

혼합 와이불 NHPP 모형에 근거한 소프트웨어 최적방출시기에 관한 연구

이 상 식* · 김 희 철**

A Study on Optimal Release Time for Software Systems based on Mixture Weibull NHPP Model

Lee, Sang Sik · Kim, Hee Cheul

〈Abstract〉

Decision problem called an optimal release policies, after testing a software system in development phase and transfer it to the user, is studied. The applied model of release time exploited infinite non-homogeneous Poisson process. This infinite non-homogeneous Poisson process is a model which reflects the possibility of introducing new faults when correcting or modifying the software. The failure life-cycle distribution used mixture which has various intensity, if the system is complicated. Thus, software release policies which minimize a total average software cost of development and maintenance under the constraint of satisfying a software reliability requirement becomes an optimal release policies. In a numerical example, after trend test applied and estimated the parameters using maximum likelihood estimation of inter-failure time data, estimated software optimal release time.

Key Words : Software Release Policies, NHPP, Mixture, Optimal Release

I. 서론

소프트웨어 방출시간에 대한 연구들은 대부분 유한 고장 NHPP(Non-Homogeneous Poisson Process)모형을 사용하였다[1-2]. 이러한 유한(Finite)고장 NHPP모형은 소프트웨어가 유한개의 고장이 있고 고장 제거 단계에서는 새로운 고장이 발생하지 않는다는 가정을 한 모형이다. 그러나 실제 고장 제거 단계에서도 새로운 고장이 발

생 할 수 있다. 따라서 본 연구에서는 무한(Infinite) 고장 NHPP 모형을 이용하여 최적 방출시기에 대한 문제를 제안하고자 한다.

이 분야에서는 Musa-Okumoto의 대수 포아송 실행시간 모형[3-4]과 로그-파우어 모형[3, 5]을 이용한 방출 문제에 대한 문제 들이 이미 연구되었고 최근까지도 이와 관련된 연구는 Yang과 Xie[6], Huang[7] 등에 의해 보증 기간을 포함한 방출시기 등에 대하여 연구되고 있다[8].

본 연구에서는 소프트웨어의 결함을 제거하거나 수정 작업 중에도 새로운 결함이 발생할 가능성이 있는 무한

* 송호대학 보건의료전자과(제1저자)

** 남서울대학교 산업경영공학과(교신저자)

고장수를 가진 비동질적인 포아송 과정에 기초하고 시스템이 복잡해지면 하나의 강도함수에만 고장이 일어나지 않고 다양한 원인에 의해 혼합되어 발생 할 수 있는 혼합모형을 이용한 방출시기에 관한 문제를 다루었다.

따라서 (1)식 과 (3)식을 연관시키고 $f(t)$ 을 확률밀도 함수, $F(t)$ 을 분포함수라고 하면 다음과 같은 관계식에 의해 NHPP의 강도함수는 $F(t)$ 의 위험함수($h(t)$)가 된다.

$$\lambda(t) = m'(t) = f(t)/(1-F(t)) = h(t) \quad (4)$$

II. 관련연구

2.1 무한고장 NHPP

NHPP 모형에서 평균값 함수 $m(t)$ (Mean value function)와 강도 함수(Intensity function) $\lambda(t)$ 는 다음과 같은 관계로 표현할 수 있다[2, 9].

$$m(t) = \int_0^t \lambda(s)ds, \quad \frac{d m(t)}{dt} = \lambda(t) \quad (1)$$

따라서 $N(t)$ 는 모수 $m(t)$ 을 가진 포아송 확률 밀도 함수(Probability density function)로 알려져 있다. 즉,

$$P\{N(t) = n\} = \frac{[m(t)]^n \cdot e^{-m(t)}}{n!}, \quad n = 0, 1, 2, \dots, \infty \quad (2)$$

이처럼 시간 관련 모형(Time domain model) 들은 NHPP에 의해서 확률 고장 과정으로 설명이 가능하다. 이러한 모형들은 고장 강도 함수 $\lambda(t)$ 가 다르게 표현됨 으로서 평균값 함수 $m(t)$ 도 역시 다르게 나타나고 이러한 NHPP 모형들은 유한 고장 모형과 무한 고장 범주로 분류한다[10].

수리 시점에서도 고장이 발생할 수도 있는 상황을 반영하기 위하여 RVS(Record Value Statistics)모형을 사용하는 NHPP 모형의 평균값함수는 다음과 같이 알려져 있다[2, 10].

$$m(t) = -\ln(1-F(t)) \quad (3)$$

시간 $(0, t]$ 까지 조사하기 위한 시간 절단(Time truncated) 모형은 n 번째 까지 고장 시점 자료를

$$x_k = \sum_{i=1}^k t_k \quad (k=1, 2, \dots, n; 0 \leq x_1 \leq x_2 \leq \dots \leq x_n) \quad (5)$$

이라고 하고 θ 을 모수공간이라고 하면 우도 함수는 다음과 같이 알려져 있다[1, 10].

$$L_{NHPP_{n:n}}(\theta | \underline{x}) = \prod_{i=1}^n \left(\frac{f(x_i)}{1-F(x_i)} \right) (1-F(x_n)) \quad (6) \\ = \left(\prod_{i=1}^n h(x_i) \right) (1-F(x_n))$$

단, $1-F(T) = e^{-m(t)}, \underline{x} = (x_1, x_2, x_3, \dots, x_n]$.

2.2 혼합 NHPP 모형

부품이나 시스템집단의 수명분포가 서로 다른 k 개의 부분집단으로 구성된 경우에 적용 가능한 모형이다[11, 12]. 예를 들어 전체집단이 결합부품이나 제조상의 결함을 지닌 초기고장 집단과 정상집단의 두 그룹으로 구성 되는 경우도 혼합모형에 해당된다. 이러한 혼합모형은 분포함수와 확률밀도함수는 여러분포의 가중합 형태로 다음과 같이 표현 된다

$$F(t|\theta, p) = \Pr(T < t) = \sum_{i=1}^k p_i F_i(t|\theta_i) \quad (7)$$

$$f(t|\Theta, p) = \sum_{i=1}^k p_i f_i(t|\theta_i) \quad (8)$$

위 식에서 $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$ 은 미지의 모수 집합을 의미하고 모집단의 $f_i(t|\theta_i)$, $i=1, \dots, k$ 는 부분집단 i 에 속하는 부품의 고장시간에 대한 확률밀도함수이며 혼합모수(Mixing parameter) $p = (p_1, \dots, p_k)$ 는 다음을 만족한다.

$$\sum_{i=1}^k p_i = 1, \quad 0 \leq p_i \leq 1 \quad (9)$$

본 논문에서는 이러한 사실로부터 NHPP의 강도함수와 평균값 함수가 $k=2$ ($p_1 = p, p_2 = 1-p$)인 경우 다음과 같이 혼합되어 있는 경우를 고려하고자 한다.

$$\lambda_{Mix}(t|\Theta, p) = p \lambda_1(t|b_1) + (1-p) \lambda_2(t|b_2) \quad (10)$$

$$m_{Mix}(t|\Theta, p) = p m_1(t|b_1) + (1-p) m_2(t|b_2) \quad (11)$$

(6)식을 이용하면 혼합 우도 함수는 다음과 같이 표현할 수 있다.

$$L_{NHPP, Mix}(\Theta | \underline{x}) = \prod_{i=1}^n (p \lambda_1(x_i) + (1-p) \lambda_2(x_i)) e^{-(p m_1(t) + (1-p) m_2(t))} \quad (12)$$

단, $\underline{x} = (x_1, x_2, x_3, \dots, x_n)$.

2.2.1 단순 모형

와이블 분포는 감마 분포와 유사하게 분포의 적용성 때문에 신뢰도에서 중요한 역할을 한다. 신뢰도 함수에 포함된 모수로서 다른 분포들을 표현하거나 데이터를 적용하기 쉽게 변형 시킬 수 있다. 이러한 와이블 분포는 $a(>0)$ 가 형상 모수, $b(>0)$ 는 척도 모수일 때, 확률 밀도 함수와 분포함수는 각각 다음과 같다[13].

$$f_{Weib}(t|a, b) = b a t^{a-1} e^{-b t^a} \quad (13)$$

$$F_{Weib}(t|a, b) = 1 - e^{-b t^a} \quad (14)$$

(13), (14)식에서 $a=1$ 인 경우는 지수모형이 되고 $a=2$ 인 경우는 Rayleigh 모형이 된다.

따라서 무한 고장 NHPP 모형의 평균값 함수 $m(t) = -\ln(1-F(t)) = \beta t^a$ 가 되고 강도함수는 다음과 같이 표현된다.

$$\lambda(t) = m'(t) = f(t)/(1-F(t)) = a b t^{a-1} \quad (15)$$

$$m(t) = -\ln(1-F(t)) = b t^a \quad (16)$$

이러한 와이블 분포는 $0 < a < 1$ 일 경우에는 (오류 수정 단계에서 발생하는) 감소하는 강도함수를 표현하고 $a=1$ 인 경우에는 (정상 운영 단계에서 발생하는) 일정한 강도함수를 나타내며 $a > 1$ 인 경우에는 (마모 단계에서 발생하는) 증가하는 강도함수를 반영 한다[13].

최우추정법(MLE)을 이용하기 위한 로그 우도 함수는 (6)식과 (16)식을 연관 하면 다음과 같이 유도된다.

$$\ln L(a, b | \underline{x}) = n \ln a + n \ln b + (a-1) \sum_{i=1}^n \ln x_i - b x_n^a \quad (17)$$

(17)식에서 a 을 고정하고 $\frac{\partial \ln L(a, b | \underline{x})}{\partial b} = 0$ 을 만족하는 \hat{b}_{MLE} 는 다음과 같이 구할 수 있다[13]

$$\hat{b}_{MLE} = \frac{n}{x_n^a} \quad (18)$$

2.2.2 지수분포와 Rayleigh분포의 혼합 모형

와이블 분포에서 형상모수 $a=1$ 인 지수분포와 $a=2$ 을 따르는 Rayleigh분포의 혼합모형에 대한 평균값 함수

는 평균값 함수 $m_{Mix}(t|\theta, p)$ 와 고장발생률(강도함수) $\lambda_{Mix}(t|\theta, p)$ 은 (10), (11)식과 관련하면 다음과 같이 표현 할 수 있다.

$$m_{Mix}(t|\theta, p) = p m_1(t|b_1) + (1-p) m_2(t|b_2) \quad (19)$$

$$\lambda_{Mix}(t|\theta, p) = p \lambda_1(t|b_1) + (1-p) \lambda_2(t|b_2) \quad (20)$$

단, p 는 혼합모수이고,

$$m_1(t|b_1) = b_1 t, \quad m_2(t|b_2) = b_2 t^2,$$

$$\lambda_1(t|b_1) = b_1, \quad \lambda_2(t|b_2) = 2b_2 t.$$

단, $p, b_1, b_2, t \geq 0$.

t 을 최종 고장시점 x_n 으로 대치하고 (12)식에 (19)식과 (20)식을 관련시킨 지수분포와 Rayleigh분포의 혼합 밀도함수를 이용한 로그우도함수를 정리하면 다음과 같이 표현된다.

$$\ln L_{Mix}(\underline{x} | p, b_1, b_2) \quad (21)$$

$$= \sum_{i=1}^n \ln(p b_1 + 2(1-p) b_2 x_i) - (p b_1 x_n + (1-p) b_2 x_n^2)$$

혼합모수 p 는 0에서부터 1까지를 고려하면 최우추정법을 이용하기 위한 로그우도함수식을 b_1 과 b_2 대하여 편미분을 하면 다음과 같은 식을 유도 할 수 있다.

$$\frac{\partial \ln L_{Mix}(\underline{x} | b_1, b_2, p)}{\partial b_1} = \quad (22)$$

$$\sum_{i=1}^n \frac{p}{p b_1 + 2(1-p) b_2 x_i} - p x_n = 0$$

$$\frac{\partial \ln L_{Mix}(\underline{x} | b_1, b_2, p)}{\partial b_2} = \quad (23)$$

$$\sum_{i=1}^n \frac{2(1-p) x_i^2}{p b_1 + 2(1-p) b_2 x_i} - (1-p) x_n^2 = 0$$

(22)식과 (23)식에서 혼합분포의 모수 p 의 값을 고려하면 비선형 연립방정식(수치 해석적 방법)을 이용하여 최우추정치 $\hat{b}_{1,MLE}, \hat{b}_{2,MLE}$ 의 값을 추정 할 수 있다.

III. 신뢰도 및 비용최소화를 고려한 방출시간

NHPP 모형에서 테스트 시점 x_n (마지막 고장시점)에서 소프트웨어 고장이 일어난다고 하는 가정 하에서 신뢰구간 $(x_n, x_n + x)$ (단, x 는 임무시간(Mission time)동안 소프트웨어의 고장이 일어나지 않을 확률인 신뢰도 $\hat{R}(x | x_n)$ 는 다음과 같이 됨이 알려져 있다[6, 10, 14].

$$\hat{R}(x | x_n) = \exp\left(-\int_{x_n}^{x_n+x} \lambda(\tau) d\tau\right) \quad (24)$$

$$= \exp[-\{m(x+x_n) - m(x_n)\}]$$

따라서 Rayleigh 분포 모형의 평균값 함수는 $m(t) = b_2 t^2$ 이므로 신뢰도는 (20)식과 (24)식에서 $t = x_n$ 라고 하면 다음과 같이 표현된다.

$$R(x | t) = \exp[-b_2(x+t)^2 + b_2(t)^2] \quad (25)$$

따라서 소프트웨어 방출시간 T_R 이 신뢰도 $R_0 = \hat{R}(x | T_R)$ 을 확보해야 한다면 다음 방정식을 만족해야 한다.

$$\ln R_0 = -b_2[(x+T_R)^2 - (T_R)^2] \quad (26)$$

단, 신뢰도 $R_0 = \hat{R}(x | T_R)$.

한편, 소프트웨어 방출시간을 T 로 표현하고 $m(T)$ 와 $m(\infty)$ 을 각각 $(0, T)$ 와 $(0, \infty)$ 의 기간에 발견된 기대

고장수라고 표현하고 $C(T)$ 을 소프트웨어 라이프사이클 (life cycle) 동안에 기대되는 소프트웨어 비용이라고 하면 $C(T)$ 는 다음과 같이 표현된다[5, 13].

$$C(T) = c_1 m(T) + c_2 [m(\infty) - m(T)] + c_3 T \quad (27)$$

위 식에서 c_1 는 테스트링 동안에 하나의 고장을 수리하는 비용이고 c_2 가동 중에 하나의 고장을 수리하는 비용 ($c_2 > c_1$), 그리고 c_3 는 단위 시간당 테스트링 비용을 나타낸다. 이와 관련하여 총비용의 최소화는 무한고장 평균값 함수를 가진 NHPP 모형에 대하여 발생 할 수 있다. 무한 수명에 대한 비용함수 $C(T)$ 식인 (27)식에서 $m(\infty)$ 은 직접 추정 할 수 없기 때문에 이 식을 사용하기 위해서는 소프트웨어 수명시간 인 T_{LC} 을 지정하여 분석한다[5]. 이러한 T_{LC} 는 소프트웨어마다 서로 다른 임의의 값이기 때문에 유한 고장 NHPP 모형이라고 할 수는 없다.

따라서 비용함수를 고려하여 소프트웨어의 모든 수명에서 총비용을 최소화함으로서 최적 테스트링 시간을 결정 할 수 있고 다음과 같은 식을 만족하면 비용함수 $C(T)$ 는 유일한 최소값을 가진다[5, 13].

$$\frac{\partial C(T)}{\partial T} = 0, \quad \frac{\partial^2 C(T)}{\partial^2 T} > 0 \quad (28)$$

결국 Rayleigh 분포에 의한 최소비용 관련 방출시간은 (27)식을 연관하여 다음과 같이 유도 된다.

$$C(T) = c_1 m(T) + c_2 [m(\infty) - m(T)] + c_3 T \quad (29)$$

$$= (c_1 - c_2)(b_2 T^2) + c_2(b_2 T_{LC}^2) + c_3 T$$

T 에 관해서 비용함수 $C(T)$ 을 미분하면 다음과 같은 방정식을 만족하는 최적방출시간 T_C 를 계산 할 수 있다.

$$2(c_1 - c_2)b_2 T_C + c_3 = 0 \quad (30)$$

위 식에서 최적방출시간은 소프트웨어 지정 수명시간 인 T_{LC} 와 의존하지 않는다는 것을 알 수 있다. 이러한 사실은 무한 고장 평균값 함수를 가진 NHPP모형들이 새로운 결점들이 발생함으로서 몇 개의 고장이 야기 될 수 있는 점을 고려한 모형으로 적합 시킬 수 있다[13].

실제로 만족할 만한 신뢰도가 부여되고 동시에 시스템 고장과 연계된 기대 총비용을 최소화시키기 위하여 필요하다면 충분한 테스트링을 계속해야 한다. 따라서 신뢰성 요구를 만족하고 총 비용을 최소화하는 상황이 최적 방출 시간 T_{OP} 는 T_R 과 T_C 에 대하여 다음을 만족한다[5].

$$T_{OP} = \text{Max}(T_C, T_R) \quad (31)$$

단, $\ln R_0 = -b_2 [(x + T_R)^2 - (T_R)^2]$,

$$2(c_1 - c_2)b_2 T_C + c_3 = 0.$$

유사한 방법으로 지수 모형에서도 T_R 과 T_C 는 다음 두 방정식에 의해서 계산 될 수 있으나 이 식으로서는 T_R 과 T_C 은 추정 할 수 없다[13].

$$\ln R_0 = -b_1 x \quad (32)$$

$$(c_1 - c_2)b_1 + c_3 = 0 \quad (33)$$

혼합모형의 경우는 T_R 과 T_C 는 다음 두 방정식에 의해서 계산 될 수 있다.

$$\ln R_0 = \quad (34)$$

$$pb_1 T_R - pb_1 (x + T_R) + (1-p)(b_2 T_R^2 - b_2 (x + T_R)^2)$$

$$(c_1 - c_2)(pb_1 + 2(1-p)b_2 T_C) + c_3 = 0 \quad (35)$$

단, 신뢰도 $R_0 = \hat{R}(x | T_R)$.

V. 수치적인 예

이 장에서는 고장 간격 시간 자료(Failure interval time data)[15]를 가지고 지수분포와 Rayleigh분포에 근거한 혼합분포 모형에 대한 최적 방출시기를 분석하고자 한다. 이 자료는 1197.945 시간단위에 41번의 고장이 발생된 자료이며 <표 1>에 나열되어 있고 제시하는 신뢰 모형들을 분석하기 위하여 우선 자료에 대한 추세 검정이 선행되어야 한다[15].

추세 분석에는 일반적으로 라플라스 추세 검정(Laplace trend test)을 사용한다. 이 검정을 실시한 결과 <그림 1>에서 라플라스 추세 검정의 결과는 라플라스 요인(Factor)이 -2와 2사이에 존재함으로써 신뢰성장(Reliability growth) 속성을 나타내고 있다. 따라서 이 자료를 이용하여 신뢰도와 소프트웨어 방출시기를 추정하는 것이 가능하다[13].

<표 2>에서는 이 자료에 대한 기초통계량이 요약되어 있다. 이 표에서 왜도와 첨도가 각각 0.86과 -0.41로 나타나 분포의 꼬리가 오른쪽에 있고 대부분의 관측 값들이 왼쪽에 분포되어 있음을 의미하고 비교적 관측 값이 넓게 분포되는 형태를 보여주고 있다.

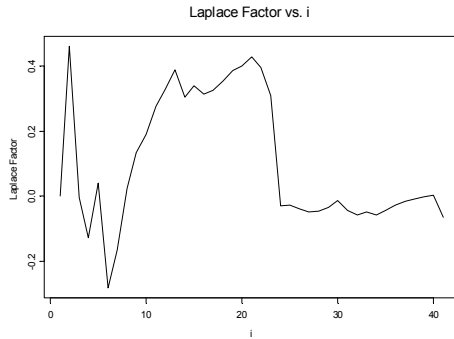
본 논문에서는 모수추정을 용이하게 하기 위하여실제자료에서수치변환된자료(failure times $\times 10^{-1}$)를 이용하였다.

혼합 모형에 대해서는 <그림 2>에서 경험적 누적분포함수 추정 결과 고장시간은 거의 와이블 분포에 접근함으로써 혼합모수 p 는 지수분포 보다는 와이블 분포에 혼합됨을 알 수 있으므로 p 값은 0.01에서부터 1사의 값을 고려하였다. 즉 p 는 지수모형의 가중치이고 $(1-p)$ 는 와이블 가중치를 의미한다. 그리고 지수모형의 모수 b_1 은 0에서부터 1까지의 값을 증가하는 패턴과 감소하는 패턴으로 분류하여 최우추정치를 추정하였다. 그 결과는 <표 4>와 <표 5>에 요약되었다. 비선형 방정식의 계산방법은 수치 해석적 기본 방법인 이분법(Bisection method)을 사용하였다. 이러한 계산은 초기값을 0.001 와 10 을, 허용 한계(Tolerance for width of interval)는 10^{-5} 을 주

<표 1> 고장 자료

Failure number	Failure Time (hours)	Failure Interval (hours)	Failure time $\times 10^{-1}$
1	5.649	5.649	0.5649
2	8.92	3.271	0.892
3	20.29	11.37	2.029
4	29.955	9.665	2.9955
5	34.715	4.76	3.4715
6	75.95	41.235	7.595
7	78.171	2.221	7.8171
8	78.625	0.454	7.8625
9	83.022	4.397	8.3022
10	89.114	6.092	8.9114
11	89.804	0.69	8.9804
12	92.86	3.056	9.286
13	93.66	0.8	9.366
14	110.655	16.995	11.0655
15	111.988	1.333	11.1988
16	122.545	10.557	12.2545
17	127.045	4.5	12.7045
18	128.712	1.667	12.8712
19	128.99	0.278	12.899
20	131.768	2.778	13.1768
21	131.829	0.061	13.1829
22	141.712	9.883	14.1712
23	164.212	22.5	16.4212
24	342.85	178.638	34.285
25	356.144	13.294	35.6144
26	399.144	43	39.9144
27	446.494	47.35	44.6494
28	476.644	30.15	47.6644
29	497.144	20.5	49.7144
30	497.661	0.517	49.7661
31	591.161	93.5	59.1161
32	665.644	74.483	66.5644
33	686.444	20.8	68.6444
34	765.944	79.5	76.5944
35	772.977	7.033	77.2977
36	774.944	1.967	77.4944
37	791.561	16.617	79.1561
38	815.978	24.417	81.5978
39	837.145	21.167	83.7145
40	861.945	24.8	86.1945
41	1197.945	336	119.7945

고 수렴성을 확인 하면서 충분한 반복 횟수인 100번을 C-언어를 이용하여 모수 추정을 수행하였다.



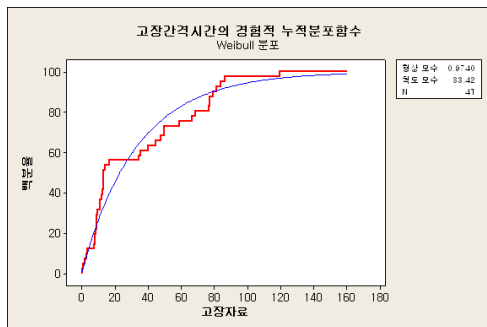
<그림 1> 라플라스 추세 검정

<표 2> 고장 시간자료의 기술통계량

기술 통계량	
평균	33.80
중앙값	13.18
첨도	-0.41
왜도	0.86
범위	119.23
관측수	41
신뢰 구간(95.0%)	(23.7035, 43.8963)

<표 3> 단순 모형의 모수 추정 값

Model	MLE
Exponential	0.342253
Rayleigh	0.002857



<그림 2> 경험적 누적분포함수

<표 4> 혼합 모형의 모수 추정 값(b_1 값이 증가 패턴)

Mixture Model		MLE
고정 값		
$p = 0.01$	$b_1 = 0.1$	$b_2 = 0.0975$
$p = 0.03$	$b_1 = 0.3$	$b_2 = 0.0994$
$p = 0.05$	$b_1 = 0.5$	$b_2 = 0.1012$
$p = 0.07$	$b_1 = 0.7$	$b_2 = 0.1030$
$p = 0.09$	$b_1 = 0.9$	$b_2 = 0.1048$

<표 6>과 <표 7>에서는 $c_1=10(\$)$, $c_2=15(\$)$ 그리고 $c_3=3(\$)$ 라고 가정하고 시스템 수명시간은 2000시간이고 임무시간 x 는 0.001이고 R_0 을 0.95(95%)를 투입하여 각 모형에 대한 추정시간의 결과와 최적방출시간이 요약되었다. 이 표에서는 단순모형보다는 혼합 모형이 길지 않아 효율적이므로 우수한 모형으로 나타나고 있다. 고려되었던 b_1 값이 증가 패턴인 경우와 감가 패턴인 경우에 있어서 는 증가 패턴이 최적방출시간이 더 효율적이고 또한 p 값이 증가 할수록 효율적으로 나타나고 있다.

<표 5> 혼합 모형의 모수 추정 값(b_1 값이 감소 패턴)

Mixture Model		MLE
고정 값		
$p = 0.01$	$b_1 = 0.9$	$b_2 = 0.0974$
$p = 0.03$	$b_1 = 0.7$	$b_2 = 0.0991$
$p = 0.05$	$b_1 = 0.5$	$b_2 = 0.1012$
$p = 0.07$	$b_1 = 0.3$	$b_2 = 0.1034$
$p = 0.09$	$b_1 = 0.1$	$b_2 = 0.1059$

결국 제안된 혼합모형은 기존의 단순 모형 비해 최적 방출시간이 효율적임을 보여주고 있다. 이 결과는 적합 시킨 자료가 달라지면 그 결과는 달라질 수 있지만 단순 모형보다 혼합모형이 우수한 모형이 됨을 확인 할 수 있다.

따라서 와이블 분포에서 정상 운영 단계에서 발생하는 일정한 강도함수 (형상모수 $a=1$)와 마모 단계에서

<표 6> 최적 방출시간 계산($R_0 = 0.95$, b_1 값이 증가 패턴)

Model	추정시간	최적방출 시간 (T_{OP})	
Exponential	$\hat{T}_R =$ 추정 할수없음 $\hat{T}_C = 1.7331$	추정 할수없음	
Rayleigh	$\hat{T}_R = 8750.4370$ $\hat{T}_C = 105.0052$	8750.4370	
M i x t u r e	$b_1 = 0.1$ $p = 0.01$	$\hat{T}_R = 7.0618$ $\hat{T}_C = 3.1028$	7.0618
	$b_1 = 0.3$ $p = 0.03$	$\hat{T}_R = 4.0612$ $\hat{T}_C = 3.0648$	4.0612
	$b_1 = 0.5$ $p = 0.05$	$\hat{T}_R = 3.1231$ $\hat{T}_C = 2.9904$	3.1231
	$b_1 = 0.7$ $p = 0.07$	$\hat{T}_R = 2.6186$ $\hat{T}_C = 2.8761$	2.8761
	$b_1 = 0.9$ $p = 0.09$	$\hat{T}_R = 2.2907$ $\hat{T}_C = 2.7210$	2.7210

발생하는 강도함수(형상모수 $a=2$) 혼합된 2가지 경우를 고려한 모형이 효율적이라고 할 수 있다.

<표 7> 최적 방출시간 계산($R_0 = 0.95$, b_1 값이 감소 패턴)

Model	추정시간	최적방출 시간 (T_{OP})	
M i x t u r e	$b_1 = 0.9$ $p = 0.01$	$\hat{T}_R = 7.0649$ $\hat{T}_C = 3.0645$	7.0649
	$b_1 = 0.7$ $p = 0.03$	$\hat{T}_R = 4.0674$ $\hat{T}_C = 3.0116$	4.0674
	$b_1 = 0.5$ $p = 0.05$	$\hat{T}_R = 3.1231$ $\hat{T}_C = 2.9904$	3.1231
	$b_1 = 0.3$ $p = 0.07$	$\hat{T}_R = 2.6135$ $\hat{T}_C = 3.0105$	3.0105
	$b_1 = 0.1$ $p = 0.09$	$\hat{T}_R = 2.2791$ $\hat{T}_C = 3.0663$	3.0663

VI. 결론

대용량 소프트웨어가 수정과 변경하는 과정에서 결점의 발생을 거의 피할 수 없는 상황이 현실이다. 실제로 만족할 만한 신뢰도가 부여되고 동시에 시스템 고장과 연계된 기대 총비용을 최소화시키기 위하여 필요하다면 충분한 테스트를 계속해야 한다. 따라서 신뢰성 요구를 만족하고 총 비용을 최소화하는 상황이 최적 방출 시간이다.

본 연구에서는 소프트웨어의 결함을 제거하거나 수정 작업 중에도 새로운 결함이 발생할 가능성이 있는 무한 고장수를 가진 비동질적인 포아송 과정에 기초하고 시스템이 복잡해지면 하나의 강도함수에만 고장이 일어나지 않고 다양한 원인에 의해 혼합되어 발생 할 수 있는 혼합모형을 이용한 최적방출시기에 관한 문제를 다루었다. 그 결과 단순모형보다 혼합모형이 우수한 모형이 됨을 확인할 수 있다. 이 연구를 통하여 소프트웨어 개발자들은 방출최적시기를 단순 보다는 혼합 모형에 대한 사전 지식을 파악 하는데 어느 정도 도움을 줄 수 있으리라 사료 된다.

참고문헌

- [1] Gokhale, S. S. and Trivedi, K. S. "A time/structure based software reliability model," Annals of Software Engineering. 8, 1999, pp. 85-121.
- [2] 김희철, "지수화 지수 분포에 의존한 NHPP 소프트웨어 신뢰성장 모형에 관한 연구," 한국 컴퓨터 정보학회 논문지, 제11권 5호, 2006, pp. 9-18.
- [3] Almering, V. and Genuchten, M, V and Cloudt, G. and Sonnemans, P. J. M, "Using Software Reliability Growth Models in Practice," IEEE SOFTWARE. 2007, pp. 82-88.

- [4] Musa, J. D and Okumoto, K, "A Logarithmic Poisson Execution Time Model for Software Reliability Measurement," Proceeding the 7th International Conference on Software engineering, 1984, pp. 230-238.
- [5] Xie, M. and Homg, G. Y, "Software release time determination based on unbound NHPP model," Proceeding of the 24th International Conference on Computers and Industrial Engineering. 1999, pp. 165-168.
- [6] Yang, B. and Xie. M, "A study of operational and testing reliability in software reliability analysis," RELIABILITY ENGINEERING & SYSTEM SAFETY, Vol, 70, 2000, pp. 323-329.
- [7] Huang, C. Y, "Cost-Reliability-optimal release policy for software reliability models incorporating improvements in testing efficiency," The journal of Systems and software. Vol, 77, 2005, pp. 139-155.
- [8] 김희철, "극값 분포 특성을 가진 소프트웨어 신뢰성 보증 모형에 관한 연구," 한국 통신학회 논문지, 34 권6호, 2009, pp. 623-629.
- [9] Lawless, J. F, "Statistical Models and Methods for Lifetime Data," John Wiley & Sons, New York, 1981.
- [10] Kuo, L. and Yang, T. Y, "Bayesian Computation of Software Reliability," Journal of the American Statistical Association, Vol. 91, 1996, pp. 763-773.
- [11] 정혜성의 2인. "신뢰성 시험분석평가," 영지문화사, 2005, pp. 99-101.
- [12] D. A. Berry and D. K. Stangl. "Bayesian Biostatistics," MARCEL DEKKER INC, 1996, pp. 601-616.
- [13] 김희철, 박형근 "와이블분포특성에 근거한 소프트웨어 최적방출시기에 관한 비교 연구," 한국 산학기술 학회 논문지, 제10권8호, 2009, pp. 1903-1910.
- [14] 김희철, 신현철, 김경수 "기록값 통계량에 기초한 무한고장 NHPP 소프트웨어 혼합 신뢰성장 모형에 관한 연구," 정보, 보안 논문지, 제7권3호, 2007, pp. 51-60.
- [15] K. Kanoun, J. C. Laprie, "Handbook of Software Reliability Engineering," M. R. Lyu, Editor, chapter Trend Analysis. McGraw-Hill New York, NY, 1996, pp. 401-437.

■ 저자소개 ■



이 상 식
Lee, Sang Sik

2001년 3월~현재
송호대학 보건의료전자과 교수
2004년 경희대학교 전자계산학과 박사
2000년 경희대학교 전자계산학과 석사
관심분야 : 소프트웨어 신뢰성공학,
멀티미디어공학
E-mail : leess@songho.ac.kr



김 희 철
Kim, Hee Cheul

2005년 3월~현재
남서울대학교 산업경영공학과 교수
1998년 2월 동국대학교 통계학과(이학박사)
1992년 2월 동국대학교 통계학과(이학석사)
관심분야 : 소프트웨어 신뢰성공학, 웹 프로그
래밍
E-mail : kim1458@nsu.ac.kr

논문접수일 : 2010년 3월 25일 수정일 : 2010년 4월 15일 게재확정일 : 2010년 4월 23일
--