A General Repairable Spare Part Demand Model Based on Quasi Birth and Death Process

HUANG Zhuo LIANG Liang GUO Bo
(Systems Engineering Department, National University of Defense Technology, Changsha 410073)
(E-mail: missyoufriend@sina.com)

Abstract This paper aims at two problems which exist in most of repairable spare part demand models at present: the exponential distribution as the basic assumption and one typical distribution corresponding to a model. A general repairable spare part demand model built on quasi birth-and-death process is developed. This model assumes that both the operational time of the unit and the maintenance time of the unit follow the continuous time phase type distributions. The first passage time distribution to be out of spares, the first mean time to be out of spares, and an algorithm to get the minimal amount of spares under certain restrictions are obtained. At the end of this paper, a numerical example is given.

Key words Reliability, repairable spare, spare demand, phase type distribution, quasi birth and death process

1 Introduction
The problem of scientifically deciding the demand of spare parts always attracts attention. To solve the problem, it is very important to select a reasonable spare part demand model. The trade-off is clear: on the one hand a large number of spare parts ties up a large amount of capital, on the other hand too little inventory may result in poor repair service or extremely costly emergency actions.

At present, most of the repairable spare part demand models are established on one of the following two assumptions: 1) the operational time and the repair time both follow exponential distributions; 2) one typical operational time distribution corresponds to one spare part demand model. Zhou Jiang-Hua [1] firstly introduced a strategic inventory system and its aging process; to analyze the aging process, Zhou split the aging process into two stages. Under the assumption that the operational and repair time follow exponential distributions, each stage was modeled by a continuous-time Markov chain. An analytical method based on such Markov chain models was provided for calculating the optimum spare part quantities of $n : k(S)$ under reliability restriction. Díaz Angel [2] established models based on $M/G/x$ queuing system for multi-echelon repairable item inventory systems with limited repair capacity. From the method that Díaz Angel used, we can see that it is assumed that the operational time follows exponential distribution. Karin [3] considered the availability of a $k$-out-of-$N$ system with identical, repairable components; in the model the operational time of component followed exponential distribution. Li Jin-Guo [4] adopted the method that one typical distribution corresponded to one spare demand model; the typical distributions included exponential distribution, normal distribution and Weibull distribution. Yang Bing-Xi [5] offered four kinds of spare demand model; the four models respectively applied to spares with exponential, Weibull, normal, and binomial life; moreover the application example of spares provisioning in the ground radar was given. Yu Jing [6] analyzed that the spares of missile weapon and equipment were not reasonably stored; a method was discussed to calculate the quantities of initial spares whose life follows exponential distribution, normal distribution and Weibull distribution, and to calculate the quantities of follow-on spares which are repairable and unrepairable. Zhang Zhi-Min [7] presented the warship spare demand model based on the assumption that the operational time of component follows the exponential distribution. Chen Song-Lin [8] considered that the test equipment spare demand arrival follows the Poisson process; Chen developed the spare demand model based on the Poisson process assumption.

The exponential distribution is easy to calculate analytically, but it is obvious that such an assumption is unreasonable under many conditions. For example, the assumption that machining time follows the exponential distribution will result in that the elapsed machining time will have nothing to do with the future spending machining time. If the life of the unit follows the exponential distribution,
the unit seems to be new all the time. However, if we assume that the operational time of the unit does not follow the exponential distribution, the quantitative analyses of the model will get into difficulty. The method that one typical probability distribution corresponds to one model is not advisable either. Developing a general spare part demand model should be a feasible approach.

Through the above analyses, two problems could be found: 1) most of the spare demand model assume that the life and repair time of the unit follow the exponential distribution; 2) one typical life distribution corresponds to a spare demand model.

To solve the two problems as mentioned above, this paper develops a general spare parts demand model, which is established on the continuous phase type distribution. In that phase type distributions are dense on \([0, +\infty)\) and convenient for computation, the model developed in this paper is not only general but also can be realized by algorithm easily.

2 Phase type distribution and its properties

Definition 1. The probability distribution \(F(\bullet)\) on \([0, +\infty)\) is a phase type distribution (PH-distribution) with representation \((\alpha, T)\) if it is the distribution of the time until absorption in a finite-state Markov process with generator

\[
\begin{pmatrix}
T & T^0 \\
0 & 0
\end{pmatrix}
\]

with initial probability vector \((\alpha, \alpha m+1)\), where \(\alpha\) is a row \(m\)-vector. Throughout this paper \(e\) denotes a column vector with all entries equal to one whose dimension is determined by the context. The matrix \(T\) of order \(m\) is non-singular with negative diagonal entries and non-negative off-diagonal entries and satisfies \(-Te = T^0 \geq 0\). The distribution \(F(\bullet)\) is given by

\[
F(x) = 1 - \alpha \exp(Tx)e, \quad x \geq 0
\]  

(1)

Every transient state of a phase-type distribution is called a phase.

Theorem 1. Phase type distributions are dense in the class of distributions defined on the nonnegative real numbers.

The theoretical meaning of dense is that any general probability distribution can be approximated by phase type distribution at will. When we study the stochastic process related to several general distributions \(F_j\) \((j = 1, 2, \cdots, n)\) defined on \([0, +\infty)\), it is very difficult to deal with the continuous generalized function \(\Phi(F_1, F_2, \cdots, F_n)\) with \(F_j\) \((j = 1, 2, \cdots, n)\). However, since phase type distribution keeps the analytical properties of exponential distribution, it is relatively easy to prove that the continuous generalized function \(\Phi\) is true about phase type distribution. If the continuous generalized function \(\Phi\) is true about phase type distribution and the proof process does not depend on the structure of phase type distribution heavily, we can assert that the continuous generalized function \(\Phi\) is also true about general distribution \(F_j\) \((j = 1, 2, \cdots, n)\).

The formalism of Kronecker product and Kronecker sum is extensively used in our analysis, which involves phase type distributions. We recall the following definitions of Kronecker product and Kronecker sum.

Definition 2. If \(A\) and \(B\) are rectangular matrices of dimensions \(k_1 \times k_2\) and \(k_1' \times k_2'\), respectively, their Kronecker product \(A \otimes B\) is defined as

\[
A \otimes B = (A_{ij}B) = \begin{bmatrix}
A_{11}B & \cdots & A_{1k_2}B \\
\vdots & \ddots & \vdots \\
A_{k_11}B & \cdots & A_{k_1k_2}B
\end{bmatrix}
\]  

(2)

A \otimes B is the matrix of dimensions \(k_1k_1' \times k_2k_2'\), written in compact form as \((A_{ij}B)\).

A useful property of this product is the following equality \((A \otimes B)(C \otimes D) = AC \otimes BD\), which holds whenever the ordinary matrix product is well defined.

Definition 3. If \(A\) and \(B\) are matrices of dimensions \(m \times m\) and \(n \times n\), respectively, their Kronecker sum \(A \oplus B\) is the matrix of dimensions \(mn \times mn\), written as

\[
A \oplus B = A \otimes I_n + I_m \otimes B
\]  

(3)

where \(I_n\) and \(I_m\) are the identity matrices of orders \(n\) and \(m\), respectively. A functional property of this sum is the following equality for the matrices above

\[
\exp(A \oplus B) = \exp(A) \otimes \exp(B)
\]
Definition 4. Consider a Markov process on the two-dimensional state space: \( \Omega = \{(k,j) : k \geq 0, j = 1, 2, \ldots, m\} \), and refer by level \( n \) to the set of states \( \{(n,1), \ldots, (n,m)\} \). Such a Markov process is called a quasi-birth-death (QBD) process when one-step transitions are restricted to states in the same level or in two adjacent levels.

For a QBD process, we order the states lexicographically, i.e.,

\[
\{(0,1), \ldots, (0,m), (1,1), \ldots, (1,m), \ldots, (n,1), \ldots, (n,m), \ldots\}
\]

and that the infinitesimal generator \( Q \) has the following block tridiagonal structure:

\[
Q = \begin{bmatrix}
A_0 & C_0 & 0 & 0 & \cdots & 0 & \cdots \\
B_1 & A_1 & C_1 & 0 & \cdots & 0 & \cdots \\
0 & B_2 & A_2 & C_2 & \cdots & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\
0 & 0 & 0 & B_n & A_n & C_n & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\end{bmatrix}
\]

(4)

\( A_i \) denotes the transitions between the same level \( i \); \( C_i \) denotes the transitions between the level \( i \) and the level \( i + 1 \); \( B_i \) denotes the transitions between the level \( i + 1 \) and the level \( i \).

3 Problem descriptions

A system comprises one certain type unit, and the unit is repairable. When the unit fails, the system fails. In order to keep the system working, several spares should be prepared for the unit. When the online unit fails, the disabled unit is replaced by one spare; then the disabled unit will be repaired; the restoration of the disabled unit will be a spare. The problem is that in order to ensure the spare provision with a probability of not less than \( P \) during the period of \( t \) how many spare parts should be prepared at least.

Assumptions of the problem

1) The operational time of the unit follows order \( m \) phase type distribution with irreducible representation \((\alpha, T)\).
2) Spares of the unit are cold standby and do not fail during the period of cold standby.
3) The repair time of the unit follows order \( n \) phase type distribution with irreducible representation \((\beta, S)\). There is only one repairman, the disabled unit is as good as new after repair and the repair rule obeys FCFS.
4) The time of replacing the disabled unit is assumed to be zero.
5) The operational time and the repair time are independent.

Assume that in order to ensure the spare provision with a probability of not less than \( P \) during the period of \( t \), \( h \) spares of the unit should be prepared at least. So when \( h + 1 \) the certain type units fails, the system fails. Then the problem is converted to how to calculate the value of \( h \) under given restrictions.

4 Model development

Let \( X(t) \) denote the number of failed units at time \( t \); then the stochastic process \( \{X(t), t \geq 0\} \) describes that the number of failed units changes with time. Due to the operational and repair time follow the phase type distributions, the stochastic process \( \{X(t), t \geq 0\} \) is not a Markov process. According to the assumption at Section 3.1, the Fig. 1 can describe the state transitions of the system.

4.1 State space analyses of system

In order to analyze the system states effectively, the system states are defined by the multidimensional state space and are discussed in three cases respectively: no failed unit, \( i (1 \leq i \leq h) \) failed units, and \( h + 1 \) failed units. In this paper the total number of failed and repairing units is called macro-state.

Throughout this paper the dimension of matrix and vector are denoted by the macro-state, that is, the dimension of the matrix is the number of the macro-state.
4.1.1 No failed unit

When there is no failed unit in the system, the system state can be described by a two dimensional vector \((0, j)\). The first entry \(0\) denotes no failed unit, so the system stays at macro-state 0; the second entry \(j\) \((1 \leq j \leq m)\) denotes the working phase of the online unit. We use that \(E_1 = \{(0, j), 1 \leq j \leq m\}\) denotes the state set at macro-state 0 of the system.

4.1.2 \(i (1 \leq i \leq h)\) failed units

When there are \(i\) \((1 \leq i \leq h)\) failed units, the system state can be denoted by a three dimensional vector \((i, j, k)\). The first entry \(i\) \((1 \leq i \leq h)\) denotes the total number of failed and repairing units, so the system stays at macro-state \(i\); the second entry \(j\) \((1 \leq j \leq m)\) denotes that the online unit occupies the work phase; the third entry \(k\) \((1 \leq k \leq n)\) denotes the repair phase occupied by the unit under repair. Currently, the system state space can be denoted by \(E_2 = \{(i, j, k), 1 \leq i \leq h, 1 \leq j \leq m, 1 \leq k \leq n\}\).

4.1.3 \(h + 1\) failed units

When there are \(h + 1\) failed units, the system will fail and the system state space can be denoted by a two dimensional vector \((h + 1, k)\). The first entry \(h + 1\) denotes that the system state occupies macro-state \(h + 1\); the second entry \(k\) \((1 \leq k \leq n)\) denotes the repair phase occupied by the unit under repair. The system state space is denoted by \(E_3 = \{(h + 1, k), 1 \leq k \leq n\}\).

4.2 System state transition matrix

On constructing the system state transition matrix, we order the states lexicographically, \(i.e.,\)

\[
(0, 1), \ldots, (0, m), (1, 1, 1), \ldots, (1, 1, n), (1, 2, 1), \ldots, (1, 2, n), \\
\ldots, (1, m, 1), \ldots, (1, m, n), \ldots, (h + 1, 1), \ldots, (h + 1, n)
\]

4.2.1 Transitions from macro-state \(i\) to macro-state \(i\)

1) \(i = 0\)

The transitions from macro-state 0 to macro-state 0 are the internal transitions. For the moment there is no failed unit in the system, so this transition can be denoted by \(T\).

2) \(1 \leq i \leq h\)

The transitions from macro-state \(i\) to macro-state \(i\) are the internal transitions. For the moment there are \(i\) failed units in the system; because the working phase and the repair phase cannot transit simultaneously, this transition can be denoted by \(T \oplus S\).

3) \(i = h + 1\)

The transitions from macro-state \(i\) to macro-state \(i\) are the internal transitions. For the moment the system fails; the transition, denoted by \(S\), occurs between the repair phases.

4.2.2 Transitions from macro-state \(i\) to macro-state \(i + 1\)

1) \(i = 0\)

When the system occupies the macro-state 0, there is no failed unit so the repairman is idle. For the moment the state transitions, denoted by \(T^0 \alpha \otimes \beta\), is from state \((0, j)\) to state \((1, j, k)\).

2) \(1 \leq i \leq (h - 1)\)

When the system occupies the macro-state \(i\) \((1 \leq i \leq h)\), there are \(i\) failed units. For the moment the state transitions, denoted by \(T^0 \alpha \otimes I\), is from state \((i, j, k)\) to state \((i + 1, j', k)\).

3) \(i = h\)

When the system occupies the macro-state \(h\), there are \(h\) failed units. For the moment the state transitions, denoted by \(T^0 \otimes I\), is from state \((h, i, j)\) to state \((h + 1, k)\).
4.2.3 Transitions from macro-state $i$ to macro-state $i - 1$

1) $i = 1$

When the system occupies the macro-state 1, there is 1 failed unit. For the moment the state transitions, denoted by $I \otimes S^0$, is from state $(1, j, k)$ to state $(0, j)$.

2) $2 \leq i \leq h$

When the system occupies the macro-state $i$, there are $i$ failed units. For the moment the state transitions, denoted by $I \otimes S^0 \beta$, is from state $(i, j, k)$ to state $(i - 1, j, k')$.

3) $i = h + 1$

When the system occupies the macro-state $h + 1$, the system fails. For the moment the state transitions, denoted by $\alpha \otimes S^0 \beta$, is from state $(i, k)$ to state $(i - 1, j, k')$.

Through the above analyses, the system state transition matrix can be denoted by $Q$.

$$
Q = \begin{bmatrix}
B_{00} & B_{01} & 0 & 0 & 0 & \cdots & 0 \\
B_{10} & A_1 & A_0 & 0 & 0 & \cdots & 0 \\
0 & A_2 & A_1 & A_0 & 0 & \cdots & 0 \\
& \vdots & \vdots & \ddots & \vdots & \ddots & \cdots & 0 \\
0 & 0 & 0 & A_2 & A_1 & A_0 & 0 \\
0 & 0 & 0 & 0 & A_2 & A_1 & B_{h,h+1} \\
0 & 0 & 0 & 0 & 0 & B_{h+1,h} & B_{h+1,h+1}
\end{bmatrix}_{(h+2) \times (h+2)}
$$

where $B_{00} = T$, $B_{01} = T^0 \alpha \otimes \beta$, $B_{10} = I \otimes S^0$, $A_0 = T^0 \alpha \otimes I$, $A_1 = T \otimes S$, $A_2 = I \otimes S^0 \beta$, $B_{h,h+1} = T^0 \otimes I$, $B_{h+1,h} = \alpha \otimes S^0 \beta$, $B_{h+1,h+1} = S$.

From the structure of matrix $Q$, we can see that matrix $Q$ has the same structure, i.e., the block tridiagonal structure, as the generator of QBD. So in this paper we call the stochastic process model \{\textbf{$X(t)$, $t \geq 0$}\}, which can be described by QBD, as repairable spare demand model based on QBD.

5 Model analyses

Because the system state transition process can be described by multidimensional Markov process based on QBD, the model analyses can be solved by multidimensional Markov process.

5.1 Important conclusions

**Theorem 1.** The first passage time distribution to be out of spares follows phase type distribution with irreducible representation $(\gamma, L)$, where $\gamma = (\alpha, 0, 0, \cdots, 0)_{1 \times (h+1)}$, and

$$
L = \begin{bmatrix}
B_{00} & B_{01} & 0 & 0 & 0 & \cdots & 0 \\
B_{10} & A_1 & A_0 & 0 & 0 & \cdots & 0 \\
0 & A_2 & A_1 & A_0 & 0 & \cdots & 0 \\
& \vdots & \vdots & \ddots & \vdots & \ddots & \cdots & 0 \\
0 & 0 & 0 & \cdots & A_2 & A_1 & A_0 \\
0 & 0 & 0 & \cdots & 0 & A_2 & A_1
\end{bmatrix}_{(h+1) \times (h+1)}
$$

**Proof.** Let random variable $T$ denote the first passage time distribution to be out of spares; then $T$ represents the first passage time distribution from macro-state 0 to macro-state $h + 1$.

If we define macro-state $h + 1$ as one dimensional absorb state, the system state transition matrix can be represented by $Q'$:

$$
Q' = \begin{bmatrix}
B_{00} & B_{01} & 0 & 0 & 0 & \cdots & 0 \\
B_{10} & A_1 & A_0 & 0 & 0 & \cdots & 0 \\
0 & A_2 & A_1 & A_0 & 0 & \cdots & 0 \\
& \vdots & \vdots & \ddots & \vdots & \ddots & \cdots & 0 \\
0 & 0 & \cdots & A_2 & A_1 & A_0 & 0 \\
0 & 0 & \cdots & 0 & A_2 & A_1 & T^0 \otimes e \\
0 & 0 & \cdots & \cdots & 0 & 0 & 0
\end{bmatrix}_{(h+2) \times (h+2)}
$$
According to the definition of the phase type distributions, we can see that $T$ follows phase type distribution with representation $(\gamma, L)$, where $\gamma = (\alpha, 0, 0, \cdots, 0)_{(1 \times (h+1))}$; $L$ results from eliminating the last row and column in $Q'$, namely $L = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & \cdots & 0 \\ B_{10} & A_1 & A_0 & 0 & \cdots & 0 \end{bmatrix}
\begin{bmatrix} 0 & A_2 & A_1 & A_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \cdots & A_2 & A_1 \end{bmatrix}_{(h+1) \times (h+1)} \begin{bmatrix} 0 & \cdots & 0 & A_2 & A_1 \end{bmatrix}_1$.

Because the system can only occupy some phase of macro-state 0 at time 0, vector $\gamma$ equals to $(\alpha, 0, 0, \cdots, 0)_{(1 \times (h+1))}$.

**Theorem 2.** Let $T_m$ denote the first mean time to be out of spares. Then $T_m \equiv \gamma(-L^{-1})e$, where $\gamma$ and $L$ are the same as Theorem 1, $e$ denotes a column vector with all entries equal to one, whose dimension is determined by $\gamma(-L^{-1})$.

**Proof.** According to Lemma 1.3.2 in [9], for order $m$ phase type distribution, let $\alpha_j$ denote the expected sojourn time at transient state $j$; then $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_m) = \alpha(-T)^{-1}$. The first mean time to be out of spares equals to the sum of the expected sojourn time at $m$ transient states, so $T_m = \gamma(-L^{-1})e$.

**Theorem 3.** Assume that in order to ensure the spare provision with a probability of not less than $P$ during the period of $t$, $h$ spares should be prepared at least. $L, t$ and $P$ satisfy the equation $\gamma \exp(L)e > P$, where $L$ is $h + 1$ dimensions macro-state matrix.

**Proof.** Let $T$ denote the first mean time to be out of spares. According to Theorem 1 we can get $P(T \leq t) = 1 - \gamma \exp(L)e$. Hence $1 - P(T \leq t) = P(T > t) = \gamma \exp(L)e$. To satisfy $P(T > t) > P$, we have $\gamma \exp(L)e > P$.

### 5.2 The least spare part demand algorithm under given restrictions

According to Theorem 3, in this section we give the least spare part demand algorithm, which can decide the least spare part demand to ensure the spare provision with a probability of not less than $P$ during a period of $t$.

The algorithm is realized as the following three steps:

**Step 1.** Let $h = 0$, $L = B_{00} = T$, $\gamma = \alpha$, and calculate $\gamma \exp(L)e$. If $\gamma \exp(L)e > P$ then it does not need any spare and the algorithm ends, else go to Step 2.

**Step 2.** Let $h = 1$, $L = \begin{bmatrix} B_{00} & B_{01} \\ B_{10} & A_1 \end{bmatrix}$, $\gamma = (\alpha, 0)$, and calculate $\gamma \exp(L)e$. If $\gamma \exp(L)e > P$ then it needs just one spare and the algorithm ends, else go to Step 3.

**Step 3.** Let $h = h + 1$, $L = L + \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & A_0 \\ 0 & \cdots & A_2 & A_1 \end{bmatrix}_{(h+1) \times (h+1)}$, $\gamma = (\alpha, 0, \cdots, 0)_{1 \times (h+1)}$. Calculate $\gamma \exp(L)e$. If $\gamma \exp(L)e > P$ then it needs $h$ spares and the algorithm ends, else execute Step 3 again.

### 6 A numerical example

Based on the previous model, a numerical example is given in this section. The phase type in this example refers to Rafael[11].

Let the operational time of the unit follows the phase type distribution with representation $(\alpha, T)$, where $\alpha = (1, 0, 0)$, $T = \begin{bmatrix} -0.0027 & 0.0027 & 0 \\ 0 & -0.008 & 0.008 \\ 0 & 0 & -0.02878 \end{bmatrix}$, $T^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Let the repair time of the unit follows the phase type distribution with representation $(\beta, S)$, where $\beta = (1, 0, 0)$, $S = \begin{bmatrix} -0.02 & 0.02 & 0 \\ 0.01 & -0.08 & 0.07 \\ 0.005 & 0 & -0.1 \end{bmatrix}$, $S^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

We use the algorithm of section 5.2 to calculate the spare part demand which can ensure the spare provision with a probability of not less than 0.95 during a period of 1500 hours. In this paper we use Matlab to implement the algorithm.
1) No spare
When there is no spare, we have \( L = T \) and \( \gamma \exp(Lt)e = 0.0290 \), that is, no spare can just ensure the spare provision with a probability of not less than 0.0290 during a period of 1500 hours.

2) One spare
One spare can ensure the spare provision with a probability of not less than 0.8995 during a period of 1500 hours.

3) Two spares
Two spares can ensure the spare provision with a probability of not less than 0.9984 during a period of 1500 hours.

From the above result, we can draw a conclusion that in order to ensure the spare provision with a probability of not less than 0.95 during a period of 1500 hours, two spares should be provided at least.

7 Conclusions
In this paper, we analyze the problems of the repairable spare demand model, propose that use phase type distribution as the basic assumption of the operational and work time distribution of the repairable unit, and develop the repairable spare part demand model based on phase type distribution. The model adapts to all kinds of the repairable unit operational and repair time probability distribution. The calculation of the model is mainly involved in matrix operation. Large matrix operations can be implemented efficiently, so the model has a better applied value.

8 Acknowledgements
We thank the two referees for their thorough reading and deep insight.

References
8 Chen S L. The spare part demand of air force missile testing equipment. Aeronautica Weapons, 2002, (6): 12~16

HUANG Zhuo Ph. D. candidate at National University of Defense Technology. His research interests include system reliability and equipment integrate support.

LIANG Liang Ph. D. candidate at National University of Defense Technology. His research interests include system reliability and flexible manufactory system.

GUO Bo Professor at National University of Defense Technology. His research interests include system reliability, equipment integrate support, system planning, and project management.