

소프트웨어 테스트 노력 함수의 파라미터 산출에 관한 연구

최 규 식* · 김 필 중**

A Study on the Parameter Estimation for Testing Effort Function of Software

Gyu Shik Che · Pil Joong Kim

Abstract

Many software reliability growth model(SRGM) have been proposed for past several decades. Most of these propositions assumed the S/W debugging testing efforts be constant or even did not consider them. A few papers were presented as the software reliability evaluation considering the testing effort was important afterwards. The testing effort forms which have been presented by this kind of papers were exponential, Rayleigh, Weibull, or Logistic functions, and one of these 4 types was used as a testing effort function depending on the S/W developing circumstances.

We consider the methodology to evaluate the SRGM using least square estimator(LSE) and maximum likelihood estimator(MLE) for those 4 functions, and then examine parameters applying actual data adopted from real field test of developing S/W.

1. 서론

소프트웨어의 신뢰도는 소프트웨어의 테스트 기간 단계중에 검출되는 관찰 결함데이터에 근거한 소프트웨어의 신뢰도 성장이라는 것으로 연구되어 왔다. 소프트웨어 신뢰도 성장 모델(SRGM : software reliability growth model)[1]은 테스트중에 검출되는 누적결함의 수 및 테스트 시간 간격과 관련이 있다. SRGM은 소프트웨어 시스템의 s-예측 초기 결함 내용, 임의의 테스트 시점에서의 s-예측 잔여 결함의 수, 소프트웨어의 신뢰도 등을 산출할 수 있다. 그동안 여러 가지 SRGM이 제안 및 연구되어 왔다.

소프트웨어 신뢰도는 규정된 시간 주기 동안 주어진 환경에서 소프트웨어가 고장 없이 동작할 확률이다. 이것은 소프트웨어의 품질에 있어서 핵심 사항이다. 소프트웨어 신뢰도는 고객의 입장에서 본 소프트웨어 품질관점이다. 이는 단순히 프로그램 설계보다는 실제적인 운전과 관련이 있다. 그러므로, 정적이라기 보다는 동적이다. 그러므로 소프트웨어 시스템의 신뢰도를 측정하고 계산하는 것이 매우 중요하다. 이들은 개발 기간동안 모든 테스트 자원을 계획하고 제어하는데 쓰일 수 있으며, 그리하여 우리에게 소프트웨어의 정확성을 확인시켜 주는 것이다. 소프트웨어 신뢰도를 측정하기 위한 공통적인 접근법은 소프트웨어 고장으로부터 구한 가용한 데이터로부터 산출된 파라미터를 가진 분석 모델을 이용하는 것이다. SRGM은 소프트웨어 신뢰도를 산출하는데 성공적이며, 그리고 소프트웨어 시스템에 잔여하고 있는 잔여결함의 수를 산출하는데도 또한 성공적이다. 일반적으로, 소프트웨어 개발에 있어서 상당한 양의 테스트 자원이 소프트웨어의 테스트에 쓰여진다. 테스트 기간에 대한 테스트 자원의 소요곡선은 테스트

노력 곡선으로 생각할 수 있다. 테스트 노력은 테스트 기간중의 실행 테스트 케이스의 수, 인력의 양, CPU시간 등이다.

소프트웨어 테스트 단계 기간 동안 소프트웨어의 신뢰도는 잠재 소프트웨어 결함을 검출하고 수정하는 데에 소요되는 개발자원의 양에 전적으로 의존한다.

야마다 등[2]은 역일시간, 테스트 노력량, 테스트 노력에 의하여 검출되는 결함의 수 사이의 관계를 명시적으로 기술하는 새롭고도 단순한 소프트웨어 신뢰도 성장모델을 제안하였다. 이는 그들이 소프트웨어 성장에의 테스트 노력의 영향과 관련된 소프트웨어 신뢰도 성장 모델을 제안한 것이다. 그들은 시간종속적 레일레이와 지수형 곡선을 이용한 테스트노력소요량 거동을 기술한다. 많은 연구가들의 연구에서는 테스트 단계의 테스트 자원의 소모율은 일정한 것으로 가정하거나 또는 그러한 테스트 노력을 고려하지도 않았다. 여러 참고 문헌에서 그 노력지수(실행시간)가 소프트웨어 신뢰도 모델링에서 역일시간보다 더 좋다는 것을 보여주고 있다. 관찰된 신뢰도 성장곡선의 형상이 테스트 노력의 시간분포에 강하게 의존하기 때문이다. 또 여러 참고문헌에서는 역일테스트, 테스트 노력량, 테스트 노력에 의하여 검출되는 결함의 수 사이의 관계를 설명하는 SRGM을 제안하였다.

소프트웨어 신뢰도 모델링 분야에서는 그 노력이 자주 전통적인 지수함수, 레일레이함수 또는 웨이블곡선으로 기술된다. 또한, 야마다[2] 등은 일반적인 테스트 노력에 대한 파라미터 산출법을 제시하고자 하였으며, 야마다는 또 다른 문헌[3]에서 웨이블 테스트노력을 적용한 경우의 신뢰도 성장모델을 이용하여 파라미터 산출법을 제시하였고, Syed A. Hossain 등[4]은 로지스틱형 함수와 고폴페르츠형 함수를 가지는 테스트 노력에 대한 파라미터 산출법을 연구하였

다. 그러나, 많은 소프트웨어 테스트 환경에서 이러한 3개의 노력 함수 곡선만으로 소프트웨어 테스트 노력 함수를 기술하는 것은 어려운 일이다. 따라서, 본 논문에서는 로지스틱 테스트 노력까지를 고려한 파라미터 산출법을 연구한다.

2항에서는 문헌에있는 기존의 테스트 노력 함수를 기술하고 추이 곡선을 그렸으며, 3항에서는 4개의 제안된 함수에 대한 파라미터를 구하는 방법을 제시하였다. 여기에 최소 자승 평가자와 최대 가능성 함수를 도입하였다. 4항에서는 실제 현장에서 구한 결함검출 데이터를 이용하여 각각의 파라미터를 구하고 이를 이용하여 목표신뢰도를 만족시키는 최적발행시기 결정법을 연구한다.

2. 테스트 노력 함수

주어진 기간에 맞추어 소프트웨어 시스템을 개발할 때 시간, 자금, 인력과 같은 자원들이 소요된다. 특히, 소프트웨어 테스트에까지 이르는 자원들이 소프트웨어 신뢰도에 상당한 영향을 미친다. 소프트웨어 개발 자원 전체의 약 40-50%가 테스트 단계에서 소요된다. Yamada 등[2]은 테스트 기간중의 테스트 노력과 소프트웨어 개발 노력 모두가 레일레이 곡선이나 웨이블곡선으로 설명될 수 있다는 것을 가정하여 소프트웨어 테스트에 쓰이는 테스트노력의 양을 고려한 소프트웨어 신뢰도 성장 모델을 제시하였다. 그들은 레일레이 곡선의 대안으로서 지수곡선도 제안하였다.

테스트 노력은 CPU 시간의 양, 실행된 테스트 케이스의 수 등으로 나타낼 수 있다. 때때로 테스트 시간은 실행시간 대신 테스트의 수로 나타낼 수도 있다. 소프트웨어 신뢰도 모델링 분야

에서는 소프트웨어의 개발 노력이 자주 전통적인 지수함수, 레일레이함수, 또는 웨이블 곡선으로 표현된다. 그러나, 많은 소프트웨어 테스트 환경에서 이러한 3개의 노력함수 곡선만으로 소프트웨어 테스트 노력함수를 기술하는 것은 어려운 일이다. 본 논문에서는 로지스틱 테스트 노력 함수도 소프트웨어 개발/테스트 노력 곡선으로 표현될 수 있다는 것을 보여주고자 한다. 실제 테스트/디버그 데이터 집합에 근거하여 실험을 수행하여 다양한 모델간의 예측 능력을 비교해본 결과, 초기 결함의 수를 산출하는 데는 로지스틱 테스트 노력 함수를 가진 SRGM이 이전의 접근법에 비하여 더 적합하다는 것을 보여준다.

2.1 평균치 함수 및 신뢰도

$R(x,t)$ 는 시각 t 에서 최종적으로 결함을 발견하여 수정한 후 x 유니트의 경과사건동안 새로운 결함이 발견되지 않을 확률이다. 소프트웨어를 개발하여 결함테스트를 하면 할수록 결함을 발견하여 수정하는 빈도가 작아지므로 신뢰도가 성장되며, 결함 수정 후 경과시간이 길어지면 질수록 결함 발견 확률이 높아지기 때문에 소프트웨어의 신뢰도는 낮아진다.

$R(x,t)$ 는 시각 t 에서 최종적으로 결함을 발견하여 수정한 후 x 유니트의 경과사건동안 새로운 결함이 발견되지 않을 확률이다. 소프트웨어를 개발하여 결함테스트를 하면 할수록 결함을 발견하여 수정하는 빈도가 작아지므로 신뢰도가 성장되며, 결함 수정 후 경과시간이 길어지면 질수록 결함 발견 확률이 높아지기 때문에 소프트웨어의 신뢰도는 낮아진다. 한편, 테스트 단계에서는 얼마나 오랜 시간동안 결함이 발견되지 않느냐가 중요한 것이 아니라, 현 단계에서 소프트웨어 내에서 발견되지 않고 잔

존하는 결함의 수가 얼마나 되는가가 더 중요하다.

소프트웨어를 시각 t에서 발행하는 경우, 발견되는 누적 분포는 평균치 함수

$$m(t) = a(1 - e^{-\gamma W(t)}) \quad (2.1)$$

이므로, 잔여 분포

$$\bar{a} = a - a(1 - e^{-\gamma W(t)}) = a \cdot e^{-\gamma W(t)}$$

이다. 이것이 소프트웨어를 발행해서 운전하는 초기 결함수이므로

$$m(t+x) = a \cdot e^{-\gamma W(t)} \cdot (1 - e^{-\gamma W(x)}) + m(t) \quad (2.2)$$

로서,

$$\begin{aligned} R(x|t) &= \exp[-m(t+x) + m(t)] \\ &= \exp[-a \cdot e^{-\gamma W(t)}(1 - e^{-\gamma W(x)})] \end{aligned} \quad (2.3)$$

이 된다.

따라서, 시각 t=T에서 발행된 소프트웨어의 신뢰도는 경과시간 x에 대해서 지수 함수적으로 감소한다. 감소비를 줄이기 위해서는 발행 시각을 늦추거나 결함검출비를 높여야 한다.

목표 신뢰도를 $R(x|T) = R_0$ 라 하면

$$W^*(t) = \frac{1}{\gamma} \log \frac{1 - e^{-\gamma W(x)}}{\frac{1}{a} \log \frac{1}{R_0}} \quad (2.4)$$

에서 $T^* = T$ 를 구한다.

2.2 웨이블형 함수

Yamada[3]가 제안한 웨이블형 테스트 노력 함수에 의하면 소프트웨어의 테스트 노력이 테스트 단계 전체에 걸쳐서 일정하다고 가정하면 비현실적이다. 그리고, 순간적인 테스트 노력이 결국은 테스트 수명주기 동안 감소한다. 그 이유는 누적 테스트 노력이 유한치에 접근하기 때문이다. 그 어떤 소프트웨어 개발회사도 소프트웨어 개발에 무한정으로 자원을 투입하지 않기 때문에 이러한 가정이 합리적이라 할 수 있다. 여러 관련 문헌에 의하면 테스트 노력이 웨이블

형 분포로 설명될 수 있고, 아래와 같은 3개의 케이스를 가진다는 것을 보여주고 있다.

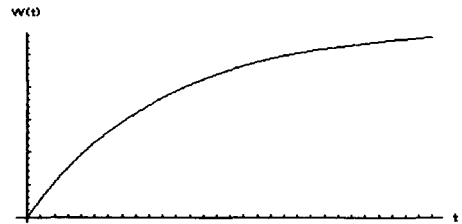
1) 지수함수곡선 : (0, t]에서 소요되는 누적 테스트 노력은

$$W(t) = M[1 - e^{-\beta t}] \quad (2.5)$$

로서 웨이블 함수의 m=1인 경우에 해당되며, 신뢰도는

$$\begin{aligned} R(x|t) &= \exp[-a \cdot e^{-\gamma M(1 - e^{-\beta t})} \\ &\cdot (1 - e^{-\gamma M(1 - e^{-\beta x})})] \end{aligned} \quad (2.6)$$

로 표현된다.

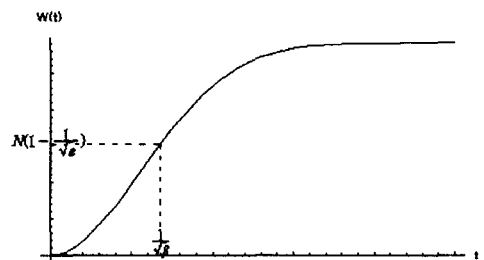


<그림 2.1> 지수함수형 누적테스트노력

2) 레일레이 곡선 : 소요되는 누적 테스트 노력은

$$W(t) = M[1 - \exp^{-\frac{\beta}{2} \cdot t^2}] \quad (2.7)$$

로서 웨이블함수의 m=2인 경우에 해당된다.



<그림 2.2> 레일레이형 누적 테스트노력 곡선

신뢰도는

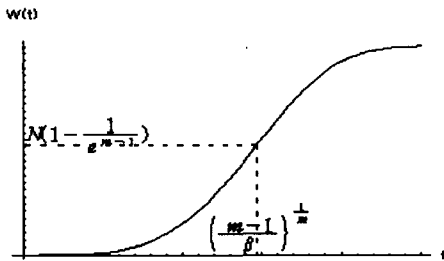
$$R(x|t) = \exp[-a \cdot e^{-rN(1-e^{-\frac{t}{\beta}})}] \cdot (1 - e^{-rN(1-e^{-\frac{t}{\beta}})}) \quad (2.8)$$

로 표현된다.

3) 웨이블 곡선 : 소요되는 누적 테스트 노력은

$$W(t) = N[1 - e^{-\beta t^m}] \quad (2.9)$$

로서 웨이블 함수의 일반적인 경우, 즉 m=3, 4, ...인 경우에 해당된다.



<그림 2.3> 웨이블형 누적테스트노력 곡선

따라서, 신뢰도는

$$R(x|t) = \exp[-a \cdot e^{-rN(1-e^{-\beta t^m})}] \cdot (1 - e^{-rN(1-e^{-\beta t^m})}) \quad (2.10)$$

로 표현된다.

웨이블형 곡선(2.9)에 대해서 m=1 또는 m=2 일 때 그 결과는 각각 지수함수 곡선이거나 레일레이 곡선이다. 따라서, 이들은 웨이블 함수의 특수한 경우에 해당된다. m=3, 4, 5일 때는 소프트웨어 개발 공정기간 동안 공칭 피크 현상(원활하지 못하게 증가하며, 소요 곡선을 열화시킴)을 나타낸다. 즉, 피크 작업량이 나타난다.[5] 이러한 현상은 실제 소프트웨어 개발/테스트 공정을 설명하는데 공통적으로 쓰이지 않으므로 현실적이지 못한 것으로 보인다. 그래서 웨이블 곡선은 형상 파라미터를 조정하여 많은 분포를 적합하게 맞출 수 있고 융통성 있게 테스트노력을 나타낼 수 있음에도 불구하고 테스트 노력

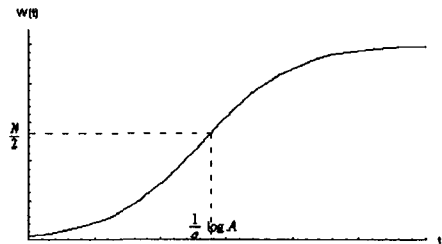
소요량을 모델링하는 데는 부적합한 것으로 사료된다.

2.3 로지스틱형 함수

실제 테스트 노력 데이터가 여러 가지 소요 패턴을 나타내므로 때때로 테스트 노력 비용을 지수함수나 레일레이 곡선만으로 설명하기는 어렵다. 웨이블형 곡선은 일반적인 소프트웨어 개발 환경 하에서 데이터에 잘 맞지만, 그리고 소프트웨어 신뢰도 모델링에 널리 쓰이지만 m>3일 때 공칭 피크현상을 가진다. 그 대안으로 제시된 것이 로지스틱형 테스트노력 함수이다. 이 함수는 실제 프로젝트 탐사에 의해서 보고된 바와 같이 매우 정확하다. (0, t]에서의 누적 테스트 노력 소요는

$$W(t) = \frac{N}{1 + A \cdot e^{-at}} \quad (2.11)$$

이고, 신뢰도는



<그림 2.4> 로지스틱형 누적테스트노력 곡선

$$R(x|t) = \exp[-a \cdot e^{-rN(\frac{1}{1+Ae^{-at}} - \frac{1}{1+A})}] \cdot (1 - e^{-rN(\frac{1}{1+Ae^{-at}} - \frac{1}{1+A})}) \quad (2.12)$$

로 표현된다.

그림의 초기점에서 웨이블형 테스트 노력 함수와 비교하여 로지스틱 테스트노력 함수인 경우 W(0)≠0 이다. 공정이 때때로 확인하기 힘든 소프트웨어 개발의 초기단계에 웨이블형 곡선과 W(t)와의 차이점이 존재한다. 그리고 적용된 테스트 노력의 양을 기록하기 위한 공식적

인 계수 공정절차가 수립되지 못한 곳에서도 차이가 난다.

이러한 두 개의 모델을 실제 고장 데이터와 잘 맞는 통계적인 테스트를 이용하여 이러한 모델 사이를 판정하는 것이 가능하다.

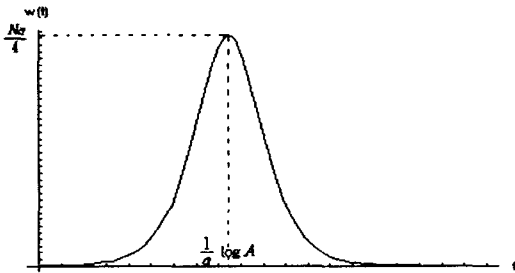
현재의 테스트 노력 소요량은 $W(t)$ 의 미분치로서

$$w(t) = \frac{dW(t)}{dt} = \frac{NAa \cdot \exp(-at)}{[1 + A \cdot \exp(-at)]^2}$$

$$= \frac{NAa}{[\exp(\frac{at}{2}) + A \cdot \exp(-\frac{at}{2})]^2}$$

(2.13)

와 같이 표현된다.



<그림 2.5> 로지스틱 테스트노력 소요량

그림에서 보는 바와 같이 이 양은

$t = \frac{1}{a} \log A$ 일 때 최대치 $\frac{NA}{4}$ 를 중심으로 좌우 대칭형이다.

3. 파라미터의 산출법

3.1 파라미터의 산출 방법

위의 각 경우에서 정의된 테스트노력함수에서 파라미터 N, A, a, β 는 최소자승법(LSE)으로 산출한다. MLE는 한 집합의 동시방정식을 풀어서 파라미터를 산출하며, s-신뢰 구간을 구동하는데 더 좋은 방법이다. 그러나, 그 방정식이

너무너무 복잡하므로 보통은 수치 해석적으로 푼다. LSE는 실제로 관찰/획득한 것과 예측한 것 사이의 차이를 제공해서 더한 총값을 최소화하는 것이다. LSE는 중간 크기의 표본에 최적인 것으로 알려지고 있으며, 최적점 산출을 제공한다. 최소자승법을 적용하기 위한 산출 공식 $S(N, A, a)$ 은 다음과 같다.

$$S(N, A, a) = \sum_{k=1}^n [W_k - W(t_k)]^2 \quad (3.1)$$

$W_k = (0, t]$ 기간동안 실제로 소요되는 누적 테스트 노력

$W(t_k) =$ 테스트 노력 함수에 의해서 산출된 누적 테스트 노력

S 를 N, A, a 에 관하여 미분하여 그 편미분치를 0으로 놓으면 그리고 이 항들을 재정비하여 이러한 형태의 비선형 최소 자승 문제를 푼다.

또한, 상기에서 정의한 테스트 노력 함수의 산출 파라미터를 이용하여 MLE에 의해서 $m(t)$ 의 신뢰도 성장 파라미터 a 와 r 을 구할 수 있다. 비동차포아송과정(NHPP)의 표준이론으로부터 임의의 $t \geq 0, x > 0$ 에서

$$\Pr\{N(t+x) - N(t) = k\}$$

$$= \frac{\{m(t+x) - m(t)\}^k}{k!} e^{-m(t+x) + m(t)} \quad (3.2)$$

이므로, 여기에 $m(t+x) = m(t_k), m(t) = m(t_{k-1}), k = m_k - m_{k-1}$ 를 대입하면

$$\Pr\{N(t_k) - N(t_{k-1}) = m_k - m_{k-1}\}$$

$$= \frac{\{m(t_k) - m(t_{k-1})\}^{m_k - m_{k-1}}}{(m_k - m_{k-1})!} e^{-m(t_k) + m(t_{k-1})} \quad (3.3)$$

주어진 시간간격 $(0, t_k) (0 < t_1 < t_2 < \dots < t_n)$ 에서 검출결합의 누적 갯수 m_k 에 대한 데이터를 관찰할 수 있다. 따라서, $m(t)$ 를 가지는 NHPP의 미지파라미터 a 와 r 에 대한 조인트 pmf 즉, MLE 함수는

$$L = \Pr\{N(t_i) = m_i, i = 1, \dots, n\}$$

$$= \prod_{k=1}^n \frac{[m(t_k) - m(t_{k-1})]^{(m_k - m_{k-1})}}{(m_k - m_{k-1})!} \cdot e^{-(m(t_k) - m(t_{k-1}))} \quad (3.4)$$

로 표시되며, 여기에 대수를 취하여 정리하면

$$\begin{aligned} \log(L) &= \sum_{k=1}^n (m_k - m_{k-1}) \\ &\cdot \log[m(t_k) - m(t_{k-1})] \\ &- \sum_{k=1}^n \log[m(t_k) - m(t_{k-1})]! \\ &- \sum_{k=1}^n \log[m(t_k) - m(t_{k-1})] \end{aligned} \quad (3.5)$$

이며, 여기서,

$$\begin{aligned} \sum_{k=1}^n [m(t_k) - m(t_{k-1})] &= m(t_n) \\ &= a[1 - e^{-rW^*(t_n)}], \\ m(t_k) - m(t_{k-1}) & \\ &= a[1 - e^{-rW^*(t_k)}] - a[1 - e^{-rW^*(t_{k-1})}] \\ &= a[e^{-rW^*(t_{k-1})} - e^{-rW^*(t_k)}], \\ \sum_{k=1}^n [m(t_k) - m(t_{k-1})] &= m(t_n) \\ &= a[1 - e^{-rW^*(t_n)}] \end{aligned}$$

를 대입하면

$$\begin{aligned} \log L &= \sum_{k=1}^n (m_k - m_{k-1}) \\ &\cdot \log a \{ e^{-rW^*(t_{k-1})} - e^{-rW^*(t_k)} \} \\ &- \sum_{k=1}^n \log(m_k - m_{k-1})! - a[1 - e^{-rW^*(t_n)}] \end{aligned} \quad (3.6)$$

이것을 a, r 에 관하여 편미분한다.

$$\begin{aligned} \frac{\partial L}{\partial a} &= \frac{1}{a} \sum_{k=1}^n \{ (m_k - m_{k-1}) - [1 - e^{-rW^*(t_k)}] \} = 0 \\ a &= \frac{m_n}{1 - e^{-rW^*(t_n)}} = \frac{m_n}{1 - \square_n} \end{aligned} \quad (3.7)$$

여기서,

$$\phi_k = e^{-rW^*(t_k)} \quad (3.8)$$

$$\begin{aligned} \frac{\partial \log L}{\partial r} &= \sum_{k=1}^n (m_k - m_{k-1}) \cdot \\ &\frac{-W^*(t_{k-1})e^{-rW^*(t_{k-1})} + W^*(t_k)e^{-rW^*(t_k)}}{e^{-rW^*(t_{k-1})} - e^{-rW^*(t_k)}} \\ &- aW^*(t_n)e^{-rW^*(t_n)} = 0 \end{aligned}$$

이므로,

$$\begin{aligned} aW^*(t_n)e^{-rW^*(t_n)} &= \sum_{k=1}^n (m_k - m_{k-1}) \\ &\cdot \frac{-W^*(t_{k-1})e^{-rW^*(t_{k-1})} + W^*(t_k)e^{-rW^*(t_k)}}{e^{-rW^*(t_{k-1})} - e^{-rW^*(t_k)}} \end{aligned} \quad (3.9)$$

$$aW^*(t_n)\phi_n$$

$$\begin{aligned} &= \sum_{k=1}^n (m_k - m_{k-1}) \\ &\cdot \frac{-W^*(t_{k-1})\phi_{k-1} + W^*(t_k)\phi_k}{\phi_{k-1} - \phi_k} \end{aligned} \quad (3.10)$$

으로 표현된다.

3.2 각 경우의 파라미터 산출

1) 지수함수형

$$S_1(N, \cdot) = \sum_{k=1}^n \{W_k - N(1 - e^{-t_k})\}^2 \quad (3.11)$$

$$\frac{\partial S_1}{\partial N} = -2 \sum_{k=1}^n \{W_k - N(1 - e^{-t_k})\} (1 - e^{-t_k}) = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-t_k})\} (1 - e^{-t_k}) = 0 \quad (3.12)$$

을 만족하는 N 값을 구한다. 또한,

$$\frac{\partial S_1}{\partial \cdot} = -2N \sum_{k=1}^n \{W_k - N(1 - e^{-t_k})\} t_k e^{-t_k} = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-t_k})\} t_k e^{-t_k} = 0 \quad (3.13)$$

을 만족하는 β 값을 구한다.

$$W(t_k) = N(1 - e^{-\beta t_k}) \text{ 에서}$$

$$a = \frac{m_n}{1 - e^{-rN(1 - e^{-\beta t_n})}} \quad (3.14)$$

를 구한다.

$$aW^*(t_n)\phi_n = \sum_{k=1}^n (m_k - m_{k-1}) \cdot \frac{-W^*(t_{k-1})\phi_{k-1} + W^*(t_k)\phi_k}{\phi_{k-1} - \phi_k} \quad (3.15)$$

$$aN(1 - e^{-\beta t_n})e^{-rN(1 - e^{-\beta t_n})} = \sum_{k=1}^n (m_k - m_{k-1}) \cdot$$

$$S_2(N, \cdot) = \sum_{k=1}^n \{W_k - N(1 - e^{-\beta/2 t_k^2})\}^2 \{-N(1 - e^{-\beta t_{k-1}})e^{-rN(1 - e^{-\beta t_{k-1}})} + N(1 - e^{-\beta t_k})e^{-rN(1 - e^{-\beta t_k})}\} / \{e^{-rN(1 - e^{-\beta t_{k-1}})} - e^{-rN(1 - e^{-\beta t_k})}\} \quad (3.16)$$

에서 r 을 구한다.

2) 레일레이형

$$\frac{\partial S_2}{\partial N} = -2 \sum_{k=1}^n \{W_k - N(1 - e^{-\beta/2 t_k^2})\} (1 - e^{-\beta/2 t_k^2}) = 0 \quad (3.17)$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-\beta/2 t_k^2})\} (1 - e^{-\beta/2 t_k^2}) = 0 \quad (3.18)$$

을 만족하는 N 값을 구한다.

$$\frac{\partial S_2}{\partial \beta} = -2N \sum_{k=1}^n \{W_k - N(1 - e^{-\beta/2 t_k^2})\} t_k e^{-\beta/2 t_k^2} = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-\beta/2 t_k^2})\} t_k^2 e^{-\beta/2 t_k^2} = 0 \quad (3.19)$$

을 만족하는 β 값을 구한다

$$W(t_k) = N(1 - e^{-\beta^2/2 t_k}) \text{ 에서}$$

$$a = \frac{m_n}{1 - e^{-rN(1 - e^{-\beta^2/2 t_n})}} \quad (3.20)$$

를 구한다.

$$aW^*(t_n)\phi_n = aN(1 - e^{-\beta^2/2 t_n})e^{-rN(1 - e^{-\beta^2/2 t_n})} \sum_{k=1}^n (m_k - m_{k-1}) \frac{-W^*(t_{k-1})\phi_{k-1} + W^*(t_k)\phi_k}{\phi_{k-1} - \phi_k} \quad (3.21)$$

$$= \sum_{k=1}^n (m_k - m_{k-1}) \cdot \{-N(1 - e^{-\beta^2/2 t_{k-1}})e^{-rN(1 - e^{-\beta^2/2 t_{k-1}})} + N(1 - e^{-\beta^2/2 t_k})e^{-rN(1 - e^{-\beta^2/2 t_k})}\} / \{e^{-rN(1 - e^{-\beta^2/2 t_{k-1}})} - e^{-rN(1 - e^{-\beta^2/2 t_k})}\} \quad (3.22)$$

에서 r 을 구한다.

3) 웨이볼형

$$S_3(N, \cdot, m) = \sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\}^2 \quad (3.23)$$

$$\frac{\partial S_3}{\partial N} = -2 \sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} (1 - e^{-t_k^m}) = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} (1 - e^{-t_k^m}) = 0 \quad (3.24)$$

을 만족하는 N 값을 구한다.

$$\frac{\partial S_3}{\partial m} = -2N \sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} t_k^{m-1} e^{-t_k^m} = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} t_k^{m-1} e^{-t_k^m} = 0 \quad (3.25)$$

을 만족하는 β 값을 구한다.

$$\frac{\partial S_3}{\partial m} = -2Nm \sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} t_k^{m-1} e^{-t_k^m} = 0$$

즉,

$$\sum_{k=1}^n \{W_k - N(1 - e^{-t_k^m})\} t_k^{m-1} e^{-t_k^m} = 0 \quad (3.26)$$

을 만족하는 m 값을 구한다.

$W(t_k) = N(1 - e^{-\beta t_k^m})$ 에서

$$a = \frac{m_n}{1 - e^{-rN(1 - e^{-\beta t_n^m})}} \quad (3.27)$$

를 구한다.

$$aW^*(t_n)\phi_n = \sum_{k=1}^n (m_k - m_{k-1}) \cdot \frac{-W^*(t_{k-1})\phi_{k-1} + W^*(t_k)\phi_k}{\phi_{k-1} - \phi_k} \quad (3.28)$$

$$\begin{aligned} & aN(1 - e^{-\beta t_n^m})e^{-rN(1 - e^{-\beta t_n^m})} \\ &= \sum_{k=1}^n (m_k - m_{k-1}) \cdot \\ & \left\{ -N(1 - e^{-\beta t_{k-1}^m})e^{-rN(1 - e^{-\beta t_{k-1}^m})} \right. \\ & \left. + N(1 - e^{-\beta t_k^m})e^{-rN(1 - e^{-\beta t_k^m})} \right\} \\ & / \left\{ e^{-rN(1 - e^{-\beta t_{k-1}^m})} - e^{-rN(1 - e^{-\beta t_k^m})} \right\} \quad (3.29) \end{aligned}$$

에서 r 을 구한다.

4) 로지스틱형

$$S_4(N, A, \textcircled{A}) = \sum_{k=1}^n \left\{ W_k - \frac{N}{1 + Ae^{-\textcircled{A} t_k}} \right\}^2 \quad (3.30)$$

$$\frac{\partial S_4}{\partial N} = 2 \sum_{k=1}^n \left\{ W_k - \frac{N}{1 + Ae^{-\textcircled{A} t_k}} \right\} \frac{1}{1 + Ae^{-\textcircled{A} t_k}} = 0$$

즉

$$N = \frac{\sum_{k=1}^n \frac{W_k}{1 + Ae^{-\textcircled{A} t_k}}}{\sum_{k=1}^n \frac{1}{\{1 + Ae^{-\textcircled{A} t_k}\}^2}} \quad (3.31)$$

을 만족하는 N 값을 구한다.

$$\frac{\textcircled{A} S_4}{\textcircled{A}} = 2N \sum_{k=1}^n \left(W_k - \frac{N}{1 + Ae^{-\textcircled{A} t_k}} \right) \frac{e^{-\textcircled{A} t_k}}{\{1 + Ae^{-\textcircled{A} t_k}\}^2} = 0$$

즉

$$\sum_{k=1}^n \left(W_k - \frac{N}{1 + Ae^{-\textcircled{A} t_k}} \right) \frac{e^{-\textcircled{A} t_k}}{\{1 + Ae^{-\textcircled{A} t_k}\}^2} = 0$$

$$(3.32)$$

를 만족하는 A 값을 구한다.

$$\begin{aligned} \frac{\partial S_4}{\partial a} &= 2NA \sum_{k=1}^n \left(W_k - \frac{N}{1 + Ae^{-at_k}} \right) \\ & \cdot \frac{t_k e^{-at_k}}{\{1 + Ae^{-at_k}\}^2} = 0 \end{aligned}$$

즉

$$\sum_{k=1}^n \left(W_k - \frac{N}{1 + Ae^{-at_k}} \right) \frac{t_k e^{-at_k}}{\{1 + Ae^{-at_k}\}^2} = 0 \quad (3.33)$$

을 만족하는 a 값을 구한다.

$$W^*(t_k) = \frac{N}{1 + Ae^{-at_k}} - \frac{N}{1 + A} \quad \text{에서}$$

$$a = \frac{m_n}{1 - e^{-rN\left(\frac{1}{1 + Ae^{at_n}} - \frac{1}{1 + A}\right)}} \quad (3.34)$$

를 구한다.

$$aW^*(t_n)\phi_n =$$

$$\sum_{k=1}^n (m_k - m_{k-1}) \frac{-W^*(t_{k-1})\phi_{k-1} + W^*(t_k)\phi_k}{\phi_{k-1} - \phi_k} \quad (3.35)$$

$$\begin{aligned} & aN\left(\frac{1}{1 + Ae^{-at_n}} - \frac{1}{1 + A}\right) e^{-rN\left(\frac{1}{1 + Ae^{-at_n}} - \frac{1}{1 + A}\right)} \\ &= \sum_{k=1}^n \{ (m_k - m_{k-1}) \cdot \end{aligned}$$

$$\begin{aligned} & \left[\left\{ -N\left(\frac{1}{1 + Ae^{-at_{k-1}}} - \frac{1}{1 + A}\right) e^{-rN\left(\frac{1}{1 + Ae^{-at_{k-1}}} - \frac{1}{1 + A}\right)} \right\} \right. \\ & \left. + \left\{ N\left(\frac{1}{1 + Ae^{-at_k}} - \frac{1}{1 + A}\right) e^{-rN\left(\frac{1}{1 + Ae^{-at_k}} - \frac{1}{1 + A}\right)} \right\} \right] \\ & / \left\{ e^{-rN\left(\frac{1}{1 + Ae^{-at_{k-1}}} - \frac{1}{1 + A}\right)} - e^{-rN\left(\frac{1}{1 + Ae^{-at_k}} - \frac{1}{1 + A}\right)} \right\} \quad (3.36) \end{aligned}$$

에서 r 을 구한다.

3.3 발행시각 결정

구한 파라미터를 적용하여 각각의 경우에 대한 최적 발행 시각을 나타내면 다음과 같다.

1) 지수함수형

발행시각을 $t=T$ 라 하고, 목표신뢰도를

$R(x|T)=R_0$ 라 하면

$$R_0 = \exp[-ae^{-rN(1-e^{-\beta T})}(1-e^{-rN(1-e^{-\beta T})})]$$

(3.37)

에서 발행시각을 구하면

$$T = -\frac{1}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right)$$

$$= -\frac{1}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right)$$

$$= -\frac{1}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log R_0}{\log R(x|0)} \right)$$

로 표현된다. 여기서,

$$R(x|0) = \exp[-a(1-e^{-rN(1-e^{-\beta x})})]$$

이다.

2) 레일레이형

$$R_0 = \exp[-ae^{-rN(1-e^{-\beta T})}(1-e^{-rN(1-e^{-\beta T})})]$$

(3.39)

에서

$$T = \left\{ -\frac{2}{\beta} \log \left[1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right] \right\}^{1/2}$$

$$= \left\{ -\frac{2}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right) \right\}^{1/2}$$

$$= \left\{ -\frac{2}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log R_0}{\log R(x|0)} \right) \right\}^{1/2} \quad (3.40)$$

이고 여기서

$$R(x|0) = \exp[-a(1-e^{-rN(1-e^{-\beta x})})]$$

이다.

3) 웨이블형

$$R_0 = \exp[-ae^{-rN(1-e^{-\beta T})}(1-e^{-rN(1-e^{-\beta T})})]$$

(3.41)

에서

$$T = \left\{ -\frac{1}{\beta} \log \left[1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right] \right\}^{1/m}$$

$$= \left\{ -\frac{1}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(1-e^{-\beta T})})} \right) \right\}^{1/m}$$

$$= \left\{ -\frac{1}{\beta} \log \left(1 + \frac{1}{rN} \log \frac{\log R_0}{\log R(x|0)} \right) \right\}^{1/m} \quad (3.42)$$

이고 여기서,

$$R(x|0) = \exp[-a(1-e^{-rN(1-e^{-\beta x})})]$$

이다.

4) 로지스틱형

$$R_0 = \exp \left[-ae^{-rN \left(\frac{1}{1+Ae^{-\beta T}} - \frac{1}{1+A} \right)} \right]$$

$$\cdot \left(1 - e^{-rN \left(\frac{1}{1+Ae^{-\beta T}} - \frac{1}{1+A} \right)} \right)$$

(3.43)

에서

$$T = -\frac{1}{\alpha} \log \frac{1}{A} \left(\frac{1}{\frac{1}{1+A} - \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(\frac{1}{1+Ae^{-\beta T}} - \frac{1}{1+A})})} - 1} \right)$$

$$T = -\frac{1}{\alpha} \log \frac{1}{A} \left(\frac{1}{\frac{1}{1+A} - \frac{1}{rN} \log \frac{\log \frac{1}{R_0}}{a(1-e^{-rN(\frac{1}{1+Ae^{-\beta T}} - \frac{1}{1+A})})} - 1} \right)$$

$$= -\frac{1}{\alpha} \log \frac{1}{A} \left(\frac{1}{\frac{1}{1+A} - \frac{1}{rN} \log \frac{\log R_0}{\log R(x|0)} - 1} \right) \quad (3.44)$$

이다.

여기서,

$$R(x|0) = \exp[-a(1-e^{-rN(\frac{1}{1+Ae^{-\beta x}} - \frac{1}{1+A})})]$$

$$= e^{-a(1-e^{-rW(x)})}$$

단, 1)-4)의 모든 경우에 대해서

$e^{-rW(T_{LC})} = p$ 라 할 때 $R^p < R(x) < R_0$ 인 경우에만 발행시각에 관한 문제가 의미를 갖게 되어 이 때 유일하고도 유한한 해가 존재한다.

3.4 적용 예

3.1-3.3항에서 연구한 NHPP 모델의 수치 적용 예를 보이기 위해서 실제적인 소프트웨어 에 레 데이터로서 [5]에서 조사된 데이터집합을 가지고 다음과 같은 값을 산출하여 적용하였다.

<표 3.1> 실제 데이터

t_k	1	2	3	4	5	6	7	8	9	10	-
m_k	8	16	25	37	41	46	52	59	64	80	-
w_k	1	1	1	1	1	1	1	3	10	1	-
tk	11	12	13	14	15	16	17	18	19	20	21
m_k	89	98	107	121	128	131	133	133	135	136	137
w_k	5	3	9	13	34	35	27	39	35	25	19

이 데이터를 이용하여 방정식 (3.11)-(3.36)에 의해서 각종 파라미터를 산출하면 표 3.2, 3.3과 같다.

<표 3.2> 파라미터 산출(오차범위 0.001인 경우)

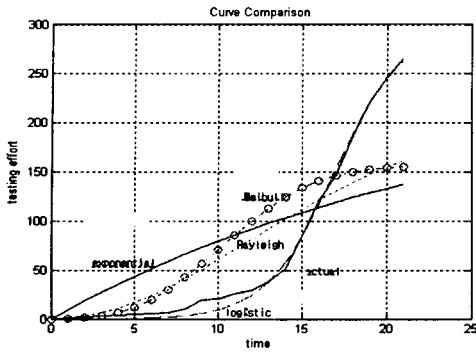
	지수함수형	레일레이형	웨이블형	로지스틱형
N	241.73	204.77	155.37	295.94
a	223.35	179.88	208.35	137.00
α	-	-	-	0.500977
β	0.003984	0.007031	0.000781	-
m	-	-	2.891602	-
A	-	-	-	4653.63
r	0.006935	0.008888	0.006935	0.040135

<표 3.3> 파라미터 산출(오차 범위 0.0001인 경우)

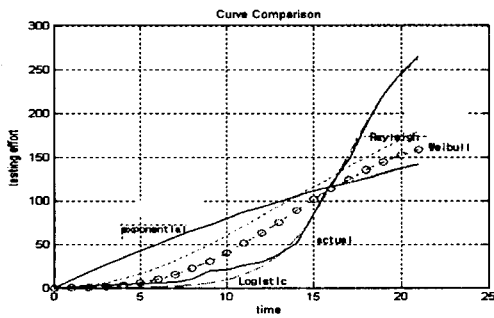
	지수함수형	레일레이형	웨이블형	로지스틱형
N	289.31	263.41	176.79	295.84
a	287.69	159.61	145.04	137.00
α	-	-	-	0.501282
β	0.003213	0.005176	0.000293	-
m	-	-	2.943054	-
A	-	-	-	4673.86
r	0.004067	0.010902	0.018226	0.039586

식을 만족하는 계수를 구하기는 일반적으로 쉽지 않다. 따라서, 수치해석적인 방법에 의하여 반복적인 연산을 수행하였으며, 이론치와 적용 데이터의 차이가 최소로 되는 시점의 값을 최적 값으로 정하였다.

표 3.2는 오차범위가 0.001 이내로 되는 값을, 표 3.3은 오차범위가 0.0001 이내로 되는 값을 수치해석적으로 구한 결과이다. 각 경우의 파라미터에 의한 테스트 노력곡선이 실제의 현상과 어떤 연관관계를 가지는가를 검토하기 위하여 추이곡선을 그렸으며, 그 결과는 그림 3.1(오차 범위 0.001 이내), 그림 3.2(오차 범위 0.0001 이내)와 같다.



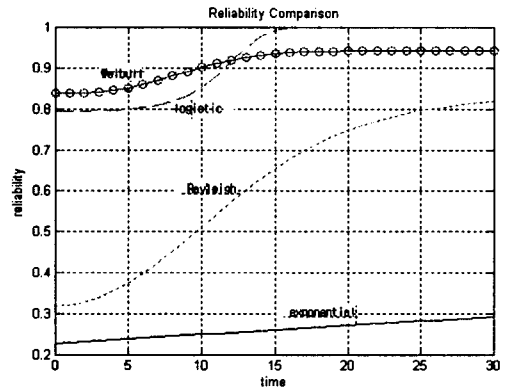
<그림 3.1> 테스트노력곡선 비교(오차 범위 0.001)



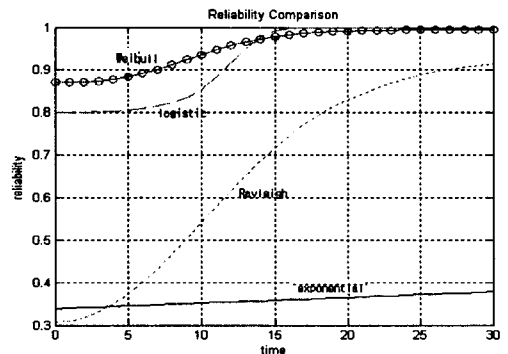
<그림 3.2> 테스트노력곡선 비교(오차 범위 0.0001)

4가지 경우의 테스트 노력 곡선에 대한 계수를 구하여 실제 관측된 데이터 곡선과 비교하였을 때, 지수함수형과 레일레이형은 현격한 차이를 나타내고 있다. 웨이블형은 상기 두 함수에 비하여 실제와 근접하나, 테스트 시간이 길어짐에 따라 커다란 차이를 보이고 있다. 이와는 대조적으로 로지스틱형은 테스트초창기만 약간 차이를 보이고 있으나, 테스트 시간이 경과함에 따라 실제 현상과 거의 근접한 것을 알 수 있다. 초창기에 차이가 나는 것은 테스트 초기에는 테스트 여건이나 테스트자의 숙련도 등이 미숙하거나 불확실한 이유 때문으로 사료된다.

또한, 이러한 각종 파라미터에 의한 테스트노력곡선을 이용하여 각 경우의 신뢰도 성장 곡선을 그리면 그림 3.3, 3.4와 같다. 단, 경과시간 $x=1$ 로 가정하였다.



<그림 3.3> SRGM 비교(오차 범위 0.001)



<그림 3.4> SRGM 비교(오차 범위 0.0001)

그림을 검토해볼 때 지수함수형이나 레일레이형인 경우 목표신뢰도에 이르기어렵게 되어 있으며, 비현실적이다. 이는 위에서 언급한 바와 같이 관찰된 데이터를 분석하는 가정이 불확실하기 때문이다. 웨이블형이나 로지스틱형은 S형 곡선을 나타내고 있으며, 실제 소프트웨어 성장모델(SRGM)곡선과 근접한 현상을 나타내고 있다. 목표신뢰도 $R0=0.9$ 로 가정하여 오차범위 0.001이내인 경우를 적용하여 각각의 경우에 대한 발행시각을 계산한 결과, 지수함수형이나

레일레이형은 성장속도가 너무 느려서 소프트웨어의 전 수명주기를 통하여 테스트를 해도 목표신뢰도를 맞출 수 없는 것으로 검토되었다. 그러나, 표 3.1-3.3에서 보는 바와 같이 초기 오류 데이터의 수가 300을 넘지 않는 상황을 감안할 때 이는 논리에 맞지 않는다. 한편, 웨이블형인 경우는 발행시각 $T=10.5$, 로지스틱형인 경우 발행시각 $T=11.5$ 로 계산되었다. 이는 그림 3.3의 곡선을 관찰해도 그 결과가 타당함을 확인할 수 있다.

4. 결 론

소프트웨어의 신뢰도를 산정하고 신뢰도를 성장시키고자 할 때 테스트 노력이 중요한 요소라는 것이 여러 연구 논문을 통하여 증명되었다. 그간 여러 문헌에서 제시된 테스트 노력함수로서는 지수함수형, 레일레이형, 웨이블형이며, 최근에는 로지스틱형도 제시되고 있다.

본 논문에서는 상기 네 가지 경우에 대한 함수의 파라미터 산출법을 연구하고, 이러한 파라미터 산출법과 함수들이 실제와 얼마나 부합되는가를 검토하기 위하여, 실제 개발 과정을 통하여 수집한 실제 데이터를 가지고 각 경우에 대한 파라미터를 산출하고 이를 상호 비교하였다. 검토결과에 의하면 지수함수형이나 레일레이형은 실제 현상과 커다란 차이를 나타내었으며, 웨이블형은 어느 정도 실제와 근접하는 특성을 나타내었다. 한편, 로지스틱형은 거의 완벽하다고 표현할 만큼 실제 현상과 잘 부합하였다.

테스트노력과 오류검출 현상이 소프트웨어 개발환경에 따라 불규칙성을 나타내는 것이 일반적이라고 할 때, 지수함수형이나 레일레이형은 이러한 현상에 융통성이 부족한 것으로 사료

된다. 웨이블형은 앞의 두 가지 형에 대한 보완형으로 제안되었다고 할 만큼 융통성이 있어서 실제 현상과 유사한 특성을 나타낸다. 단, 웨이블형인 경우는 시간의 지수가 3 이상이면 문제가 있는 것으로 검토되었으므로 이를 유의할 필요가 있다.

최근에는 로지스틱형이 제안되고 있는데, 본 연구에서 비교 연구한 바에 의하면 실제 현상에 가장 부합하는 것으로 판명되었다.

그리고, 목표신뢰도를 정하고 산출된 파라미터에 의해서 각 경우의 발행시각을 계산해본 결과, 지수함수형과 레일레이형은 소프트웨어 전 수명기간동안 목표신뢰도를 만족시킬 수 없는 큰 문제점을 나타내고 있다. 이는 곡선의 융통성이 적을 뿐만 아니라, 테스트 초창기의 불규칙하고 불확실한 테스트 노력 현상 때문에 산출된 파라미터의 신뢰성이 저하되었기 때문으로 사료된다. 웨이블형이나 로지스틱형은 발행시각이 각각 10.5, 11.5로서 비교적 정확한 분석 결과이다.

본 논문에서는 목표 신뢰도만을 연구하였으나, 추후 소프트웨어의 개발->테스트->발행->운영의 전 기간에 걸쳐서 소요되는 소프트웨어의 비용을 고려한 발행시각을 결정하는 연구가 계속되어야 할 것이다.

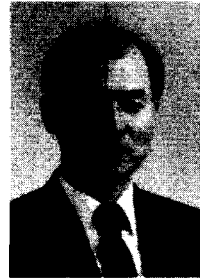
참 고 문 헌

- [1] C. V. Ramamoorthy, F. B. Bastani, "Software reliability - Status and perspectives", IEEE Trans. on Software Eng., vol. SE-8, pp354-371, 1982 Aug.
- [2] S. Yamada, H. Ohtera, H. Narihisa, "Software reliability growth models with testing-efforts", IEEE Trans. Reliability,

vol. R-35, pp19-23, 1986 Apr.

- [3] S. Yamada, J. Hishitani, S. Osaki, "Software - Reliability Growth with a Weibull Test-Effort : A Model & Application", IEEE Trans. Reliability, vol. 42, no.1, pp100-106, 1993 March.
- [4] Syed A. Hossain, Ram C. Dahiya, "Estimating the Parameters of a Non-homogeneous Poisson-Process Model for Software Reliability", IEEE Trans. Reliability, vol. 42, pp604-612, no.4, 1993 Dec.
- [5] Chin-Yu Huang, Sy-Yen Kuo, "Analysis of Incorporating Logistic Testing-Effort Function into Software Reliability Modeling", IEEE Trans. on Reliability, vol.51, pp261-270, 2002, Sep.

■ 저자소개



최 규 식

서울대학교 공과대학 전기과를 졸업하고, 뉴욕공과대학에서 석사학위를 받았으며, 명지대학교 전기과에서 박사학위를 받았다. OPC 중앙연구소와 KOPEC

중앙연구소에서 근무하였다.

현재는 건양대학교 정보통신학과 교수이며, 관심분야는 무선통신 및 소프트웨어 신뢰도이다.



김 필 중

1981년 고려대학교 물리학과 졸업(학사)
1991년 (미)조지아 스테이트대 졸업(전산과학 석사)
1999년 충남대학교 전산학과 졸업(이학박사)

1994년~현재 건양대학교 전산학과 부교수

관심분야 : 컴퓨터그래픽스, 인공지능, 패턴인식