9-21-2006

Analysis of a k-out-of-N System with Spares, Repairs, and a Probabilistic Rule

Srinivas R. Chakravarthy
Kettering University, schakrav@kettering.edu

ANALYSIS OF A k-OUT-OF-N SYSTEM WITH SPARES, REPAIRS, AND A PROBABILISTIC RULE

SRINIVAS R. CHAKRAVARTHY

Received 16 April 2006; Revised 21 September 2006; Accepted 21 September 2006

We consider a k-out-of-N reliability system with identical components having exponential lifetimes. There is a single repairman who attends to failed components on a first-come first-served basis. The repair times are assumed to be of phase type. The system has K spares that can be used according to a probabilistic rule to extend the lifetime of the system. The system is analyzed using Markov chain theory and some interesting results are obtained. A few illustrative numerical examples are discussed.

1. Introduction and model description

The k-out-of-N reliability system is one of the most popular and widely used systems in practice. These systems have been studied extensively in the literature in the context of computing the reliability, optimizing the system, common cause failures, and with repair facility for fixing the failed components [5]. The k-out-of-N reliability systems have been studied in certain situations where redundancy is of importance. Redundancy is required not only to extend the functioning of the system but also to achieve a certain reliability of the system. Classical examples of redundancy requirement occur in aircrafts, space shuttles, nuclear plants, satellites, electric generators, design of VLSI (very large scale integrated) circuits, and computer systems. The k-out-of-N systems can be classified into (a) active redundant systems in which all N components are active even though only k components are required for the proper functioning of the system; (b) cold standby systems in which $N-k$ components will not be active and upon failure of one of the k active components, cold standby component will instantaneously replace the failed component; (c) warm standby systems in which $N-k$ components will have a smaller failure rate compared to the active ones; (d) hot standby systems in which $N-k$ components will have a higher failure rate as compared to the k active ones. Furthermore, the k-out-of-N systems

Hindawi Publishing Corporation
Journal of Applied Mathematics and Stochastic Analysis
Volume 2006, Article ID 39093, Pages 1–23
DOI 10.1155/JAMSA/2006/39093

can have a repair facility containing one or more servers who will fix the failed components. The literature on k-out-of-N systems is quite extensive. Thus, we refer the reader to [5] for a comprehensive review of the k-out-of-N systems including the setting up of several optimization problems and their solution techniques. This book includes several references to published articles on the k-out-of-N systems. Recently, the papers [1, 2] integrate the spares, the repairs, and the maintenance policy so as to prolong the life of a k-out-of-N system. While these papers deal with a maintenance policy along with spares and repairs, our focus here is mainly on the reliability system with spares and repairs.

In this paper we consider a k-out-of-N system with active redundancy. There is a reliable single repairman who attends to failed components on a first-come first-served basis. There is an inventory consisting of K spares. These units are called upon using a probabilistic rule. Suppose that the system has r, $k \le r \le N$, functioning components. When a component fails, with probability p_r, a spare unit, if one can be made available, will be used to replace the failed one which is immediately sent for repair. The motivation for introducing the probability structure, $\{p_r\}$, comes from a scenario where in spite of having spares, the immediate dispatch of them requires operators or machines to be available at that instant; or a switch that is used for instantaneous transfer of the unit may malfunction with probability p_r. Hence, in these cases the availability of the spares is described only through some probability structure. Note that this system will include cold standby system by taking $p_r = 1$, $k \le r \le N$. By taking $p_r = 0$, for $k < r \le N$ and $p_k = 1$ we can model situations where the spare units are used only at a time when the system failure can be avoided. Here we analyze the k-out-of-N system with the above-mentioned probability structure and establish some interesting results that we believe are not noticed in the literature.

Before we list the basic assumptions of the model, we set up some notations. By $\mathbf{e}$ we will denote a column vector (of appropriate dimension) of 1's; $\mathbf{e}_i$ will denote a unit column vector (of appropriate dimension) with 1 in the ith position and 0 elsewhere; and I an identity matrix (of appropriate dimension). We will display the dimension should there be a need to emphasize it. The notation "$'$" will stand for the transpose of a matrix and the symbol $\otimes$ denotes the Kronecker product of matrices. For details and properties on Kronecker products we refer the reader to [4].

Model description. (i) The system has N components and requires at least k of these components to function.

(ii) The components work independently of each other and each component is assumed to have a lifetime that is exponentially distributed with parameter λ.

(iii) There are K spares in the system. These spares are used as follows. When the system is functioning with i, $k \le i \le N$, components, a spare component will be used upon a failure of component with probability p_i. With probability $q_i = 1 - p_i$, a spare component will not be used when a failure occurs. We will take $p_k = 1$ so that the system will not fail when at least one spare component is available. Note that this explicitly assumes that the transfer of the spare unit will occur without any problem to prevent the system from failing when a spare is in the inventory. However, the current model can easily be modified to consider the case when $p_k < 1$. The details are outlined in a separate section.

(iv) There is a single repairman who will attend to failed components on a first-come first-served basis. The repair times are assumed to follow a (continuous) phase-type distribution (PH-distribution) with representation $(\boldsymbol{\beta}, S)$ of order m. The mean service rate, μ, is given by $\mu = [\boldsymbol{\beta}(-S)^{-1}\mathbf{e}]^{-1}$. A continuous PH-distribution is obtained as the time until absorption in a finite-state continuous time Markov chain with one absorbing state. PH-distributions play an important role in stochastic modeling. Erlang, generalized Erlang, exponential, and hyperexponential are all special cases of PH-distributions. For details on PH-distribution we refer the reader to [6, 7].

(v) Repaired items are considered as new and are sent to the inventory of spares only if the system has N working components. Otherwise the repaired items are sent directly to the system.

(vi) The failure times and the repair times are assumed to be independent.

2. Markov process description

The reliability system outlined in Section 1 can be studied as a continuous time Markov chain. To see this, we first define $J_1(t)$ to be the number of components working at time t, $J_2(t)$ to be the number of failed components (including the one under repair) at time t, and $J_3(t)$ to be the phase of the repair process at time t, ($J_3(t)$ is not defined when the repairman is idle and will be denoted by $*$). The reliability system can be studied as a finite state nonhomogeneous quasi-birth-and-death (QBD) process with state space given by $\Omega = \{(N,0,*)\} \bigcup \{(N,r,j), 1 \le r \le K, 1 \le j \le m\} \bigcup \{(i,r,j) : k \le i \le N-1, N-i \le r \le N+K-i, 1 \le j \le m\} \bigcup \{(k-1, N+1+K-k, j) : 1 \le j \le m\}$.

Note that the state (i,r,j) corresponds to the case when the system has i working components and r components are under repair with the phase of the current repair in j.

Denote by level $\mathbf{N}$ the set of states given by $\{(N,0,*)\} \bigcup \{(N,r,j) : 1 \le r \le K, 1 \le j \le m\}$; by $\mathbf{i}$, for $k-1 \le i \le N-1$, the set of states given by $\{(i,r,j) : N-i \le r \le N+K-i, 1 \le j \le m\}$. Note that in the case when $p_k = 1$, the level $\mathbf{k}-1$ which is of dimension m takes the form $\{(k-1, N+1+K-k, j) : 1 \le j \le m\}$. In general the levels, $\mathbf{i}$, $k-1 \le i \le N$, are of dimension $(K+1)m$. The generator of the Markov chain, $\{(J_1(t), J_2(t), J_3(t)) : t \ge 0\}$, is then given by

$$Q = \begin{pmatrix} A_N & B_N & & & & & \\ C_{N-1} & A_{N-1} & B_{N-1} & & & & \\ & I \otimes \mathbf{S}^0\boldsymbol{\beta} & A_{N-2} & B_{N-2} & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & & I \otimes \mathbf{S}^0\boldsymbol{\beta} & A_k & B_k \\ & & & & & C_{k-1} & A_{k-1} \end{pmatrix}, \tag{2.1}$$

where the matrices appearing in (2.1) are defined as follows:

$$A_N = \begin{pmatrix} -N\lambda & N\lambda p_N \boldsymbol{\beta} & & & & \\ \mathbf{S}^0 & S - N\lambda I & N\lambda p_N I & & & \\ & \mathbf{S}^0\boldsymbol{\beta} & S - N\lambda I & N\lambda p_N I & & \\ & \ddots & \ddots & \ddots & & \\ & & & \mathbf{S}^0\boldsymbol{\beta} & S - N\lambda I & N\lambda p_N I \\ & & & & \mathbf{S}^0\boldsymbol{\beta} & S - N\lambda I \end{pmatrix}, \tag{2.2}$$

$$A_i = \begin{pmatrix} S - i\lambda I & i\lambda p_i I & & \\ & S - i\lambda I & i\lambda p_i I & \\ & \ddots & \ddots & \\ & & S - i\lambda I & i\lambda p_i I \\ & & & S - i\lambda I \end{pmatrix}, \quad k \le i \le N-1, \tag{2.3}$$

$$B_N = \begin{pmatrix} N\lambda q_N \boldsymbol{\beta} & & & & \\ & N\lambda q_N I & & & \\ & & \ddots & & \\ & & & N\lambda q_N I & \\ & & & & N\lambda I \end{pmatrix}, \tag{2.4}$$

$$B_i = \begin{pmatrix} i\lambda q_i I & & & \\ & \ddots & & \\ & & i\lambda q_i I & \\ & & & i\lambda I \end{pmatrix}, \quad k \le i \le N-1, \tag{2.5}$$

$$C_{N-1} = \begin{pmatrix} \mathbf{S}^0 & \\ & I \otimes \mathbf{S}^0\boldsymbol{\beta} \end{pmatrix}, \tag{2.6}$$

and the matrices C_{k-1} and A_{k-1} are defined as

$$\begin{aligned} C_{k-1} &= \begin{cases} \mathbf{e}'_K(K+1) \otimes \mathbf{S}^0\boldsymbol{\beta}, & p_k = 1, \\ I \otimes \mathbf{S}^0\boldsymbol{\beta}, & p_k < 1, \end{cases} \\ A_{k-1} &= \begin{cases} S, & p_k = 1, \\ I \otimes S, & p_k < 1. \end{cases} \end{aligned} \tag{2.7}$$

However, when $p_k = 1$, B_k is given by

$$B_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ k\lambda I \end{pmatrix}. \tag{2.8}$$

3. Steady-state analysis

In this section we perform a steady-state analysis of the k-out-of-N reliability system with K spares including the derivation of the probability density functions of (a) the time until the failure of the system; (b) the time between failures of the system; (c) the idle time of the repairman; (d) the busy period of the repairman; and (e) the analysis of a series (k-out-of-k) system. We also list some selected key system performance measures.

3.1. Steady-state probability vector. Let $\mathbf{x}$, partitioned as $\mathbf{x} = ((\mathbf{x}(N)),\dots,\mathbf{x}(k),\mathbf{y}(k-1))$, denote the steady-state probability vector of Q. That is, $\mathbf{x}$ satisfies

$$\mathbf{x}Q = 0, \qquad \mathbf{xe} = 1. \tag{3.1}$$

We further partition $\mathbf{x}(i)$, $k \le i \le N-1$, as $\mathbf{x}(i) = (\mathbf{y}_0(i),\dots,\mathbf{y}_K(i))$, and $\mathbf{x}(N) = (y_0(N), \mathbf{y}_1(N),\dots,\mathbf{y}_K(N))$. Note that $y_0(N)$ is a scalar and $\mathbf{y}_j(i)$'s are vectors of dimension m. The system of equations given in (3.1) can be solved exploiting the special structure of the coefficient matrices. To illustrate this, the steady-state equations satisfying (3.1) can be written as follows:

$$\mathbf{y}_1(N)\mathbf{S}^0 + \mathbf{y}_0(N-1)\mathbf{S}^0 - N\lambda y_0(N) = 0, \tag{3.2}$$

$$N\lambda p_N y_0(N)\boldsymbol{\beta} + \mathbf{y}_2(N)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_1(N-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_1(N)(S - N\lambda I) = 0, \tag{3.3}$$

$$N\lambda p_N \mathbf{y}_{j-1}(N) + \mathbf{y}_{j+1}(N)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_j(N-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_j(N)(S - N\lambda I) = 0, \quad 2 \le j \le K-1, \tag{3.4}$$

$$N\lambda p_N \mathbf{y}_{K-1}(N) + \mathbf{y}_K(N-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_K(N)(S - N\lambda I) = 0, \tag{3.5}$$

$$N\lambda q_N y_0(N)\boldsymbol{\beta} + \mathbf{y}_0(N-2)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_0(N-1)[S - (N-1)\lambda I] = 0, \tag{3.6}$$

$$\begin{aligned} N\lambda q_N \mathbf{y}_j(N) + (N-1)\lambda p_{N-1}\mathbf{y}_{j-1}(N-1) + \mathbf{y}_j(N-2)\mathbf{S}^0\boldsymbol{\beta} \\ + \mathbf{y}_j(N-1)[S - (N-1)\lambda I] = 0, \quad 1 \le j \le K-1, \end{aligned} \tag{3.7}$$

$$N\lambda \mathbf{y}_K(N) + (N-1)\lambda p_{N-1}\mathbf{y}_{K-1}(N-1) + \mathbf{y}_K(N-2)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_K(N-1)[S - (N-1)\lambda I] = 0, \tag{3.8}$$

and for $k+1 \leq i \leq N-2$,

$$(i+1)\lambda q_{i+1}\mathbf{y}_0(i+1)+\mathbf{y}_0(i-1)\mathbf{S}^0\boldsymbol{\beta}+\mathbf{y}_0(i)[S-i\lambda I]=0, \tag{3.9}$$

$$\begin{aligned}(i+1)\lambda q_{i+1}\mathbf{y}_j(i+1)+i\lambda p_i\mathbf{y}_{j-1}(i)+\mathbf{y}_j(i-1)\mathbf{S}^0\boldsymbol{\beta}+\mathbf{y}_j(i)[S-i\lambda I]=0,&\\ 1\leq j\leq K-1,&\end{aligned} \tag{3.10}$$

$$(i+1)\lambda\mathbf{y}_K(i+1)+i\lambda p_i\mathbf{y}_{K-1}(i)+\mathbf{y}_K(i-1)\mathbf{S}^0\boldsymbol{\beta}+\mathbf{y}_K(i)[S-i\lambda I]=0, \tag{3.11}$$

$$(k+1)\lambda q_{k+1}\mathbf{y}_0(k+1)+\mathbf{y}_0(k)[S-k\lambda I]=0, \tag{3.12}$$

$$(k+1)\lambda q_{k+1}\mathbf{y}_j(k+1)+i\lambda p_k\mathbf{y}_{j-1}(k)+\mathbf{y}_j(k)[S-k\lambda I]=0, \quad 1\leq j\leq K-1, \tag{3.13}$$

$$(k+1)\lambda\mathbf{y}_K(k+1)+k\lambda p_k\mathbf{y}_{K-1}(k)+\mathbf{y}(k-1)\mathbf{S}^0\boldsymbol{\beta}+\mathbf{y}_K(i)[S-i\lambda I]=0, \tag{3.14}$$

$$k\lambda\mathbf{y}_K(k)+\mathbf{y}(k-1)S=0, \tag{3.15}$$

with the normalizing condition

$$y_0(N)+\sum_{j=1}^{K}\mathbf{y}_j(N)\mathbf{e}+\sum_{i=k}^{N-1}\sum_{j=0}^{K}\mathbf{y}_j(i)\mathbf{e}+\mathbf{y}(k-1)\mathbf{e}=1. \tag{3.16}$$

There is a variety of methods such as (block) Gauss-Seidel, aggregate/disaggregate available for solving the system equations given in (3.2)–(3.16) and one can refer to [9] for full details. An alternate method due to Gaver et al. [3] is highly suitable and efficient especially when the elements of B_i, $k+1\leq i\leq N$, are neither too small nor too large. Very briefly the algorithm using the technique outlined in [3] is Algorithm 3.1. First for notational convenience we define $C_i = I\otimes\mathbf{S}^0\boldsymbol{\beta}$, $k\leq i\leq N-2$.

ALGORITHM 3.1

Step 1. Determine D_i, $0\leq i\leq N-k+1$, recursively as follows:

$$D_0=A_N, \quad D_i=A_{N-i}+C_{N-i}\left(-D_{i-1}\right)^{-1}B_{N-i+1}, \quad 1\leq i\leq N-k+1. \tag{3.17}$$

Step 2. Solve $\mathbf{x}(k-1)D_{N-k+1}=0$, $\mathbf{x}(k-1)\mathbf{e}=1$, and recursively compute $\mathbf{x}(i)$, $k\leq i\leq N$ as

$$\mathbf{x}(i)=\mathbf{x}(i-1)C_{i-1}\left(-D_{N-i}\right)^{-1}, \quad k\leq i\leq N. \tag{3.18}$$

Step 3. Renormalize $\mathbf{x}(i)$, $k-1\leq i\leq N$, such that the normalizing condition $\mathbf{xe}=1$ is satisfied.

Note. When K or m is large, one can exploit the special structure of the matrices appearing in Q to find the inverses in the algorithm.

3.2. The time until the failure of the system. Suppose that X_{TTF} denotes the time until the first failure of the system starting with N working components and K spares. The following theorem shows that X_{TTF} is of phase type.

Theorem 3.2. *X_{TTF} follows a PH-distribution with representation $(\mathbf{e}_1', T)$ of dimension $r = m[K(N-k+1)+N-k]$, where*

$$T = \begin{pmatrix} A_N & B_N & & & & \\ C_{N-1} & A_{N-1} & B_{N-1} & & & \\ & I\otimes \mathbf{S}^0\boldsymbol{\beta} & A_{N-2} & B_{N-2} & & \\ & \ddots & \ddots & \ddots & & \\ & & & I\otimes \mathbf{S}^0\boldsymbol{\beta} & A_{k+1} & B_{k+1} \\ & & & & I\otimes \mathbf{S}^0\boldsymbol{\beta} & A_k \end{pmatrix}. \tag{3.19}$$

Proof. The proof follows immediately on noting that (a) starting the system with all N components working and K spares remaining to be used corresponds to the Markov chain with generator Q as given in (2.1), starting in state $(N,0)$; and (b) the failure of the system corresponds to the Markov chain visiting level $\mathbf{k}-1$ for the first time. Thus, the system failure time is modeled as the time until absorption in a finite-state Markov chain with an absorbing state. □

Remark 3.3. The mean, μ_{TTF}, of the time until the failure of the system is given by

$$\mu_{\mathrm{TTF}} = \mathbf{e}_1'(-T)^{-1}\mathbf{e}. \tag{3.20}$$

However, due to the large dimension of T, an efficient way to calculate the mean is done as follows (see, e.g., [7]).

Let $\boldsymbol{\delta}^{(\mathrm{TTF})}$ denote the steady-state probability vector of the irreducible generator $T+\mathbf{T}^0\mathbf{e}_1'(r)$,where $\mathbf{T}^0$ is such that $T\mathbf{e}+\mathbf{T}^0=0$.Then $\mu_{\mathrm{TTF}} = [\boldsymbol{\delta}^{(\mathrm{TTF})}\mathbf{T}^0]^{-1}$. The vector $\boldsymbol{\delta}^{(\mathrm{TTF})}$ can be solved similar to the vector $\mathbf{x}$ appearing in (3.1) and the details are omitted.

3.3. The downtime of the system. Let X_{DT} denote the duration during which the system is down. Then we have the following.

Theorem 3.4. *X_{DT} follows a PH-distribution with representation $((1/\mathbf{y}_K(k)\mathbf{e})\mathbf{y}_K(k), S)$ of dimension m.*

Proof. Observing that (a) X_{DT} is the duration that the Markov process with generator Q spends in level $(\mathbf{k}-1)$ before hitting level $\mathbf{k}$; and (b) the jth component of the vector $ck\lambda\mathbf{y}_K(k)$, where c is the normalizing constant, gives the (conditional) probability that the repair process is in phase j at the time of the system failure, the stated result follows immediately. □

Remark 3.5. (a) The mean, μ_{DT}, downtime of the system is then given by

$$\mu_{\mathrm{DT}} = \frac{1}{\mathbf{y}_K(k)\mathbf{e}}\mathbf{y}_K(k)(-S)^{-1}\mathbf{e}. \tag{3.21}$$

(b) In the case of exponential repairs, (3.21) reduces to $\mu_{\mathrm{DT}} = 1/\mu$. This is intuitively obvious due to the memoryless property of the repair times.

3.4. The time between failures of the system. Suppose that X_{TBF} denotes the time between two successive failures of the system. First note that any time after a failure of the system always becomes functional with only k working components. The random variable X_{TBF} is of the form $X_{\mathrm{TBF}} = Z + X_{\mathrm{DT}}$, where Z is the failure time of the reliability system that started functioning with k working components.

Theorem 3.6. *Z follows a PH-distribution with representation $(\boldsymbol{\eta}, T)$ of dimension $m[K(N-k+1)+N-k]$, where $\boldsymbol{\eta} = (0, \boldsymbol{\beta})$ and T is as given in (3.19).*

Proof. The proof is very similar to Theorem 3.2 and the details are omitted. □

Remark 3.7. The mean, μ_{TBF}, of the time between two successive failures of the system is given by

$$\mu_{\mathrm{TBF}} = \mu_Z + \mu_{\mathrm{DT}} = \boldsymbol{\eta}(-T)^{-1}\mathbf{e} + \mu_{\mathrm{DT}}. \tag{3.22}$$

The following result gives simple expressions for μ_{DT}, μ_Z, and μ_{TBF} in terms of the steady-state probability vector $\mathbf{x}$.

Theorem 3.8. *The means μ_Z, μ_{DT}, and μ_{TBF} are calculated as*

$$\begin{aligned} \mu_{\mathrm{DT}} &= \frac{\mathbf{y}(k-1)\mathbf{e}}{k\lambda \mathbf{y}_K(k)\mathbf{e}}, \\ \mu_Z &= \frac{1-\mathbf{y}(k-1)\mathbf{e}}{k\lambda \mathbf{y}_K(k)\mathbf{e}}, \\ \mu_{\mathrm{TBF}} &= \frac{1}{k\lambda \mathbf{y}_K(k)\mathbf{e}}. \end{aligned} \tag{3.23}$$

Proof. From the definition of $\boldsymbol{\eta}$ and T, we can write the generator Q of (2.1) as

$$Q = \begin{pmatrix} T & \mathbf{e}_{N-k+1}(N-k+1) \otimes B_k \\ \mathbf{S}^0\boldsymbol{\eta} & S \end{pmatrix}. \tag{3.24}$$

Partitioning the steady-state probability vector $\mathbf{x}$ as $\mathbf{x} = (\mathbf{u}, \mathbf{y}(k-1))$, the equations in (3.1) reduce to

$$\begin{gathered} \mathbf{u}T + \mathbf{y}(k-1)\mathbf{S}^0\boldsymbol{\eta} = 0, \\ \mathbf{u}\big(\mathbf{e}_{N-k+1}(N-k+1) \otimes B_k\big) + \mathbf{y}(k-1)S = 0, \end{gathered} \tag{3.25}$$

with the normalizing equation

$$\mathbf{u}\mathbf{e} + \mathbf{y}(k-1)\mathbf{e} = 1. \tag{3.26}$$

From equations in (3.25) we can immediately deduce that

$$\begin{aligned} \mathbf{u} &= k\lambda \mathbf{y}_K(k)\mathbf{e}\boldsymbol{\eta}(-T)^{-1}, \\ \mathbf{y}(k-1) &= k\lambda \mathbf{y}_K(k)(-S)^{-1}. \end{aligned} \tag{3.27}$$

The stated result follows immediately by using (3.26) in (3.27). □

3.5. The busy period of the repairman. The busy period of the repairman is defined as the interval that starts with the repairman getting busy and ends when for the first time the repairman becomes idle. Let X_{BPR} denote the busy period of the repairman. Then we have the following result.

Theorem 3.9. *The random variable X_{BPR} follows a PH-distribution with representation $(\boldsymbol{\theta}, L)$ of dimension $m(K+1)(N-k+1)$, where $\boldsymbol{\theta} = (p_N \mathbf{e}_1'(K) \otimes \boldsymbol{\beta}, q_N \mathbf{e}_1'(K+1) \otimes \boldsymbol{\beta}, 0)$ and L is given by*

$$L = \begin{pmatrix} \widetilde{A}_N & \widetilde{B}_N & & & & & \\ \widetilde{C}_{N-1} & A_{N-1} & B_{N-1} & & & & \\ & I \otimes \mathbf{S}^0\boldsymbol{\beta} & A_{N-2} & B_{N-2} & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & I \otimes \mathbf{S}^0\boldsymbol{\beta} & A_{k+1} & B_{k+1} & \\ & & & & I \otimes \mathbf{S}^0\boldsymbol{\beta} & A_k & B_k \\ & & & & & C_{k-1} & A_{k-1} \end{pmatrix}, \tag{3.28}$$

where $\widetilde{A}_N$ is obtained from A_N by deleting its first row and first column; $\widetilde{B}_N$ is obtained from B_N by removing its first row; and $\widetilde{C}_{N-1}$ is obtained from C_{N-1} by deleting its first column.

Proof. The proof follows immediately on noting that the busy period of the repairman can start when the reliability system gets into either (a) the set of states $\{(N, r, j), 1 \le r \le K, 1 \le j \le m\}$ with probability p_N or (b) the level $\mathbf{N} - 1$ with probability q_N. □

Remark 3.10. On noting that

$$Q = \begin{pmatrix} -N\lambda & N\lambda\boldsymbol{\theta} \\ \mathbf{L}^0 & L \end{pmatrix}, \tag{3.29}$$

where $\mathbf{L}^0$ is such that $L\mathbf{e} + \mathbf{L}^0 = 0$, the mean, $\mu_{\mathrm{BPR}} = \boldsymbol{\theta}(-L)^{-1}\mathbf{e}$, of the busy period of the repairman can be evaluated as

$$\mu_{\mathrm{BPR}} = \frac{1 - y_0(N)}{N\lambda y_0(N)}. \tag{3.30}$$

The above equation is intuitively clear as $1 - y_0(N)$ is the probability that the server is busy and $1 - y_0(N) = \mu_{\mathrm{BPR}}/(\mu_{\mathrm{BPR}} + (1/N\lambda))$.

3.6. Series system. In the special case when $N = k$ we get a closed-form solution and interesting limiting results. Furthermore these results are useful as accuracy checks in our numerical computation. For the current case, noting that (a) p_i, $k+1 \le i \le N$, plays no role and that $p_k = 1$, and (b) the steady state probability vector $\mathbf{x}$ can be written by suppressing N as

$$\mathbf{x} = \left(y_0, \mathbf{y}_1, \ldots, \mathbf{y}_K, \mathbf{y}(k-1)\right), \tag{3.31}$$

the following theorem gives an explicit expression for $\mathbf{x}$.

Theorem 3.11. *For a series system with K spares, for $K \ge 1$,*

$$\begin{aligned} \mathbf{y}_j &= y_0 \boldsymbol{\beta} R^j, \quad 1 \le j \le K, \\ \mathbf{y}(k-1) &= k\lambda y_0 \boldsymbol{\beta} R^K (-S)^{-1}, \end{aligned} \tag{3.32}$$

where y_0 is the normalizing constant and R is given by

$$R = k\lambda(k\lambda I - k\lambda \mathbf{e}\boldsymbol{\beta} - S)^{-1}. \tag{3.33}$$

Proof. The proof follows immediately on noting that $k\lambda \mathbf{y}_j \mathbf{e} = \mathbf{y}_{j+1}\mathbf{S}^0$, $1 \le j \le K-1$, and $k\lambda \mathbf{y}_K \mathbf{e} = \mathbf{y}(k-1)\mathbf{S}^0$. $\square$

Remark 3.12. (1) It is easy to verify that the inverse appearing in (3.33) does indeed exist and is nonnegative.

(2) When there are no spares in the system (i.e., when $K = 0$), it can be verified that $y_0 = \mu/(\mu + k\lambda)$ and $\mathbf{y}(k-1) = (k\lambda\mu/(\mu + k\lambda))\boldsymbol{\beta}(-S)^{-1}$.

The following result shows that the mean downtime of the system approaches a limit as the number of spares increases. In the sequel ρ denotes the spectral radius of R and σ^2 is the variance of the repair times.

Theorem 3.13. *For a series system with k components, as $K \to \infty$,*

$$\mu_{\mathrm{DT}} \longrightarrow \begin{cases} \dfrac{\rho(\mu - k\lambda)}{k\lambda\mu(1-\rho)}, & \mu \ne k\lambda, \\ 0.5k\lambda\left(\sigma^2 + \dfrac{1}{\mu^2}\right), & \mu = k\lambda. \end{cases} \tag{3.34}$$

Proof. First note that R is nonnegative and irreducible. From Perron-Frobenius theory of nonnegative matrices (see, e.g., [8]) we observe that

(i) the spectral radius, ρ, of R is simple and positive,

(ii) the left eigenvector, $\mathbf{u}$, and the right eigenvector, $\mathbf{v}$, for the spectral radius ρ are positive and can be chosen so that $\mathbf{ue} = 1$ and $\mathbf{uv} = 1$,

(iii) the matrix $((1/\rho)R)^j \to \mathbf{vu}$ as $j \to \infty$.

Now $\mathbf{u}R = \rho\mathbf{u}$ and the form of the matrix R as given in (3.33) implies

$$k\lambda(\rho - 1)\mathbf{u} = k\lambda\rho\boldsymbol{\beta} + \rho\mathbf{u}S. \tag{3.35}$$

Post-multiplying (3.35) by $(-S)^{-1}\mathbf{e}$ and using the facts that $\mu = [\boldsymbol{\beta}(-S)^{-1}\mathbf{e}]^{-1}$ and $\mathbf{ue} = 1$, we get

$$k\lambda(\rho - 1)\mathbf{u}(-S)^{-1}\mathbf{e} = \rho\left[\frac{k\lambda}{\mu} - 1\right]. \tag{3.36}$$

Since $\mathbf{u}(-S)^{-1}\mathbf{e} > 0$, (3.36) shows that $\rho = 1$ if and only if $\mu = k\lambda$. The stated result, for the case $\mu \neq k\lambda$, follows immediately from the form of μ_{DT} as given in (3.23) and the fact that $\boldsymbol{\beta}\mathbf{v}$ is positive.

In the case when $\mu = k\lambda$, (3.35) yields

$$\mathbf{u} = k\lambda\boldsymbol{\beta}(-S)^{-1}. \tag{3.37}$$

The stated result, for the case $\mu = k\lambda$, follows by observing $\sigma^2 = 2\boldsymbol{\beta}(-S)^{-2}\mathbf{e} - 1/\mu^2$.

The following theorem establishes limiting results for various probabilities. □

Theorem 3.14. *For a series system with k components, as $K \to \infty$,*

$$y_0 \longrightarrow \begin{cases} \dfrac{\mu - k\lambda}{\mu}, & \mu > k\lambda, \\ 0, & \mu \leq k\lambda, \end{cases} \tag{3.38}$$

$$\mathbf{y}(k-1)\mathbf{e} \longrightarrow \begin{cases} 0, & \mu > k\lambda, \\ \dfrac{k\lambda - \mu}{k\lambda}, & \mu \leq k\lambda, \end{cases} \tag{3.39}$$

$$\sum_{j=1}^{K}\mathbf{y}_j(k)\mathbf{e} \longrightarrow \begin{cases} \dfrac{k\lambda}{\mu}, & \mu > k\lambda, \\ \dfrac{\mu}{k\lambda}, & \mu \leq k\lambda. \end{cases} \tag{3.40}$$

Proof. Let $a_K = \sum_{j=0}^{K}\boldsymbol{\beta}R^j\mathbf{e}$, $c_K = a_K/\rho^K$, and $b = \boldsymbol{\beta}\mathbf{v}$. The proof follows immediately by noting that

$$\begin{gathered} k\lambda\boldsymbol{\beta}R^K(-S)^{-1}\mathbf{e} = 1 + \frac{k\lambda - \mu}{\mu}\sum_{j=0}^{K}\boldsymbol{\beta}R^j\mathbf{e}, \\ a_K \longrightarrow \frac{\mu}{\mu - k\lambda}, \quad \text{for } \mu > k\lambda, \qquad a_K \longrightarrow \infty, \quad \text{for } \mu \leq k\lambda, \\ c_K \longrightarrow \frac{b\rho}{(\rho - 1)}, \quad \text{for } \mu < k\lambda. \end{gathered} \tag{3.41}$$

□

Remark 3.15. In the case when $\mu > k\lambda$, the limiting results in (3.38) and (3.40) can be viewed as, respectively, the idle and the busy probabilities of the server (repairman) in the $M/G/1$ queue.

3.7. The case when $p_k \le 1$. While $p_k = 1$ is of particular interest in this paper, one can easily incorporate the case when $p_k < 1$. In this section we will sketch only the minimum details.

(i) First note that the matrix B_k as given in (2.5) for the current case takes the form

$$B_k = \begin{pmatrix} q_k k\lambda I & & & \\ & \ddots & & \\ & & q_k k\lambda I & \\ & & & k\lambda I \end{pmatrix}, \tag{3.42}$$

and $\mathbf{e}'_K(K+1) \otimes \mathbf{S}^0\boldsymbol{\beta}$ and S appearing in the generator given in (2.1) are replaced, respectively, with $I \otimes \mathbf{S}^0\boldsymbol{\beta}$ and $I \otimes S$. Due to these changes in the entries of the generator Q governing the reliability system under study and noting that $\mathbf{y}(k-1) = (\mathbf{y}_0(k-1),\ldots,\mathbf{y}_K(k-1))$ is now a vector of dimension $(K+1)m$, the steady-state (3.10)–(3.13) are replaced with

$$(k+1)\lambda q_{k+1}\mathbf{y}_0(k+1) + \mathbf{y}_0(k-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_0(k)[S - k\lambda I] = 0, \tag{3.43}$$

$$(k+1)\lambda q_{k+1}\mathbf{y}_j(k+1) + i\lambda p_k\mathbf{y}_{j-1}(k) + \mathbf{y}_j(k-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_j(k)[S - k\lambda I] = 0, \quad 1 \le j \le K-1, \tag{3.44}$$

$$(k+1)\lambda\mathbf{y}_K(k+1) + k\lambda p_k\mathbf{y}_{K-1}(k) + \mathbf{y}(k-1)\mathbf{S}^0\boldsymbol{\beta} + \mathbf{y}_K(i)[S - i\lambda I] = 0, \tag{3.45}$$

$$q_k k\lambda\mathbf{y}_j(k) + \mathbf{y}_j(k-1)S = 0, \quad 0 \le j \le K-1, \tag{3.46}$$

$$k\lambda\mathbf{y}_K(k) + \mathbf{y}_K(k-1)S = 0. \tag{3.47}$$

(ii) The statement of Theorem 3.4 now reads as follows. X_{DT} follows a PH-distribution with representation $(cq_k\sum_{j=0}^{K-1}\mathbf{y}_j(k) + c\mathbf{y}_K(k), S)$ of dimension m, where c is the normalizing constant given by $c = [q_k\sum_{j=0}^{K-1}\mathbf{y}_j(k)\mathbf{e} + \mathbf{y}_K(k)\mathbf{e}]^{-1}$.

(iii) In Theorem 3.6, the vector $\boldsymbol{\eta}$ is replaced by $\boldsymbol{\eta} = (0, c\mathbf{y}(k-1)(I \otimes \mathbf{S}^0\boldsymbol{\beta}))$, where c is the normalizing constant.

(iv) The term $k\lambda\mathbf{y}_K(k)\mathbf{e}$ appearing in (3.23) is replaced by $k\lambda[q_k\sum_{j=0}^{K-1}\mathbf{y}_j(k)\mathbf{e} + \mathbf{y}_K(k)\mathbf{e}]$.

(v) Theorem 3.9 holds good with two modifications in the entries of L matrix and these are similar to the ones carried out for Q.

(vi) Similar explicit expressions appearing in Theorem 3.11 for the series system can be derived for the current case. The expressions are as follows:

$$\begin{aligned} \mathbf{y}_j(k) &= y_0\boldsymbol{\beta}R^j, \quad 1 \le j \le K, \\ \mathbf{y}_0(k-1) &= q_k k\lambda y_0\boldsymbol{\beta}(-S)^{-1}, \\ \mathbf{y}_j(k-1) &= q_k k\lambda y_0\boldsymbol{\beta}R^j(-S)^{-1}, \quad 1 \le j \le K, \\ \mathbf{y}_K(k-1) &= k\lambda y_0\boldsymbol{\beta}R^K(-S)^{-1}, \end{aligned} \tag{3.48}$$

where y_0 is the normalizing constant and R is given by

$$R = p_k k\lambda(k\lambda I - k\lambda \mathbf{e}\boldsymbol{\beta} - S)^{-1}. \tag{3.49}$$

(vii) Theorem 3.13 is generalized for the case $p_k < 1$ as follows. First we let ρ_k denote the spectral radius of R as given in (3.44) with corresponding left and right eigenvectors denoted by $\mathbf{u}$ and $\mathbf{v}$, respectively. From Perron-Frobenius theory, we can take $\mathbf{u}$ and $\mathbf{v}$ to be positive vectors such that $\mathbf{ue} = 1$ and $\mathbf{uv} = 1$. In the following we let $a_K = \sum_{j=0}^{K} \boldsymbol{\beta} R^j \mathbf{e}$, $a = \boldsymbol{\beta}(I - R)^{-1}\mathbf{e}$ (in the case when $\rho_k < 1$), and $b = \boldsymbol{\beta}\mathbf{v}$. It is easy to verify the following facts by routine calculations.

(1) From the equations in (3.48) and the form of R matrix given in (3.49), we have

$$\sum_{j=0}^{K} \mathbf{y}_j(k-1)\mathbf{e} = y_0\left[1 + a_K \frac{k\lambda - \mu}{\mu}\right]. \tag{3.50}$$

(2) The mean downtime, μ_{DT}, is rewritten as

$$\mu_{\mathrm{DT}} = \frac{\mu + a_K(k\lambda - \mu)}{\mu k\lambda\left[q a_K + p(a_K - a_{K-1})\right]}. \tag{3.51}$$

(3) $a_K \to a$ when $\rho_k < 1$, and $a_K \to \infty$ when $\rho_k \geq 1$. However, when $\rho_k > 1$, $(a_K/\rho_k^K) \to (b\rho_k/(\rho_k - 1))$.

Then as $K \to \infty$, we have

$$\mu_{\mathrm{DT}} \longrightarrow \begin{cases} \dfrac{\mu + a(k\lambda - \mu)}{ak\lambda\mu}, & \rho_k < 1, \\ \dfrac{(k\lambda - \mu)}{k\lambda\mu(1 - p_k)}, & \rho_k = 1, \\ \dfrac{\rho_k(k\lambda - \mu)}{k\lambda\mu(\rho_k - p_k)}, & \rho_k > 1. \end{cases} \tag{3.52}$$

(viii) The results in Theorem 3.14 take the following form. As $K \to \infty$, we have

$$y_0 \longrightarrow \begin{cases} \dfrac{\mu}{\mu + ak\lambda}, & \rho_k < 1, \\ 0, & \rho_k \geq 1, \end{cases} \tag{3.53}$$

$$\sum_{j=0}^{K} \mathbf{y}_j(k-1)\mathbf{e} \longrightarrow \begin{cases} \dfrac{\mu + a(k\lambda - \mu)}{\mu + ak\lambda}, & \rho_k < 1, \\ \dfrac{k\lambda - \mu}{k\lambda}, & \rho_k \geq 1, \end{cases} \tag{3.54}$$

$$\sum_{j=1}^{K} \mathbf{y}_j(k)\mathbf{e} \longrightarrow \begin{cases} \dfrac{\mu(a-1)}{\mu + ak\lambda}, & \rho_k < 1, \\ \dfrac{\mu}{k\lambda}, & \rho_k \geq 1. \end{cases} \tag{3.55}$$

4. Illustrative numerical examples

In this section we discuss some interesting numerical examples that qualitatively describe the model under study. The correctness and the accuracy of the code are verified by a number of accuracy checks. For example, we obtained the numerical solution for the exponential repairs in its simple form. Next, we implemented the general algorithm, but using the following PH-representation. Let S be an irreducible, stable matrix with eigenvalue of maximum real part $-\tau < 0$. Let ς denote the corresponding left eigenvector, normalized by $\varsigma\mathbf{e} = 1$. Taking $\boldsymbol{\beta} = \varsigma$, the PH-representation reduces to the exponential repairs with intensity rate τ. The general algorithm does not utilize this fact in any manner, but the numerical results agreed very much. We also used the limiting results as another set of accuracy checks.

For the repair times, we consider the following three PH-distributions

(1) *Erlang* (ERL),

$$\boldsymbol{\beta} = (1,0,0,0,0), S = \begin{pmatrix} -5 & 5 & 0 & 0 & 0 \\ 0 & -5 & 5 & 0 & 0 \\ 0 & 0 & -5 & 5 & 0 \\ 0 & 0 & 0 & -5 & 5 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix}; \tag{4.1}$$

(2) *exponential* (EXP),

$$\boldsymbol{\beta} = (1), \qquad S = (-5); \tag{4.2}$$

(3) *hyperexponential* (HEX),

$$\boldsymbol{\beta} = (0.9, 0.1), \qquad S = \begin{pmatrix} -100 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.3}$$

All these three PH-distributions will be normalized so as to have a specific repair rate. Note that these are qualitatively different distributions having different variances. The ratio of the standard deviations of these three distributions with respect to ERL are, respectively, 1.0, 2.236068, and 8.901896.

In addition to various means listed above, we consider the following system performance measures for our numerical examples:

(i) the probability, P_{idle}, that the repairman is idle is calculated as

$$P_{\text{idle}} = y_0(N); \tag{4.4}$$

(ii) the probability, P_{down}, that the system is down is given by

$$P_{\text{down}} = \mathbf{y}(k-1)\mathbf{e}; \tag{4.5}$$

(iii) the probability mass function, f_j^{NR}, of the number of units under repair is obtained as

$$f_j^{\text{NR}} = \begin{cases} y_0(N), & j = 0, \\ \sum_{i=0}^{j} \mathbf{y}_{j-i}(N-i)\mathbf{e}, & 1 \le j \le N-k, \\ \sum_{i=0}^{N-k} \mathbf{y}_{j-i}(N-i)\mathbf{e}, & N-k+1 \le j \le N+K-k, \\ \mathbf{y}(k-1)\mathbf{e}, & j = N+1+K-k. \end{cases} \tag{4.6}$$

Note that from (4.6) we can compute the mean, μ_{NR}, number of units under repair.

Example 4.1. In this example, we consider a 15-out-of-20 system by first fixing $p_i = p$, $k+1 \le i \le N$, $p_k = 1$, $\lambda = 1$, and varying the parameters K, μ, p, and the repair time distribution. The purpose here is to see the impact of varying these parameters on the measures, μ_{TBF} and μ_{BPR}. The graphs of these measures for various combinations are displayed in Figures 4.1 and 4.2. Next we compare the two cases when $p_k = 1.0$ with $p_k < 1$ when all other parameters are fixed. Towards this end, we take $\mu = 20$, and vary K, p_k and the repair time distribution. The graphs of μ_{TTF}, μ_{NR} and the fraction of spares used are displayed for ERL and HEX repairs in Figures 4.3 through 4.5. The fraction of spares used is defined as the ratio of the mean number of spares used and K. Some interesting results observed from these figures are summarized below.

(i) As μ increases, μ_{TBF} appears to increase in all cases. This is to be expected since increasing the repair rate makes more components available for functioning.

(ii) While we see a very significant improvement in this measure for ERL and EXP repairs, there seems to be only a marginal improvement for the HEX repairs.

(iii) The effect of the parameter p is noticeable only for small values p (between 0 and 0.4). This can be intuitively explained as follows. When p is small, the spares are used sporadically unless the system is about to fail in which case we always use a spare (when available). This on the average will result in an increase in the mean. However, if p is large, the spares will be called upon as and when a unit fails (irrespective of whether the system failure occurs at a failure or not). In this case not only the failure rate of the system increases but also the spares deplete faster.

(iv) The impact of K on the mean is seen either when p is small or when μ is large. Furthermore, after certain value for K (which depends on the type of repair distribution) the impact becomes significant. As is to be expected, the smaller the repair rate the more the impact of K on the mean.

(v) With respect to the measure μ_{BPR}, we notice that ERL repair appears to have the largest value in all cases as compared to the other two repair times.

(vi) As K increases, we notice that the dependence of μ_{BPR} on p appears to increase in all cases. [Note: we confirmed this result in the case of HEX repairs also even though the graph in Figure 4.2 does not reflect this as K goes only up to 5, and for HEX case K has to be even larger.]

(vii) Looking at Figures 4.3 through 4.5, we see that μ_{TTF} and μ_{NR} appear to increase with p_k. For K up to a certain value, the increase in these measures (as functions of p_k) appears to be very insignificant. However for K reasonably large the significance of these measures can be seen even for moderate values of p_k.

(viii) With regards to the fraction of spares used by the system, we notice that p_k appears to be play no significant role in the case of ERL repairs for all values of K; however, for HEX repairs p_k plays a significant role for almost all values of K. A similar behavior is seen for the mean number of units under repair.

In Example 4.1 we assumed p_i, $k+1 \le i \le N$, to be a fixed number (not depending on i). But in some situations it is necessary to consider p_i as a decreasing function of i or p_i varying progressively. For example, as the number of working components (starting with N) decreases, the probability of requesting (or delivering) a spare should increase. In some situations, the uncertainty to deliver a spare remains the same irrespective of when a spare is requested. We consider these scenarios in the next two examples.

Example 4.2. In this example, we let p_i vary as a function of i and compare it to the case when p_i is constant. All other parameters are identical to the ones considered in Example 4.1. The purpose here is to see the impact of the varying p_i as opposed to keeping it fixed. We look at the following scenarios.

(i) Case 1 (fixed): let $p_i = p$, $15 \le i \le 20$.

(ii) Case 2 (linear): let $p_i = p - (0.2p - 0.01)(i - 15)$, $15 \le i \le 20$, where $0 < p \le 1$. In this case, $p_{15} = p$ and $p_{20} = 0.05$, and $0.05 < p_i < p$, $15 < i < 20$.

(iii) Case 3 (piecewise linear): given a set of 3 probabilities, say $p^{(j)}$, $1 \le j \le 3$, with $p^{(1)} < p^{(2)} < p^{(3)}$, we take $p_{15} = p_{16} = p^{(1)}$, $p_{17} = p_{18} = p^{(2)}$, $p_{19} = p_{20} = p^{(3)}$.

By fixing $p^{(1)} = 0.1$, $p^{(2)} = 0.5(0.1 + p)$, and $p^{(3)} = p$, we vary p from 0.3 to 1.0. Looking at the graph of μ_{TTF} as shown in Figure 4.6, we notice the following observations.

(i) There appears to be a cut-off point for p, say p^*, such that for $p < p^*$, fixed values for p_i outperform the other two scenarios. For $p > p^*$, the other two scenarios outperform the fixed value scenario.

(ii) The value of p^* appears to be larger for HEX repairs as compared to ERL case. This illustrates the role of variability in the repairs played in determining the strategy of when to request for spares in the presence of uncertainty in the delivery.

Example 4.3. In this example, we consider the data given in Example 4.1 and we consider probability structures that progressively request for spares. Denoting $\mathbf{p} = (p_{15}, \ldots, p_{20})$, the six probability structures considered are $\mathbf{p}^{(1)} = (p, 0, 0, 0, 0, 0)$, $\mathbf{p}^{(2)} = (p, p, 0, 0, 0, 0)$, $\mathbf{p}^{(3)} = (p, p, p, 0, 0, 0)$, $\mathbf{p}^{(4)} = (p, p, p, p, 0, 0)$, $\mathbf{p}^{(5)} = (p, p, p, p, p, 0)$, and $\mathbf{p}^{(6)} = (p, p, p, p, p, p)$. In Figure 4.7, we display the probability structure that yields the maximum value for the mean time to failure as p is varied. It is clear from this graph that in the case of

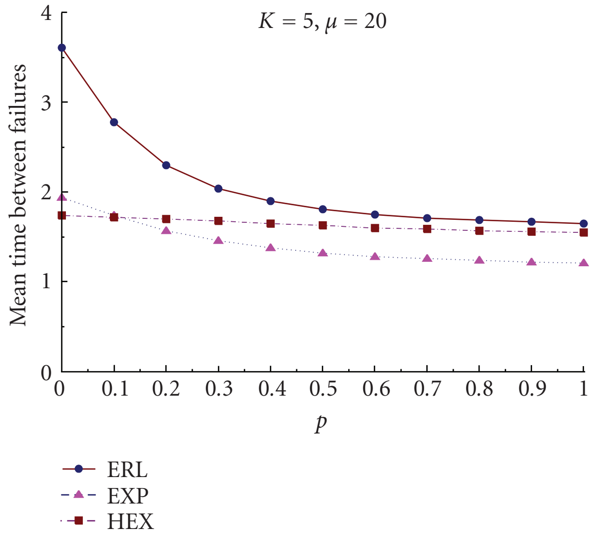

Figure 4.1. Mean time between failures of a 15-out-of-20 system.

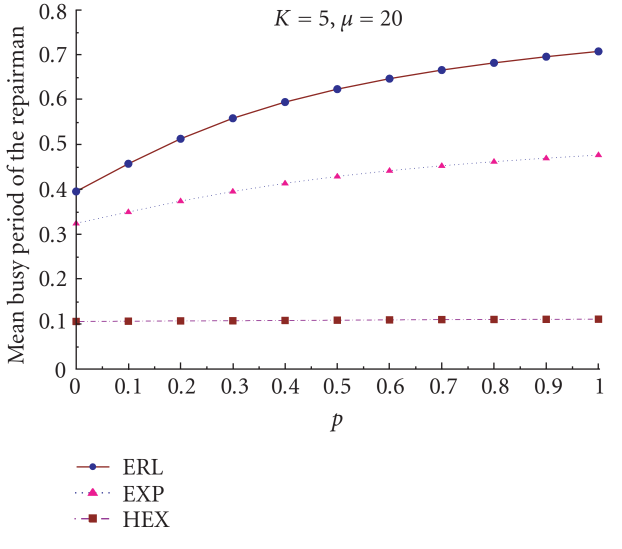

Figure 4.2. Mean busy period of the repairman of a 15-out-of-20 system.

smaller values of p, requests for a spare should be made as and when a component fails; as p increases to 1, the requests for a spare can be postponed closer to system failure, so as to maximize the mean time to failure. Furthermore, the larger the variation in the repair times is, the much earlier requests should be made to system failure even when p is reasonably large.

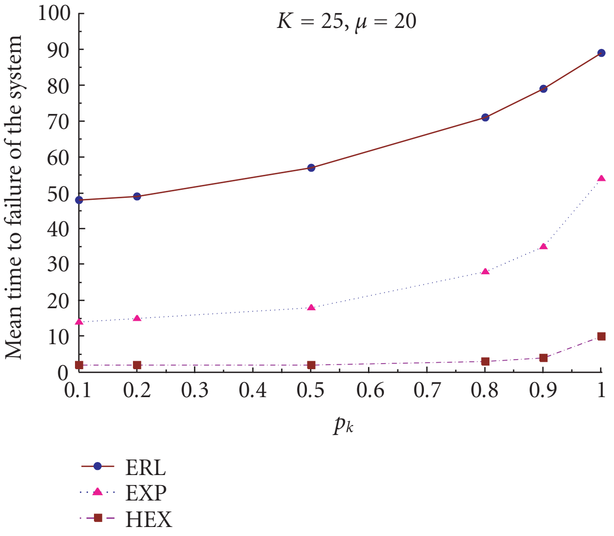

Figure 4.3. Mean time to failure of a 15-out-of-20 system.

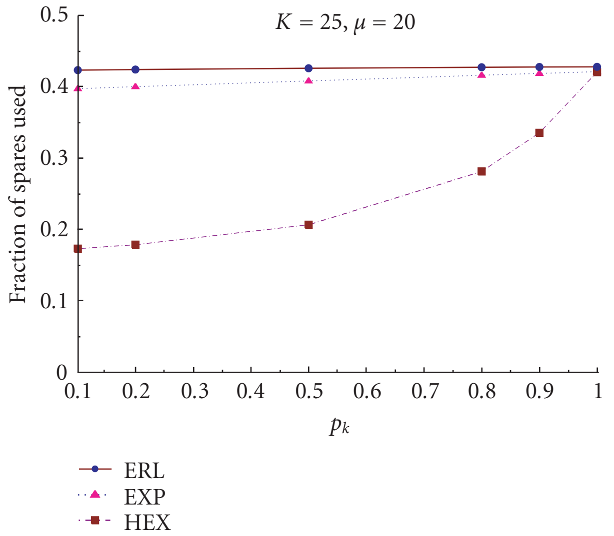

Figure 4.4. Fraction of spares used in a 15-out-of-20 system.

Conclusions and recommendations. Based on our numerical experimentation, we make the following conclusions and offer some recommendations for practicing engineers.

(i) As is to be expected, having $p_k = 1$ and $p_i = 0$, $k+1 \leq i \leq N$, is the ideal situation in that it increases (a) the mean time to failure; (b) the mean time between failures;

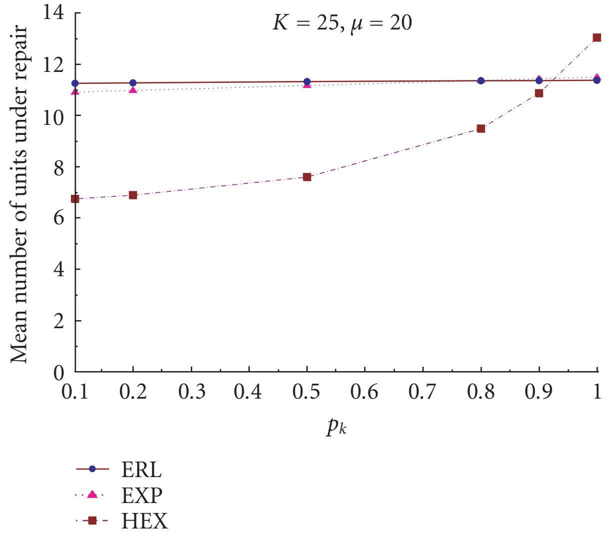

Figure 4.5. Mean number of units under repair in a 15-out-of-20 system.

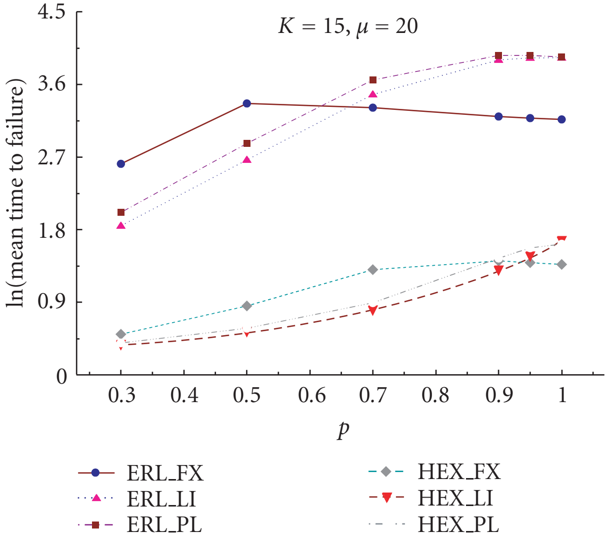

Figure 4.6. MTTF of a 15-out-of-20 system for three probabilistic structures.

(c) the probability that the server is idle; and the following measures decrease: (i) the probability that the system is down; (ii) the mean number of units under repair; (iii) the mean number of spares used; and (iv) the mean busy period of the repairman.

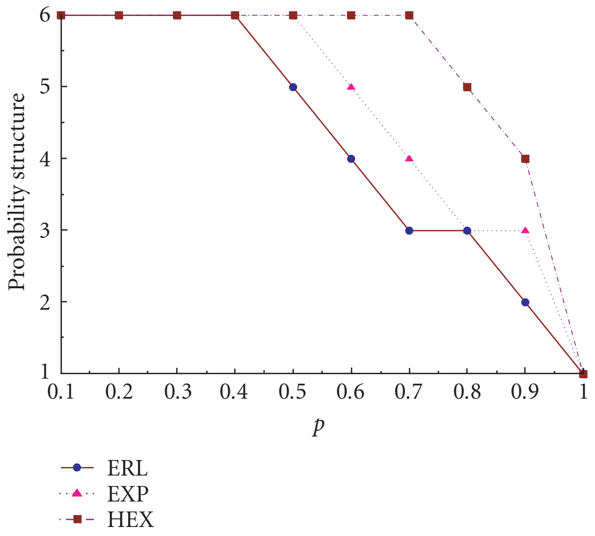

Figure 4.7. Optimum point to start requesting a spare in a 15-out-of-20 system.

(ii) However, it is not always the case that a spare unit is made available with certainty when requested. As pointed out earlier if there is a problem (which is described in some probabilistic sense) in replacing a failed component with a spare, how and when should a spare be requested so as to, say, increase the mean time to failure? Of course, the answer to this question depends on a number of factors such as the repair time distribution, the rate of repairs, and the probability structure, $\{p_i\}$, for making the requests. If there is uncertainty in the delivery of spares (namely, when $p_i < 1$) it is advisable to make a request for spares way ahead of the system failure, for all repair time distributions. In particular, whenever the repair times have a larger variability, it is better to request for spares as early as possible. Given the input parameters of the reliability model, one can use the results of this paper to determine the strategy for requesting spares.

(iii) In practice, one has to use historical data to estimate the probabilities, $\{p_i\}$.

5. An optimization problem

A number of optimization problems of practical interest can be proposed and solved once the reliability model is analyzed. Here we will propose an optimization problem that takes into account various costs and arrive at optimum values for N and K. Specifically, we define

(i) c_1 = cost per unit when the system is functional;
(ii) c_2 = cost of system failure;
(iii) c_3 = cost per spare unit made available during the operation of the system;
(iv) c_4 = repair cost per unit;
(v) c_5 = cost of using a spare unit.

The optimization problem of interest is to find the optimum values, N^* and K^*, of N and K by fixing k, λ, μ, p, and the repair time distribution, that will minimize the expected cost rate per unit of time. That is,

$$\min_{(N,K)} \frac{1}{\mu_{\mathrm{TTF}}}\left[c_1 N + c_2 + c_3 K + c_4 \mu_{\mathrm{NR}}^c + c_5 \mu_{NS}^c\right], \tag{5.1}$$

where μ_{NR}^c and μ_{NS}^c are, respectively, the conditional mean number of units repaired and the conditional mean number of spares used, conditioned on the fact that the system was functioning. These can be obtained from the steady-state probability vector $\mathbf{x}$ (see, e.g., (3.3)) and the details are omitted.

Example 5.1. For the optimization problem we take $\lambda = 1$, $k = 15$, $c_1 = 1.0$, $c_2 = 1000.0$, $c_3 = 0.1$, $c_4 = 0.2$, $c_5 = 1.0$, and find (N^*, K^*, F^*) by restricting $k \le N \le 20$ and $0 \le K \le 10$, and F^* is the value of the objective function at the optimum point, for various combinations including the two cases: $p_k = 0.5$ and $p_k = 1$. Table 5.1 lists the optimum values after running all possible combinations. Note that when the number of combinations becomes prohibitively large, one can develop a suitable heuristic method to look for a local or a global optimum. An examination of the table reveals the following observations.

(i) $F^*(p)$ as a function p is such that this function is smallest when $p = 0$ and the largest when $p = 1$ when all other parameters are fixed. This is as expected.

(ii) While $F^*(\mu)$ is a decreasing function of μ (when all other parameters are fixed), there is an interesting trend for N^* as μ is increased. For example, in the case of ERL repairs, we notice that the optimum value of N^* appears to decrease as μ approaches $k\lambda$ and then increases as μ increases further away from $k\lambda$.

(iii) The rate of decrease in $F^*(\mu)$ as μ increases (when all other parameters are fixed) is high for ERL and EXP repairs while the rate is much smaller for HEX repairs.

(iv) Comparing the results for $p_k = 0.5$ and $p_k = 1.0$ cases, we notice the following.

  (1) The objective function value increases as p_k is decreased. This is intuitively clear as reducing this probability makes the system fail earlier.

  (2) When $p_k = 1.0$, all combinations resulted in an optimum value for K^* to be 10, the maximum that is allowed. However when $p_k = 0.5$ for some combinations (especially when p is small) optimum value for K is less than 10. This is again due to the fact that not using the spares regularly leads to more failures of the system and thus results in a higher cost.

(v) There appears to be two cut-off points, say μ_1 and μ_2, such that for $\mu < \mu_1$, F^* appears to be decreasing with increasing variability of the repair times; for $\mu > \mu_2$, F^* appears to increase with increasing variability of the repairs; for $\mu_1 \le \mu \le \mu_2$, it appears that EXP repairs dominate the other two distributions before HEX becomes the dominator. In interpreting this, one has to keep in mind the possible changes in N^* and K^*.

6. Concluding remarks

In this paper we considered a k-out-of-N reliability system with a single repairman offering phrase-type services to failed components. We introduced a decision rule (based on

Table 5.1. Values of (N^*, K^*, F^*) when $\lambda = 1$, $p_k = 0.5$, and $k = 15$.

μ	p	$p_k = 1$			$p_k = 0.5$		
		ERL	EXP	HEX	ERL	EXP	HEX
1	0.0	(20,10,977.47)	(20,10,953.54)	(20,10,692.60)	(20,7,2473.70)	(20,7,2372.47)	(20,6,1586.05)
	0.1	(20,10,983.74)	(20,10,960.10)	(20,10,702.45)	(20,7,2296.80)	(20,7,2196.90)	(20,8,1428.69)
	0.2	(20,10,991.71)	(20,10,968.47)	(20,10,714.75)	(20,9,2107.29)	(20,9,2009.92)	(20,10,1270.87)
	0.5	(20,10,1036.24)	(20,10,1014.96)	(20,10,767.68)	(20,10,1486.65)	(20,10,1415.94)	(20,10,918.52)
	0.8	(20,10,1128.72)	(20,10,1104.58)	(20,10,822.80)	(20,10,1163.66)	(20,10,1134.73)	(20,10,835.80)
	0.9	(20,10,1158.33)	(20,10,1131.30)	(20,10,835.25)	(20,10,1170.30)	(20,10,1142.72)	(20,10,843.15)
	1.0	(20,10,1181.19)	(20,10,1151.33)	(20,10,844.44)	(20,10,1191.95)	(20,10,1161.82)	(20,10,852.15)
5	0.0	(20,10,737.32)	(20,10,726.87)	(20,10,567.81)	(20,7,1971.18)	(20,7,1876.65)	(20,7,1054.21)
	0.1	(20,10,746.57)	(20,10,736.70)	(20,10,572.44)	(20,8,1767.93)	(20,8,1691.37)	(20,8,1007.59)
	0.2	(20,10,758.94)	(20,10,749.41)	(20,10,577.88)	(20,10,1551.91)	(20,10,1495.83)	(20,9,955.91)
	0.5	(20,10,828.52)	(20,10,813.41)	(20,10,603.05)	(20,10,1012.26)	(20,10,1006.85)	(20,10,768.45)
	0.8	(20,10,911.81)	(20,10,888.87)	(16,10,630.39)	(20,10,920.74)	(20,10,899.93)	(20,10,657.24)
	0.9	(20,10,927.44)	(20,10,904.95)	(15,10,636.05)	(20,10,933.57)	(20,10,911.12)	(19,10,659.21)
	1.0	(20,10,937.59)	(20,10,916.34)	(15,10,636.05)	(20,10,943.70)	(20,10,922.36)	(16,10,648.97)
10	0.0	(20,10,444.54)	(20,10,455.14)	(20,10,452.96)	(20,10,446.10)	(20,10,456.79)	(20,10,453.87)
	0.1	(20,10,462.57)	(20,10,470.96)	(20,10,457.16)	(20,10,464.19)	(20,10,472.68)	(20,10,458.11)
	0.2	(20,10,485.86)	(20,10,490.10)	(20,10,462.25)	(20,10,487.57)	(20,10,491.89)	(20,10,463.25)
	0.5	(15,10,557.90)	(17,10,552.52)	(20,10,486.76)	(15,10,559.54)	(17,10,554.37)	(20,10,487.96)
	0.8	(15,10,557.90)	(15,10,566.89)	(16,10,513.57)	(15,10,559.54)	(15,10,568.40)	(16,10,514.48)
	0.9	(15,10,557.90)	(15,10,566.89)	(16,10,519.60)	(15,10,559.54)	(15,10,568.40)	(16,10,520.55)
	1.0	(15,10,557.90)	(15,10,566.89)	(15,10,520.59)	(15,10,559.54)	(15,10,568.40)	(15,10,521.42)
14	0.0	(20,10,227.47)	(20,10,250.73)	(20,10,370.96)	(20,7,794.61)	(20,7,898.31)	(20,9,849.29)
	0.1	(15,10,228.51)	(19,10,268.83)	(20,10,375.07)	(20,10,596.73)	(20,10,730.55)	(20,9,802.88)
	0.2	(15,10,228.51)	(17,10,278.09)	(20,10,380.09)	(20,10,457.38)	(20,10,580.63)	(20,10,750.69)
	0.5	(15,10,228.51)	(15,10,287.57)	(20,10,404.34)	(20,10,358.98)	(20,10,388.46)	(20,10,563.06)
	0.8	(15,10,228.51)	(15,10,287.57)	(17,10,431.92)	(18,10,338.01)	(18,10,368.17)	(20,10,453.62)
	0.9	(15,10,228.51)	(15,10,287.57)	(16,10,438.14)	(17,10,319.36)	(17,10,363.07)	(19,10,455.44)
	1.0	(15,10,228.51)	(15,10,287.57)	(15,10,440.45)	(16,10,272.50)	(16,10,317.27)	(16,10,448.70)
15	0.0	(15,10,164.68)	(20,10,203.50)	(20,10,352.05)	(20,7,689.25)	(20,7,812.50)	(20,9,829.21)
	0.1	(15,10,164.68)	(18,10,217.70)	(20,10,356.13)	(20,10,502.35)	(20,10,651.50)	(20,9,782.62)
	0.2	(15,10,164.68)	(17,10,224.25)	(20,10,361.12)	(20,10,378.03)	(20,10,509.79)	(20,10,730.29)
	0.5	(15,10,164.68)	(15,10,232.18)	(20,10,385.23)	(19,10,295.72)	(20,10,333.16)	(20,10,542.79)
	0.8	(15,10,164.68)	(15,10,232.18)	(17,10,412.94)	(18,10,271.63)	(18,10,310.29)	(20,10,433.91)
	0.9	(15,10,164.68)	(15,10,232.18)	(16,10,419.38)	(17,10,250.67)	(17,10,304.65)	(19,10,435.82)
	1.0	(15,10,164.68)	(15,10,232.18)	(15,10,422.03)	(16,10,204.11)	(16,10,259.33)	(16,10,429.78)
20	0.0	(20,10,12.17)	(20,10,40.42)	(20,10,267.47)	(20,8,308.40)	(20,8,475.30)	(20,9,735.42)
	0.1	(18,10,15.58)	(19,10,47.72)	(20,10,271.32)	(20,10,193.43)	(20,10,355.36)	(20,10,688.22)
	0.2	(17,10,17.14)	(18,10,52.60)	(20,10,276.04)	(20,10,116.61)	(20,10,253.43)	(20,10,635.54)
	0.5	(15,10,18.19)	(16,10,60.53)	(20,10,298.93)	(20,10,60.08)	(20,10,115.24)	(20,10,450.00)
	0.8	(15,10,18.19)	(15,10,63.92)	(17,10,327.41)	(18,10,42.77)	(18,10,99.38)	(20,10,344.16)
	0.9	(15,10,18.19)	(15,10,63.92)	(16,10,334.86)	(18,10,41.11)	(18,10,91.92)	(19,10,346.64)
	1.0	(15,10,18.19)	(15,10,63.92)	(15,10,339.12)	(16,10,23.59)	(16,10,72.06)	(16,10,344.46)

a probability structure) for replacing a failed component with a spare one, and analyzed the reliability system in steady state, and obtained some interesting results.

Acknowledgments

The author expresses his sincere thanks to the three anonymous referees and the editor for their useful suggestions that improved the presentation of the paper.

References

[1] K. S. de Smidt-Destombes, M. C. van der Heijden, and A. van Harten, *On the availability of a k-out-of-N system given limited spares and repair capacity under a condition based maintenance strategy*, Reliability Engineering and System Safety **83** (2004), no. 3, 287–300.

[2] ______, *On the interaction between maintenance, spare part inventories and repair capacity for a k-out-of-N system with wear-out*, European Journal of Operational Research **174** (2006), no. 1, 182–200.

[3] D. P. Gaver, P. A. Jacobs, and G. Latouche, *Finite birth-and-death models in randomly changing environments*, Advances in Applied Probability **16** (1984), no. 4, 715–731.

[4] A. Graham, *Kronecker Products and Matrix Calculus: With Applications*, Ellis Horwood, Chichester, 1981.

[5] W. Kuo and M. J. Zuo, *Optimal Reliability Modeling - Principles and Applications*, John Wiley & Sons, New York, 2003.

[6] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Pennsylvania; American Statistical Association, Virginia, 1999.

[7] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach*, Dover, New York, 1994, (originally published by John Hopkins University Press in 1981).

[8] E. Seneta, *Non-Negative Matrices and Markov Chains*, Springer, New York, 1981.

[9] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, New Jersey, 1994.

Srinivas R. Chakravarthy: Department of Industrial and Manufacturing Engineering,
Kettering University, Flint, MI 48504, USA
E-mail address: schakrav@kettering.edu