Approximate aggregation and applications to reliability

Yannick Lefebvre
Université de Marne-la-Vallée, EDF R&D
EDF R&D, dépt. MRI, 1 av. du Général De Gaulle
92141 Clamart cedex, France
yannick.lefebvre@edf.fr

Abstract

In this paper, we present an approximate aggregation principle to reduce the state space of an homogeneous Markov process. This method is then used to compute a bound of the asymptotic unavailability of a Markovian system. Finally, we show that calculating this dependability measure is of much interest even in different contexts, since the unavailability at time t of several ageing systems can be efficiently approximated by the asymptotic unavailability of an homogeneous Markov process.

1 Introduction

Homogeneous Markov processes are commonly used to assess dependability measures of complex systems. A well-known issue in such models is the combinatorial explosion when the system is too large. One of the possible solutions is to avoid the construction of the state graph by exploring it in terms of paths. This idea already led to efficient algorithms for the evaluation of reliability (Bon & Collet 1994) and asymptotic unavailability (Bouissou & Lefebvre 2002), implemented in a tool called Figseq.

An alternative solution is to compare the system with another one evolving in a reduced state space (Mahevas & Rubino 2001), which can be done using exact aggregation, also called strong lumpability, in the most favourable case. The comparison of the two processes may then be carried out using stochastic ordering (Stoyan 1983) or coupling techniques (Doisy 2000). In this paper, we present an "approximate aggregation" principle, and use a martingale technique to give an upper bound of the asymptotic unavailability.

Another drawback of Markov models is that the transition rates are supposed to be constant, which is not a reasonable assumption if the system is deteriorating with the passing of time. Nevertheless, we prove in the final section that the distribution at time t of an ageing system is close to the stationary distribution of a non-ageing one, actually if the ageing is not too fast.

2 The reduced state space

Let $(S_t)_{t\geq 0}$ be a homogeneous Markov process, with values in a finite state space $\mathcal{E}$. Our aim is to compare the process with another one, say $(S'_t)_{t\geq 0}$, which takes values in a smaller state space $\mathcal{F}$. More precisely, we suppose that there exists a surjective function $f : \mathcal{E} \to \mathcal{F}$ such that:

$$\mathcal{E} = \bigcup_{i\in\mathcal{F}} \mathcal{E}_i \ , \ \eta \in \mathcal{E}_i \Leftrightarrow f(\eta) = i$$

In other words, the state space of $(S_t)_{t\geq 0}$ is divided in subsets, each one corresponding to a single state of $\mathcal{F}$.

As an example, let us consider a system made of N components with constant failure and repair rates, and with two possible states (the working state 1 and the failure state 0). Suppose that these components can be classified in k different families (for instance, two components of the same family have similar failure and repair rates), and let us denote N_i the number of components of the i^{th} family. Then, we can use the function $f(\eta) = (n_1, \ldots, n_k)$, where n_i denotes the number of components of family i that are failed when the system is in the state η. In that case, the reduced state space $\mathcal{F}$ is equal to:

$$\mathcal{F} = \Big\{ (n_1, \ldots, n_k) \Big| \ 0 \le n_i \le N_i, \ \forall \ 1 \le i \le k \Big\}$$

Note that for this example, the respective sizes of $\mathcal{E}$ and $\mathcal{F}$ can be roughly estimated to:

$$|\mathcal{E}| \simeq 2^N \ , \ |\mathcal{F}| \simeq \prod_{i=1}^{k} (N_i + 1)$$

If $N = 10$, $k = 5$, and $N_i = 2$ for all i, then $|\mathcal{E}| \simeq 1024$ whereas $|\mathcal{F}| \simeq 243$.

3 Upper bound of the asymptotic unavailability

Let us denote $\mathcal{P}$ the subset of failure states of $(S_t)_{t \ge 0}$, and let η_0 be a recurrent state of the process. General theorems of renewal theory ensure that the asymptotic unavailability is equal to the ratio of the mean failure time between two passages in η_0 (MFT) over the mean duration of a cycle (MDC).

$$\overline{A}_S(\infty) = \lim_{t \to +\infty} \mathbb{P}(S_t \in \mathcal{P}) = \frac{\mathbb{E}\Big(\int_0^{\tau_0} \mathbf{1}_{\{S_t \in \mathcal{P}\}} \, dt \ \Big| \ S_0 = \eta_0 \Big)}{\mathbb{E}\big(\tau_0 \ \big| \ S_0 = \eta_0 \big)} = \frac{\text{MFT}}{\text{MDC}}$$

where $\mathbf{1}$ denotes the indicator function, and τ_0 denotes the first time $(S_t)_{t \ge 0}$ returns in the state η_0. To calculate an upper bound of $\overline{A}_S(\infty)$, we will respectively give an upper and a lower bound of MFT and MDC that depend on the following quantities:

$$\text{MFT}' = \mathbb{E}\Big(\int_0^{\tau_0'} \mathbf{1}_{\{S_t' \in \mathcal{P}'\}} \, dt \ \Big| \ S_0' = i_0 \Big) \ , \ \text{MDC}' = \mathbb{E}\big(\tau_0' \ \big| \ S_0' = i_0 \big)$$

where $\mathcal{P}'$ is the subset of failure states of $(S_t')_{t \ge 0}$, i_0 is a recurrent state of the process, and τ_0' is the first time $(S_t')_{t \ge 0}$ returns in i_0.

3.1 Mean failure time over a cycle

Let $\mathcal{A}$ be the infinitesimal generator (also called transition rate matrix) of the Markov process $(S_t)_{t \ge 0}$. To compare MFT and MFT$'$, we use the following lemma, which can be easily proved using the martingale characterization of $\mathcal{A}$.

Lemma 1 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two separate subsets of $\mathcal{E}$, and let V be a positive, bounded function defined on $\mathcal{E}$. Suppose that there exists a constant $\Delta > 0$ such that for all $\eta \notin \mathcal{C}_1$, $\mathcal{A}V(\eta) \le -\Delta \mathbf{1}_{\{\eta \in \mathcal{C}_2\}}$. Then for all $\eta \notin \mathcal{C}_1$ we have:*

$$\mathbb{E}\Big(\int_0^{\tau_{\mathcal{C}_1}} \mathbf{1}_{\{S_t \in \mathcal{C}_2\}} \, dt \ \Big| \ S_0 = \eta \Big) \le \frac{V(\eta)}{\Delta}$$

where $\tau_{\mathcal{C}_1}$ denotes the first time $(S_t)_{t \ge 0}$ enters $\mathcal{C}_1$.

The idea is then to consider $\mathcal{C}_1 = \{\eta_0\}$, $\mathcal{C}_2 = \mathcal{P}$ and the following function defined on $\mathcal{F}$:

$$V_1(i) = \mathbb{E}\Big(\int_0^{\tau_0'} \mathbf{1}_{\{S_t' \in \mathcal{P}'\}} \, dt \ \Big| \ S_0 = i \Big)$$

If we introduce a partial ordering relation $\preceq$ on the space $\mathcal{F}$, we find explicit conditions on the processes $(S_t)_{t \ge 0}$ and $(S_t')_{t \ge 0}$ so that the lemma applies with $V = V_1 \circ f$.

(H1) The state i_0 is the unique solution to $f(\eta_0) = i$. Moreover, i_0 is the smallest element of $\mathcal{F}$ for the partial ordering relation $\preceq$ (for all i in $\mathcal{F}$, $i_0 \preceq i$).

(H2) Let $\eta \in \mathcal{E}_i$ and $\xi \in \mathcal{E}_j$. If i and j cannot be compared via the partial ordering relation $\preceq$, then $\mathcal{A}(\eta, \xi) = 0$.

(H3) The function V_1 is increasing with respect to the partial ordering relation $\preceq$ (if $i \preceq j$, then we have $V_1(i) \leq V_1(j)$).

(H4) The subsets $\mathcal{P}$ and $\mathcal{P}'$ satisfy $f(\mathcal{P}) \subset \mathcal{P}'$.

(H5) Let Q and Q' denote the transitions matrices of the embedded Markov chains of $(S_t)_{t\geq 0}$ and $(S'_t)_{t\geq 0}$, and let η be a state of $\mathcal{E}$. For all η in $\mathcal{E}$ and i in $\mathcal{F}$ such that $i \preceq f(\eta)$, then $Q\big(\eta, \mathcal{E}_i\big) \geq Q'\big(f(\eta), i\big)$, where $Q\big(\eta, \mathcal{E}_i\big)$ is shorthand for:

$$Q\big(\eta, \mathcal{E}_i\big) = \sum_{\xi \in \mathcal{E}_i} Q(\eta, \xi)$$

Moreover, for all i in $\mathcal{F}$ such that $i \succ f(\eta)$, then $Q\big(\eta, \mathcal{E}_i\big) \leq Q'\big(f(\eta), i\big)$.

Assuming that these conditions are checked, we get $\mathrm{MFT} \leq \mathrm{MFT}'/\Delta$, where Δ_1 is a constant depending on the transition rates of $(S_t)_{t\geq 0}$ and $(S'_t)_{t\geq 0}$. Roughly speaking, Δ_1 indicates the degree of aproximation in the aggregation: it is equal to 1 in a strong lumpability situation, and is far from 1 when there is a big difference between the states that compose a subset $\mathcal{E}_i$.

If we come back to the example given in the previous section (k different families of components), a natural partial ordering relation may be defined as follows.

$$(n_1, \ldots, n_k) \preceq (m_1, \ldots, m_k) \Leftrightarrow n_i \leq m_i \ \forall i \in \{1, \ldots, k\}$$

In that case, conditions **(H1)** and **(H2)** are fulfilled if we consider η_0 equal to the perfect working state (all the components are in the state 1). Hypothesis **(H3)** is also natural: the higher the number of failed components in each family, the worst the state of the system. Finally, **(H4)** and **(H5)** simply mean that giving an upper bound of the mean failure time over a cycle requires to be pessimistic in the approximate aggregation.

3.2 Mean duration of a cycle

The method used in the previous paragraph can be adapted to get a lower bound of MDC. But unfortunately, this leads to a condition **(H5)'** that is exactly opposite to **(H5)**. So, we have two possibilities. The first one is to build a third process $(S''_t)_{t\geq 0}$ that satisfies this new assumption, which yields $\mathrm{MDC} \geq \mathrm{MDC}''/\Delta_2$, where Δ_2 is the "degree of approximation" of this new aggregation. Finally:

$$\overline{A}_S(\infty) \leq \frac{\Delta_2}{\Delta_1} \frac{\mathrm{MFT}'}{\mathrm{MDC}''}$$

On the other hand, we can give a relation between MDC and MDC' using a more direct method. First, remark that:

$$\mathrm{MDC} \geq 1/\big|\mathcal{A}(\eta_0, \eta_0)\big|$$

This may seem to be a rough lower bound, since this inequality means that we only consider the mean duration before $(S_t)_{t\geq 0}$ leaves η_0. But if η_0 is the perfect working state, and if the system studied is highly available ($\lambda \ll \mu$ for all the components), this bound remains close to the actual duration of the cycle. Now, suppose that the following condition is checked:

(H6) Let $\mathcal{A}'$ be the infinitesimal generator of $(S'_t)_{t\geq 0}$. Then $\big|\mathcal{A}(\eta_0, \eta_0)\big| = \big|\mathcal{A}'(i_0, i_0)\big|$.

Hence, $\mathrm{MDC} \geq \mathrm{MDC}'/\Delta_3$, with Δ_3 equal to the ratio of MDC' over $1/\big|\mathcal{A}'(i_0, i_0)\big|$. This time, Δ_3 measures the "degree of availability" of the process $(S'_t)_{t\geq 0}$. Intuitively, Δ_3 tends to 1 when λ/μ tends to 0 for all the components. Finally:

$$\overline{A}_S(\infty) \leq \frac{\Delta_3}{\Delta_1} \frac{\mathrm{MFT}'}{\mathrm{MDC}'} = \frac{\Delta_3}{\Delta_1} \overline{A}_{S'}(\infty)$$

4 Unavailability of some ageing systems

From now on, $(S_t)_{t\geq 0}$ is not supposed to be a homogeneous Markov process, which allows us to consider ageing systems. More precisely, the transion rates at time t are supposed to depend on continuous parameters $C_t \in \mathbb{R}^d$ that typically indicate the degree of ageing of the components.

$$\mathbb{P}\big(S_{t+\Delta_t} = \xi \big| S_t = \eta, C_t = x\big) = a(x,\eta,\xi)\Delta_t + o(\Delta_t)$$

Our aim is now to prove that the distribution of S_t, say $\mathbf{P}_t$, can be efficiently approximated by the stationary distribution of an homogeneous Markov process. Note that the case $C_t = t$ corresponds to a non-homogeneous Markov process, for which Massey & Whitt (1998) have already studied this approximation.

First, we suppose that $(S_t, C_t)_{t\geq 0}$ is an homogeneous Markov process. It follows from the martingale characterization of its infinitesimal generator that $\mathbf{P}_t$ is solution to $\mathbf{P}'_t = \mathbf{P}_t\,\mathcal{B}_t$, where $\mathbf{P}_t$ is written as a row vector, and $\mathcal{B}_t$ denotes the "mean transion rate matrix" at time t.

$$\mathbf{P}_t(\eta) = \mathbb{P}\big(S_t = \eta\big)\ ,\ \mathcal{B}_t(\eta,\xi) = \mathbb{E}\big(a(C_t,\eta,\xi)\big|S_t = \eta\big)$$

The idea is then to compare the evolution of $(S_t)_{t\geq 0}$ to that of a convenient homogeneous Markov process over a period $[t-\Delta, t]$, assuming that both processes starts from the same state at time $t-\Delta$. Intuitively, if the homogeneous process reaches its stationary distribution before Δ units of time, and if the behaviour of the two processes is not too different over this period, then $\mathbf{P}_t$ is close to the stationary distribution. The following proposition formalizes this idea, using the theory of perturbated differential equations, and the spectral gap of $\mathcal{B}_t$ to control the convergence rate to the stationary distribution.

Proposition 1 *Let $|\lambda_{\mathcal{B}_t}|$ denote the spectral gap of $\mathcal{B}_t$, that is to say the magnitude of the second eigenvalue of the matrix. Then, for all $t > 0$ and $\Delta < t$, there exists a constant c such that:*

$$\| \mathbf{P}_t - \pi_{\mathcal{B}_t} \|_2 \leq\ c\exp\Big(-\big|\lambda_{\mathcal{B}_t}\big|\Delta\Big) + \left(\sup_{s\in[t-\Delta,t]} \frac{\| \mathcal{B}_s - \mathcal{B}_t \|_1}{\| \mathcal{B}_t \|_1}\right)\left(\exp\big(\Delta \| \mathcal{B}_t \|_1\big) - 1\right)$$

where $\pi_{\mathcal{B}_t}$ denotes the stationary distribution of the homogeneous Markov process with generator $\mathcal{B}_t$, $\| v \|_2$ denotes the euclidian norm of the row vecor v, and $\| M \|_1$ denotes the matrix norm induced by the ℓ^1-norm on row vectors.

The spectral gap of matrix is not really an intuitive quantity, but empirically, the convergence to the stationary distribution often proves to be fast for highly available systems. As a consequence, it seems that $\mathbf{P}_t$ is close to $\pi_{\mathcal{B}_t}$ if the mean transition matrix is varying slowly. In other words, this approximation is valid for slowly ageing systems.

References

Bon, J. L., Collet, J. (1994). An algorithm in order to implement reliability exponential approximations. *Rel. Eng. Sys. Saf. 43*, 263–268.

Bouissou, M., Lefebvre, Y. (2002). A path-based algorithm to evaluate asymptotic unavailability for large Markov systems. *Proc. 48th RAMS conference*.

Doisy, M. (2000). A coupling technique for stochastic comparison of functions of Markov processes. *J. Appl. Math. Decis. Sci. 4*(1), 39–64.

Mahévas, S., Rubino,G. (2001). Bound computation of dependability and performance measures. *IEEE Trans. Comp. 50*(5), 399–413.

Massey, W. A., Whitt, W. (1987). Uniform acceleration expansions for Markov chains with time-varying rates. *Ann. Appl. Prob. 8*(4).

Stoyan, D. (1983). *Comparison methods for queues and others stochastic models.* New-York: John Wiley & Sons.