Note on a Class of Admission Control Policies for the Stochastic Knapsack Problem

Adriana F. Gabor[1] and Jan-Kees C.W. van Ommeren[2]

[1] EURANDOM and Faculty of Mathematics and Computer Science, TUE, P.O.Box 513, 5600 MB, Eindhoven, The Netherlands
a.f.gabor@tue.nl

[2] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands
j.c.w.vanommeren@ewi.utwente.nl

Abstract. In this note we discuss a class of exponential penalty function policies recently proposed by Iyengar and Sigman for controlling a stochastic knapsack. These policies are based on the optimal solution of some related deterministic linear programs. By finding explicitly their optimal solution, we reinterpret the exponential penalty function policies and show that they belong to the class of threshold policies. This explains their good practical behavior, facilitates the comparison with the thinning policy, simplifies considerably their analysis and improves the bounds previously proposed.

1 Introduction

Recently, Iyengar and Sigman [1] proposed an exponential penalty function policy for controlling a loss network. A loss network is a network of resources, each with a known capacity. Requests for using the capacity are divided into classes, corresponding to arrival rates, service duration, resource requirements, and the profit they will bring for the network. There is no waiting room, so at every arrival of a request, it must be decided whether to accept the request or not. An admitted request occupies the allocated resource for the service duration and releases all the resources when it leaves the network. The objective is to design an admission policy that optimizes an appropriate performance measure of the revenue.

A major part of [1] is dedicated to the stochastic knapsack, which is a loss network with only one resource. For a review on other policies proposed for controlling the stochastic knapsack, see [5].

In this note we will focus on the exponential penalty function policy proposed in [1] for controlling a stochastic knapsack. This policy is based on the solution to a linear program. By solving this LP explicitly, we will show that for the stochastic knapsack, this policy reduces to a threshold policy. From the optimal solution of this LP we will derive an index, called the "threshold" index, which will divide the classes of different indices into two groups: one that will always be rejected, and one that will be accepted if there is enough capacity to accommodate them. The requests belonging to the class with the threshold index, are

S.-W. Cheng and C.K. Poon (Eds.): AAIM 2006, LNCS 4041, pp. 207–219, 2006.

accepted only if they satisfy an extra condition, given by the penalty function. By interpreting the exponential penalty function policy as a threshold policy, we are then able to improve the bounds on the expected reward rate obtained in [1] and to compare the exponential policy with the thinning policy proposed in [3].

This note closely follows [1] and is organized as follows. In Section 2 we present in detail the stochastic knapsack problem. With the exception of the last section, we will present the analysis for exponential service times. In Section 3 we discuss bounds for the expected reward rate achieved by admission policies in a stochastic knapsack. We start by discussing the upper bound proposed in [1] on the expected reward rate achieved by a policy and tighten it. Then we focus on the exponential penalty function policy and show that it is a threshold policy. This will lead to improved lower bounds for the expected reward rate. We continue by discussing the bounds in the "steady-state" regime. In Section 4 we will compare the exponential penalty function policy and the thinning policy (for the stochastic knapsack). In Section 5 we generalize the results presented in the previous sections to service times with a general distribution. We conclude with some remarks on the exponential penalty function policy for the stochastic knapsack and discuss why the results presented in this note are not easily generalized to loss networks.

2 Admission Control in the Stochastic Knapsack

The stochastic knapsack problem is a special case of a loss network problem and can be formulated as follows. There is a knapsack (network) of capacity $b \in R_+$. Requests for using the network belong to m independent Poisson arrival classes. Class i requests have an arrival rate λ_i and a service duration S_i which is exponentially distributed with rate μ_i (with the exception of Section 5), i.e., $S_i \sim exp(\mu_i)$. The requests in class i need a capacity b_i and pay r_i per unit time during their service duration. There is no waiting room in the system, therefore, each arriving request must either be accepted to the system and assigned a capacity allocation or rejected. When an accepted request departs after service completion, it releases all the allocated resources simultaneously. For simplicity, we will assume that the system is initially empty (all results easily generalize to the case when the system is not initially empty).

Let $T_{(i,n)}$, $i = 1, ..., m, n \geq 1$ denote the arrival epoch of the nth class i request. Since all admission decisions are made at arrival epochs, a feasible admission control policy π can be described as a collection of random variables $\pi = \{\pi_{(i,n)}, i = 1, ..., m, n \geq 1\}$, with $\pi_{(i,n)} = 0$ denoting that request of type i arriving at the epoch $T_{(i,n)}$ is rejected and $\pi_{(i,n)} = 1$ denoting that the request is accepted.

Let $x_i^\pi(t)$ be the number of class i requests in the system at time t under policy π. Define $x^\pi(t) = (x_1(t), ..., x_m(t))$. A request class i can be accepted only if there is sufficient capacity to accommodate it, that is

$$\sum_{i'=1}^{m} b_{i'} x_{i'}(t) + b_i \leq b.$$

The system controller is permitted to reject requests even if there is sufficient capacity.

The instantaneous reward rate $R^\pi(t)$ under policy π at time t is given by

$$R^\pi(t) = \sum_{i=1}^{m} r_i x_i^\pi(t).$$

The objective of the controller is to choose a policy π that maximizes a certain function of the reward rate process $\{R^\pi(t), t \geq 0\}$. Common performance measures for finite time horizon problems are either the expected total reward $E[\int_0^T R^\pi(s)ds]$ or the expected discounted reward $E[\int_0^T e^{-\beta s} R^\pi(s)ds]$, with $\beta > 0$; for infinite horizon problems, appropriate measures are either the discounted reward $E[\int_0^\infty e^{-\beta s} R^\pi(s)ds]$, $\beta > 0$ or the long-run average reward limit $\lim_{T\to\infty} \frac{1}{T} E[\int_0^T R^\pi(s)ds]$.

In [1] the authors construct feasible policies that perform well both in the transient period and in steady state. They first establish an upper bound $R^*(t)$ on the achievable expected reward rate $E[R^\pi(t)]$ and then construct a feasible policy $\bar{\pi}$ with expected reward rate $\mathbf{E}[R^\pi(t)] \simeq R^*(t)$. Thus, the policy $\bar{\pi}$ satisfies

$$E[\int_0^T e^{-\beta s} R^\pi(s)ds] \leq \int_0^T e^{-\beta s} R^*(s)ds \simeq E[\int_0^T e^{-\beta s} R^{\bar{\pi}}(s)ds],$$

for $\beta > 0$, which means that $\bar{\pi}$ is approximately optimal for any finite time horizon, and

$$\lim_{T\to\infty} \frac{1}{T} E[\int_0^T R^\pi(s)ds] \leq \lim_{T\to\infty} \frac{1}{T} \int_0^T R^*(s)ds \simeq \lim_{T\to\infty} \frac{1}{T} E[\int_0^T R^{\bar{\pi}}(s)ds],$$

that is, $\bar{\pi}$ is approximately optimal in steady state as well.

In the next sections we will discuss the admission policy $\bar{\pi}$ proposed in [1]. We will prove that it is a threshold policy, *i.e.*, only classes of a certain index are admitted to the network. This will also lead to improved bounds and an analytical comparison with the thinning policy proposed by Kelly.

3 Control Policies for the Stochastic Knapsack

3.1 Upper Bound on the Achievable Reward Rate

In this section we discuss the upper bound on the achievable reward at time t proposed in [1] and show a simple way of calculating it.

Let π denote any feasible control policy for the single resource model. Let $x_i^\pi(t)$ denote the number of class i requests at time t. Since $(x_i^\pi(t))_{i=\overline{1,m}}$ is feasible,

$$\sum_{i=1}^{m} b_i E[x_i^\pi(t)] \leq b.$$

Clearly, $E[x_i^\pi(t)] \leq E[q_i(t)]$, where $q_i(t)$ is the number of class i requests at time t in a corresponding $M/M/\infty$ system. Since the system is initially empty, $E[q_i(t)] = \rho_i(1 - e^{-\mu_i t})$, where $\rho_i = \frac{\lambda_i}{\mu_i}$ (see e.g. [6] page 75). Hence,

$$\alpha = (\frac{E[x_1^\pi(t)]}{\rho_1}, ..., \frac{E[x_m^\pi(t)]}{\rho_m})$$

is feasible for the following linear program:

$$\begin{array}{lll} & maximize & \sum_{i=1}^{m} r_i \rho_i \alpha_i \\ & P(t) & s.t. \sum_{i=1}^{m} b_i \rho_i \alpha_i \leq b, \\ & & 0 \leq \alpha_i \leq 1 - e^{-\mu_i t}. \end{array}$$

Let $\alpha^*(t)$ denote an optimal solution of the linear program P(t) and let $R^*(t)$ denote its optimal value. Clearly,

$$E[R^\pi(t)] = \sum_{i=1}^{m} r_i E[x_i^\pi(t)] \leq R^*(t).$$

In [1] the authors find an upper bound on $E[R^\pi(t)]$ by finding an upper bound on $R^*(t)$. Next we show how the exact value of $R^*(t)$ can be directly calculated.

Note that the problem P(t) is a continuous knapsack problem (see e.g. [4]). Thus, an optimal solution can be found as follows. Suppose from now on that the classes are indexed in decreasing order of the profit to capacity ratio, i.e.,

$$\frac{r_1}{b_1} \geq ... \geq \frac{r_m}{b_m}.$$

Let $k^*(t)$ be the index with the following property:

$$\sum_{i=1}^{k^*(t)-1} b_i \rho_i (1 - e^{-\mu_i t}) \leq b \text{ and } \sum_{i=1}^{k^*(t)} b_i \rho_i (1 - e^{-\mu_i t}) > b. \tag{1}$$

Then, the optimal solution of P(t) is given by:

$$\alpha_i^*(t) = \begin{cases} 1 - e^{-\mu_i t}, \text{ for } i < k^*(t) - 1 \\ \frac{b - \sum_{i=1}^{k^*(t)} b_i \rho_i (1 - e^{-\mu_i t})}{b_{k^*(t)} \rho_{k^*(t)} (1 - e^{-\mu_{k^*(t)} t})}, \text{ for } i = k^*(t) \\ 0, \text{ for } i > k^*(t). \end{cases} \tag{2}$$

Hence, we have obtained the following upper bound.

Theorem 1. *The reward rate $R^\pi(t)$ of any feasible policy π satisfies*

$$E[R^\pi(t)] \leq R^*(t) = \sum_{i=1}^{k^*(t)} r_i \rho_i \alpha_i^*(t),$$

where $R^*(t)$ is the optimal value of (P) and $\alpha_i^*(t)$ is given by (2).

3.2 The Exponential Penalty Function Policy

In this section we describe the penalty function policy proposed in [1] and show that it is a threshold policy. This leads to improved lower bounds for the expected reward obtained by the penalty function policy and facilitates the comparison with the thinning policy proposed by Kelly [3].

Next we introduce two linear programs which play an essential role in describing and analyzing the penalty policy.

Define the "steady state" version of $P(t)$ as

$$\begin{array}{lll} maximize & & \sum_{i=1}^{m} r_i \rho_i \alpha_i \\ P & s.t. & \sum_{i=1}^{m} b_i \rho_i \alpha_i \le b, \\ & & 0 \le \alpha_i \le 1. \end{array}$$

Since P is a continuous knapsack problem, it's optimal solution α^* has the following structure:

$$\alpha_i^* = \begin{cases} 1, \text{ for } i < k^* \\ \frac{b - \sum_{i=1}^{k^*} b_i \rho_i}{b_{k^*} \rho_{k^*}}, \text{ for } i = k^* \\ 0, \text{ for } i > k^*, \end{cases} \quad (3)$$

where k^* is the index for which

$$\sum_{i=1}^{k^*-1} b_i \rho_i \le b \text{ and } \sum_{i=1}^{k^*} b_i \rho_i > b.$$

Consider the following perturbation of the program P.

$$\begin{array}{lll} maximize & & \sum_{i=1}^{m} r_i \rho_i \alpha_i \\ P_\epsilon & s.t. & \sum_{i=1}^{m} b_i \rho_i \alpha_i \le \frac{b}{1+4\epsilon}, \\ & & 0 \le \alpha_i \le 1. \end{array}$$

The optimal solution α^ϵ of P_ϵ is : $\alpha_i^\epsilon = 1$, for $i \le k^\epsilon$, $\alpha_{k^\epsilon}^\epsilon \in (0,1)$ and $\alpha_i^\epsilon = 0$, for $i \ge k^\epsilon$, where k^ϵ is the index for which

$$\sum_{i=1}^{k^\epsilon-1} b_i \rho_i \le \frac{b}{1+4\epsilon} \text{ and } \sum_{i=1}^{k^\epsilon} b_i \rho_i > \frac{b}{1+4\epsilon}.$$

Denote by R^*, respectively R_ϵ^*, the optimal value of P, respectively P_ϵ. By comparing the feasibility regions and the optimal solutions of the problems $P(t)$, P and P_ϵ, we obtain the following relationships among them.

Lemma 1. *a)* $k_\epsilon \leq k^* \leq k^*(t)$
b) $R^*_\epsilon \leq R^* \leq R^*(t)$.

In our analysis, we will also make use of the dual problems D, respectively D_ϵ, of P, respectively P_ϵ:

$$\begin{array}{llll} & minimize & ub + \sum_{i=1}^{m} v_i & \\ D & s.t. & v_i + b_i\rho_i u \geq r_i\rho_i, i = 1, ..., m \\ & & \mathbf{v} \geq 0, \mathbf{u} \geq 0 \end{array} \qquad \begin{array}{llll} & minimize & u\frac{b}{1+4\epsilon} + \sum_{i=1}^{m} v_i \\ D_\epsilon & s.t. & v_i + b_i\rho_i u \geq r_i\rho_i, i = 1, ..., m \\ & & \mathbf{v} \geq 0, \mathbf{u} \geq 0 \end{array}$$

The next lemma will prove useful in the analysis of the exponential penalty policy.

Lemma 2. *If* (u^*, v^*) *is optimal solution for both* D *and* D_ϵ*, then* $\frac{r_{k^*}}{b_{k^*}} = \frac{r_{k^\epsilon}}{b_{k^\epsilon}}$.

Proof. From the complementary slackness conditions follows that the optimal solutions (u^*, v^*) of D and (u_ϵ, v_ϵ) of D_ϵ are equal to:

$$u^* = \frac{r_{k^*}}{b_{k^*}}$$
$$v^* = \begin{cases} (r_i - \frac{r_{k^*}}{b_{k^*}} b_i)\rho_i, \text{ for } i = 1, ..., k^* \\ 0, \text{ for } i \geq k^* + 1 \end{cases}$$

and

$$u_{k^\epsilon} = \frac{r_{k^\epsilon}}{b_{k^\epsilon}}$$
$$v_{k^\epsilon} = \begin{cases} (r_i - \frac{r_{k^\epsilon}}{b_{k^\epsilon}} b_i)\rho_i, \text{ for } i = 1, ..., k^\epsilon \\ 0, \text{ for } i \geq k^\epsilon + 1. \end{cases}$$

Hence, for (u^*, v^*) to be optimal for D_ϵ, it is necessary that $\frac{r_{k^*}}{b_{k^*}} = \frac{r_{k^\epsilon}}{b_{k^\epsilon}}$.

The penalty function policy $\bar{\pi}$ proposed in [1] can be described as follows. The classes of requests that may be accepted by the penalty function policy are restricted to the ones with $\alpha^*_i \neq 0$, where α^* is the optimal solution of P. Hence, only the classes of index at most k^* are considered.

"Construct" an "augmented network" as follows. Additional to the initial system, called system 0, consider a fictitious infinite capacity system, called system 1. The state of the augmented network (formed by system 0 and system 1 together) at time t is $s(t) = (x(t), y(t)) \in Z^{2m}$, where $x_i(t), i \in \{1, ..., m\}$ denotes the number of class i requests in system 0 at time t and $y_i(t), i \in \{1, ..., m\}$ denotes the number of class i requests being served in system 1 at time t. System 0 is initially empty, and system 1 is initialized with $y^0_i(0^-) = (1 - \alpha^\epsilon_i)\rho_i$, for $i = 1, k_\epsilon$. Note that $y^0_i(0^-) = 0$, for $i < k^\epsilon$.

For each class i, define the following penalty function $\Psi_i(s)$:

$$\Psi_i(s(t)) = exp(\beta \frac{b_i x_i(t)}{c_i^0}) + exp(\beta \frac{b_i y_i(t)}{c_i^1}).$$

An incoming request of type i is accepted in server 0 if it fits into the knapsack and the following condition holds

$$\frac{\partial \Psi_i(s(t))}{\partial x_i} \leq \frac{\partial \Psi_i(s(t))}{\partial y_i}, \tag{4}$$

otherwise it is sent to server 1, where it stays its service time and then leaves the network. The constants c_i^0, c_i^1 and β are defined as

$$c_i^0 = (1 + 4\epsilon)\alpha_i^\epsilon b_i \rho_i \text{ and } c_i^1 = (1 + 4\epsilon)(1 - \alpha_i^\epsilon) b_i \rho_i, \tag{5}$$

$$\beta \leq \epsilon \min\{\frac{c_i^0}{b_i}, \frac{c_i^1}{b_i} : c_i^0 \neq 0 \text{ and } c_i^1 \neq 0\} \tag{6}$$

where α^ϵ is the optimal solution of (P_ϵ).

Remark 1. Condition (4) is equivalent with:

$$\frac{x_i(t)}{c_i^0} \leq \frac{y_i(t)}{c_i^1} + \frac{1}{\beta b_i} log(\frac{c_i^0}{c_i^1}).$$

Interpretation of the penalty policy. Based on the exact expression of α^ϵ, we are now able to reinterpret the penalty policy $\bar{\pi}$ as follows:

ACCEPT all the requests of classes $i < k^\epsilon$ that fit into the knapsack,
REJECT all the requests of classes $i > k^\epsilon$
ACCEPT the requests of class k^ϵ if

$$\frac{x_{k^\epsilon}(t)}{\alpha_{k^\epsilon}^\epsilon} \leq \frac{y_{k^\epsilon}(t)}{1 - \alpha_{k^\epsilon}^\epsilon} + \frac{(1 + 4\epsilon)\rho_{k^\epsilon}}{\beta} log(\frac{\alpha_{k^\epsilon}^\epsilon}{1 - \alpha_{k^\epsilon}^\epsilon})$$

The rejected requests are sent to system 1, where they remain for the duration of their service time.

Remark 2. Since all requests of class $i, i < k^\epsilon$, are accepted as long as there is capacity, we conclude that the exponential penalty policy proposed in [1] is a threshold policy with the threshhold index k^ϵ.

3.3 On a Lower Bound on the Expected Reward Achieved by $\bar{\pi}$

In this section we will show how the analysis of the exponential penalty function policy presented in [1] can be simplified and improved by interpreting the policy as a threshold policy. We will first summarize the results obtained in [1].

Let $\xi_i(t)$, respectively $\eta_i(t)$ be the number of class i requests in system 1 at time t that were rejected by the penalty function, respectively by the capacity constraints.

Clearly, for each $i \leq m$,

$$\begin{aligned} E[x_i(t)] &= E[q_i(t)] + E[y_i^0(t)] - E[y_i(t)] \\ &\geq E[q_i(t)] + E[y_i^0(t)] - (E[\xi_i(t)] + E[\eta_i(t)]). \end{aligned} \tag{7}$$

Hence, one way to obtain lower bounds for $E[x_i(t)]$, is to obtain upperbounds for $E[\xi_i(t)]$, respectively $E[\eta_i(t)]$.

These upper bounds are obtained by comparison with $\tilde{x}_i(t)$, respectively $\tilde{y}_i(t)$, the number of requests of type i present at time t in system 0, respectively system 1, in the network if the capacity was infinite.

Between $x_i(t)$, $\tilde{x}_i(t)$, $\xi_i(t)$ and $\eta_i(t)$, the following relationships exist (see [1] for the proofs):

Lemma 3. *a)For each $i \leq m$, $x_i(t) \overset{d}{\leq} \tilde{x}_i(t)$ and $\tilde{y}_i(t) \overset{d}{\leq} y_i(t)$, where $X \overset{d}{\leq} Y$ denotes the fact that, for all $u \geq 0$, $P(X \geq u) \leq P(Y \geq u)$.*

b)For each $i \leq m$, $E[e^{\frac{\beta b_i \tilde{x}_i(t)}{b}}] \leq \left(2e^{(1-\frac{\epsilon}{2})\beta}\right)^{\frac{c_i^0}{b}}$ and $E[\tilde{y}_i(t)] \leq (1+\varsigma)(1-\alpha_i^\epsilon)\rho_i$, where $\varsigma = \left(\frac{log(2)}{\beta} + 1 - \frac{\epsilon}{2}\right)(1+4\epsilon) - 1$.

c) For each $i \leq m$, $E[\xi_i(t)] \leq E[\tilde{y}_i(t)]$.

d) For each $i \leq m$, $E[\eta_i(t)] \leq 2\rho_i e^{-\frac{\epsilon}{2}(\beta-4)}(1-e^{-\mu_i t})$.

Substituting the bounds obtained in Lemma 3 in formula (7), one can lower bound the expected reward achieved by policy $\bar{\pi}$:

Theorem 2. *For $\epsilon < \frac{1}{4}$,*

$$E[\bar{R}(t)] \geq \max\{\sum_{i=1}^{m} r_i\rho_i(1-e^{-\mu_i t})(\alpha_i^\epsilon - 2e^{-\frac{\epsilon}{2}(\beta-4)}) - \varsigma(1-\alpha_i^\epsilon)), 0\}, \tag{8}$$

where $\varsigma = \left(\frac{log(2)}{\beta} + 1 - \frac{\epsilon}{2}\right)(1+4\epsilon) - 1$ and c_i^0, c_i^1, β are given by (5) and (6).

We proceed now with the tightening of the bound in Theorem 2.

First, remark that for $i > k_\epsilon$, $x_i(t) = 0$, hence these types of requests will not bring any profit. Therefore, in the remainder of this note, we will omit from the analysis the classes of index higher then k^ϵ. Moreover, the definition of $\tilde{x}_i(t)$, together with the fact that k^ϵ is the threshold index, implies that for $i < k^\epsilon$, $\tilde{x}_i(t) \overset{d}{=} q_i(t)$. Hence, for these classes $E[e^{\frac{\beta b_i \tilde{x}_i(t)}{b}}]$ can be obtained exactly, namely:

$$E[e^{\frac{\beta b_i \tilde{x}_i(t)}{b}}] = E[e^{\frac{\beta b_i q_i(t)}{b}}] = e^{\rho_i(1-e^{-\mu_i t})(e^{\frac{\beta b_i}{b}}-1)} \leq e^{\rho_i(\epsilon+1)(1-e^{-\mu_i t})\frac{\beta b_i}{b}}, \tag{9}$$

where for the last inequality we have used the fact that for $x \in (0,1)$, $e^x \leq x + x^2$ and that $\frac{\beta b_i}{b} \leq \epsilon$.

Also, for $i < k^\epsilon$, $\tilde{y}_i(t) \stackrel{d}{=} \xi_i(t) \stackrel{d}{=} 0$.

Consider now $E[\eta_i(t)]$. For $i \leq k^\epsilon$,

$$E[\eta_i(t)] = \int_0^t \lambda_i P(\sum_{i=1}^{k^\epsilon} b_i x_i(u) \geq b - b_i) e^{-\mu_i(t-u)} du \tag{10}$$

$$\leq \int_0^t \lambda_i P(\sum_{i=1}^{k^\epsilon} b_i \tilde{x}_i(u) \geq b - b_i) e^{-\mu_i(t-u)} du \tag{11}$$

$$\leq e^{-\beta(1-\frac{b_i}{b})} \int_0^t \lambda_i E[e^{\sum_{i=1}^{k^\epsilon} \frac{\beta b_i}{b} \tilde{x}_i(u)}] e^{-\mu_i(t-u)} du \tag{12}$$

$$= e^{-\beta(1-\frac{b_i}{b})} \int_0^t \lambda_i \prod_{i=1}^{k^\epsilon} E[e^{\frac{\beta b_i}{b} \tilde{x}_i(u)}] e^{-\mu_i(t-u)} du, \tag{13}$$

where in (11) we have used Lemma 3 a), in (12) we have used Markov's inequality and in (13) we have used the independency of the $\tilde{x}_i$'s.

By substituting in (13) the expression for $E[e^{\frac{\beta b_i \tilde{x}_i(t)}{b}}]$ obtained in (9) for indices $i < k^\epsilon$ and the bound given in Lemma 3 b) for $i = k^\epsilon$, we obtain that:

$$E[\eta_i(t)] \leq 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)} (1 - e^{-\mu_i t}). \tag{14}$$

Finally, by combining (7), the bound in Lemma 3 b) and c) for $i = k^\epsilon$ (for $i \neq k^\epsilon$, $\xi_i(t) = 0$), and (14) , we improve the lower bounds on the expected number of requests of each type in the network at time t and on the expected reward achieved by policy $\bar{\pi}$ as follows.

Theorem 3. *a) For $i < k^\epsilon$,*

$$E[x_i(t)] \geq \rho_i (1 - e^{-\mu_i t}) max\{1 - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}, 0\}.$$

For $i = k^\epsilon$,

$$E[x_{k^\epsilon}(t)] \geq \rho_{k^\epsilon} max\{(1 - e^{-\mu_{k^\epsilon} t})(\alpha_{k^\epsilon} - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}) - \varsigma(1 - \alpha^\epsilon_{k^\epsilon}), 0\}.$$

b) For $\epsilon < \frac{1}{4}$, the average return $E[\bar{R}(t)]$ obtained by policy $\bar{\pi}$ can be bounded from below as follows:

$$E[\bar{R}(t)] \geq \sum_{i=1}^{k^\epsilon - 1} r_i \rho_i (1 - e^{-\mu_i t}) \max\{1 - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}, 0\} +$$
$$+ r_{k^\epsilon} \rho_{k^\epsilon} \max\{(1 - e^{-\mu_{k^\epsilon} t})(\alpha^\epsilon_{k^\epsilon} - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}) - \varsigma(1 - \alpha^\epsilon_{k^\epsilon}), 0\}. \tag{15}$$

Remark 3. From the comparison of the upper bound on the achievable reward $R^*(t)$ and the lower bound given in Theorem 3, we conclude that if k^ϵ is close to $k^*(t)$, and if $\beta >> 1$, the quality of the bounds is very good. However, if ϵ is chosen such that $\frac{\epsilon^2}{2} \leq \frac{b_{k^\epsilon}}{b}$, then it can be proven that $1 < 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}$, which implies that the lower bound given in the previous theorem is 0.

3.4 Bounds of the Exponential Penalty Policy in a Limiting Regime

In this section we will discuss the behaviour of policy $\bar{\pi}$ when $t \to \infty$ and the influence of the choice of ϵ on the policy in this regime.

Denote by $L(t)$ the lower bound in Theorem 3. Clearly, the following relation holds:

$$\lim_{t\to\infty} \frac{L(t)}{R^*} = \max\{1 - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}, 0\}(1 - \frac{\mathbf{I}_{\{\mathbf{k}^\epsilon < \mathbf{k}^*\}} \sum_{i=k^\epsilon}^{k^*-1} r_i\rho_i + r_{k^*}\rho_{k^*}\alpha_{k^*}}{\sum_{i=1}^{k^*-1} r_i\rho_i + r_{k^*}\rho_{k^*}\alpha_{k^*}}) - \max\{(1 - e^{-\mu_{k^\epsilon} t})(\alpha^\epsilon_{k^\epsilon} - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta-4)}) - \varsigma(1 - \alpha^\epsilon_{k^\epsilon}), 0\} \frac{r_{k^\epsilon}\rho_{k^\epsilon}}{\sum_{i=1}^{k^*-1} r_i\rho_i + r_{k^*}\rho_{k^*}\alpha_{k^*}}, \quad (16)$$

where $\mathbf{I}_{\{\mathbf{k}^\epsilon < \mathbf{k}^*\}} = 1$ if $k^\epsilon < k^*$ and 0 otherwise.

Note that the classes that cause the bound in (16) to deviate from 1 are the ones that are admitted in the knapsack problem P but are not admitted in the perturbed knapsack problem P_ϵ. It is then intuitive that by restricting the number of such classes, the bound improves. This is exactly what happens by choosing e.g. ϵ such that $\epsilon < \max\{\epsilon_0, \frac{1}{4}\}$, with $\epsilon_0 = \max\{\epsilon :$ D and D_ϵ have the same optimal solution$\}$, as in [1], Corollary 1. From Lemma 1 follows that if $\epsilon < \epsilon_0$, for each k such that $k^\epsilon \le k \le k^*$, $\frac{r_k}{b_k} = \frac{r_{k^*}}{b_{k^*}}$. If for each $k \ne k^*$, $\frac{r_k}{b_k} \ne \frac{r_{k^*}}{b_{k^*}}$, then the classes admitted into the knapsack in problem P and P_ϵ coincide ($k^\epsilon = k^*$). The only difference is that in P_ϵ, a lower fraction of class k^* is admitted.

4 On the Penalty Function Policy and the Thinning Policy

The thinning policy was proposed by Kelly in [3]. In [1], the authors compare experimentally the exponential penalty function policy with the thinning policy and conclude that the first policy performs better in the transient period and the second in steady state. In this section we will see that by interpreting both policies as threshold policies, one can explain to a certain extent their behaviour.

The thinning policy, which we will denote by $\tilde{\pi}$, is based on α^*, the optimal solution of the "steady state program" P. It accepts a request of type i with probability α^*_i if it fits into the knapsack and if it does not fit, it rejects it. Based on the exact calculation of α^*, we conclude that the thinning policy can be described as follows:

ACCEPT a request of type $i < k^*$ if it fits into the knapsack,
REJECT all requests of types $i > k^*$,
ACCEPT a request of type k^* with probability α_{k^*}.

Note that the definitions of the problems P and P_ϵ imply that $k^\epsilon < k^*$. Hence, the exponential penalty policy and the thinning policy treat the classes $i < k^\epsilon$ and $i > k^*$ in the same way. The only difference between the two policies consists in the way they treat the classes $k^\epsilon \leq i \leq k^*$. The superior behavior of the exponential penalty policy on the thinning policy in transient period, observed experimentally in [1], may be due to the fact that by rejecting "some less profitable" classes, i.e., the classes of index $k^\epsilon < i \leq k^*$, there will be more space in the knapsack for "the more profitable" ones.

5 General Service

In this subsection we assume that the service duration S_i has a general distribution with mean $\frac{1}{\mu_i}$, $i = 1, ..., m$. Let g_i denote the density and G_i denote the cumulative distribution function (CDF) of the service duration $i = 1, ..., m$.

Since the LP's P and D only depend on the mean service time, they will remain the same. The program $P(t)$ changes as follows. For the number of users $q_i(t)$ in service at time t in an $M/G/\infty$ system, it is known that $E[q_i(t)] = \rho_i(1 - G_i^e(t))$, where $G_i^e(t)$ is the tail of the equilibrium CDF of the class i service time distribution (see e.g. [6]). Thus, the only change in $P(t)$ is that the tail $e^{-\mu_i t}$ is replaced by $G_i^e(t)$.

Denote this new LP by $\tilde{P}(t)$, by $\tilde{\alpha}$ his optimal solution and by $\tilde{R}(t)$ the optimal value of $\tilde{P}(t)$. Again, $\tilde{P}(t)$ is a continuous knapsack problem, so the optimal solution is $0-1$, but for at most one class. Let $\tilde{k}(t)$ be the index of this class.

Theorem 1 can be easily generalized for the case where the service times have a genral distribution.

Theorem 4. *For general service times, the reward rate $R^\pi(t)$ of any feasible policy π satisfies*

$$E[R^\pi(t)] \leq \tilde{R}(t) = \sum_{i=1}^{\tilde{k}(t)-1} r_i\rho_i + r_{\tilde{k}(t)}\rho_{\tilde{k}(t)}\tilde{\alpha}_{\tilde{k}(t)}.$$

Consider next the exponential penalty policy $\bar{\pi}$. For finding similar lower bounds to the one in Theorem 3, in [1] extra assumptions on G_i are introduced. Let g_i^t and G_i^t be the density and the CDF of the remaining service time of a class i request conditioned on that it has been in service for time t units. Then, the tail

$$\bar{G}_i^t(s) = 1 - G_i^t(s) = \frac{G_i^e(t+s) - G_i^e(s)}{G_i^e(t)}$$

and

$$g_i^t(s) = -\frac{d\bar{G}_i^t(s)}{ds} = \frac{g_i^e(s) - g_i^e(t+s)}{G_i^e(t)}.$$

Assumption 1. The function $g_i^t(s)$ is a decreasing function of t for all $i = 1, ..., m$, *i.e.*, $g_i^t(0) \geq \lim_{tu\to\infty} g_i^u(0) = g_i^e(0) = \mu_i$, for all $i = 1, ..., m$.

Note that, since for classes $i < k^\epsilon$ one can obtain better bounds by estimating the number of users of class i accepted in the knapsack with the number of users in service at time t in an $M/G/\infty$ queue, the assumption above is not necessary only for the class k^ϵ (see also Remark 9). However,unless the class k^ϵ is fixed from before (e.g. equal to k^*), we cannot renounce at the assumption above for all classes $i < k^\epsilon$. Since $\epsilon < \frac{1}{4}$, we can though assume general service times for the classes $i < k_{\frac{1}{4}}$ (the classes accepted into the knapsack when the total capacity is $\frac{b}{2}$). Also, since the classes of index $i, i > k^*$ are never admitted into the knapsack, we can assume general service time for them as well.

Under Assumption 1 for the classes $k_{\frac{1}{4}} < i < k^*$, Theorem 3 has the following equivalent.

Theorem 5. *For $\epsilon < \frac{1}{4}$, the average return $E[\bar{R}(t)]$ obtained by policy $\bar{\pi}$ can be bounded from below as follows:*

$$E[\bar{R}(t)] \geq \sum_{i=1}^{k^\epsilon - 1} r_i \rho_i \max\{1 - \bar{G}_i^e(t) - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta - 4)}(1 - e^{-\mu_i t}), 0\} +$$
$$+ r_{k^\epsilon} \rho_{k^\epsilon} \max\{(1 - \bar{G}_{k^\epsilon}(t) + \varsigma)\alpha^\epsilon_{k^\epsilon} - 2^{\frac{c^0_{k^\epsilon}}{b}} e^{-\frac{\epsilon}{2}(\beta - 4)}(1 - e^{-\mu_{k^\epsilon}})$$
$$+ \bar{G}_{k^\epsilon}(t) - \bar{G}^e_{k^\epsilon}(t) - \varsigma, 0\}.$$

5.1 Concluding Remarks

In this note we have shown, based on the optimal solution of some continuous knapsack problems, that the exponential penalty function policy proposed in [1] for controlling loss networks reduces to a threshold policy in the case of the stochastic knapsack. Thus, all requests up to a certain index (the "threshold" index) are accepted if there is enough space in the knapsack. Only for accepting the requests of the class with the threshold index one makes use of the penalty function. As a consequence, the question whether the exponentiality of the penalty functions is necessary is reduced to one single class, namely the class with the "threshold" index. Furthermore, we were able to improve the bounds proposed in [1] and to compare the exponential penalty policy with the thinning policy proposed in [3].

In the last section of [1], the authors generalize the penalty approach to control loss networks and to problems in which the constraints in the LP characterizing the "steady state" define a general polytope. Since the optimal solution of this LP's is not as structured as the optimal solution of continuous knapsack problems, the simplified analysis and the improved bounds presented in this note do not extend to the general case.

References

1. Iyengar, G. and Sigman K. (2004). Exponential penalty function control of loss networks, *The Annals of Applied Probability*, 2004, Vol. 14, No. 4, 1698-1740.
2. Halfin, S. and Whitt, W. (1981) Heavy-traffic limits for queues with many exponential servers. *Operations research* 29, 567-588.

3. Kelly, F.P. (1991). Loss Networks, *Annals of Applied Probability* 1, 319-378.
4. Martello, S. and Toth, P. (1990) Knapsack Problems: Algorithms and Computer Implementations. Wiley, Chichester, West Sussex, England.
5. Ross, K.W. (1995) *Multiservice Loss Models for Broadband Telecommunication Networks*, Springer-Verlag.
6. Wolff, R.W. (1989) *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, Englewood Cliffs, NJ.