Performability analysis of an unreliable M/M/1-type queue system

Ruslan Krenzler[1] Prof. Hans Daduna [1]

[1]Uni Hamburg

8. GI/ITG - Workshop des Fachausschusses „Messung, Modellierung und Bewertung von Rechensystemen (MMB)" 2015

Problem description

- Unreliable $M/M/1/\infty$ queueing system.
- Can fail. Repaired after failure.
- Number of services affects failure rate.
- Maintained after N services.
- If blocked (repair or maintenance): no service, no new customers.

Subject to optimization

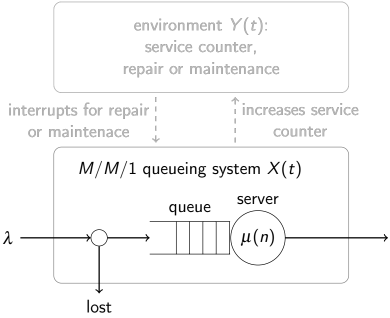
N - number of services, after which the system needs to be maintained.

Mathematical representation

- Markov-process $(X(t), Y(t) : t \in \mathbb{R}_0^+)$
- System state $(n, k) \in \mathbb{N}_0 \times K$
  - $n \in \mathbb{N}_0$ number of customers
  - $k \in K := \{\underbrace{0, 1, \ldots N-1}_{\in K_W}, \underbrace{b_m, b_r}_{\in K_B}\}$
    - $0, 1, \ldots, N-1$ service counter
    - b_m **b**locked because of **m**aintenance
    - b_r **b**locked because of **r**epair
- Transition rates in generator
  $Q := (q((n,k),(n',k')) : n, n' \in \mathbb{N}_0, k, k' \in K)$

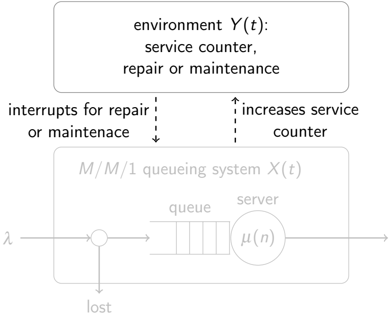
Component: $M/M/1/\infty$ queueing system

- Process $(X(t) : t \in \mathbb{R}_0^+)$, $X(t) \in \mathbb{N}_0$.
- $X(t)$ describes number of customers at time t.
- Poisson input with rate λ.
- Exponential service rates $\mu(n)$.
- Waiting area of infinite size.
- FIFO.
- Lost customers when repaired or maintained.



Component: Environment

- Process $(Y(t) : t \in \mathbb{R}_0^+)$, $Y(t) \in K$.
- Environment states $K = K_W \uplus K_B$.
  - $K_W = \{0, 1, ..., N-1\}$ service counter.
  - $K_B = \{b_m, b_r\}$, states for maintenance and repair.
- Failure rates ν_k, $k \in K_W$.
- Maintenance rate ν_m.
- Repair rate ν_r.



Generator

- System state $(n,k) \in \mathbb{N}_0 \times K$,
  - $n \in \mathbb{N}_0$ number of customers,
  - $k \in K$ service counter (K_W), repair or maintenance (K_B).
- Transition rates $Q := (q((n,k),(n',k')) : n, n' \in \mathbb{N}_0, k, k' \in K)$

$$
\begin{aligned}
q((n,k),(n+1,k)) &= \lambda, & k &\in K_W, \\
q((n,k),(n-1,k+1)) &= \mu(n) & k &\leq N-2, n \geq 1, \\
q((n,N-1),(n-1,b_m)) &= \mu(n), & n &\geq 1, \\
q((n,k),(n,b_r)) &= \nu_k \in \mathbb{R}_0^+, & k &\in K_W, n \geq 1, \\
q((n,b_m),(n,0)) &= \nu_m \in \mathbb{R}_0^+, & n &\geq 0, \\
q((n,b_r),(n,0)) &= \nu_r \in \mathbb{R}_0^+, & n &\geq 0.
\end{aligned}
$$

Cost function

- c_m maintenance costs per unit of time.
- c_r repair costs per unit of time.
- c_b costs of non-availability per unit of time.
- c_w waiting costs per customer per unit of time.

Cost function per time unit and state

$$f(n,k) = \begin{cases} c_w \cdot n + c_b + c_m, & k = b_m, \\ c_w \cdot n + c_b + c_r, & k = b_r, \\ c_w \cdot n, & k \in K \setminus \{b_m, b_r\}. \end{cases}$$

n – customers, k – environment

Average costs

The asymptotic average costs for an ergodic system can be calculated as

$$\frac{1}{T}\int_0^T f\left(X_t(\omega), Y_t(\omega)\right) dt \xrightarrow{T\to\infty} \sum_{(n,k)} f(n,k)\pi(n,k) := \bar{f}(N), \qquad P-a.s.$$

- $f(n,k)$ cost of the system state (n,k) per unit of time.
- $\pi(n,k)$ steady state probability of the system state (n,k).

Steady state distribution

- Given generator $Q := (q((n,k),(n',k')) : n,n' \in \mathbb{N}_0, k,k' \in K)$

Find

$\pi(n,k)$ solution of

$$\pi Q = 0 \qquad\qquad ||\pi||_1 = 1$$

Steady state distribution

- $n \in \mathbb{N}_0$ number of customers
- $k \in K$ service counter, repair, or maintenance

$$P(X=n, Y=k) := \pi(n,k) = \xi(n)\theta(k) \qquad \text{with}$$

$$\boxed{\xi(n) = \prod_{i=1}^{n} \frac{\lambda}{\mu(i)} \xi(0)}$$

and θ, which is a stochastic solution of

$$\boxed{\theta \tilde{Q} = 0 \text{ with } \tilde{Q} \in \mathbb{R}^{K \times K} - \text{easy to calculate.}}$$

Steady state distribution of the environment

$$\theta_N(k) := \prod_{i=1}^{k} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \theta_N(0), \qquad 0 \leq k \leq N-1$$

$$\theta_N(b_m) := \frac{\lambda}{\nu_m} \theta(N-1) = \frac{\lambda}{\nu_m} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \theta_N(0)$$

$$\theta_N(b_r) := \left(\frac{(\nu_0 + \lambda)}{\nu_r} - \frac{\lambda}{\nu_r} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \right) \theta_N(0)$$

$$\theta_N(0) := \left(\sum_{k=0}^{N-1} \prod_{i=1}^{k} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i + \frac{\lambda}{\nu_m} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \right.$$

$$\left. + \frac{(\nu_0 + \lambda)}{\nu_r} - \frac{\lambda}{\nu_r} \prod_{i=1}^{N-1} \left(\frac{\lambda}{\nu_i + \lambda} \right)^i \right)^{-1}$$

Cost function

The asymptotic average costs for an ergodic system

$$\bar{f}(N) = \sum_{(n,k)} f(n,k)\pi(n,k) \qquad P-a.s.$$

- $\pi(n,k) = \xi(n)\theta(k)$ steady state probability of the system state (n,k).

$$\bar{f}(N) = \underbrace{(c_b + c_m)\,\theta_N(b_m) + (c_b + c_r)\,\theta_N(b_r)}_{=:g(N)} + \underbrace{c_w \sum_{n=1}^{\infty} n\xi(n)}_{\text{independent of } N}$$

optimize $g(N)$

Other parameters

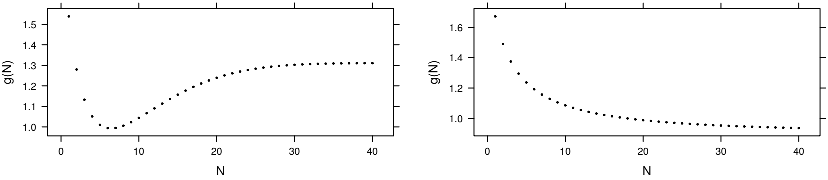
- The average number of failures

$$\theta_N(b_r)\nu_r$$

- The average number of maintenances

$$\theta_N(b_m)\nu_m$$

Total costs

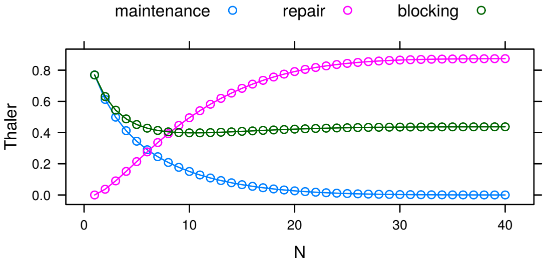


(a) linear $\nu_k = 0.01 \cdot k$ (b) constant $\nu_k \equiv 0.0436$

Figure : Cost functions g with $\lambda = 1$, $c_m = 1$, $c_b = 1$, $c_r = 2$, $\nu_m = 0.3$, $\nu_r = 0.1$, $\max(N) = 40$

- Optimal maintenance for linear failure rate: after 6 services.
- Optimal maintenance for constant failure rate: as late as possible.

Separated costs



$$g(N) = (c_m + c_b)\,\theta_N(b_m) + (c_r + c_b)\,\theta_N(b_r)$$

Software rejuvenation

Software rejuvenation is a proactive and preventive solution to handle transient software failures.

Model

- Software system as $M/M/1/\infty$ queue.
- Restart after N finished tasks.
- No failures ($\nu_k \equiv 0$) before restart.

Previous results

Steady state solution $\pi(n,k)$

- Product form $\pi(n,k)=\xi(n)\theta(k)$.
- Known $\xi(n)=\prod_{i=1}^{n}\frac{\lambda}{\mu(i)}\xi(0)$.
- θ easier to calculate as solution of $\theta\tilde{Q}=0$ than $\pi Q=0$.

$$\tilde{Q}\in\mathbb{R}^{K\times K}\quad vs\quad Q\in\mathbb{R}^{\mathbb{N}_0\times K\times\mathbb{N}_0\times K}$$

- ξ does not depend on the environment.
- θ does not depend on service rate μ.

General mathematical model

Two dimensional process $(X(t), Y(t) : t \geq 0)$ which describes an ergodic $M/M/1/\infty$ system in a random environment

- system state $(n, k) \in \mathbb{N}_0 \times K$
- input rate λ, service rates $\mu(n)$
- $K_B \subset K$, "blocking states" (I_W mask)
- $V \in \mathbb{R}^{K \times K}$, transition rates of environment
- $R \in [0, 1]^{K \times K}$, after customer leaves queue, environment changes from k to m with probability R_{km}
- no service, no new customers when blocked (lost customers)

Steady state results

System is described by: input rate λ, service rates $\mu(n)$, environment state set K, blocking states K_B (I_W mask), generator $V \in \mathbb{R}^{K\times K}$, stochastic matrix $R \in [0,1]^{K\times K}$ with steady state distribution, **lost** customers when blocked.

$$P(X=n, Y=k) := \pi(n,k).$$

- Product form $\pi(n,k) = \xi(n)\theta(k)$.
- Known $\xi(n) = \prod_{i=1}^{n} \frac{\lambda}{\mu(i)}\xi(0)$.
- θ easier to calculate as solution of $\theta\tilde{Q} = 0$ than $\pi Q = 0$:

$$\tilde{Q} \in \mathbb{R}^{K\times K} \quad vs \quad Q \in \mathbb{R}^{\mathbb{N}_0\times K\times\mathbb{N}_0\times K}$$

$$\tilde{Q} = \lambda I_W(R-I) + V.$$

- ξ does not depend on the environment.
- θ does not depend on service rates $\mu(n)$.

Network: Mathematical model

$(X(t), Y(t) : t \geq 0)$, Jackson network with J nodes in a random environment

Similar results

Talk: *Networks of Queues in a Random Environment: A Survey of Product Form Results* by Hans Daduna

BOINC

BOINC (Berkeley Open Infrastructure for Network Computing)
in cooperation with Alexander Rumyantsev (Karelian Research Centre of the RAS)

- Network of workstations.
- Breakdowns.
- Some tasks:
  - Queueing-network model.
  - Verify model.

In general not that easy when

- input rate $\lambda(n)$ depends on customer number n.
- $M/M/1/N$ queues.
- Non-exponential service times $(M/G/1/\infty)$.
- No lost customers due to blocking.

Approximation

Previous results

Lost customers when blocked $\Longrightarrow$ steady state $\pi(n,k) = \xi(n)\theta(k)$. With $\xi(n)$ and $\theta(k)$ easier to calculate.

No lost customers

What if the customer are **not** lost?
$\Longrightarrow$ no easy solution, but, maybe, an approximation can help.

Approximation

Two dimensional process $(X(t), Y(t) : t \geq 0)$ which describes an ergodic $M/M/1/\infty$ system in a random environment, **NO** customer loss:

- Same parameter μ, V, R, K_B as system with lost customers.
- No lost customers when blocked $\Rightarrow$ most likely no product form.

Approximation

There exists a queueing system in a random environment with **lost** customers with the same μ, V, R, and the same throughput.

Future research

- BOINC - models.
- Approximation of non-product systems.
- Numerical bounds and starting values for system with non-product steady state.

Thank you for your attention!

Bibliography

Ruslan Krenzler and Hans Daduna.
Loss systems in a random environment: steady state analysis.
Queueing Systems, 80(1-2):127–153, 2015.

Ruslan Krenzler and Hans Daduna.
Performability analysis of an unreliable M/M/1-type queue.
In *Achter Workshop Leistungs-, Zuverlaessigkeits- und Verlaesslichkeitsbewertung von Kommunikationsnetzen und verteilten Systemen 2015 (MMBnet2015)*, Hamburg, Germany, September 2015.