Performance of Faulty Loss Systems with Persistent Connections

Brendan Patch*†‡, Thomas Taimre‡, and Yoni Nazarathy‡§

ABSTRACT

We consider a finite capacity Erlang loss system that alternates between active and inactive states according to a two state modulating Markov process. Work arrives to the system as a Poisson process but is blocked from entry when the system is at capacity or inactive. Blocked jobs cost the owner a fixed amount that depends on whether blockage was due to the system being at capacity or due to the system being inactive. Jobs which are present in the system when it becomes inactive pause processing until the system becomes active again.

A Laplace transform expression for the expected undiscounted revenue lost in $[0, t]$ due to blocking is found. Further, an expression for the total time discounted expected lost revenue in $[0, \infty)$ is provided. We also derive a second order approximation to the former that can be used when the computing power to invert the Laplace transform is not available. These expressions can be used to ascribe a value to four alternatives for improving system performance: (i) increasing capacity, (ii) increasing the service rate, (iii) increasing the repair rate, or (iv) decreasing the failure rate.

Keywords Erlang loss system, failure, transient analysis

1. INTRODUCTION

Loss networks model systems for which jobs arrive randomly throughout time and simultaneously utilise some set of the available resources for a period of random length, before departure from the system. In the loss network model, if a job arrives to the system and there are not enough resources available for it to begin processing, then it is lost. Landline telephone connections between cities are a classic example. Motivated by this, the resources are called *links* and the arriving jobs are *calls*. A classic review of loss networks is [4], and a more recent review is [5].

This paper introduces a study that is focused on loss networks where links are prone to failure. Arrivals occur according to a Poisson process of rate λ and, if possible, immediately begin service at rate μ. Links which are *active* become *inactive* at rate α and inactive links repair at rate β. Arrivals to a link when it is inactive experience *failure blocking*, and arrivals when the link is active but at full capacity experience *capacity blocking*. Blocked arrivals result in a loss of revenue — the size of which depends on the reason for blockage. In this preliminary exposition we focus on a single link network where existing connections on the link *persist* during link failure — continuing their service upon link reactivation. Our model and analysis is useful for managing communication systems which are subject to sabotage or adverse environmental (e.g. weather) conditions.

Our work extends results in [1] and [2]. In [2] a Laplace transform expression in terms of Charlier polynomials is given for the undiscounted value of an additional unit of capacity during the planning horizon $[0, t]$ on a link which never fails. This expression acts as a performance measure for the system. The authors also demonstrate how the expression can be used to ascribe a value to an extra unit of capacity on the link over the planning horizon. In [1] a second order approximation to the inverted expression is given.

Our methodological contribution is two-fold. First, we extend the model and results of [1] and [2] to allow for link failure. Second, we obtain an expression for infinite horizon total discounted lost revenue when interest is compounded continuously at rate r. As opposed to the expression for finite horizon undiscounted lost revenue, this expression does not require numerical inversion.

The expressions that we obtain can be used to ascribe a value to four alternatives for improving system performance: (i) increasing capacity, (ii) increasing the service rate, (iii) increasing the repair rate, or (iv) decreasing the failure rate. The set of *control parameters*, denoted by $\mathcal{X} \stackrel{\text{def}}{=} \{C, \mu, \beta, \alpha\}$, is used to invoke these changes. We envisage that a system manager can vary the control parameters through mechanisms such as equipment purchases, training programs, and wages.

2. MODEL

Consider the Markov process $\{(N(t), J(t))\}_{t\in\mathbb{R}_+}$, where $N(t)$ takes values in $\{0, 1, \ldots, C\}$ and represents the number of connections in use at time t, and $J(t)$ takes the value 1 if the link is active at time t and 0 otherwise. More precisely, this process has state space

$$\big\{(n, j) : n \in \{0, 1, \ldots, C\}, j \in \{0, 1\}\big\}$$

with evolution governed by the transition rates as given in Table 1. When the process is in state $(C, 1)$ or any state

*Corresponding author. Email: b.patch@uq.edu.au
†School of Mathematics and Physics, The University of Queensland, Qld 4072, Australia.
‡Korteweg-de Vries Institute for Mathematics, University of Amsterdam, The Netherlands.
§Yoni Nazarathy is supported by Australian Research Council (ARC) grants DP130100156 and DE130100291.

IFIP WG 7.3 Performance 2015, October 19-21, Sydney, Australia.

$(n, 0)$, $n \in \{0, 1, \dots, C\}$, potential calls continue to arrive according to a Poisson process of rate λ. These calls are blocked and losses are recorded. Let $\theta > 0$ denote the revenue lost if a call is blocked when the system is in state $(C, 1)$ (capacity blocking) and $\overline{\theta} > 0$ denote the revenue lost if a call is blocked when the system is in a state $(n, 0)$ (failure blocking). When the system is in state $(C,\ 1)$ during a time interval $[a,\ b)$, an expected loss of $\lambda\,\theta\,(b-a)$ is incurred due to the Poisson call arrival process of rate λ. Similarly, if the system is in a state $(n,\ 0)$ during a time interval $[a,\ b)$ then a loss of $\lambda\,\overline{\theta}\,(b-a)$ is expected. Therefore, the expected lost revenue during $[0,\ t]$ for a link with $N(0) = n$ and $J(0) = j$ can be written as

Table 1: Transition rates of our faulty loss system.

Transition	Rate	States
$(n,\,1) \to (n,\,0)$	α	$0 \le n \le C$
$(n,\,0) \to (n,\,1)$	β	$0 \le n \le C$
$(n,\,1) \to (n+1,\,1)$	λ	$0 \le n < C$
$(n,\,1) \to (n-1,\,1)$	$n\,\mu$	$0 < n \le C$

$$R^{(j)}_{n,\mathcal{X}}(t) \stackrel{\text{def}}{=} \mathbb{E}\left[\int_0^t \lambda\left(\overline{\theta}\ \mathrm{I}_{\{J(\tau)=0\}} + \theta\,\mathrm{I}_{\{N(\tau)=C\}}\,\mathrm{I}_{\{J(\tau)=1\}}\right)\mathrm{d}\tau \,\middle|\, \big(N(0), J(0)\big) = (n, j)\right]. \quad (1)$$

The function defined in (1) is analogous to the capacity value function of [2]. The value over the planning horizon of $[0,\ t]$ of a parameter adjustment from $\mathcal{X}$ to $\widetilde{\mathcal{X}} \stackrel{\text{def}}{=} \{\widetilde{C},\ \widetilde{\mu},\ \widetilde{\beta},\ \widetilde{\alpha}\}$ is

$$\Delta R^{(j)}_{n,t}(\mathcal{X}, \widetilde{\mathcal{X}}) \stackrel{\text{def}}{=} R^{(j)}_{n,\mathcal{X}}(t) - R^{(j)}_{n,\widetilde{\mathcal{X}}}(t)\,. \quad (2)$$

We call this the *finite horizon performance value function*.

Furthermore, the expected rate at which revenue is lost from the system is $\lambda\,\theta$ when it is in state $(C,\ 1)$ and $\lambda\,\overline{\theta}$ when it is in a state $(n, 0)$. So the expected rate at which revenue is lost at time t is

$$\ell^{(j)}_{n,\mathcal{X}}(t) \stackrel{\text{def}}{=} \lambda\,\mathbb{E}\Big[\overline{\theta}\ \mathrm{I}_{\{J(t)=0\}} + \theta\,\mathrm{I}_{\{N(t)=C\}}\,\mathrm{I}_{\{J(t)=1\}} \,\Big|\, \big(N(0), J(0)\big) = (n, j)\Big]\,.$$

Assuming that interest is compounded continuously at rate r, the discounted value of the lost revenue during $[0,\ \infty)$ is

$$\mathscr{L}^{(j)}_{n,\mathcal{X}}(r) \stackrel{\text{def}}{=} \int_0^\infty \ell^{(j)}_{n,\mathcal{X}}(t)\,\mathrm{e}^{-r\,t}\mathrm{d}t\,. \quad (3)$$

Note that this functional is equivalently the Laplace transform.

Similar to (2),

$$\Delta\mathscr{L}^{(j)}_{n,r}(\mathcal{X},\ \widetilde{\mathcal{X}}) \stackrel{\text{def}}{=} \mathscr{L}^{(j)}_{n,\mathcal{X}}(r) - \mathscr{L}^{(j)}_{n,\widetilde{\mathcal{X}}}(r) \quad (4)$$

gives the difference in total time discounted value obtained from variations in the control parameters. We call this the *discounted performance value function*.

It is straightforward to combine (2) or (4) with a budget constraint to obtain an optimization problem that can be solved, and thus help direct the manager of the system on how best to vary the control parameters.

3. RESULTS

In this section we give explicit expressions for (1) and (3) that can be used to calculate (2) and (4). Let $R^{(j)}_{n,\mathcal{X}}(t\,|\,x)$ be $R^{(j)}_{n,\mathcal{X}}(t)$ conditional on the fact that the first time that the link departs from state $(n,\ j)$ is x. Now,

$$R^{(1)}_{n,\mathcal{X}}(t\,|\,x) = \begin{cases} 0\,, & n < C,\ t < x\,, \\ \theta\lambda t\,, & n = C,\ t < x\,, \\ \frac{n\,\mu R^{(1)}_{n-1,\mathcal{X}}(t-x) + \lambda R^{(1)}_{n+1,\mathcal{X}}(t-x) + \alpha R^{(0)}_{n,\mathcal{X}}(t-x)}{n\,\mu+\lambda+\alpha}\,, & n < C,\ t \ge x\,, \\ \theta\lambda x + \frac{C\,\mu\,R^{(1)}_{C-1,\mathcal{X}}(t-x) + \alpha\,R^{(0)}_{C,\mathcal{X}}(t-x)}{C\,\mu+\alpha}\,, & n = C,\ t \ge x\,, \end{cases} \quad (5)$$

and

$$R^{(0)}_{n,\mathcal{X}}(t\,|\,x) = \begin{cases} \overline{\theta}\,\lambda\,t\,, & t < x\,, \\ \overline{\theta}\lambda x + R^{(1)}_{n,\mathcal{X}}(t-x)\,, & t \ge x\,. \end{cases} \quad (6)$$

For $(N(0), J(0)) = (n, j)$, $j \in \{0, 1\}$ let $X^{(j)}_n$ be the time until the first transition, with distribution $F^{(j)}_n$, so that we obtain the Riemann–Stieltjes integral

$$R^{(j)}_{n,\mathcal{X}}(t) = \int_0^\infty R^{(j)}_n(t\,|\,x)\,\mathrm{d}F^{(j)}_n(x)\,. \quad (7)$$

Since $\{(N(t),\ J(t))\}_{t\in\mathbb{R}_+}$ is a continuous time Markov chain, $X^{(1)}_n \sim \mathsf{Exp}(\lambda + n\,\mu + \alpha)$ when $n < C$, $X^{(1)}_n \sim \mathsf{Exp}(C\,\mu + \alpha)$ when $n = C$, and $X^{(0)}_n \sim \mathsf{Exp}(\beta)$ when $0 \le n \le C$. Upon substitution of (5) and (6) into (7), and application of the Laplace transform

$$\widetilde{R}^{(j)}_{n,\mathcal{X}}(s) \stackrel{\text{def}}{=} \int_0^\infty R^{(j)}_{n,\mathcal{X}}(t)\,\mathrm{e}^{-s\,t}\,\mathrm{d}t\,,$$

one obtains a system of equations that can be solved in a similar way to the method used in [2]. This relies on the use of classic results for Meixner, Charlier, and Laguerre polynomials (see e.g. [3]). After some computations we obtain an explicit expression for the Laplace transform of the finite horizon performance value function:

$$\widetilde{R}^{(1)}_{n,\mathcal{X}}(s) = \frac{P_n(s)\,\theta\,\lambda}{s\left(s\,P_C(s) + C\,\mu\left(P_C(s) - P_{C-1}(s)\right)\right)} + B(s)$$

and $\widetilde{R}^{(0)}_{n,\mathcal{X}}(s) = \big(\beta\,\widetilde{R}^{(1)}_{n,\mathcal{X}}(s) + \overline{\theta}\,\lambda/s\big)/(s+\beta)$, where

$$P_n(s) = \sum_{k=0}^{n}\binom{n}{k}\lambda^{-k}\prod_{i=0}^{k-1}A_i(s)\,, \quad (8)$$

$$A_i(s) = s + \alpha - \alpha\,\beta\,(s+\beta)^{-1} + i\,\mu\,, \text{ and} \quad (9)$$

$$B(s) = \frac{\alpha\,\overline{\theta}\,\lambda}{s\,(s+\beta)\big(s+\alpha-\alpha\,\beta\,(s+\beta)^{-1}\big)}\,. \quad (10)$$

This generalizes the result in [2] for a loss system that never fails. Taking $\alpha = 0$ (no failures) or $\beta \to \infty$ (instantaneous repairs) recovers equation (15) of [2].

While this expression seems nice and compact, it does require a summation involving binomial coefficients, making inversion numerically cumbersome. In this case approximations to the inverted expression are a sensible alternative. We will now study approximations for a system with $J(0) = 1$, it is a simple extension to examine a system with $J(0) = 0$.

A first order linear approximation is given by

$$R^{(1)}_{n,\mathcal{X}}(t) = a\,t + \mathrm{o}(t)\,, \quad (11)$$

where $a = \lambda\,(\theta\,\pi + \overline{\theta}\,\overline{\pi})$, $o(t)/t \to 0$ as $t \to \infty$,

$$\pi = \left(\frac{\varrho^C}{C!}\right)\left((1+\psi)\sum_{n=0}^{C}\frac{\varrho^n}{n!}\right)^{-1},$$

and $\overline{\pi} = \psi/(1+\psi)$, with $\varrho = \lambda/\mu$, and $\psi = \alpha/\beta$.

The values π and $\overline{\pi}$ represent the equilibrium probability that $\{(N(\cdot),\, J(\cdot))\}$ is in state $(C,\,1)$ or, respectively, a state $(n,\,0)$. Combined with the Poisson call arrival process of rate λ it is clear that a is the equilibrium rate of loss. Note the similarity of π to Erlang's B formula, which is retrieved for $\alpha = 0$ or $\beta \to \infty$.

The error introduced to (11) as it transitions to equilibrium can be corrected by a second order linear approximation. To obtain this more accurate approximation we apply a standard Tauberian theorem (final value) to

$$\widetilde{R}^{(1)}_{n,\mathcal{X}}(s) - a/s^2\,.$$

This yields,

$$R^{(1)}_{n,\mathcal{X}}(t) = a\,t + b + o(1)\,,$$

with $o(1) \to 0$ as $t \to \infty$, where our newly found correction term is

$$b = \frac{\gamma_1 + 2\,\theta\,\lambda + \gamma_2\,\gamma_3 - 2\,a\,\gamma_4}{2\,\beta\,\gamma_4\,(1+\psi)}$$

with

$$\gamma_1 = 2\,(1+\psi)\,\beta\,\theta\,\lambda\,g_1(n)\,,$$
$$\gamma_2 = \alpha\,\overline{\theta}\,\lambda - (1+\psi)\,\beta\,a\,,$$
$$\gamma_3 = -2\,\psi/\beta + 2\,(1+\psi)\,g_1(C) + C\,\mu\,\big(g_2(C) - g_2(C-1)\big)\,,$$
$$\gamma_4 = 1 + \psi + C\,\mu\,\big(g_1(C) - g_1(C-1)\big)\,,$$
$$g_1(n) = \frac{\psi+1}{\mu}\sum_{k=1}^{n}\varrho^{-k}\binom{n}{k}(k-1)!\,,\ \text{and} \tag{12}$$
$$g_2(n) = \frac{2\,(\alpha+\beta)^2}{\beta^2\,\mu^2}\sum_{k=2}^{n}\varrho^{-k}\binom{n}{k}(k-1)!\sum_{m=1}^{k-1}\frac{1}{m} - \frac{2\,\psi}{\beta\,\mu}\sum_{k=1}^{n}\varrho^{-k}\binom{n}{k}(k-1)!\,. \tag{13}$$

This result generalizes Theorem 4.1 in [1]. Again, their result can be retrieved by taking $\alpha = 0$ or $\beta \to \infty$.

Finally, observe that $\ell^{(j)}_{n,\mathcal{X}}(t) = \partial R^{(j)}_{n,\mathcal{X}}(t)/\partial t$. Hence since $R^{(j)}_{n,\mathcal{X}}(0) = 0$, using the properties of the Laplace transform,

$$\mathscr{L}^{(j)}_{n,\mathcal{X}}(r) = r^{-1}\,\widetilde{R}^{(j)}_{n,\mathcal{X}}(r)\,.$$

4. ILLUSTRATION

Consider a system where $\mathcal{X} = \{6,\,3,\,0.5,\,\cdot\}$, $\lambda = 3$, $\theta = 1$, $\overline{\theta} = 2$, and $J(0) = 1$. Figure 1 displays the expected lost revenue of this system over planning horizons $t \in [0,\,7]$. We see that the second order approximation (dashed) converges to the numerically-inverted Laplace transform (solid). The left panel is a system where there is no failure ($\alpha = 0$) and in the right panel ($\alpha = 0.002$) the system is expected to fail approximately as often as 1500 calls arrive and then take six calls worth of time to repair. It can be seen from comparing the left and right panels that without accounting for these faults a (potentially substantial) error in the evaluation of expected losses can occur.

Now consider the same system but with $\alpha = 0.5$. Figure 2 shows the discounted ($r = 0.1$) performance function (i.e. increase in revenue) that occurs when either the failure rate is decreased (left) or the repair rate is increased (right). We see that, for this system, increasing β increases revenue at a decreasing rate, while decreasing α increases revenue at an increasing rate. As $\beta \to \infty$ the right graph will asymptote to the vertical intercept of the left graph.

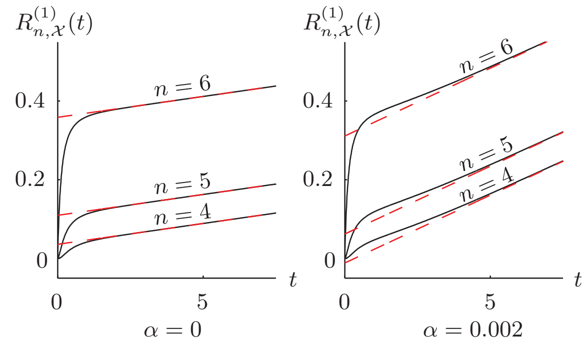

Figure 1: Expected lost revenue in $[0,\,t]$ without failure (left) and with failure (right).

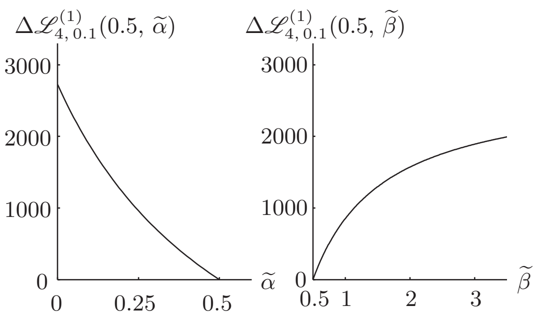

Figure 2: Discounted performance for decreases in failure rate (left) or increases in repair rate (right).

5. OUTLOOK

The performance value functions introduced here could play a role in more complex networks, in which a routing decision plays a role. Relaxing the persistent connections assumption to permit disconnection when the system is inactive would be both interesting and practical. It would also be useful to generalize these results by replacing the exponential distributions used with phase-type distributions.

6. REFERENCES

[1] B. A. Chiera, A. E. Krzesinski, and P. G. Taylor. Some Properties of the Capacity Value Function. *SIAM Journal on Applied Mathematics*, 65(4):1407–1419, 2005.

[2] B. A. Chiera and P. G. Taylor. What is a unit of capacity worth? *Probability in the Engineering and Informational Sciences*, 16(4):513–522, 2002.

[3] T. Chihara. *An Introduction to Orthogonal Polynomials*. Gordon and Breach, New York, 1978.

[4] F. P. Kelly. Loss Networks. *The Annals of Applied Probability*, 1(3):319–378, 1991.

[5] S. Zachary and I. Ziedins. Loss networks. In R. J. Boucherie and N. M. van Dijk, editors, *Queueing Networks*, volume 154 of *International Series in Operations Research & Management Science*, pages 701–728. Springer, 2011.