1.1 Related Work

Buffer sharing is a classic model in telecommunications. Early works enjoyed limited success in identifying the structure property of the policy which is optimal among a subset of the stationary policies. In [FG83,FGH81], Foschini et al. studied buffer sharing with two and three classes of jobs. They showed that the optimal coordinate convex policy that maximizes the average throughput or average buffer utilization is of threshold type. Jordan and Varaiya [JV94] showed that when the objective is maximizing average buffer utilization, the optimal policy is of generalized threshold type. The Internet booming makes an urgent call for a more practical approach to buffer sharing problem. Choudhury and Hahne [CH98] for example, proposed a scheme called dynamic threshold that combines the simplicity of static threshold and the adaptivity of push-out policies. In that scheme, the threshold for individual classes changes as unused buffer varies.

Buffer sharing can be thought of as admission control of multiple independent queues with constraints on the aggregated queue size which in turn is a special case of stochastic control for multiple independent random processes with the sample path constraints on the state space. A closely related category of stochastic control model is restless bandits in which the constraint is on the sample path of actions. The restless bandits model was initially proposed by Whittle [Whi88] as an extension to the classic multi-armed bandits model. In their inspiring work [BNM00], Bertsimas and Niño-Mora proposed a hierarchy of linear programming relaxations for the restless bandit problem based on the polyhedron projection representation idea described in [LS91]. They also proposed a primal-dual heuristic with exceptional performance. The application of projection representation idea to buffer sharing model in this work is to a large extend encouraged by their success.

1.2 Results

In this work, we obtain a hierarchy of increasingly stronger linear programming (LP) relaxations for the buffer sharing model with a class of reward functions including weighted sum of throughput and weighted sum of buffer utilization as special cases. The number of hierarchies is corresponding to the number of job classes. The last relaxation in the hierarchy is exact and corresponds to exponential size LP formulation of the problem as a Markov decision process. Intuitively, the first order relaxation is obtained by relaxing the constraint that no buffer overflow in every sample path to the constraint that no buffer overflow on average. The nature of this relaxation can also be viewed from the achievable performance region perspective. In the exact formulation of the buffer sharing model, the number of variables and the number of constraints to describe the achievable performance polytope is exponential in the number of classes. We then construct a polytope that contains the original performance region but has fewer variables and constraints. The maximization over this enlarged but simpler polytope gives a tight upper bound for the original formulation. A hierarchy of increasingly stronger linear programming relaxations can then be easily identified from this perspective.

Based on the first order LP relaxation, we also propose a heuristic policy which is applicable to a wide range of reward objectives including throughput maximization and buffer utilization maximization. The time complexity to obtain and the space complexity to store the proposed heuristic is polynomial in the product of buffer size and number of

The rate of state transition when the system is in state $\mathbf{i}$ is $\nu_{\mathbf{i}} = \sum_{k\in\mathscr{K}} \lambda_k + \mu_k 1\{i_k > 0\}$.

When the action is $\mathbf{a}$, the probability that a transition from state $\mathbf{i} \in \mathscr{S}$ to state $\mathbf{j} \in \mathscr{S}$ is $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}}$ and $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \lambda_k a_k/\nu_{\mathbf{i}}$ if $\mathbf{j} = \mathbf{i} + \mathbf{e}_k$; $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \mu_k/\nu_{\mathbf{i}}$ if $\mathbf{j} = \mathbf{i} - \mathbf{e}_k$; $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \sum_{k\in\mathscr{K}} \lambda_k(1 - a_k)/\nu_{\mathbf{i}}$ if $\mathbf{j} = \mathbf{i}$; $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = 0$ otherwise.

Thus the buffer sharing problem (discounted case) is in essence an α-discounted semi-Markov decision process with state space $\mathscr{S}$, the action set $\mathscr{A}(\mathbf{i})$, $\mathbf{i} \in \mathscr{S}$, the rate of state transition $\nu_{\mathbf{i}}$, $\mathbf{i} \in \mathscr{S}$, the state transition probabilities $q^{\mathbf{a}}_{\mathbf{i},\mathbf{j}}$, $\mathbf{i}, \mathbf{j} \in \mathscr{S}, \mathbf{a} \in \mathscr{A}(\mathbf{i})$, the initial state distribution Θ and the reward function $r_{\mathbf{i}}, \mathbf{i} \in \mathscr{S}$.

2.2 Equivalent Discrete time Markov Decision Process

For every α-discounted Semi-Markov decision process described above, we can construct a β-discounted discrete time Markov decision process with the same state space $\mathscr{S}$, the same action set $\mathscr{A}(\mathbf{i})$, $\mathbf{i} \in \mathscr{S}$, but different state transition probabilities $p^{\mathbf{a}}_{\mathbf{i},\mathbf{j}}$, $\mathbf{i}, \mathbf{j} \in \mathscr{S}, \mathbf{a} \in \mathscr{A}(\mathbf{i})$ and a different reward function $R_{\mathbf{i}}, \mathbf{i} \in \mathscr{S}$. Let $\nu = \sum_{k\in\mathscr{K}}(\lambda_k + \mu_k)$. When $\beta = \frac{\nu}{\alpha+\nu}$, $R_{\mathbf{i}} = r_{\mathbf{i}}/(\alpha + \nu)$, and $p^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \lambda_k a_k/\nu$ if $\mathbf{j} = \mathbf{i} + \mathbf{e}_k$; $p^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \mu_k/\nu$ if $\mathbf{j} = \mathbf{i} - \mathbf{e}_k$; $p^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = \sum_{k\in\mathscr{K}} (\lambda_k(1 - a_k) + \mu_k 1\{i_k = 0\})/\nu$ if $\mathbf{j} = \mathbf{i}$; $p^{\mathbf{a}}_{\mathbf{i},\mathbf{j}} = 0$ otherwise. The constructed discrete time model is equivalent to the continuous one in the sense that the optimal policy and the quantity of optimal objective for both models are the same. See [Ber95, §5] for the details and the proof of such a construction for Semi-Markov decision processes in general.

The equivalent discrete time model is more pleasant to work with because of the common denominator ν in the state transition probability. Therefore, the rest of the paper only deals with the equivalent discrete time Markov decision process.

3 THE EXACT LP FORMULATION FOR THE BUFFER SHARING MODEL

3.1 LP Formulation from the Achievable Performance Region Perspective

We first introduce the performance measure of the discounted average number of times that the system is in state $\mathbf{i}$ and action $\mathbf{a}$ is applied: $x^{\mathbf{a}}_{\mathbf{i}}(u) = \mathsf{E}^u_\Theta\left[\sum_{t=0}^{\infty} \beta^t I^{\mathbf{a}}_{\mathbf{i}}(t)\right]$where $u \in \mathscr{U}$ and $\mathscr{U}$ is the class of Markovian policies which selects the current action as a function (possibly randomized of the current state and time), $\Theta = (\theta_{\mathbf{i}})_{\mathbf{i}\in\mathscr{S}}$ is the initial state distribution, $I^{\mathbf{a}}_{\mathbf{i}}(t)$ is 1 if at time t, the system is in state $\mathbf{i}$ and action is $\mathbf{a}$.

The optimal performance objective can thus be written as

$$Z^* = \max_{u\in\mathscr{U}} \sum_{(\mathbf{i},\mathbf{a})\in\mathscr{C}} R_{\mathbf{i}} x^{\mathbf{a}}_{\mathbf{i}}(u).$$

Note that the reward $R_{\mathbf{i}}$ is only dependent on the state and independent of action.

Denote the vector $\mathbf{x}(u) = (x^{\mathbf{a}}_{\mathbf{i}}(u))_{(\mathbf{i},\mathbf{a})\in\mathscr{C}}$. The performance region spanned by the performance vector under all admissible policies $u \in \mathscr{U}$ is $\mathscr{X} = \{\mathbf{x}(u) \ : \ u \in \mathscr{U}\}$. Hence

$$Z^* = \max_{\mathbf{x}\in\mathscr{X}} \sum_{(\mathbf{i},\mathbf{a})\in\mathscr{C}} R_{\mathbf{i}} x^{\mathbf{a}}_{\mathbf{i}}$$

4.1 The First Order LP Relaxation

The first order LP relaxation can alternatively be viewed like this: For each class we allocate a buffer of size B with the constraint that the aggregated buffer utilization on average must be less than B. However for higher order relaxations, we found that the polyhedron projection viewpoint is more helpful.

For convenience, we introduce a few subset sets of the original state space, action space and state-action pair space, $\mathscr{S}_i^k = \{\mathbf{i} \in \mathscr{S} \ : \ i_k = i\}$, $\mathscr{A}_a^k(\mathbf{i}) = \{\mathbf{a} \in \mathscr{A}(\mathbf{i}) \ : \ a_k = a\}$ and $\mathscr{C}_{j,a}^k = \left\{(\mathbf{i}, \mathbf{a}) \ \in \mathscr{C} : \ \mathbf{i} \in \mathscr{S}_j^k, \mathbf{a} \in \mathscr{A}_a^k(\mathbf{i})\right\}$.

When we isolate class k from others, we need new notations for the state space, action set corresponding class k jobs: $S^k = \{i \ \in \mathbb{Z}_+ : \ b_k i \leq C\}$,
$A^k(i) = \left\{a \in \{0,1\} \ : \ \ i + a \in S^k\right\}$, and $C^k = \left\{(i,a) \ : \ i \in S^k, a \in A^k(i)\right\}$.

Introduce new parameters $\theta_i^k = \sum_{\mathbf{i} \in \mathscr{S}_i^k} \theta_{\mathbf{i}}, \forall i \in S^k, k \in \mathscr{K}$; $\nu_k = \lambda_k + \mu_k, \forall k \in \mathscr{K}$; $\beta_k = \frac{\nu_k}{\alpha + \nu_k}, \forall k \in \mathscr{K}$ $R_{i_k}^k = r_{i_k}^k / (\alpha + \nu_k)$ and $\left(p_{i_k, j_k}^{k, a_k}\right)_{i_k, j_k \in S^k, a_k \in A^k(i_k), k \in \mathscr{K}}$ as follows, $p_{i_k,j_k}^{k,a_k} = \lambda_k a_k / \nu_k$, if $j_k = i_k + 1$; $p_{i_k,j_k}^{k,a_k} = \mu_k / \nu_k$, if $j_k = i_k - 1$;
$p_{i_k,j_k}^{k,a_k} = \left(\lambda_k (1 - a_k) + \mu_k 1\left\{i_k = 0\right\}\right) / \nu_k$, if $j_k = i_k$ and $p_{i_k,j_k}^{k,a_k} = 0$ otherwise. Notice that ν_k and p_{i_k,j_k}^{k,a_k} are actually the normalized transition rate and the state transition probability when the class k jobs are isolated from others.

Introduce new variables $\left(x_{i_k}^{k,a_k}\right)_{(i_k,a_k) \in C^k, k \in \mathscr{K}}$. Geometrically speaking, we draw a new set of axes $\left(x_{i_k}^{k,a_k}\right)_{(i_k,a_k) \in C^k, k \in \mathscr{K}}$ in the achievable performance space $\mathscr{P}$ whose axes are $(x_{\mathbf{i}}^{\mathbf{a}})_{(\mathbf{i},\mathbf{a}) \in \mathscr{C}}$. We related the new axes to the old ones by the following linear transformation.

$$x_{i_k}^{k,a_k} = \frac{1 - \beta}{1 - \beta_k} \sum_{(\mathbf{i},\mathbf{a}) \in \mathscr{C}_{i_k,a_k}^k} x_{\mathbf{i}}^{\mathbf{a}}, \quad \forall (i,a) \in C^k, k \in \mathscr{K}.$$

The performance objective in (1) can be represented by the newly introduced variables

Lemma 4.1. $\sum_{(\mathbf{i},\mathbf{a}) \in \mathscr{C}} R_{\mathbf{i}} x_{\mathbf{i}}^{\mathbf{a}} = \sum_{k \in \mathscr{K}} \sum_{(i_k,a_k) \in C^k} R_{i_k}^k x_{i_k}^{k,a_k}$

We can also project the $|\mathscr{S}|$ number of facts of the polytope $\mathscr{P}$ into the subspace formed by new axes.

Lemma 4.2. *For all* $k \in \mathscr{K}$, $j_k \in S^k$, $\sum_{a_k \in A^k(j_k)} x_{j_k}^{k,a_k} = \theta_{j_k}^k + \beta_k \sum_{(i_k,a_k) \in C^k} x_{i_k}^{k,a_k} p_{i_k j_k}^{a_k}$.

Notice that the number of newly constructed equality constraints is $\sum_{k \in \mathscr{K}} \left|S^k\right|$.

Lemma 4.3. $B \geq \sum_{k \in \mathscr{K}} (1 - \beta_k) \sum_{(i_k,a_k) \in C^k} i_k b_k x_{i_k}^{k,a_k}$

Define the polytope $\mathscr{Q}^{(1)}$ as follows

$$\mathscr{Q}^{(1)} = \left\{ \left(x_{i_k}^{k,a_k}\right)_{(i_k,a_k) \in C^k, k \in \mathscr{K}} \in \mathbb{R}_+^{\sum_{k \in \mathscr{K}} \left|C^k\right|} : \ B \geq \sum_{k \in \mathscr{K}} (1 - \beta_k) \sum_{i_k \in S^k} i_k b_k \sum_{a_k \in A^k(i)} x_{i_k}^{k,a_k}, \right.$$
$$\left. \sum_{a_k \in A^k(j_k)} x_{j_k}^{k,a_k} = \theta_{j_k}^k + \beta_k \sum_{(i_k,a_k) \in C^k} x_{i_k}^{k,a_k} p_{i_k j_k}^{a_k}, \quad \forall j_k \in S^k, k \in \mathscr{K} \right\}.$$

5 A HEURISTICS BASED ON THE FIRST ORDER LP RELAXATION

The dual of the LP formulation in (2) is

$$Z^{(1)} = \min \sum_{k \in \mathscr{K}} \sum_{j_k \in S^k} \theta^k_{j_k} y^k_{j_k} + z \ B$$

subject to

$$y^k_{i_k} \geq R^k_{i_k} - z(1-\beta_k) i_k b_k + \beta_k \sum_{j_k \in S^k} p^{k,a_k}_{i_k j_k} y^k_{j_k}, \quad \forall (i_k, a_k) \in C^k, k \in \mathscr{K}$$

and $z \geq 0$.

Let $\left(\bar{x}^{k,a_k}_{i_k}\right)_{(i_k,a_k) \in C^k, k \in \mathscr{K}}$ and $\left(\left(\bar{y}^k_{i_k}\right)_{i_k \in S^k, k \in \mathscr{K}}, \bar{z}\right)$ be the optimal primal and dual solution pair to the first order relaxation and its dual. Let $\left(\gamma^{k,a_k}_{i_k}\right)_{(i_k,a_k) \in C^k, k \in \mathscr{K}}$ be the corresponding optimal reduced cost coefficients, for all $(i_k, a_k) \in C^k, \ k \in \mathscr{K}$,

$$\gamma^{k,a_k}_{i_k} = \bar{y}^k_{i_k} + (1-\beta_k) i_k b_k \bar{z} - R^k_{i_k} - \beta_k \sum_{j_k \in S^k} p^{k,a_k}_{i_k j_k} \bar{y}^k_{j_k}.$$

Notice that $\gamma^{k,a_k}_{i_k} \geq 0$ by definition. By the complementary slackness, we have $x^{k,a_k}_{i_k} \gamma^{k,a_k}_{i_k} = 0$. Therefore if $x^{k,1}_{i_k} > 0$ and $x^{k,0}_{i_k} = 0$ then $\gamma^{k,1}_{i_k} = 0 \leq \gamma^{k,0}_{i_k}$. The reduced cost is the rate of decrease in the objective value of the primal LP per unit increase in the value of variable $x^{k,a_k}_{i_k}$. This motives us to admit class k jobs when either one of the conditions below is satisfied:

1. If $x^{k,1}_{i_k} > 0$ and $x^{k,0}_{i_k} = 0$ when there are i_k class-k jobs in the buffer and when there is enough capacity left, i.e. $\sum_{l \in \mathscr{K}} i_l b_l + b_k \leq B$.
2. If $x^{k,1}_{i_k} > 0$, $x^{k,0}_{i_k} > 0$ and $\gamma^{k,1}_{i_k} \leq \gamma^{k,0}_{i_k}$($\sum_{j_k \in S^k} p^{k,1}_{i_k j_k} \bar{y}^k_{j_k} \geq \sum_{j_k \in S^k} p^{k,0}_{i_k j_k} \bar{y}^k_{j_k}$ after simplification) when there are i_k class-k jobs in the buffer and when there is enough capacity left.

As our model allows a wide class of reward functions including the weighted sum of throughput and the weighted sum of buffer utilization, the proposed heuristic is, therefore, more applicable than those heuristics with specialized objectives.

6 NUMERICAL EXAMPLES

In this section, we report a series of numerical experiments aimed at investigating the tightness of the relaxations and the performance of our heuristic proposed in this paper. Due to the limited space here, we only show the results for the throughput maximization case. There are in total six problem instances, A1, A2, A3, B1, B2 and B3 as shown in Table1. Problems A and problems B are different in the system configurations: the number of classes and the number of classes. Problems A1, A2 and A3 differ in arrival rates and

Table 2. Numerical results for throughput maximization.

	α	Z_{cs}	Z_{cp}	Z_{h}	Z^*	$Z^{(2)}$	$Z^{(1)}$
A1	0.01	*98.66*	*95.39*	*98.15*	98.99	98.99	98.99
	0.1	*7.74*	*7.87*	*8.33*	9.09	9.09	9.09
A2	0.01	*375.95*	*366.78*	*375.97*	378.68	380.36	384.81
	0.1	*30.88*	*30.96*	*31.04*	32.47	32.50	32.55
A3	0.01	*424.48*	*488.93*	*485.35*	494.58	498.21	503.55
	0.1	*42.61*	*43.53*	*43.03*	45.19	45.46	45.61
B1	0.01	*445.92*	*444.92*	*445.46*	-	-	446.00
	0.1	*39.92*	*38.40*	*39.88*	-	-	41.37
B2	0.01	*1717.60*	*1683.42*	*1713.08*	-	-	1718.49
	0.1	*144.81*	*143.68*	*145.38*	-	-	145.81
B3	0.01	*1622.65*	*2098.57*	*2104.72*	-	-	2119.88
	0.1	*187.29*	*192.32*	*187.90*	-	-	194.40

10% of the optimal for problem instances A1, A2 and A3 and with 5% of the optimal for problem instances B1, B2 and B3.

2. Regarding the performance of relaxation, the first order relaxation already provides upper bounds within 2% of the optimal Z^* for problem instances A1, A2 and A3 and within 4% of the optimal for problem instance B1, B2 and B3.

References

[Ber95] Dimitri P. Bertsekas. *Dynamic programming and optimal control*, volume 2. Athena Scientific, 1995.

[BNM00] Dimitri Bertsimas and José Niño-Mora. Restless bandits, linear programming relaxations and primal-dual index heuristic. *Operations Research*, 48(1), 2000.

[CH98] Abhijit K. Choudhury and Ellen L. Hahne. Dynamic queue length thresholds for shared-memory packet switches. *IEEE/ACM Transactions on Networking*, 6(2), 1998.

[FG83] Gerard J. Foschini and B. Gopinath. Sharing memory optimally. *IEEE Transactions on Communications*, COM-31(3), 1983.

[FGH81] Gerard J. Foschini, B. Gopinath, and J . F. Hayes. Optimal allocation of servers to two types of competing customers. *IEEE Transactions on Communications*, COM-29(7), 1981.

[HS04] Daniel P. Heyman and Matthew J. Sobel. *Stochastic model in operations research*, volume 2. Dover edition, 2004.

[JV94] S. Jordan and P.P. Varaiya. Control of multiple service, multiple resource communication networks. *Communications, IEEE Transactions on*, 42(11):2979–2988, 1994.

[LS91] Lászó Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1:166–190, 1991.

[Whi88] Peter Whittle. Restless bandits: activity allocation in a changing world. *J. Appl. Prob*, 25A:287–298, 1988.