Reliability and Availability Engineering: Modeling, Analysis, Applications

Chapter 14 - Semi-Markov and Markov Regenerative Models

Kishor Trivedi and Andrea Bobbio

Department of Electrical and Computer Engineering
Duke University
Dipartimento di Scienze e Innovazione Tecnologica
Università del Piemonte Orientale

Jan 2017

Outline

Semi Markov Processes: An introduction

A semi-Markov process satisfies the Markov property only at the time of entry into a state. All the transitions into a new state should satisfy the Markov property.

Let $\{Z(t) | t \geq 0\}$ be a SMP.

Let $t_0 = 0,\ t_1,\ t_2,\ \ldots,\ t_n,\ \ldots$ be the time instances at which $Z(t)$ undergoes a state transition.

The sequence of states $\{X_n = Z(t_n) | n \geq 0\}$ is a DTMC characterized by its one-step transition probability matrix (TPM) $\boldsymbol{P} = [p_{ij}]$.

Let $H_i(t)$ be the sojourn time distribution in state i. $H_i(t)$ may be generally distributed.

Steady state solution

Steady state analysis

Think of transitions as occurring in two stages:

- First stage: the SMP stays in state i for an amount of time with distribution function $H_i(t)$, that is the sojourn time distribution in state i.
- Second stage: the SMP moves from state i to state j with probability p_{ij}.
- The SMP is thus described by the transition probability matrix $\boldsymbol{P} = [p_{ij}]$ and the vector of sojourn time distributions $H_i(t)$.

Steady state analysis

The steady state probability vector $\boldsymbol{\nu}$ of the embedded DTMC is obtained by solving the linear system of equations:

$$\boldsymbol{\nu} = \boldsymbol{\nu}\,\boldsymbol{P} \quad \text{subject to} \quad \sum_i \nu_i = 1$$

where ν_i is the steady-state probability of the embedded DTMC in state i. The mean sojourn time in state i is:

$$h_i = \int_0^\infty (1 - H_i(t))dt$$

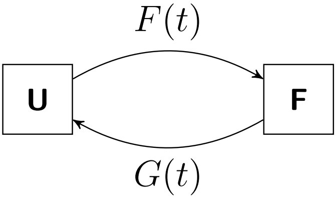
Then the steady state probability π_i for the SMP state i is given by:

$$\pi_i = \frac{\nu_i h_i}{\sum_j \nu_j h_j}$$

A 2-state SMP - 1

Consider a 2-state SMP modeling the failure-repair of a single component.

$F(t)$ and $G(t)$ denote the distributions of time to failure and time to repair.



We construct the TPM of the embedded DTMC.

$$\boldsymbol{P} = \begin{array}{c c} & \begin{array}{cc} U & D \end{array} \\ \begin{array}{c} U \\ D \end{array} & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

The DTMC is periodic and hence limiting state probabilities do not exist as the DTMC alternates between two states.

A 2-state SMP - 2

However, the stationary probability vector can still be computed from

$$\boldsymbol{\nu} = \boldsymbol{\nu}\,\boldsymbol{P} \quad , \qquad \nu_U + \nu_D = 1 \qquad \text{to obtain} \qquad \nu_U = \frac{1}{2} = \nu_D$$

Furthermore, $$h_U = \int_0^\infty (1 - F(t))dt = \text{MTTF}$$

and, $$h_D = \int_0^\infty (1 - G(t))dt = \text{MTTR}$$

Finally, $$\pi_U = \frac{\frac{1}{2}\text{MTTF}}{\frac{1}{2}\text{MTTF} + \frac{1}{2}\text{MTTR}} = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} = \left[1 + \frac{\text{MTTR}}{\text{MTTF}}\right]^{-1}$$

which also gives the steady state availability of the component.

A 2-state SMP - 3

Quite often in practice, state D will not be considered a down state unless the sojourn time in the state exceeds a threshold, say, τ.

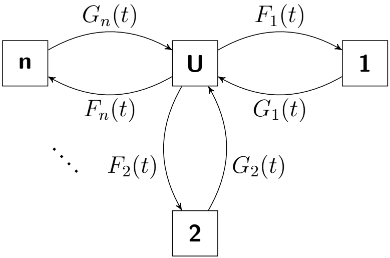
The probability of an individual sojourn in state D not exceeding the threshold is given by $G(\tau)$.

By assigning $G(\tau)$ as a reward rate to the state D and a reward rate 1 to state U, we get the steady state availability

$$A(\tau) = \pi_U + G(\tau)\,\pi_D$$

Component subject to different types of failures

Next we consider a single component subject to different types of failures and corresponding repairs.



We assume that TTF and TTR for failure type i $(i = 1, 2, \ldots, n)$ are respectively distributed as $F_i(t)$ and $G_i(t)$.

Y. Cao, H. Sun, K. Trivedi, and J. Han, "System availability with non-exponentially distributed outages," IEEE Transactions on Reliability, vol. 51, no. 2, pp. 193-198, Jun 2002.

Component subject to different types of failures - 1

Model assumptions:

- Only one type of failure can occur at a time and during the recovery / repair from a failure, another failure cannot occur.
- Failures occur independently.
- Each time the repair completion brings the system to as good as new state.
- Age clock of each type of failure is thus reset upon the completion of a repair.

With these assumptions, the underlying stochastic process is an SMP with $n + 1$ states $\{U, 1, \ldots, n\}$.

Component subject to different types of failures - 2

To derive the entries of the TPM of the embedded DTMC, we observe that

$$p_{i,U} = 1$$

$$p_{U,i} = \int_0^\infty \prod_{j=1, j\neq i}^{n} (1 - F_j(x))\, dF_i(x) \qquad \text{with} \qquad \sum_{i=1}^{n} p_{U,i} = 1$$

Furthermore, $$H_U(t) = 1 - \prod_{j=1}^{n} (1 - F_j(t))$$

$$H_i(t) = G_i(t)$$

Hence, $$h_U = \int_0^\infty \prod_{j=1}^{n} (1 - F_j(t))dt = \text{MTTF}$$

and $$h_i = \int_0^\infty (1 - G_i(t))dt = \text{MTTR}_i \qquad i = 1, 2, \ldots, n$$

Component subject to different types of failures - 3

The TPM of the embedded DTMC can be seen to be

$$\boldsymbol{P} = \begin{array}{c c} & \begin{array}{cccc} U & 1 & \cdots & n \end{array} \\ \begin{array}{c} U \\ 1 \\ \vdots \\ n \end{array} & \begin{bmatrix} 0 & p_{U,1} & \cdots & p_{U,n} \\ p_{1,U} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,U} & 0 & \cdots & 0 \end{bmatrix} \end{array}$$

It is then easy to see that the solution to

$$\boldsymbol{\nu} = \boldsymbol{\nu}\,\boldsymbol{P} \quad \text{subject to} \quad \nu_U + \sum_{j=1}^{n} \nu_j = 1$$

is given by $\quad \nu_U = \dfrac{1}{2} \quad$ and $\quad \nu_i = \dfrac{p_{U,i}}{2}$

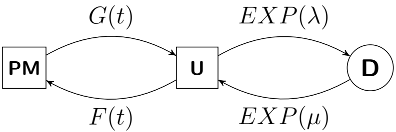
Component subject to different types of failures - 4

The steady state availability is:

$$
\begin{aligned}
A = \pi_U &= \frac{h_U}{h_U + \sum_{j=1}^{n} p_{U,j} h_j} \\
&= \left[1 + \sum_{j=1}^{n} p_{U,j} \frac{h_j}{h_U} \right]^{-1} \\
&= \left[1 + \sum_{j=1}^{n} p_{U,j} \frac{\mathrm{MTTR}_j}{\mathrm{MTTF}} \right]^{-1}
\end{aligned}
$$

Component subject planned and unplanned outages - 1

We specialize previous example to the case of planned and unplanned outages, denoted by *PM* and *D* respectively.



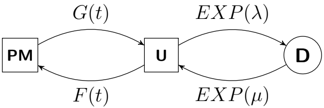
The planned outage may be for system upgrades, configuration changes, maintenance and so on. In many practical situations, the planned downtime is often larger than unplanned downtime.

Assume that TTF and TTR for planned outages have respective distributions $F(t)$ and $G(t)$.

The TTF and TTR for unplanned outages are assumed to follow exponential distribution with respective rates λ and μ.

Component subject planned and unplanned outages - 2

We specialize previous result, to get:



$$A = \left[1 + p_{U,D}\frac{h_D}{h_U} + p_{U,PM}\frac{h_{PM}}{h_U}\right]^{-1}$$

where,

$$p_{U,PM} = \int_0^\infty e^{-\lambda t} dF(t) = \lambda \int_0^\infty e^{-\lambda t} F(t)\, dt$$

$$p_{U,D} = \int_0^\infty (1 - F(t))\lambda e^{-\lambda t} dt$$

$$h_U = \int_0^\infty (1 - F(t))e^{-\lambda t} dt \quad , \quad h_{PM} = \int_0^\infty (1 - G(t))dt$$

$$h_D = \int_0^\infty e^{-\mu t} dt$$

Component subject planned and unplanned outages - 3

by denoting,

$$\alpha(\lambda) = \int_0^\infty e^{-\lambda x}\, dF(x) \quad \text{and} \quad \theta(\lambda) = \frac{\alpha(\lambda)}{1-\alpha(\lambda)}$$

the result of the previous example, when specialized to this case, yields:

$$h_D = \frac{1}{\mu} \;\; ; \;\; h_{PM} = \int_0^\infty (1 - G(t))dt = \frac{1}{\mu_2} \;\; ; \;\; h_U = (1-\alpha(\lambda))\,\frac{1}{\lambda}$$

$$p_{U,D} = \lambda\, h_U$$

We finally obtain: $A = \left[1 + \frac{\lambda}{\mu} + \theta(\lambda)\frac{\lambda}{\mu_2}\right]^{-1}$

Component subject planned and unplanned outages - 4

The ratio of the downtime due to update to the downtime due to failure is given by $\frac{\mu}{\mu_2}\theta(\lambda)$.

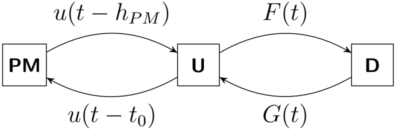
In case the time to trigger update is deterministic equal to τ, $F(t) = u(t - t_0)$ and the above ratio becomes

$$\frac{\mu}{\mu_2} \frac{e^{-\lambda t_0}}{1 - e^{-\lambda t_0}}$$

Time-based preventive maintenance - 1

Now specialize the general model to the case of time-based preventive maintenance.

We assume that the TTF and TTR for unplanned outages have respective general distributions $F(t)$ and $G(t)$.



The planned outage now will be for preventive maintenance which is carried out after a deterministic duration t_0 and hence the corresponding distribution is the unit step function $u(t - t_0)$.

D. Chen and K. Trivedi, and J. Han, "Analysis of periodic preventive maintenance with general system failure distribution," in Proc. Pacific Rim International Symposium on Dependable Computing (PRDC), 2001, pp. 103-107.

Time-based preventive maintenance - 2

For the time to carry out preventive maintenance, we need only its mean which we assume to be given by h_{PM}. Without loss of generality, we assume this distribution to be $u(t - h_{PM})$.
It follows that

$$\begin{aligned} p_{U,D} &= P\{\text{failure occurs before the PM trigger}\} = F(t_0) \\ p_{U,PM} &= 1 - F(t_0) \\ p_{PM,U} &= 1 = p_{D,U} \end{aligned}$$

$$h_U(t_0) = \int_0^{t_0} (1 - F(t))dt \ , \quad \text{and} \qquad h_D = h = \int_0^{\infty} (1 - G(t))dt$$

Hence the steady state availability is given by,

$$\begin{aligned} A = \pi_U &= \frac{h_U(t_0)}{h_U(t_0) + p_{U,PM} h_{PM} + p_{U,D} h} \\ &= \frac{h_U(t_0)}{h_U(t_0) + (1 - F(t_0))h_{PM} + F(t_0)h} \end{aligned}$$

Time-based preventive maintenance - 3

Given the nature of the distribution function $F(t)$, we can derive the expression for $h_U(t_0)$ and then compute the steady state availability.

Note that $F(t)$ should be an Increasing Failure Rate (IFR) distribution in order for PM to yield positive results.

We assume:

- HYPO(λ_1, λ_2)
- WEIB(λ, α), $\alpha > 1$
- log-normal

Time-based preventive maintenance - 4

Case 1) - The time to failure F(t) is two-stage hypo-exponentially distributed

$$F(t) = 1 - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_2 t}$$

Then,

$$h_U = \int_0^{t_0} (1 - F(t))dt$$
$$= \frac{\lambda_1}{\lambda_2(\lambda_2 - \lambda_1)} e^{-\lambda_2 t_0} - \frac{\lambda_2}{\lambda_1(\lambda_2 - \lambda_1)} e^{-\lambda_1 t_0} + \frac{\lambda_1 + \lambda_2}{\lambda_1 \lambda_2}$$

From the above relations, we can get the final expression for the steady state availability.

Time-based preventive maintenance - 5

Case 2) - The time to failure $F(t)$ is Weibull distributed

$$F(t) = 1 - \exp^{-\alpha t^{\beta}} \qquad (1)$$

and,

$$h_U = \int_0^{t_0} (1 - F(t))dt = \int_0^{t_0} e^{-\alpha t^{\beta}} dt$$

$$= \frac{1}{\beta \alpha^{1/\beta}} \int_0^{\alpha t_0^{\beta}} e^{-u} u^{\frac{1}{\beta}-1} du \qquad (2)$$

$$= \frac{\alpha^{-1/\beta}}{\beta} \Gamma\left(\frac{1}{\beta}\right) G\left(\alpha t_0^{\beta}, \frac{1}{\beta}\right)$$

where $G(x, a) = \frac{1}{\Gamma(a)} \int_0^x e^{-u} u^{a-1} du$ is the incomplete Gamma function.

From the above relations, we can get the final expression for the steady state availability.

Time-based preventive maintenance - 6

Case 3) - The time to failure $F(t)$ is Log-normal distributed
The corresponding density and distribution functions are given by:

$$f(t) = \frac{1}{\sigma t\sqrt{2\pi}} e^{\frac{-(ln(t)-\mu)^2}{2\sigma^2}}$$

$$F(t) = \frac{1}{2}\left[1 + erf\left(\frac{ln(t)-\mu}{\sigma\sqrt{2}}\right)\right]$$

Further analysis is similar to the earlier case of Weibull distributed TTF:

$$h_U = \int_0^{t_0} (1 - F(t))dt$$

$$h_U = \frac{1}{2}\int_0^{t_0} erfc\left(\frac{ln(t)-\mu}{\sigma\sqrt{2}}\right) dt$$

Where *erf* is the error function and *erfc* is the complementary error function.

Optimal preventive maintenance interval - 1

We wish to determine the optimal value of the preventive maintenance trigger interval t_0.

We can compute the optimal PM trigger interval t_0 by taking the derivative of expression for A with respect to t_0 and equate it to 0.

$$\frac{\partial A}{\partial t_0} =$$
$$\frac{h'_U(t_0)[h_U(t_0)+h_{PM}(1-F(t_0))+hF(t_0)]-h_U(t_0)[h'_U(t_0)-h_{PM}F'(t_0)+hF'(t_0)]}{[h_U(t_0)+h_{PM}(1-F(t_0))+hF(t_0)]^2}$$
$$= 0$$

Equating the numerator of the previous equation to 0 and expanding,

$$h_U(t_0)h'_U(t_0)+h_{PM}h'_U(t_0)-h_{PM}h'_U(t_0)F(t_0)+hh'_U(t_0)F(t_0)$$
$$= h_U(t_0)h'_U(t_0)-h_{PM}h_U(t_0)F'(t_0)+hh_U(t_0)F'(t_0)$$

and then,

$$h_{PM}h'_U(t_0)+(h-h_{PM})F(t_0)h'_U(t_0) = (h-h_{PM})F'(t_0)h_U(t_0)$$

Optimal preventive maintenance interval - 2

Case 2) - The time to failure $F(t)$ is Weibull distributed

$$F'(t_0) = f(t_0) = \alpha\beta(t_0)^{\beta-1}e^{-\alpha t_0^\beta}$$

and $$h'_U(t_0) = \frac{\partial}{\partial t_0}\int_0^{t_0} e^{-\alpha t^\beta}dt = e^{-\alpha t_0^\beta}$$

Substituting the expressions for $F'(t_0)$ and $h'_0(t_0)$,

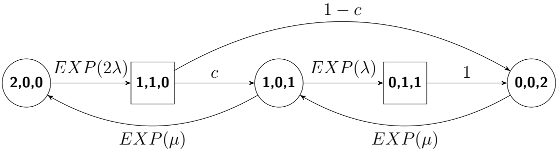
$$h_{PM}e^{-\alpha t_0^\beta} + (h - h_{PM})(1 - e^{-\alpha t_0^\beta})(e^{-\alpha t_0^\beta})$$
$$= (h - h_{PM})\alpha\beta(t_0)^{\beta-1}e^{-\alpha t_0^\beta}h_U(t_0)$$

Thus the general non-linear equation for the optimal t_0 becomes

$$(h + (h_{PM} - h)e^{-\alpha(t_0)^\beta}) + \beta\alpha(h_{PM} - h)(t_0)^{\beta-1}h_U(t_0) = 0$$

Two Repairable Components with imperfect coverage - 1

The two-component system with imperfect coverage was modeled using a GSPN in Chapter 12 whose ERG can easily be viewed as an SMP, as presented in the Figure.



This SMP consists of two kinds of states, those in which the sojourn time is exponentially distributed (the tangible states), and those in which the sojourn time is zero (the vanishing states).

We show the solution of the ERG with *preservation*, whereby the vanishing states are preserved, and the resulting ERG can be solved as an SMP.

Two Repairable Components with imperfect coverage - 2

The corresponding TPM of the embedded DTMC is given as follows:

$$\boldsymbol{P} = \begin{array}{c c} & \begin{array}{ccccc} 200 & 110 & 101 & 011 & 002 \end{array} \\ \begin{array}{c} 200 \\ 110 \\ 101 \\ 011 \\ 002 \end{array} & \left[\begin{array}{ccccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 1-c \\ \frac{\mu}{\lambda+\mu} & 0 & 0 & \frac{\lambda}{\lambda+\mu} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] \end{array}$$

Solving the steady state Equations for the embedded DTMC, we get,

$$\nu_{200} = \nu_{110} = \frac{\mu}{(4-c)\mu + 3\lambda} \quad , \quad \nu_{101} = \frac{\lambda + \mu}{(4-c)\mu + 3\lambda}$$

$$\nu_{011} = \frac{\lambda}{(4-c)\mu + 3\lambda} \quad , \quad \nu_{002} = \frac{(1-c)\mu + \lambda}{(4-c)\mu + 3\lambda}$$

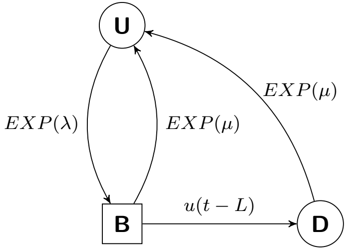
Noting the mean sojourn times in the SMP states: $h_{2,0,0} = 1/(2\lambda)$, $h_{1,0,1} = 1/(\lambda + \mu)$, $h_{0,0,2} = 1/(\mu)$, $h_{1,1,0} = h_{0,1,1} = 0$, we can derive the steady state probabilities of the SMP.

Availability of an Uninterruptible power supply (UPS)

Upon failure of the main power supply the battery supplies power to the device.

However, batteries drain out after the stored charge is exhausted - this will last for a fixed amount of time, say L.

Thus if the power supply is restored before the battery is fully discharged, the system will experience UPS.



L. Yin, R. Fricks, and K. Trivedi, "Application of semi-markov process and ctmc to evaluation of UPS system availability," in Proceedings Reliability and Maintainability Symposium, 2002, pp. 584-591

Availability of an Uninterruptible power supply (UPS)

$$p_{B,U} = \int_0^L \mu e^{-\mu t} dt = 1 - e^{-\mu L}$$

$$p_{B,D} = e^{-\mu L}$$

$$p_{U,B} = 1$$

$$p_{D,U} = 1$$

$$\boldsymbol{P} = \begin{array}{c c} & \begin{array}{ccc} U & B & D \end{array} \\ \begin{array}{c} U \\ B \\ D \end{array} & \begin{bmatrix} 0 & 1 & 0 \\ 1 - e^{-\mu L} & 0 & e^{-\mu L} \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

solving $\boldsymbol{\nu} = \boldsymbol{\nu P}$

then $\nu_U = \nu_B(1 - e^{-\mu L}) + \nu_D \, , \quad \nu_B = \nu_U \, , \quad \nu_D = \nu_B e^{-\mu L}$

and $\nu_U = \dfrac{1}{2 + e^{-\mu L}} \, , \quad \nu_B = \dfrac{1}{2 + e^{-\mu L}} \, , \quad \nu_D = \dfrac{e^{-\mu L}}{2 + e^{-\mu L}}$

Availability of an Uninterruptible power supply (UPS)

State sojourn times and SMP steady-state probabilities:

$$H_B(t) = \begin{cases} 1 - e^{-\mu t} & , t < L \\ 1 & , t \geq L \end{cases}$$

$$h_U = \frac{1}{\lambda} \quad , \quad h_B = \int_0^L e^{-\mu t} dt = \frac{1 - e^{-\mu L}}{\mu} \quad , \quad h_D = \frac{1}{\mu}$$

$$\pi_U = \frac{\mu}{\lambda + \mu} \quad , \quad \pi_B = \frac{\lambda(1 - e^{-\mu L})}{\lambda + \mu} \quad , \quad \pi_D = \frac{\lambda e^{-\mu L}}{\lambda + \mu}$$

Finally, the steady state availability of the UPS:

$$A = \pi_U + \pi_B = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu}(1 - e^{-\mu L})$$

Security attributes of a computer system - 1

In order to analyze the security attributes of an intrusion tolerant system, we need to consider the actions undertaken by an attacker as well as the system's response to an attack.

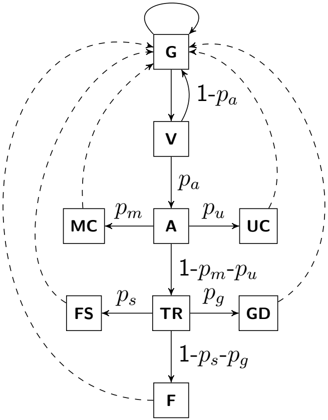
We will assume that the system satisfies the assumption of an SMP.

The embedded DTMC of this SMP is shown in the Figure. The dashed lines represents transitions for recovering the full services after an attach and returning to the good state by manual intervention.

A complete description of this SMP model requires the knowledge of various parameters, viz. mean sojourn time in each state and the transition probabilities indicated in the Figure.

Security attributes of a computer system - 2

A generic model that enables multiple intrusion tolerance strategies to exist and supports tolerance of intrusions with different impacts (e.g. compromise of confidentiality, compromise of data integrity, and DoS attacks), is sketched in the Figure.



B. B. Madan, K. Vaidyanathan, and K. S. Trivedi, "A method for modeling and quantifying the security attributes of intrusion tolerant systems," Performance Evaluation, vol. 56, 2004, pp. 167-186

Security attributes of a computer system - 3

The states have the following meaning:

Table: State description

G	good state
V	vulnerable state
A	active attack state
MC	masked compromised state
UC	undetected compromised state
TR	triage state
FS	fail-secure state
GD	graceful degradation state
F	failed state

Recovering the full services after an attack and returning to the good state by manual intervention is represented by transitions denoted with dashed lines.

Security attributes of a computer system - 4

$$
\mathbf{P} = \begin{array}{c} \\ G \\ V \\ A \\ MC \\ UC \\ TR \\ FS \\ GD \\ F \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} G & V & A & MC & UC & TR & FS & GD & F \end{array} \\
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1-p_a & 0 & p_a & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & p_m & p_u & 1-p_m-p_u & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p_s & p_g & 1-p_s-p_g \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

Solving for the steady-state probabilities of the DTMC, we get:

$$
\nu_G = \nu_V(1-p_a) + \nu_{MC} + \nu_{UC} + \nu_{FS} + \nu_{GD} + \nu_F, \;\; \nu_V = \nu_G, \;\; \nu_A = \nu_V p_a = \nu_G p_a,
$$

$$
\nu_{MC} = \nu_A p_m, \;\; \nu_{UC} = \nu_A p_u, \;\; \nu_{TR} = \nu_A(1-p_m-p_u) = \nu_G p_a(1-p_m-p_u)
$$

$$
\nu_{FS} = \nu_{TR} p_s, \;\; \nu_{GD} = \nu_{TR} p_g, \;\; \nu_F = \nu_{TR}(1-p_s-p_g)
$$

Security attributes of a computer system - 5

From the above equations, along with the normalization condition ($\sum_i \nu_i = 1, i \in \{$G,V,A, MC, UC, TR, FS, GD, F$\}$) we get

$$\nu_G = \frac{1}{2 + p_a(3 - p_m - p_u)}$$

Once the mean sojourn times in all states are given, we can derive expressions for the SMP steady state probabilities.

From these, we can get steady state system availability after recognizing that states F, FS and UC are system down states.

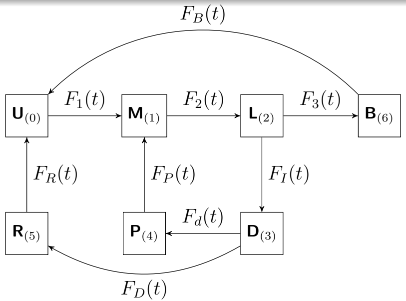
$$A = 1 - (\pi_{FS} + \pi_F + \pi_{UC})$$

For computing other security attributes such as the confidentiality and integrity, we need to consider specific attack scenarios. For several types of attacks, states UC and F will mean a loss of confidentiality while states FS will not. Hence steady state confidentiality is given by:

$$C = 1 - (\pi_F + \pi_{UC})$$

Two level rejuvenation model -1

Analysis of software aging and rejuvenation, with rejuvenation actions offered at two levels.



- U is the highly efficient and highly robust software execution phase.
- M is the medium-efficient software execution phase.
- L is alert phase. A rejuvenation action is needed or a crash may happen.
- D state where it is determined which level of rejuvenation is appropriate.
- P is the partial rejuvenation state (level 1 rejuvenation).
- R is the full rejuvenation state (level 1 rejuvenation).
- B state where the system is recovering from a crash failure.

W. Xie, Y. Hong, and K. Trivedi, "Analysis of a two-level software rejuvenation policy," Reliability Engineering & System Safety, vol. 87, no. 1, pp. 13-22, 2005.

Two level rejuvenation model - 2

All the distributions $F_i(t)$ are assumed to be non-exponential.
The time to trigger rejuvenation is typically a fixed duration, i.e. its cdf has the form of $F_I(t) = u(t-\tau)$. The non-zero entries of matrix $\boldsymbol{P}$ are:

$$p_{0,1} = p_{1,2} = p_{4,1} = p_{5,0} = p_{6,0} = 1$$

$$\begin{array}{ll} p_{2,3} = 1 - F_3(\tau) \ , & p_{2,6} = F_3(\tau) \ , \\ p_{3,4} = \int_0^\infty (1 - F_D(t)) dF_d(t) \ , & p_{3,5} = 1 - p_{3,4} \ , \end{array}$$

Solving the equations

$$\boldsymbol{\nu} = \boldsymbol{\nu}\,\boldsymbol{P} \qquad \text{subject to} \qquad \boldsymbol{\nu}\,\boldsymbol{e}^T = 1$$

we obtain the steady state probabilities of the EMC for the SMP as

$$\boldsymbol{\nu} = \frac{1}{D(\tau,p)}(1-p(1-F(\tau)), 1, 1, (1-F(\tau)), p(1-F(\tau)), (1-F(\tau))-p(1-F(\tau)), F_3(\tau))$$

in which $D(\tau,p) = 5 - p + pF_3(\tau) - F_3(\tau)$ and $p = p_{3,4}$.

Two level rejuvenation model - 3

The expected sojourn times can be computed as:

$$h_0 = \int_0^\infty (1 - F_1(t))dt \ ,$$

$$h_1 = \int_0^\infty (1 - F_2(t))dt \ ,$$

$$h_2 = \int_0^\infty (1 - F_I(t))(1 - F_3(t))dt = \int_0^\tau (1 - F_3(t))dt \ ,$$

$$h_3 = \int_0^\infty (1 - F_d(t))(1 - F_D(t))dt \ ,$$

$$h_4 = \int_0^\infty (1 - F_P(t))dt \ ,$$

$$h_5 = \int_0^\infty (1 - F_R(t))dt \ ,$$

$$h_6 = \int_0^\infty (1 - F_B(t))dt \ ,$$

Two level rejuvenation model - 4

The steady-state probability of each SMP state is

$$\pi_i = \frac{\nu_i \, h_i}{\sum_{j=0}^{6} \nu_j h_j}$$

Finally, the steady state availability of the system is:

$$A = \pi_0 + \pi_1 + \pi_2 = \frac{S(\tau, p)}{S(\tau, p) + V(\tau, p)}$$

where

$$S(\tau, p) = (1 - p + pF_3(\tau))h_0 + h_1 + \int_0^{\tau} (1 - F_3(t))dt$$
$$V(\tau, p) = (p - pF_3(\tau))h_4 + (1 - p - F_3(\tau) + pF_3(\tau))h_5 + F_3(\tau)h_6.$$

Protocol in vehicular ad-hoc networks - 1

This example regards the SMP model of the protocol used for safety messages DSRC (Dedicated Short Range Communication) in vehicular ad-hoc networks.

This SMP characterizes the packet transmission by a tagged vehicle. The vehicle is in idle state if its queue is empty.

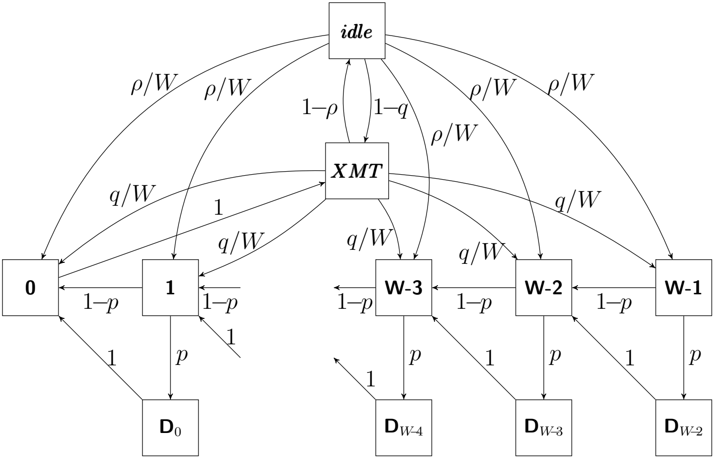
Once a packet is ready to be transmitted the vehicle first senses for channel activity for a duration known as DIFS (Distributed InterFrame Space).

If the channel is detected to be idle during this period (associated probability $1 - q$), the vehicle transits to XMT state implying that the packet is being transmitted.

X. Yin, X. Ma, and K. Trivedi, "An interacting stochastic models approach for the performance evaluation of dsrc vehicular safety communication," IEEE Transactions on Computers, vol. vol. 62, no. 5, pp. 873-885, May 2013.

Protocol in vehicular ad-hoc networks - 2

The SMP model capturing channel contention and backoff behavior is shown in Figure



Protocol in vehicular ad-hoc networks - 3

In the case the channel is sensed to be busy during the DIFS period, the vehicle will backoff with the backoff counter set randomly in the range $[0, W-1]$ where W is the backoff window size.

The backoff counter is decremented by 1 each time the channel is detected idle during a time slot of duration σ (this happens with probability $1-p$) corresponding SMP transition is from state $W-i$ to $W-i-1$.

If another vehicle is transmitting and hence the channel is sensed busy during the backoff time slot σ (i.e., another vehicle is transmitting a packet), the backoff counter of the tagged vehicle is suspended and deferred for the duration of packet transmission, that is a period of length T.

Protocol in vehicular ad-hoc networks - 4

The corresponding transition in the SMP occurs from state $W - i$ to D_{W-i-1} with probability p.

Once the backoff counter reaches the value 0, the packet will be transmitted while SMP transits from state 0 to state XMT.

Upon transmission, if the queue of the tagged vehicle is empty (with probability $1 - \rho$) the SMP transits to the idle state.

Else if there are packets in the queue (with probability ρ), the vehicle will sense the channel for a DIFS period and randomly chooses a backoff counter.

Protocol in vehicular ad-hoc networks - 5

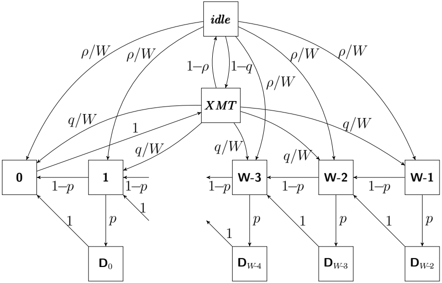
From the state diagram and corresponding TPM, we can obtain the stationary probability vector of the embedded DTMC:

$$\begin{cases} \nu_j &= (W-j)\,\nu_{W-1}, \qquad j = 0, 1, \ldots, W-1 \\ \nu_{D_j} &= (W-j-1)\,p\,\nu_{W-1}, \quad j = 0, 1, \ldots, W-2 \\ \nu_{\text{XMT}} &= \dfrac{W}{\rho + q\,(1-\rho)}\nu_{W-1} \\ \nu_{\text{idle}} &= \dfrac{(1-\rho)W}{\rho + q\,(1-\rho)}\nu_{W-1} \end{cases}$$

Finally using the normalization condition, we get :

$$\nu_{W-1} = \frac{2[\rho + q\,(1-\rho)]}{[W+1+p(W-1)][\rho + q\,(1-\rho)]W + 2(2-\rho)W}$$

Protocol in vehicular ad-hoc networks - 6



The mean sojourn times in the states of the SMP are given by:

$$h_j = \begin{cases} \sigma & j = 0, 1, \ldots, W-1 \\ T & j = D_0, D_1, \ldots, D_{W-2} \\ T & j = \text{XMT} \\ \frac{1}{\lambda} + \text{DIFS} & j = \text{idle} \end{cases}$$

From the above the steady state probabilities of SMP states can be computed

SMP with Absorbing States

SMP with Absorbing States

SMP with one or more absorbing states. The TPM of the embedded DTMC now can be partitioned as:

$$\boldsymbol{P} = \left[\begin{array}{c|c} \boldsymbol{P}_u & \boldsymbol{c}^T \\ - & - \\ \boldsymbol{0} & 1 \end{array} \right]$$

$\boldsymbol{P}_u$ is the $(n-1 \times n-1)$ partition of matrix $\boldsymbol{P}$ over the transient states and is a sub-stochastic matrix (row sum less than 1);

$\boldsymbol{c}^T$ is a column vector grouping the transition probabilities from any transient state to the absorbing state.

Note that the absorbing state needs a self loop with probability 1 in order to make the TPM a stochastic matrix so that all row sums are equal to 1.

Mean Time to Absorption

The expected number of visits to state j until absorption is:

$$V_j = \alpha_j + \sum_{i=1}^{n-1} V_i p_{ij} \qquad\qquad j = 1, 2, \ldots, n$$

where α_j is the initial probability of state j and n is the absorbing state.

In vector-matrix form: $\boldsymbol{V} = \boldsymbol{\alpha} + \boldsymbol{V}\,\boldsymbol{P}_u$

where $\boldsymbol{\alpha} = [\alpha_j]$ is the initial probability vector and $\boldsymbol{V} = [V_j]$ the vector of the expected number of visits in each state.

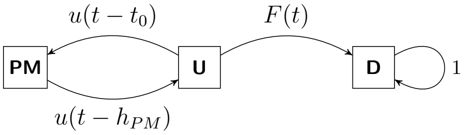
The mean time to absorption is written as

$$\text{MTTA} = \sum_{j=1}^{n-1} V_j\, h_j$$

where h_j is the mean sojourn time in state j and $V_j\, h_j$ the total expected time spent in state j before absorption.

MTTA: periodic preventive maintenance - 1

We compute mean time to absorption by disallowing repair from the failure state F.



The TPM for this case is

$$
\boldsymbol{P} = \begin{array}{c} \\ U \\ PM \\ D \end{array}
\begin{array}{c}
\begin{array}{ccc} U & PM & D \end{array} \\
\left[\begin{array}{cc|c}
0 & 1 - F(t_0) & F(t_0) \\
1 & 0 & 0 \\
\hline
0 & 0 & 1
\end{array}\right]
\end{array}
$$

We assume, as initial probability, $\alpha_U = 1$.

MTTA: periodic preventive maintenance - 2

Then: $V_U = 1 + V_{PM} \quad , \quad V_{PM} = [1 - F(t_0)]\, V_U$

Hence: $V_U = \dfrac{1}{F(t_0)} \quad , \quad V_{PM} = \dfrac{1 - F(t_0)}{F(t_0)}$

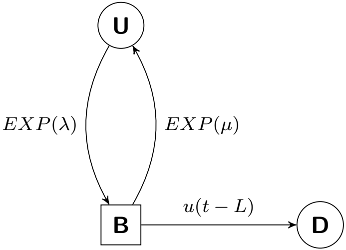
Finally,
$$\begin{aligned} \text{MTTA} &= \frac{h_U}{F(t_0)} + \frac{1 - F(t_0)}{F(t_0)} h_{PM} \\ &= \frac{h_U + h_{PM}}{F(t_0)} - h_{PM} \end{aligned}$$

During time to absorption, the system is down for PM multiple times. So if we attach reward rate 1 to U and 0 to PM, we can compute the expected accumulated reward till absorption as mean capacity till failure (MCTF):

$$\text{MCTF} = \frac{h_U}{F(t_0)} = \frac{\int_0^{t_0} (1 - F(t))dt}{F(t_0)}$$

MTTA: UPS Example - 1

By dropping the repair transition, we have modified the UPS system Example, to have a SMP model with absorbing state.



The TPM for this case is

$$
\boldsymbol{P} = \begin{array}{c} \\ U \\ B \\ D \end{array}
\begin{array}{c}
\begin{array}{ccc} U & B & D \end{array} \\
\left[\begin{array}{cc|c}
0 & 1 & 0 \\
1 - e^{-\mu L} & 0 & e^{-\mu L} \\
\hline
0 & 0 & 1
\end{array}\right]
\end{array}
$$

Then assuming as initial probability $\alpha_U = 1$, we have

$$
\begin{aligned}
V_U &= 1 + V_B(1 - e^{-\mu L}) \\
V_B &= V_U
\end{aligned}
$$

Hence $V_U = 1 + V_U(1 - e^{-\mu L})$ or $V_U = e^{\mu L}$.

MTTA: UPS Example - 2

MTTA, which in this case is also MTTF, is given by :

$$
\begin{aligned}
\text{MTTF} &= \frac{V_U}{\lambda} + V_B \left(\frac{1 - e^{-\mu L}}{\mu} \right) \\
&= \frac{e^{+\mu L}}{\lambda} + \frac{e^{+\mu L}}{\mu} - \frac{1}{\mu} \\
&= e^{+\mu L} \left(\frac{1}{\lambda} + \frac{1}{\mu} \right) - \frac{1}{\mu}
\end{aligned}
$$

Note that if $L = 0$, MTTF reduces to $1/\lambda$ since as soon as the power supply is down, the system goes down without a battery backup.

RF channel in a cellular wireless network - 1

Consider a model of an RF (radio frequency) channel in a cellular wireless network. The quality of a channel is determined by its signal to noise ratio (SNR).

An instantaneous drop in SNR below a threshold does not necessarily lead to the occurrence of an outage event.

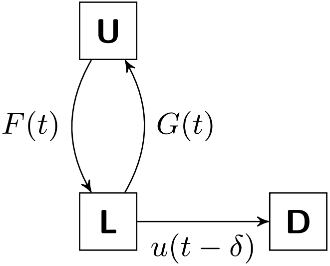
A channel outage event is determined by the duration of time that the SIR stays below a threshold.

To reflect the impact of this time duration, an RF channel outage event is said to occur when the SNR stays below a threshold for a duration longer than the “minimum duration", δ.

This leads to a SMP model.

Y. Ma, J. Han, and K. Trivedi, "Transient analysis of minimum duration outage for rf channel in cellular systems," in IEEE Vehicular Technology Conference, vol. 2, Jul 1999, pp. 1698-1702

RF channel in a cellular wireless network - 2



The TPM of the embedded DTMC is given by

$$
\boldsymbol{P} = \begin{array}{c} \\ U \\ L \\ D \end{array}
\begin{array}{c}
\begin{array}{cc|c} \;\;U\;\; & \;\;L\;\; & \;\;D\;\; \end{array} \\
\left[\begin{array}{cc|c}
0 & 1 & 0 \\
G(\delta) & 0 & 1-G(\delta) \\
\hline
0 & 0 & 1
\end{array}\right]
\end{array}
$$

The mean sojourn times are given by:

$$
h_U = \int_0^{\infty} (1 - F(t))dt \quad , \quad h_L = \int_0^{\delta} (1 - G(t))dt
$$

RF channel in a cellular wireless network - 3

Assuming, as initial probability, $\alpha_U = 1$, we obtain the average number of visits:

$$V_U = 1 + V_L G(\delta) \ , \ \ V_L = V_U$$

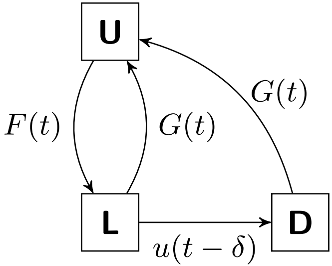
Hence $V_U = 1 + V_U G(\delta)$ and $V_U = \frac{1}{1-G(\delta)}$.

Thence MTTA which is also MTTF in this case is given by

$$\begin{aligned} \text{MTTF} &= \frac{h_U}{1-G(\delta)} + \frac{h_L}{1-G(\delta)} \\ &= \frac{\int_0^\infty (1-F(t))dt}{1-G(\delta)} + \frac{\int_0^\delta (1-G(t))dt}{1-G(\delta)} \end{aligned}$$

RF channel in a cellular wireless network - 4

We can also now study the steady state availability of the channel by allowing recovery from state D as in the Figure.



We assume that once channel has failed, recovery activity, if any, that was started prior to failure needs to be restarted. Now substitute MTTF from above and MTTR $= \int_0^\infty (1 - G(t))dt$ to get:

$$A = \frac{MTTF}{MTTF + MTTR} = \left(1 + \frac{MTTR}{MTTF}\right)^{-1}$$

Probability of Absorptions

SMP with multiple absorbing states. Assume that $1 \leq m < n$ absorbing states in an n-state SMP. Then the TPM can be partitioned so that

$\boldsymbol{Q}$ is an $(n-m)$ by $(n-m)$ sub-stochastic matrix, $\boldsymbol{C}$ is a rectangular $(n-m)$ by m matrix.

$$\boldsymbol{P} = \left[\begin{array}{c|c} \boldsymbol{P}_u & \boldsymbol{C} \\ -- & -- \\ \boldsymbol{0} & \boldsymbol{I} \end{array} \right]$$

Matrix $\boldsymbol{B} = [b_{ij}]$ is defined so that b_{ij} is the probability of being absorbed in state j given that the DTMC started in state i. It can be shown that:

$$\boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{P}_u)^{-1}\boldsymbol{C}$$

Note that $(\boldsymbol{I} - \boldsymbol{P}_u)^{-1}$ is known as the fundamental matrix of the embedded DTMC.

Security attributes of a computer system - 1

We return to the security quantification Example.

We identify those states that are security compromised states and make them absorbing states.

Subsequently the MTTA in this case can be christened as mean time to security failure (MTTSF).

For example, for SUN web server bulletin board vulnerability (Bugtrack ID 1600), states UC, FS, GD and F will form the set of absorbing states.

B. B. Madan, K. Vaidyanathan, and K. S. Trivedi, "A method for modeling and quantifying the security attributes of intrusion tolerant systems," Perform. Eval, vol. 56, 2004, pp. 167-186

Security attributes of a computer system - 2

Assuming that $\alpha_G = 1$, we find:

$V_G = \frac{1}{p_a(1-p_m)} = V_V$, $V_A = \frac{1}{1-p_m}$, $V_{MC} = \frac{p_m}{1-p_m}$, $V_{TR} = \frac{1-p_m-p_u}{1-p_m}$

and $$\mathrm{MTTSF} = \frac{h_G p_a^{-1} + h_V p_a^{-1} + h_A + h_{MC} p_m + h_{TR}(1 - p_m - p_u)}{1 - p_m}$$

Furthermore, we can find the absorption probabilities as:

$$\begin{aligned}
b_{G,F} &= \frac{(1 - p_s - p_g)(1 - p_m - p_u)}{1 - p_m} \\
b_{G,FS} &= \frac{p_s(1 - p_m - p_u)}{1 - p_m} \\
b_{G,GD} &= \frac{p_g(1 - p_m - p_u)}{1 - p_m} \\
\text{and} \quad b_{G,UC} &= \frac{p_u}{1 - p_m}
\end{aligned}$$

Transient solution

Transient analysis

We denote by $\boldsymbol{V}(t) = [V_{ij}(t)]$ the matrix of the transient solution for the conditional transition probability, where each entry $V_{ij}(t)$ is defined as:

$$V_{ij}(t) = P\{Z(t) = j \,|\, Z(0) = i\}$$

We define the kernel matrix of the SMP as

$$\boldsymbol{K}(t) = [k_{ij}(t)].$$

For this purpose, define the sequence of time points at which $\{Z(t) \,|\, t \geq 0\}$ makes state transitions as

$$T_0 = 0,\ T_1,\ T_2,\ \ldots,\ T_n,\ T_{n+1},\ \ldots, \qquad \text{and} \qquad X_n = Z(T_n).$$

$$\text{Then,} \qquad k_{ij}(t) = P\{X_{n+1} = j,\, T_{n+1} - T_n \leq t \,|\, X_n = i\}$$

$k_{ij}(t)$ is the conditional probability that, given the process has entered state i at time T_n, the next transition occurs at time t toward state j. We assume that $k_{ij}(t)$ satisfy the homogeneity property so that they don't depend on n.

Transient analysis

Based on our earlier characterization, it can be seen that

$$k_{ij}(t) \quad = \quad p_{ij}\, H_i(t), \qquad \text{or, alternatively}$$

$$p_{ij} \; = \; lim_{t\to\infty} \; k_{ij}(t) \quad \text{and} \quad \boldsymbol{P} \; = \; lim_{t\to\infty} \; \boldsymbol{K}(t) \; ,$$

$$\text{and} \qquad H_i(t) = \sum_j k_{ij}(t)$$

where $H_i(t)$ is the distribution of the sojourn time in state i.

The transient solution for the conditional probabilities $V_{ij}(t)$ can be shown to satisfy the equations:

$$V_{ij}(t) = [1 - H_i(t)]\delta_{ij} + \sum_k \int_0^t V_{kj}(t-\tau)\, dk_{ik}(\tau)$$

where δ_{ij} is the Kronecker delta equal to 1 if $i = j$ and is 0 otherwise.

Transient analysis

The coupled set of Volterra equations of the second kind is a *Markov renewal equation* and can be solved in L-S domain.

$$\begin{aligned} \boldsymbol{V}^{\sim}(s) &= \boldsymbol{E}^{\sim}(s) + \boldsymbol{K}^{\sim}(s)\,\boldsymbol{V}^{\sim}(s) \\ &= [\boldsymbol{I} - \boldsymbol{K}^{\sim}(s)]^{-1}\boldsymbol{E}^{\sim}(s) \end{aligned}$$

where: $\boldsymbol{E}^{\sim}(s) = \int_0^{\infty} e^{-st} d\boldsymbol{E}(t)$ is a diagonal matrix with entries $E_{ii}(t) = 1 - H_i(t)$,

$\boldsymbol{K}^{\sim}(s) = \int_0^{\infty} e^{-st} d\boldsymbol{K}(t)$.

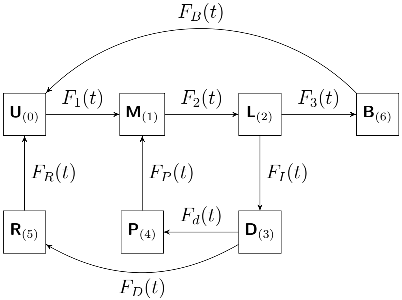
After solving the equation and taking the inverse L-S transform of $\boldsymbol{V}^{\sim}(s)$, the unconditional state probabilities $\boldsymbol{\pi}(t) = [\pi_i(t)]$ becomes

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\,\boldsymbol{V}(t)$$

where $\boldsymbol{\pi}(0)$ the initial state probability vector.

Two level rejuvenation model: Kernel matrix

The transient analysis of the two level rejuvenation model can be carried out by determining the kernel matrix $\boldsymbol{K}(t)$ as follows:



$$\boldsymbol{K}(t) = \begin{bmatrix} 0 & k_{0,1}(t) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_{1,2}(t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_{2,3}(t) & 0 & 0 & k_{2,6}(t) \\ 0 & 0 & 0 & 0 & k_{3,4}(t) & k_{3,5}(t) & 0 \\ 0 & k_{4,0}(t) & 0 & 0 & 0 & 0 & 0 \\ k_{5,0}(t) & 0 & 0 & 0 & 0 & 0 & 0 \\ k_{6,0}(t) & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Two level rejuvenation model: Kernel matrix

where the non-zero elements of $\boldsymbol{K}(t)$ could be derived as follows

$$k_{0,1}(t) = F_1(t), \qquad k_{1,2}(t) = F_2(t),$$

$$k_{2,3}(t) = \int_0^t (1 - F_3(t))dF_I(t), \qquad k_{2,6}(t) = \int_0^t (1 - F_I(t))dF_3(t),$$

$$k_{3,4}(t) = \int_0^t (1 - F_D(t))dF_d(t), \qquad k_{3,5}(t) = \int_0^t (1 - F_d(t))dF_D(t),$$

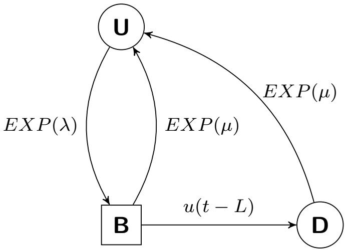
$$k_{4,1}(t) = F_P(t), \quad k_{5,0}(t) = F_R(t), \quad k_{6,0}(t) = F_B(t),$$

We can take the limit as $t \to \infty$ and compare with the previous example.

Transient analysis of the UPS model

To carry out the transient analysis, first we develop the kernel matrix

$$\boldsymbol{K}(t) = \begin{bmatrix} 0 & k_{UB}(t) & 0 \\ k_{BU}(t) & 0 & k_{BD}(t) \\ k_{DU}(t) & 0 & 0) \end{bmatrix}$$



$$k_{UB}(t) = P\{\text{power supply fails before or at time t}\}$$
$$k_{BU}(t) = P\{\text{repair of power supply completed before or at time } t \text{ and } t < L\}$$
$$k_{BD}(t) = P\{\text{repair of power supply not completed at time } t \text{ and } t \geq L\}$$
$$k_{DU}(t) = P\{\text{power supply is repaired before or at time } t\}$$

Transient analysis of the UPS model

From the above definitions, we can determine the entries in the kernel matrix both in the time domain and in the LST domain:

$$
\begin{aligned}
k_{UB}(t) &= 1 - e^{-\lambda t} &, \quad k^{\sim}_{UB}(s) &= \frac{\lambda}{s+\lambda} \\
k_{BU}(t) &= \begin{cases} 1 - e^{-\mu t} & , \quad t < L \\ 1 - e^{-\mu L} & , \quad t \geq L \end{cases} &, \quad k^{\sim}_{BU}(s) &= (1 - e^{-(s+\mu)L}) \left(\frac{\mu}{s+\mu} \right) \\
k_{BD}(t) &= \begin{cases} 0 & , \quad t < L \\ e^{-\mu L} & , \quad t \geq L \end{cases} &, \quad k^{\sim}_{BD}(s) &= e^{-(s+\mu)L} \\
k_{DU}(t) &= 1 - e^{-\mu t} &, \quad k^{\sim}_{DU}(s) &= \frac{\mu}{s+\mu}
\end{aligned}
$$

Transient analysis of the UPS model

Further,

$$E_{ii}(t) = 1 - H_i(t) = 1 - \sum_j k_{ij}(t) \qquad ; \quad i, j = U, B, D$$

Hence, the individual entries for the diagonal matrix $\boldsymbol{E}(t)$ are, in the time domain and in LST domain:

$$\begin{aligned}
E_{UU}(t) &= e^{-\lambda t} &, \quad E^{\sim}_{UU}(s) &= \frac{s}{s+\lambda} \\
E_{BB}(t) &= \begin{cases} e^{-\mu t} & , \quad t < L \\ 0 & , \quad t \geq L \end{cases} &, \quad E^{\sim}_{BB}(s) &= \frac{1 - e^{-(s+\mu)L}}{s+\mu} \\
E_{DD}(t) &= e^{-\mu t} &, \quad E^{\sim}_{DD}(s) &= \frac{s}{s+\mu}
\end{aligned}$$

Assuming as initial probability $\pi_U(0) = 1$, the availability can be expressed as $A(t) = 1 - \pi_D(t) = 1 - V_{UD}(t)$.

Transient analysis of the UPS model

Inverting $\boldsymbol{V}^{\sim}(s)$ to get $\boldsymbol{V}(t)$ is a difficult problem that usually precludes closed-form solutions. In this case a closed-form solution has been derived:

$$\begin{aligned} A(t) &= 1 - \pi_D(t) \\ &= 1 - \frac{\lambda}{\lambda + \mu}\left[e^{-\mu L} - e^{-(\lambda+\mu)t+\lambda L}\right] u(t - L) \end{aligned} \tag{3}$$

The steady state availability A can be determined either by taking the limit of Equation (3) as $t \to \infty$ or by simply considering that

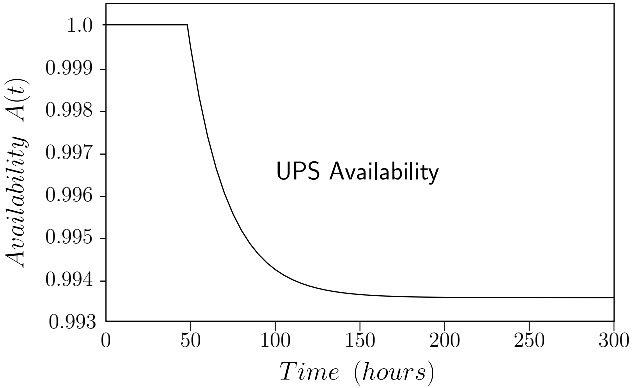
$$\boldsymbol{V}(\infty) = \lim_{t\to\infty} \boldsymbol{V}(t) = \lim_{s\to 0} \boldsymbol{V}^{\sim}(s)$$

and determining A directly from $\boldsymbol{V}^{\sim}(s)$ without the inverse Laplace transform:

$$A = 1 - \frac{\lambda}{\lambda + \mu} e^{-\mu L}$$

Transient analysis of the UPS model

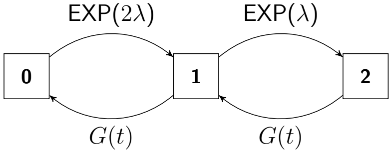
Assuming as numerical values $L = 48\ h$, $\lambda = 1/480\ h^{-1}$ and $\mu = 1/24\ h^{-1}$, the availability $A(t)$ is plotted as a function of time in the Figure.



Markov Regenerative Process

Markov Regenerative Processes (MRGP): An introduction

Consider a system with two components each with the time to failure distribution EXP(λ) that share a single repair-person whose repair time follows a general distribution $G(t)$.



In a realistic model, the repair continues until completion even if a second failure occurs.

In this case, the instant of entrance in State 2 is not a regeneration point for the process, since a general distribution is active.

In fact, the systems enters State 2 only if a failure occurs in state 1 before repair is completed, and thus repair continues and must be completed with the same distribution after the entrance in state 2.

In this model, the instants of entrance in states 0 and 1 are regeneration points for the process, but not the instants of entrance in state 2.

MRGP: An introduction

A Markov regenerative process is a generalization of many stochastic processes.

- In a CTMC (either Homogeneous or non Homogeneous) any time epoch t satisfies the Markov property;
- In a NHCTMC transition rates are time dependent but share the same global clock;
- in a SMP the Markov property is satisfied at all the time instances at which the process undergoes a state transition;
- in a MRGP the Markov property holds only when the process enters a subset of specific states called *Regeneration states*.
  In a MRGP the Markov property is satisfied in a sequence of embedded time points (called *Regeneration Time Point - RTP*).

MRGP: An introduction

More formally, for a MRGP $\{Z(t),\ t \geq 0\}$ there are time instances $T_0,\ T_1,\ ...,\ T_n,\ ...$ such that the states of the process at those time points ($Y_0,\ Y_1,\ ...,\ Y_n,\ ...$, respectively) satisfy the Markov property, ($\forall\ i_0, i_1, ..., i_n \in \Omega$)

$$P\{Y_n = i_n \mid Y_{n-1} = i_{n-1}, ..., Y_1 = i_1, Y_0 = i_0\} = P\{Y_n = i_n \mid Y_{n-1} = i_{n-1}\}.$$

Therefore it does not matter what states the process $\{Z(t),\ t \geq 0\}$ has visited until reaching Y_n at time T_n.

The state $Z(T_n)$ is the only needed information for the future development of $Z(T_n + t),\ t \geq 0$.

MRGP: An introduction

The embedded time points $\{T_n, n \geq 0\}$ are the **regeneration time points** (RTP): the process behaves in the same way each time it passes in a regeneration time point.

$$P\{Z(T_n + t),\ t \geq 0 \mid Z(T_n) = i_n\} = P\{Z(t),\ t \geq 0 \mid Z(0) = i_n\}$$

The evolution of the process between two consecutive RTPs is called the *subordinated process*.

Transient and steady-state analysis of a MRGP

The two matrices that must be defined for the transient analysis of a MRGP are commonly referred to as the *global kernel* and *local kernel*.

The global kernel is a matrix $\boldsymbol{K}(t) = [K_{ij}(t)]$ that describes the occurrence of the next *RTP*:

$$K_{ij}(t) = P\{Y_1 = j\,,\; T_1 \le t \,|\, Y_0 = i\}$$

where Y_1 is the right continuous state hit by $Z(t)$ at the next *RTP*.

The local kernel describes the behaviour of the system inside a subordinated process.

The local kernel is a matrix $\boldsymbol{E}(t) = [E_{ij}(t)]$ that gives the state transition probabilities in a regeneration interval, before the next *RTP* occurs:

$$E_{ij}(t) = P\{Z(t) = j,\; T_1 > t \,|\; Y_0 = i\}$$

The transition probability matrix of a MRGP

Let $\boldsymbol{V}(t) = [V_{ij}(t)]$ define the transition probability matrix over $(0, t]$:

$$V_{ij}(t) = P\{Z(t) = j \mid Z(0) = Y_0 = i\} \qquad\qquad i, j \in \Omega$$

Based on the global and the local kernels the transient analysis can be carried out in the time domain by solving the following generalized Markov renewal equation:

$$V_{ij}(t) \;=\; E_{ij}(t) + \sum_{\Omega} \int_0^t \, d\,K_{ik}(y)\; V_{kj}(t-y)$$

In matrix form:

$$\boldsymbol{V}(t) \;=\; \boldsymbol{E}(t) + \boldsymbol{K} * \boldsymbol{V}(t)$$

where $*$ is the convolution integral symbol.

The transition probability matrix of a MRGP

In the LST domain:

$$\boldsymbol{V}^{\sim}(s) = [\boldsymbol{I} - \boldsymbol{K}^{\sim}(s)]^{-1}\, \boldsymbol{E}^{\sim}(s) \tag{4}$$

where the superscript $\sim$ indicates the Laplace-Stieltjes transform (LST) and s the complex transform variable of t

(remainder: $F^{\sim}(s) = \int_0^{\infty} e^{-st} dF(t)$).

In order to use the above equations, we need to specify $\boldsymbol{K}(t) = [K_{ij}(t)]$ and $\boldsymbol{E}(t) = [E_{ij}(t)]$ matrices and their LST.

A time domain solution for $\boldsymbol{V}(t)$ can be obtained by numerically integrating the convolution equation.

Alternatively, starting from the LST equation, a combination of symbolic and numeric computation is needed to obtain measures in the time domain.

Steady-state analysis of a MRGP

For the purpose of the steady-state analysis of a MRGP, the following two matrices $\boldsymbol{\alpha} = [\alpha_{ij}]$ and $\boldsymbol{\phi} = [\phi_{ij}]$ should be evaluated.

The two matrices are defined as:

$$
\begin{aligned}
\alpha_{ij} &= \int_{t=0}^{\infty} E_{ij}(t)\, dt = \lim_{s \to 0} \frac{1}{s} E_{ij}^{\sim}(s) \\
\phi_{ij} &= \lim_{t \to \infty} K_{ij}(t) = \lim_{s \to 0} K_{ij}^{\sim}(s)
\end{aligned}
$$

α_{ij} is the expected time that the subordinated process starting from state i spends in state j.

Matrix $\boldsymbol{\phi} = [\phi_{ij}]$ is the one-step transition probability matrix of the DTMC embedded at the *RTP*s and hence ϕ_{ij} gives the probability that the subordinated process is followed by a regeneration interval starting from state j.

We assume, here, that all the subordinated processes have a finite mean sojourn time, so that the above defined measures exist and are finite.

Steady-state analysis of a MRGP

Once matrices $\boldsymbol{\alpha} = [\alpha_{ij}]$ and $\boldsymbol{\phi} = [\phi_{ij}]$ have been generated from the model definition, the steady-state analysis of a MRGP requires the following three steps:

Step 1: Compute:

$$\alpha_i = \sum_{j \in \Omega} \alpha_{ij}$$

α_i is the expected duration of the subordinated process starting from state i, before the next *RTP*.

Step 2: Evaluate the state probability vector $\boldsymbol{\nu} = [\nu_i]$, whose elements are the unique solution of:

$$\boldsymbol{\nu} = \boldsymbol{\nu}\boldsymbol{\phi} \; ; \qquad \sum_{\Omega} \nu_i = 1$$

$\boldsymbol{\nu}$ is the stationary state probability vector of the *DTMC* embedded at the *RTP*s.

Steady-state analysis of a MRGP

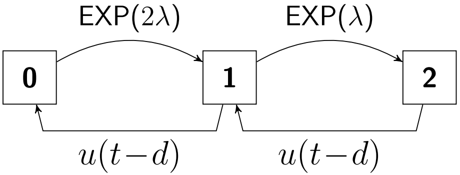
Step 3: The steady-state probabilities of the MRGP are given by:

$$\pi_j = \lim_{t \to \infty} P\{Z(t) = j\} = \frac{\displaystyle\sum_{k \in \Omega} \nu_k \, \alpha_{kj}}{\displaystyle\sum_{k \in \Omega} \nu_k \, \alpha_k}$$

Example: The M/D/1/2/2 Queue

The M/D/1/2/2 queueing model represents two customers in the system each of which submits a job at exponentially distributed interval of rate λ, and one server with deterministic service times of duration *d*.

The size of the queues for arriving jobs is two.

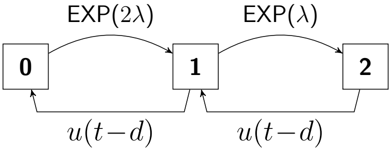


The model of the figure fits also the 2-component parallel system with shared repair when the repair transitions are assumed deterministic with service time *d*.

Hoon Choi and V.G. Kulkarni and K. Trivedi, "Transient analysis of deterministic and stochastic Petri nets," Proc. 14-th Int. Conf. Application and Theory of Petri Nets, LNCS-691, 1993, pp. 166-185

Example: The M/D/1/2/2 Queue

In State 0 both costumers are thinking and the arrival rate of a job to the queue is 2λ.



In State 1 two activities are in competition: the arrival of a second job at rate λ and the service of the job in the queue with a deterministic time d.

In State 2 two jobs are in the queue and the only action is the completion of one of them with a deterministic time d.

States 0 and 1 are regeneration states.
We want to compute the state probability vector $\boldsymbol{\pi}(t) = [\pi_j(t)]$, $(j = 0, 1, 2)$ at time t assuming the system starts from state 0 with probability 1, so that $\pi_j(t) = V_{0j}(t)$.

Example: The M/D/1/2/2 Queue

We derive the global kernel $\boldsymbol{K}(t) = [K_{ij}(t)]$ $(i, j = 0, 1, 2)$ of this model row by row.

$$K_{00}(t) = 0 \quad ; \quad K_{01}(t) = 1 - e^{-2\lambda t} \quad ; \quad K_{02}(t) = 0$$

$$K_{10}(t) = \begin{cases} 0 & t < d \\ e^{-\lambda d} & t \geq d \end{cases} \quad ; \quad K_{11}(t) = \begin{cases} 0 & t < d \\ 1 - e^{-\lambda d} & t \geq d \end{cases} \quad ; \quad K_{12}(t) = 0$$

$$K_{20}(t) = 0 \quad ; \quad K_{21}(t) = \begin{cases} 0 & t < d \\ 1 & t \geq d \end{cases} \quad ; \quad K_{22}(t) = 0$$

Example: The M/D/1/2/2 Queue

Then we derive the local kernel $\boldsymbol{E}(t) = [E_{ij}(t)]$ $(i, j = 0, 1, 2)$ row by row, as:

$$E_{00}(t) = e^{-2\lambda t} \quad ; \quad E_{01}(t) = 0 \quad ; \quad E_{02}(t) = 0$$

$$E_{10}(t) = 0 \quad ; \quad E_{11}(t) = \begin{cases} e^{-\lambda t} & t < d \\ 0 & t \geq d \end{cases} \quad ; \quad E_{12}(t) = \begin{cases} 1 - e^{-\lambda t} & t < d \\ 0 & t \geq d \end{cases}$$

$$E_{20}(t) = 0 \quad ; \quad E_{21}(t) = 0 \quad ; \quad E_{22}(t) = \begin{cases} 1 & t < d \\ 0 & t \geq d \end{cases}$$

Solution is obtained by deriving the LST's of $\boldsymbol{K}(t)$ and $\boldsymbol{E}(t)$, replacing the obtained expressions in the equation for $\boldsymbol{V}^{\sim}(s)$ and then by numerically inverting the LST.

Example: The M/D/1/2/2 Queue

For the numerical solution of the above model and further readings see:

H. Choi, V. Kulkarni, and K. Trivedi, "Transient analysis of deterministic and stochastic petri nets," in *Application and Theory of Petri Nets 1993*, ser. Lecture Notes in Computer Science, M. Ajmone Marsan, Ed., vol. 691. Springer, 1993, pp. 166–185.

A. Bobbio and M. Telek, "Computational restrictions for SPN with generally distributed transition times," in *First European Dependable Computing Conference (EDCC-1), Lecture Notes in Computer Sciencee - LNCS 852*, D. H. K. Echtle and D. Powell, Eds. Springer Verlag, 1994, pp. 131–148.

H. Choi, V. Kulkarni, and K. Trivedi, "Markov regenerative stochastic Petri nets," *Performance Evaluation*, vol. 20, pp. 337–357, 1994.

R. German, D. Logothetis, and K. Trivedi, "Transient analysis of Markov Regenerative Stochastic Petri Nets: a comparison of approaches," in *6-th International Conference on Petri Nets and Performance Models - PNPM95*. IEEE Computer Society, 1995, pp. 103–112.

A. Bobbio, A. Puliafito, and M. Telek, "A modeling framework to implement preemption policies in non-Markovian SPN," *IEEE Transactions Software Engineering*, vol. 26, pp. 36–54, 2000.