On lifetime optimization of Boolean parallel systems with Erlang repair distributions

Alexander Gouberman and Markus Siegle

We assume a Boolean parallel system with exponentially distributed component failure times. In order to maximize the lifetime of the system we consider a repairman with Erlang-k distributed repair time. By extending the classical exponential case $k = 1$ to $k \geq 2$ different repair semantics arise in this context. In the case of restart repair semantics we show that a repairman should have few Erlang phases in order to maximize mean time to failure (MTTF). In the case $k \geq 2$ with full memory semantics an optimal repairman policy shows counter-intuitive repair behaviour dependent on the mean repair time.

1 Introduction

The field of research for system optimization covers a wide range of models and specialized optimization algorithms. For the subclass of coherent Boolean systems, with which we are concerned here, in [6, 1] the authors maximized the availability of a series system by assigning the repairman to the most reliable failed component (MRC policy). In [2] a generalization of this result to K-out-of-N systems and a repair team consisting of several repairmen was established. The assumption of both exponential component failure rate and repair rate makes it possible to provide an optimal policy which depends only on the order of failure resp. repair rates and not on their concrete values: the fastest repairman should repair the most reliable component. While this assumption may hold for components due to the memoryless property of the exponential distribution (if no concrete component model is known), it is used for the repairman in general just for reasons of model simplification. Repair time distributions which follow the typical "S"-shape like Weibull or Lognormal distributions or even deterministic repair time are known to be more realistic because they often fit to empirical repairman data [3]. Erlang distributions

Alexander Gouberman, Markus Siegle – Fakultät für Informatik, Universität der Bundeswehr München, e-mail: alexander.gouberman@unibw.de, markus.siegle@unibw.de

with a small number of phases can be used to approximate these repair distributions. We extend the exponential repair model to Erlang repair distributions, whereby first of all some repair semantics concerning the Erlang phases have to be chosen. We concentrate on parallel systems and show that for the "restart repair semantics" the fewer phases an Erlang distribution has, the better it maximizes the lifetime of the system. In the case of a "full memory semantics" model we show by a small case study that the optimal policy which assigns the repairman to a failed component depends on the concrete values of the component failure rates and mean Erlang repair time in a somewhat counter-intuitive manner.

2 Markov Decision Model

A parallel system with N components where the lifetime of the i-th component is exponentially distributed with parameter $\mu_i > 0$ can be described by a CTMC with state space $S = \{0,1\}^N$. A system state is given by $x = (x_1,\dots,x_N)$ where $x_i = 1$ or 0 depending whether the i-th component is functioning or not. Define

$$C_0(x) := \{i \mid x_i = 0\} \quad \text{and} \quad C_1(x) := \{i \mid x_i = 1\}$$

for a state $x \in S$ as the sets of nonfunctioning resp. functioning components. From a state x there are $|C_1(x)|$ transitions given by $x \xrightarrow{\mu_k} (0_k,x),\ k \in C_1(x)$ with rate μ_k, where $(\delta_k,x) := (x_1,\dots,x_{k-1},\delta,x_{k+1},\dots,x_N),\ \delta \in \{0,1\}$ denotes the state, where the entry corresponding to the k-th component is set to δ. The single absorbing state of the CTMC is $(0,\dots,0)$ which represents the failure of the whole parallel system. As a reliability measure for the lifetime of the system we analyze in the following the mean time to failure (MTTF) and assume that the system starts in state $(1,\dots,1)$. In order to maximize MTTF by assigning a single repairman with exponential repair time distribution $Exp(\lambda)$ a continuous time Markov decision process (CTMDP) is induced, where the action set in state $x \in S$ is given by $Act(x) := \{r_i \mid i \in C_0(x)\} \cup \{nr\}$, r_i representing the choice to repair the failed component i and nr not to repair any component. The transitions of this CTMDP are given by

$$x \xrightarrow{r_i,\ \lambda} (1_i,x), \text{ for } i \in C_0(x) \quad \text{and} \quad x \xrightarrow{a,\ \mu_j} (0_j,x), \text{ for } j \in C_1(x),\ a \in Act(x),$$

meaning that by choosing the repair action r_i the i-th component can be repaired, and by choosing any action a working component can fail.
The reward which measures the mean sojourn time in a state x when choosing action $a \in Act(x)$ is given by $R(x,a) = \frac{1}{E(x,a)}$, where $E(x,a)$ is the exit rate in state x provided action a is chosen, i.e.

$$E(x,a) = \sum_{j \in C_1(x)} \mu_j + \delta(a)\lambda, \quad \delta(a) \in \{0,1\},\ \delta(a) = 0 \Leftrightarrow a = nr.$$

This reward definition implies that the optimal MTTF from state x to the system failure state $(0,\dots,0)$ can be found by solving the Bellman equation induced by this CTMDP. For a detailed discussion on stochastic dynamic programming and especially CTMDPs we refer to [5]. In order to compare exponential and Erlang repair time distributions with the same expected values we propose two different (extremal) repair model semantics:

- Model 1 (”restart repair semantics”): If during repair of a component a further component fails then the reached phase of the Erlang-k distribution is reset and the repairman can choose any failed component to repair beginning from phase 0 again.
- Model 2 (”full memory semantics”): The Erlang phases of all partially repaired components are remembered and the repairman can be assigned in each state to another failed component (even if no further component failed during the repair procedure). Moreover the repairman continues the repair process from the Erlang phase reached so far for that failed component.

These semantics are indeed “extremal”, because one could define other possible repair semantics in between these two. For example, if during a repair procedure of component C a further component fails and the repairman changes to another failed component $\neq C$, then the repair phase of C could get lost. If he does not change he could continue the repair of C from the phase already reached so far.
We now describe the state spaces S_i of model i. In order to compose the system state space S together with an Erlang-k repair state space $E_k := \{0,1,\ldots,k-1\}$ we remember the repair phase for each component, s.t. the composed state space can be described by a subset of $\hat{S} := S \times E_k^N$. For both models 1 and 2 there are states in $\hat{S}$ which are not reachable, more precisely for model 1

$$S_1 = \left\{(x,e) \in \hat{S} \mid x_i = 1 \Rightarrow e_i = 0,\ \exists^{\leq 1} i:\ e_i > 0\right\},$$

meaning that working components cannot be repaired and there is at most one component i which has a positive Erlang repair phase e_i. In the following, we denote a system state by $x \in S$, active resp. failed components by $j \in C_1(x)$ resp. $i \in C_0(x)$ and the Erlang-k parameter by τ. The transitions for model 1 are given by

$$(x,(0,\ldots,0)) \xrightarrow{nr,\mu_j} ((0_j,x),(0,\ldots,0)) \quad \text{and}$$

$$(x,(0,\ldots,e_i,\ldots,0)) \xrightarrow{r_i,\tau} \begin{cases} (x,(0,\ldots,e_i+1,\ldots,0)) & \text{if } e_i < k-1 \\ ((1_i,x),(0,\ldots,0)) & \text{if } e_i = k-1 \end{cases}$$

$$(x,(0,\ldots,e_i,\ldots,0)) \xrightarrow{r_i,\mu_j} ((0_j,x),(0,\ldots,0))$$

In model 2 with full memory semantics there are more reachable states from $\hat{S}$ because Erlang phases are used to model the partially reached repair stage for a failed component. The corresponding state space is given by

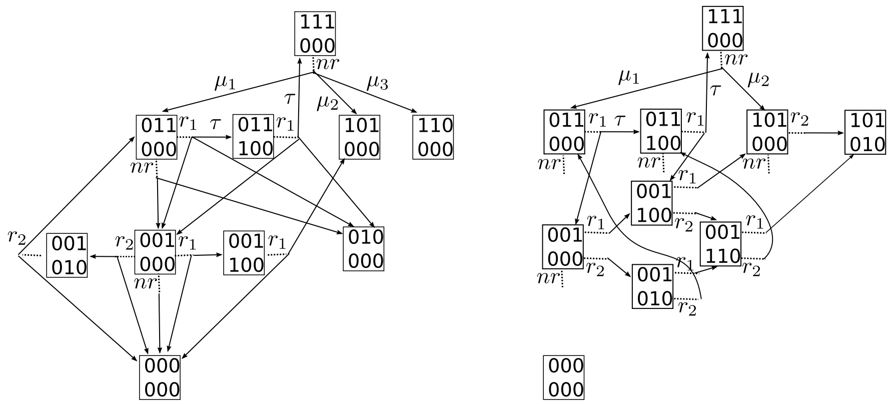
$$S_2 = \left\{(x,e) \in \hat{S} \mid x_i = 1 \Rightarrow e_i = 0\right\}$$

with transitions

$$(x,e) \xrightarrow{a,\mu_j} ((0_j,x),e), \text{ for any } a \in Act(x) \quad \text{and}$$

$$(x,e) \xrightarrow{r_i,\tau} \begin{cases} (x,(e_1,\ldots,e_i+1,\ldots,e_N)) & \text{if } e_i < k-1 \\ ((1_i,x),(e_1,\ldots,e_{i-1},0,e_{i+1},\ldots,e_N)) & \text{if } e_i = k-1 \end{cases}$$

Figure 1 shows an excerpt of the CTMDP for a parallel system with $N = 3$ components and Erlang-2 repairman. The upper line represents a system state $x \in S$ and the number below each component state x_i is its Erlang phase $e_i \in E_k$.



(a) Model 1 (restart repair semantics) (b) Model 2 (full memory semantics)

Fig. 1 Excerpt of the CTMDP for different Erlang-2 repair semantics for a parallel system of 3 components ((b) showing especially transitions for repair procedure which differ from (a)). Dashed lines represent decisions and solid lines exponential transitions.

Let us assume, that the repairman does not change to another failed component before a further functioning component fails (as it is always the case in model 1), see Fig. 2. Then the following proposition holds.

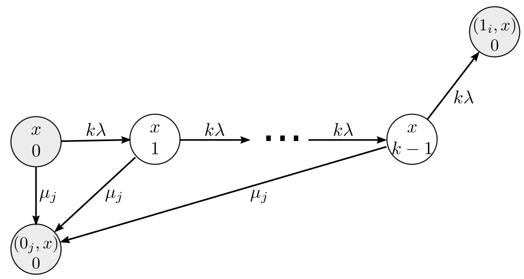


Fig. 2 Local excerpt of the induced CTMC by choosing a repair action r_i in state $x = (0_i, x)$. The lower number indicates the Erlang phase of the i-th component. The shaded states correspond to states in the case of exponential repair time.

Proposition:

Let $x \in S$ and consider Erlang-k distributed repair times $T_k \sim Erl_k(k\lambda)$ with same expected values $\mathbb{E}(T_k) = \frac{1}{\lambda}\ \forall k \in \mathbb{N}$.

(a) The probability to repair a failed component $i \in C_0(x)$ without further failure of another component during the repair procedure is decreasing with k.

(b) The mean time to repair a failed component, under the condition that no further failure of another component during the repair procedure occurs, is increasing with k.

Proof: The probability for following the path

$$\sigma_k := (x,0) \to (x,1) \to \cdots \to (x,k-1) \to ((1_i,x),0)$$

is given by $P_k = \left(\frac{k\lambda}{k\lambda+\mu}\right)^k$, where $\mu = \sum_{j\in C_1(x)} \mu_j$. It is known that $P_k = \left(1+\frac{\mu/\lambda}{k}\right)^{-k}$ is monotonically decreasing (and converging to $e^{-\mu/\lambda}$), thus (a) holds. The mean time to repair a failed component on the path σ_k can be computed by summing up the sojourn times in the states along σ_k: $\mathbb{E}(\sigma_k) = k \cdot \frac{1}{k\lambda+\mu} = \frac{1}{\lambda}\frac{k\lambda}{k\lambda+\mu}$. But since $f(x) := \frac{x}{x+\mu}$, $x \in \mathbb{R}$ is strictly monotonic increasing, statement (b) also holds. □

Special cases: Among all Erlang repair distributions with identical expected values the following holds.

1. An exponentially distributed repair time ($k = 1$) maximizes the probability to repair a failed component and minimizes the mean time to repair it before a further component fails.
2. In case of deterministic repair time (for $k \to \infty$) this probability is minimal $\left(e^{-\mu/\lambda} = P(\text{sojourn time in state } x \text{ is lower than repair time } \frac{1}{\lambda})\right)$ and the corresponding repair time is maximal.

For the restart repair model, exponentially distributed repair time is the best among all Erlang repair time distributions. But since the practical issue of this model is low (and only introduced for comparison with the exponential repairman case) we adhere to the more practical model 2 with full memory semantics, since Erlang phases can approximately describe repair stages during repair of a failed component. For the case $k = 1$ Katehakis and Melolidakis showed, that the MRC policy which assigns the repairman to the failed component with least failure rate maximizes MTTF [2]. We show that for the case $k \geq 2$ in model 2 this is not the case, since repair phases are remembered. Figure 2 shows optimal policies with regard to MTTF maximization in model 2 with $k = 2$ Erlang repair phases for a parallel system with 3 components and component failure rates $\mu_1 = 1,\ \mu_2 = 2,\ \mu_3 = 3$. In this case the optimal policy differs only for state $010-001$ in a noncoherent manner: Relatively ”slow” and ”fast” repairmen should continue to repair the third component, but there is also a region, in which it is better to change repair to component one which is in repair phase 0. The results were computed by implementing the policy iteration method [5] in Wolfram Mathematica 7 and apply it to the CTMDP of model 2.

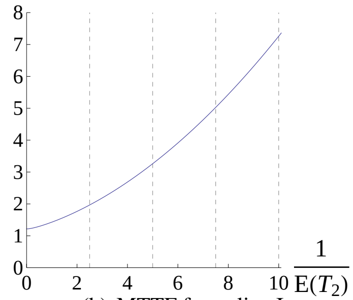
3 Conclusion

We have generalized the classical exponential repairman models to Erlang-k repairman models in order to maximize the MTTF of a Boolean parallel system. In this context we discussed two different semantics which lead to different extensions of the system state space. We showed by optimizing the CTMDP for a toy example with full memory semantics that classical results on optimal repair strategies do not hold any more for Erlang-k repairmen. It would be interesting to see whether avail-

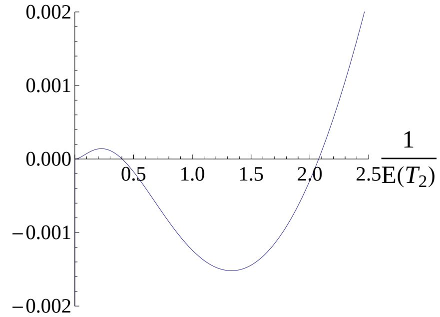
State (x,e)	action set $Act(x,e)$	policy I	policy II
111-000	nr	nr	nr
011-000	nr, r_1	r_1	r_1
011-100	nr, r_1	r_1	r_1
101-000	nr, r_2	r_2	r_2
101-010	nr, r_2	r_2	r_2
110-000	nr, r_3	r_3	r_3
110-001	nr, r_3	r_3	r_3
001-000	nr, r_1, r_2	r_1	r_1
001-100	nr, r_1, r_2	r_1	r_1
001-010	nr, r_1, r_2	r_2	r_2
001-110	nr, r_1, r_2	r_1	r_1
010-000	nr, r_1, r_3	r_1	r_1
010-100	nr, r_1, r_3	r_1	r_1
010-001	$\mathbf{nr, r_1, r_3}$	$\mathbf{r_3}$	$\mathbf{r_1}$
010-101	nr, r_1, r_3	r_1	r_1
100-000	nr, r_2, r_3	r_2	r_2
100-010	nr, r_2, r_3	r_2	r_2
100-001	nr, r_2, r_3	r_3	r_3
100-011	nr, r_2, r_3	r_2	r_2

mean repair time $\mathbb{E}(T_2)$	4	1	0.4	0.1
optimal MTTF	1.244	1.429	1.966	7.261
optimal policy	I	II	I	I

(a) Optimal policies with mean repair time $\mathbb{E}(T_2)$



(b) MTTF for policy I



(c) Difference of MTTF values for policies I and II

Fig. 3 Optimal policies for a parallel system with 3 different components and Erlang-2 distributed repair time T_2 with full memory semantics

able measures for component importance (like Barlow-Proschan or Natvig [4]) lead to good heuristics or even optimal repair policies.

References

1. Michael N. Katehakis and Cyrus Derman. Optimal repair allocation in a series system. *Mathematics of Operations Research*, 9(4):615–623, 1984.
2. Michael N. Katehakis and Costis Melolidakis. Dynamic repair allocation for a k-out-of-n system maintained by distinguishable repairmen. *Probability in the Engineering and Informational Sciences*, 2:51–62, 1988.
3. R.M. Kieckhafer, M.H. Azadmanesh, and Y. Hui. On the sensitivity of NMR unreliability to non-exponential repair distributions. In *Fifth IEEE International Symposium on High Assurance Systems Engineering*, pages 293–300, 2000.
4. Bent Natvig and Jorund Gasemyr. New results on the Barlow-Proschan and Natvig Measures of Component Importance in Nonrepairable and Repairable Systems. *Methodology and Computing in Applied Probability*, 11:603–620, 2009.
5. Martin L. Puterman. *Markov Decision Processes*. John Wiley & Sons INC., 1994.
6. Donald R. Smith. Optimal repair of a series system. *Operations Research*, 26(4):653–662, 1978.