LONG-RUN AVAILABILITY OF A PRIORITY SYSTEM:
A NUMERICAL APPROACH

EDMOND J. VANDERPERRE AND STANISLAV S. MAKHANOV

Received 29 June 2004

We consider a two-unit cold standby system attended by two repairmen and subjected to a priority rule. In order to describe the random behavior of the twin system, we employ a stochastic process endowed with state probability functions satisfying coupled Hokstad-type differential equations. An explicit evaluation of the exact solution is in general quite intricate. Therefore, we propose a numerical solution of the equations. Finally, particular but important repair time distributions are involved to analyze the long-run availability of the \( T \)-system. Numerical results are illustrated by adequate computer-plotted graphs.

1. Introduction

Standby provides a powerful tool to enhance the reliability, availability, quality, and safety of operational plants, see, for example, [3, 7, 14]. However, in practice, standby systems are often subjected to an appropriate priority rule. For instance, the external power supply station of a technical plant has usually overall priority in operation with regard to an internal (local) power generator kept in cold or warm standby [3]. The local generator is only deployed if the external power station is down.

Cold or warm standby systems subjected to a priority rule and attended by a repair facility have received considerable attention in the current literature, see, for example, [1, 2, 4, 5, 8, 9, 10, 11, 12, 13, 16, 17, 18, 19, 20, 21].

As a variant, we consider a twin system composed of a priority unit (the \( p \)-unit) and a nonpriority unit (the \( n \)-unit) kept in cold standby. The \( p \)-unit has overall (break-in) priority in operation with regard to the \( n \)-unit, that is, the \( n \)-unit is only used when the \( p \)-unit is down. In order to avoid undesirable delays in repairing failed units, we suppose that the entire system (henceforth called the \( T \)-system) is attended by two different repairmen. The \( T \)-system satisfies the usual conditions, that is, independent identically distributed random variables, instantaneous and perfect switch [3], and perfect repair [6]. Each repairman has his own particular task. Repairman \( N \) is skilled in repairing the \( n \)-unit, whereas repairman \( P \) is an expert in repairing the \( p \)-unit. Both repairmen are jointly busy, if and only if, both units (\( p \)-unit and \( n \)-unit) are down. In any other case, at least one repairman is idle.
Numerical availability of a priority system

In order to describe the random behavior of the T-system, we employ a stochastic process endowed with transition probability functions satisfying steady-state Høkstad-type differential equations. Unfortunately, the exact solution procedure is quite intricate (see, [21, page 359] and Remark 4.1). Therefore, we propose a numerical solution of the equations.

Finally, current repair time distributions (such as the Weibull-Gnedenko distribution) are involved to compute the long-run availability of the T-system. The results are illustrated by adequate computer-plotted graphs.

2. Formulation

Consider a T-system satisfying the usual conditions. The p-unit has a constant failure rate \( \lambda > 0 \) and a general repair time distribution \( R(\cdot), R(0) = 0 \), with mean \( \rho \). The operative n-unit has a constant failure rate \( \lambda_s > 0 \), but a zero failure rate in standby (the so-called cold standby state) and a general repair time distribution \( R_S(\cdot), R_S(0) = 0 \), with mean \( \rho_s \).

In order to describe the random behavior of the T-system, we introduce a stochastic process \( \{N_t, t \geq 0\} \) with arbitrary discrete state space \( \{A, B, C, D\} \subseteq [0, \infty) \), characterized by the following mutually exclusive events:

(i) \( \{N_t = A\} \): “the p-unit is operative and the n-unit is in cold standby at time \( t \),”
(ii) \( \{N_t = B\} \): “the n-unit is operative and the p-unit is under repair at time \( t \),”
(iii) \( \{N_t = C\} \): “the p-unit is operative and the n-unit is under repair at time \( t \),”
(iv) \( \{N_t = D\} \): “both units are simultaneously down at time \( t \).”

State \( D \) is called the system-down state.

Figures 2.1, 2.2, 2.3, and 2.4 display a functional block diagram of the T-system operating in states A, B, C, and D.

Observe that the process \( \{N_t, t \geq 0\} \) is non-Markovian. A Markov characterization of the process is piecewise and conditionally defined by:

(i) \( \{N_t\} \), if \( N_t = A \) (i.e., if the event \( \{N_t = A\} \) occurs),
(ii) \( \{(N_t, X_t)\} \), if \( N_t = B \), where \( X_t \) denotes the remaining repair time of the p-unit under progressive repair at time \( t \),
(iii) \( \{(N_t, Y_t)\} \), if \( N_t = C \), where \( Y_t \) denotes the remaining repair time of the n-unit under progressive repair at time \( t \),
(iv) \( \{(N_t, X_t, Y_t)\} \), if \( N_t = D \).

The state space of the underlying piecewise linear (vector) Markov process is given by

\[
A \cup \{(B, x); x \geq 0\} \cup \{(C, y); y \geq 0\} \cup \{(D, x, y); x \geq 0; y \geq 0\}. \tag{2.1}
\]

Next, we consider the T-system in stationary state (the so-called ergodic state) with invariant measure \( \{p_K; K = A, B, C, D\}, \sum_K p_K = 1 \), where

\[
p_K := \lim_{t \to \infty} P\{N_t = K | N_0 = A\}. \tag{2.2}
\]
Figure 2.1. Functional block diagram of the T-system operating in state A.

Figure 2.2. Functional block diagram of the T-system operating in state B.

Figure 2.3. Functional block diagram of the T-system operating in state C.

Figure 2.4. Functional block diagram of the T-system in state D.
It can be demonstrated that the invariant measure exists for arbitrary \( R \) and \( R_S \) with finite mean. However, in order to keep the analysis as simple as possible, we henceforth assume that \( R \) and \( R_S \) have bounded densities on \([0, \infty)\), denoted by \( r \) and \( r_S \). Finally, we introduce the measures

\[
p_B(x)dx := \lim_{t \to \infty} P\{N_t = B, X_t \in (x, x + dx]|N_0 = A\},
\]

\[
p_C(y)dy := \lim_{t \to \infty} P\{N_t = C, Y_t \in (y, y + dy]|N_0 = A\},
\]

\[
p_D(x,y)dx\,dy := \lim_{t \to \infty} P\{N_t = D, X_t \in (x, x + dx], Y_t \in (y, y + dy]|N_0 = A\}.
\]

Note that, for instance, \( p_D = \int_0^\infty \int_0^\infty p_D(x,y)dx\,dy \).

### 3. Long-run availability

We recall that the \( T \)-system is only available (functioning) in states \( A, B, \) and \( C \). Therefore, the long-run availability of the operational plant, denoted by \( \mathcal{A} \), is given by \( \mathcal{A} = 1 - p_D \). Invoking the substitutions \( p_B(\cdot) = p_A \varphi_B(\cdot), \ p_C(\cdot) = p_A \varphi_C(\cdot), \ p_D(\cdot, \cdot) = p_A \varphi_D(\cdot, \cdot) \) and the law \( \sum_K p_K = 1 \) entails that \( p_A = 1/(1 + \Phi_B + \Phi_C + \Phi_D) \), where \( \Phi_B := \int_0^\infty \varphi_B(x)dx \), \( \Phi_C := \int_0^\infty \varphi_C(y)dy \) and \( \Phi_D := \int_0^\infty \int_0^\infty \varphi_D(x,y)dx\,dy \). Hence,

\[
\mathcal{A} = \frac{1 + \Phi_B + \Phi_C}{1 + \Phi_B + \Phi_C + \Phi_D}.
\]

### 4. Differential equations

In order to determine the \( \varphi \)-functions, we first construct a system of coupled steady state-type differential equations based on a time-independent version of Hokstad’s supplementary variable technique (see, e.g., [22, page 526] for further details). For \( x > 0, \ y > 0 \), we obtain

\[
\lambda = \varphi_B(0) + \varphi_C(0),
\]

\[
(\lambda_s - \frac{d}{dx}) \varphi_B(x) = \varphi_D(x,0) + \lambda r(x),
\]

\[
(\lambda - \frac{d}{dy}) \varphi_C(y) = \varphi_D(0,y),
\]

\[
\left(- \frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) \varphi_D(x,y) = \lambda_s \varphi_B(x)r_s(y) + \lambda \varphi_C(y)r(x).
\]

**Remark 4.1.** A particular but important family \( \mathcal{F} \) of current repair time distributions with nonrational characteristic functions, such as the Weibull-Gnedenko and Lognormal distributions, are fairly suitable to model repair times. Unfortunately, if both \( R \) and \( R_S \) belong to \( \mathcal{F} \), an explicit evaluation of the exact solution of (4.1), (4.2), (4.3), and (4.4) in terms of finite linear combinations of known algebraic and/or transcendental functions is as good as excluded (see [21, page 361] for further details). Therefore, we propose a numerical solution of the equations.
5. Numerical scheme

In order to construct an appropriate numerical procedure, we first remark that the \( \phi \)-functions are vanishing at infinity *irrespective* of the asymptotic behavior of the repair time density functions! Therefore, a numerical procedure to solve the equations in the region \((0, \infty) \times (0, \infty)\) may be converted into a numerical solution procedure in the *truncated* region \((0, L) \times (0, L)\), for some \(L > 0\), with prescribed boundary conditions \( \varphi_B(L) = \varphi_C(L) = \varphi_D(L, \cdot) = \varphi_D(\cdot, L) = 0 \). Let \( \varphi_{B,i} := \varphi_B(x_i), \varphi_{C,j} := \varphi_C(y_j), \varphi_{D,i,j} := \varphi_D(x_i, y_j) \), where \(x_i := i\Delta, y_j := j\Delta, i = 0, \ldots, N + 1; j = 0, \ldots, N + 1; \Delta := L/N\). We propose the following numerical scheme. Let \(k\) be the iteration number. Given \( \varphi_{D,i,N+1}^{k+1} = 0, \varphi_{D,N+1,j}^{k+1} = 0, \varphi_{B,N+1}^{k+1} = 0, \varphi_{C,N+1}^{k+1} = 0, \) and the values of \( \varphi_{B,i}^k \) and \( \varphi_{C,j}^k \), we compute \( \varphi_{D,i,j}^{k+1} \) by means of the two-point first-order approximation of (4.4), namely,

\[
\varphi_{D,i,j}^{k+1} = \frac{1}{2} (\varphi_{D,i,j+1}^k + \varphi_{D,i+1,j}^k) + \Delta (\lambda_s \varphi_{B,i}^k \delta_{j} + \lambda \varphi_{C,j}^k \gamma_{i}), \quad (5.1)
\]

\(i = N, N - 1, \ldots, 0\) and \(j = N, N - 1, \ldots, 0\).

Next, we calculate \( \varphi_{B,i}^{k+1} \) and \( \varphi_{C,j}^{k+1} \) by means of the first-order approximations of (4.2) and (4.3) given by

\[
\varphi_{B,i}^{k+1} = \frac{1}{\gamma_B} (\frac{\varphi_{B,i+1}^k}{\Delta} + \varphi_{D,i,0}^k + \lambda \gamma \delta_{i}), \quad (5.2)
\]

\[
\varphi_{C,j}^{k+1} = \frac{1}{\gamma_C} (\frac{\varphi_{C,j+1}^k}{\Delta} + \varphi_{D,0,j}^k),
\]

where \( \gamma_B := \lambda_s + 1/\Delta \) and \( \gamma_C := \lambda + 1/\Delta \). Finally, in order to satisfy (4.1) we use the normalizing procedure

\[
\varphi_{C,j}^{k+1,\text{new}} = \lambda \frac{\varphi_{C,j}^{k+1}}{\varphi_{C,0}^{k+1} + \varphi_{B,0}^{k+1}}, \quad (5.3)
\]

\[
\varphi_{B,i}^{k+1,\text{new}} = \lambda \frac{\varphi_{B,i}^{k+1}}{\varphi_{B,i}^{k+1} + \varphi_{D,i}^{k+1}}.
\]

**Remarks 5.1.** Let \( \varphi_\Delta \) denote a numerical solution obtained with the space-step \( \Delta \). The relevant numerical error is then evaluated on a nested grid by \( \varepsilon := |\varphi_\Delta - \varphi_{\Delta/2}| \). However, such an estimate is only accurate if \(L\) is large enough. Roughly speaking, if \( \max(r(x), r(y)) \) at \( x = L \) is small, then (most likely) this particular \( L \) is appropriate. However, such a “brutal force” approach may require a large number of grid points and is therefore rarely applicable. We illustrate the phenomenon by comparing the exact and the numerical solution in the most simple case, that is, let \( R(x) = 1 - e^{-x}, R_S(y) = 1 - e^{-y} \). Then, \( \varphi_D(x, y) = l_D e^{-r(x+y)}, \varphi_C(y) = l_C e^{-y}, \varphi_B(x) = l_B e^{-x} \), where \( l_D := \lambda_s (\lambda + 1) / (\lambda_s + \lambda + 2), l_C := \lambda_s \lambda / (\lambda_s + \lambda + 2), l_B := \lambda / (\lambda + 2) / (\lambda_s + \lambda + 2) \).

Figure 5.1 shows the numerical error

\[
\varepsilon_M := \max \{ |\varphi_D^{\text{exact}} - \varphi_D|, |\varphi_C^{\text{exact}} - \varphi_C|, |\varphi_B^{\text{exact}} - \varphi_B| \}
\]

versus the grid size for various \( L \).
Observe that, if $L$ is not large enough, $\varepsilon_M$ does not decrease as $\Delta$ decreases (see Figure 5.1). On the other hand, too large $L$ (consequently, too large $\Delta$) lead to large numerical errors. For instance, the error with $L = 30$ is larger than $2.5 \cdot 10^{-2}$ for any $N \in [20, 100]$, whereas the error with $L = 4$ is less than $2.5 \cdot 10^{-2}$. There could be multiple options too. For instance, an error less than $2.5 \cdot 10^{-2}$ is achieved either with $L = 4$, $N = 15$, or $L = 6$, $N = 22$, or $L = 10$, $N = 38$.  

Figure 5.2 shows a two-dimensional spatial distribution of the error $\varepsilon_D := |\varphi_D^{\text{exact}} - \varphi_D|$ for various $L$. Clearly, $\varepsilon_D$ could be increasing near the origin as $L$ increases. However, the error decreases for large $x$ and $y$.  

---

**Figure 5.1.** The horizontal axis denotes the logarithm of the numerical error, the vertical axis denotes the number of the grid points, (1) $L = 0.4$; (2) $L = 1.0$; (3) $L = 1.5$; (4) $L = 2$; (5) $L = 4$; (6) $L = 6$; (7) $L = 10$; (8) $L = 50$.  

**Figure 5.2.** Spatial distribution of $\varepsilon_D$, (1) $L = 1.5$; (2) $L = 3$; (3) $L = 6$.  

---
6. Trial-and-error procedure

The complicated behavior of the numerical error requires an adaptive choice of $\Delta$ and $L$. Therefore, we introduce the subordinate errors $\varepsilon_1 := |\varphi_{\Delta L} - \varphi_{\Delta/2,L}|$ and $\varepsilon_2 := |\varphi_{\Delta L} - \varphi_{\Delta/2,L}|$, where $\varepsilon_1$ characterizes the numerical error caused by truncation of the infinite region and $\varepsilon_2$ the numerical error related to the first-order approximants. In order to find the optimal pair $(L, \Delta)$, we first specify the required accuracy $\varepsilon$. Next, we propose the following trial-and-error procedure: we vary $L$ until $\varepsilon_1 < \varepsilon$ and then $\Delta$ until $\varepsilon_2 < \varepsilon$. Finally, we introduce the following.

7. Application. The Weibull-Gnedenko distribution

We consider the particular but important case of Weibull-Gnedenko repair time distributions, that is, let $R(x) = 1 - e^{-x^{\beta_1}}$, $R_S(y) = 1 - e^{-y^{\beta_2}}$. Obviously, the optimal pair $(L, \Delta)$ depends on $\lambda, \lambda_s, \beta_1, and \beta_2$. We demonstrate the trial-and-error procedure applied to the particular case $\lambda = 1, \lambda_s = 0.1, \beta_1 = 2, \beta_2 = 3$. However, no restrictions are imposed on the analysis of $\mathcal{A}$ for arbitrary values of $\lambda, \lambda_s, \beta_1$ and $\beta_2$. Let $\varepsilon = 10^{-3}$.

First, we vary $L$, as shown in Table 7.1, until $\varepsilon_1 < \varepsilon$. Next, we vary $\Delta$, as shown in Table 7.2, until $\varepsilon_2 < \varepsilon$. A spatial distribution of $\varepsilon_1$ and $\varepsilon_2$ is depicted in Figures 7.1 and 7.2.

<table>
<thead>
<tr>
<th>Table 7.1. The $L$ trials.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7.2. The $\Delta$ trials.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
</tbody>
</table>

Figure 7.1. Spatial distribution of $\varepsilon_1$, (1) $L = 3$; (2) $L = 6$. 
Numerical availability of a priority system

Figure 7.2. Spatial distribution of $\varepsilon_2$ for $N = 320$.

Figure 7.3. Numerically generated: (1) $p_B(x)/1.5$, (2) $p_C(x)$, (3) $p_D(x, y)$, $\lambda_s = 0.3$ $\lambda = 1.0$. Note that $p_B$ is divided by 1.5 due to scaling.

Figure 7.4. Numerically generated: (1) $p_B(x)/1.5$, (2) $p_C(x)$, (3) $p_D(x, y)$, $\lambda_s = 0.7$ $\lambda = 1.0$. Note that $p_B$ is divided by 1.5 due to scaling.

Figure 7.3 displays $p_B(\cdot)$, $p_C(\cdot)$, and $p_D(\cdot, \cdot)$ for $\lambda = 1$, $\lambda_s = 0.3$ and Figure 7.4 for $\lambda = 1$, $\lambda_s = 0.5$. Figure 7.5 shows $p_D(x, y)$ for various $\lambda_s$. Let $\lambda_{\beta_1, \beta_2}(\lambda, \lambda_s)$ denote the long-run availability as a function of $\lambda$ and $\lambda_s$. 
Figure 7.5. Numerically generated $p_D(x, y), \lambda = 1.0, (1) \lambda_s = 0.1, (2) \lambda_s = 0.3, (3) \lambda_s = 0.7$.

Figure 7.6. Numerically generated long-run availability.

Figure 7.6 shows that the long-run availability exhibits a nonlinear behavior for sufficiently large $\lambda$ and $\lambda_s$ (see also Table 7.3). Finally, Figure 7.7 displays the deviations $d_1 := |A_{2,2} - A_{2,4}|$, $d_2 := |A_{2,2} - A_{4,2}|$, $d_3 := |A_{2,2} - A_{4,4}|$. The plot reveals that $A$ is fairly insensitive for $\beta$-variations.
8. Conclusion

An effective statistical analysis of the T-system requires the solution of coupled Hokstad-type differential equations. Our numerical solution procedure, endowed with a simple and robust algorithm, allows to compute and to analyze the long-run availability for a general class of current repair time distributions with tangible engineering applications.

References


Edmond J. Vanderperre: Department of Quantitative Management, University of South Africa, P.O. Box 392, Pretoria 0003, South Africa

E-mail address: edmondvanderperre@hotmail.com

Stanislav S. Makhanov: Faculty of Information Technology, Sirindhorn International Institute of Technology, Thammasat University, P.O. Box 22, Patumthani 12121, Thailand

E-mail address: makhanov@siit.tu.ac.th