. and Bernhard, P., *H^∞- Optimal Control and Related Minimax Design Problems: a dynamic game approach*, Springer, 2008.

[11] Callegari, C., Giordano, S., Pagano, M. and Pepe, T., "A survey of congestion control mechanisms in linux TCP." *Distributed Computer and Communication Networks*, 28-42, 2014.

[12] Chiu, D. and Jain, R., "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks." *Computer Networks and ISDN Systems*, v. 17, pp.1-14, 1989.

[13] Corless, M., King, C., Shorten, R. and Wirth, F., *AIMD dynamics and distributed resource allocation,* SIAM, Philadelphia, 2016.

[14] Crisostomi, E., Liu, M., Raugi, M. and Shorten, R., "Plug-and-play distributed algorithms for optimized power generation in a microgrid." *IEEE Trans. Smart Grid*, 5(4), 2145-2154, 2014.

[15] Fan, Z., "A distributed demand response algorithm and its application to PHEV charging in smart grids." *IEEE Trans. Smart Grid* 3(3), 1280-1290, 2012.

[16] Hermanns, H., and Wiechmann, H., "Demand-response management for dependable power grids." *Embedded Syst. Smart Appliances Energy Manag.* 3, 122, 2013.

[17] Hespanha, J. P., Bohacek, S., Obraczka, K. and Lee, J., "Hybrid modeling of TCP congestion control." *International Workshop on Hybrid Systems: Computation and Control.* Springer, 2001.

[18] Hollot, C. V., Misra, V., Towsley, D. and Gong, W. B., "A control theoretic analysis of RED." *Proceedings of INFOCOM 2001: 20th Joint Conference of the IEEE Computer and Communications Societies* Vol. 3, 2001.

[19] Jacko, P. and Sanso, B., "Optimal anticipative congestion control of flows with time-varying input stream." *Performance Evaluation*, 69(2), 86-101, 2012.

[20] Jacobson, V., "Congestion avoidance and control." *ACM SIGCOMM Computer Communication Review*, 18(4), 314-329, ACM, 1988.

[21] Keshav, S. and Rosenberg, C., "How internet concepts and technologies can help green and smarten the electrical grid." *ACM SIGCOMM Computer Communication Review*, 41(1), 109-114, 2011.

[22] Larranaga, M., Ayesta, U. and Verloop, I.M., "Stochastic and fluid index policies for resource allocation problems." *Computer Communications (INFOCOM), 2015 IEEE Conference on.* IEEE, 2015.

[23] Li, Y. T., Leith, D. and Shorten, R. N., "Experimental evaluation of TCP protocols for high-speed networks." *IEEE/ACM Transactions on Networking*, 15(5), 1109-1122, 2007.

[24] Low, S. H., Paganini, F., and Doyle, J. C., "Internet congestion control." *IEEE Control Systems*, 22(1), 28-43, 2002.

[25] Miller, A., "Using methods of stochastic control to prevent overloads in data transmission networks." *Automation and Remote Control*, 71(9), 1804-1815, 2010.

[26] Miller, B., Avrachenkov, K., Stepanyan, K. and Miller, G., "Flow control as a stochastic optimal control problem with incomplete information." *Problems of Information Transmission*, 41(2), 150-170, 2005.

[27] Naoum-Sawaya, J., Crisostomi, E., Liu, M., Gu, Y. and Shorten, R., "Smart procurement of naturally generated energy (sponge) for plug-in hybrid electric buses." *IEEE Transactions on Automation Science and Engineering.* Accepted, in press. 2016.

[28] Osipova, N., Blanc, A. and Avrachenkov, K., "Improving TCP fairness with the MarkMax policy." *Telecommunications, 2008. ICT 2008. International Conference on.* IEEE, 2008.

[29] Papadimitriou, C. H. and Tsitsiklis J. N., "The complexity of optimal queuing network control." *Mathematics of Operations Research* 24(2), 293-305, 1999.

[30] Puterman, M.L., *Markov decision processes: Discrete stochastic dynamic programming.* John Wiley & Sons, 2014.

[31] Srikant, R., *The mathematics of internet congestion control*. Springer. 2012.

[32] Stüdli, S., Crisostomi, E., Middleton R., and Shorten R., "A flexible distributed framework for realising electric and plug-in hybrid vehicle charging policies." *International Journal of Control* 85(8), 1130-1145, 2012.

[33] Sundaram, R.K. *A first course in optimization theory*. Cambridge university press, 1996.

[34] Verloop, I.M., "Asymptotically optimal priority policies for indexable and nonindexable restless bandits." *The Annals of Applied Probability* 26(4), 1947-1995, 2016.

[35] Weber, R.R., and Weiss,G., "On an index policy for restless bandits." *Journal of Applied Probability* 27.3, 637-648, 1990.

[36] Whittle, P., "Restless bandits: Activity allocation in a changing world." *Journal of Applied Probability*, 25.A, 287-298, 1988.