테스트 노력 함수의 파라미터 산출

Parameter Estimation of Testing Effort

최규식
* 건양대학교 정보통신학과, E-mail : che@konyang.ac.kr

Abstract : 지난 수십년간 많은 SRGM이 제안되었다. 이러한 모델들 중 많은 부분에서 암암리에 소프트웨어 테스트 전 단계를 거쳐서 테스트노력이 상수인 것으로 가정하거나 또는 아예 고려하지도 않았다. 그 후 몇몇 논문을 통하여 테스트노력을 고려한 소프트웨어의 신뢰도 평가가 중요한 인자인 것으로 발표되었다. 이들에 의해 지금까지 제안된 형태를 보면 지수함수형, 레일레이형, 웨이불형, 로지스틱형 테스트노력함수로서 경우에 따라 이 중 하나의 적합한 형태를 사용해왔다.

본 논문에서는 이 네 가지 형태에 대해서 최소자승평가자(LSE)와 최대가능성평가자(MLE)를 써서 신뢰도 성장 파라미터를 구하는 방법에 대해서 고찰하고, 실제의 데이터를 적용하여 각각의 경우에 대한 파라미터를 구하고 이를 이용하여 목표신뢰도에 맞는 발행시기를 결정하는 문제를 연구하였다.

1. 서론

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.

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 웨이불곡선으로 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 분야에서는 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

소프트웨어를 시각 t에서 발행하는 경우, rm 신뢰도는

$$R(x|t)=\exp[-m(t+x)+m(t)]$$

$$=\exp[-a\cdot e^{-\gamma W^{*}(t)}(1-e^{-\gamma W^{*}(x)})]$$

이 된다.

따라서, 목표 신뢰도를 $R(x|t)=R_0$ 라 하면

$$W^{*}(t)=\frac{1}{r}\log\frac{1-e^{-rW^{*}n(x)}}{\frac{1}{a}\log\frac{1}{R_0}}$$

2.2 웨이불형 함수

Yamada가 제안한 웨이불형 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 웨이불형 분포로 설명될 수 있고, 아래와 같은 3개의 케이스를 가진다는 것을 보여주고 있다.

1) 지수함수곡선 : (0, t]에서 소요되는 누적 테스트 노력은

$W(t)=N[1-e^{-\beta t}]$ 로서 웨이불 함수의 m=1인 경우에 해당되며, 신뢰도는

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

로 표현된다.

2) 레일레이 곡선 : 소요되는 누적 테스트 노력은

$W(t)=N[1-\exp^{-\frac{\beta}{2}\cdot t^2})]$ 로서 웨이불 함수의 m=2인 경우에 해당된다.

신뢰도는

$$R(x|t)=\exp[-a\cdot e^{-rN(1-e^{-\frac{\beta}{2}t^2})}\cdot(1-e^{-rN(1-e^{-\frac{\beta}{2}x^2})})]$$

로 표현된다.

3) 웨이불 곡선 : 소요되는 누적 테스트 노력은

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

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

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

로 표현된다.

2.3 로지스틱형 함수

실제 테스트 노력 데이타가 여러 가지 소요 패턴을 나타내므로 때때로 테스트 노력 비용을 지수함수나 레일레이 곡선만으로 설명하기는 어렵다. 웨이불형 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

$$W(t))=\frac{N}{1+A\cdot\exp^{-\alpha t}}$$ 이다.

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

로 표현된다.

그림의 초기점에서 웨이불형 테스트 노력 함수와 비교하여 로지스틱 테스트노력 함수인 경우 $W(0)\neq 0$ 이다. 공정이 때때로 확인하기 힘든 소프트웨어 개발의 초기단계에 웨이불형 곡선과 W(t)와의 차이점이 존재한다. 그리고 적용된 테스트 노력의 양을 기록하기 위한 공식적인 계수 공정절차가 수립되지 못한 곳에서도 차이가 난다.

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

현재의 테스트 노력 소요량은 W(t)의 미분치로서

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

와 같이 표현된다.

그림에서 보는 바와 같이 이 양은 $t=\frac{1}{\alpha}\log A$ 일 때 최대치 $\frac{N\alpha}{4}$ 를 중심으로 좌우 대칭형이다.

3. 파라미터의 산출법

3.1 테스트노력의 파라미터의 산출

위의 각 경우에서 정의된 테스트노력함수에서 파라미터 N, A, α, β는 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 제곱해서 더한 총값을 최소화하는 것이다. LSE는 중간 크기의 표본에 최적인 것으로 알려지고 있으며, 최적점 산출을 제공한다. 최소자승법을 적용하기 위한 산출 공식S(N, A, α)은 다음과 같다.

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

W_k = (0, t]기간동안 실제로 소요되는 누적 테스트 노력

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

S를 N, A, α에 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 a와 r을 구할 수 있다. 미지파라미터 a와 r에 대한 조인트 pmf 즉, MLE 함수는

$$L=\Pr\{N(t_i)=m_{i,\ i=1,\cdots,n}\}$$

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

표시되며, 여기에 대수를 취하여 정리하면

$$\log L=\sum_{k=1}^{n}(m_k-m_{k-1})\log a\{e^{-rW^*(t_{k-1})}-e^{-rW^*(t_k)}\}$$

$$-\sum_{k=1}^{n}log(m_k-m_{k-1})!-a[1-e^{-rW^*(t_n)}]$$

이것을 a, r에 관하여 편미분한다.

$$aW^*(t_n)e^{-rW^*(t_n)}$$

$$=\sum_{k=1}^{n}(m_k-m_{k-1})\frac{-W^*(t_{k-1})e^{-rW^*(t_{k-1})}+W^*(t_k)e^{-rW^*(t_k)}}{e^{-rW^*(t_{k-1})}-e^{-rW^*(t_k)}}$$

3.2 각 경우의 파라미터 산출

1) 지수함수형

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

을 만족하는 N값을 구한다. 또한,

$W(t_k)=N(1-e^{-\beta t_k})$ 에서

$$a=\frac{m_n}{1-e^{-rN(1-e^{-\beta t_n})}}$$ 를 구한다.

2) 레일레이형

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

을 만족하는 N값을 구한다.

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

을 만족하는 β값을 구한다

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

를 구한다.

3) 웨이블형

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

을 만족하는 N값을 구한다.

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

즉,

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

을 만족하는 β값을 구한다.

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

을 만족하는 m값을 구한다.

$W(t_k)=N(1-e^{-\beta t_k^m})$ 에서

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

를 구한다.

$$aN(1-e^{-\beta t_n^m})e^{-rN(1-e^{-\beta t_n^m})}$$

$$=\sum_{k=1}^{n}(m_k-m_{k-1})\frac{-N(1-e^{-\beta t_{k-1}^m})e^{-rN(1-e^{-\beta t_{k-1}^m})}+N(1-e^{-\beta t_k^m})e^{-rN(1-e^{-\beta t_k^m})}}{e^{-rN(1-e^{-\beta t_{k-1}^m})}-e^{-rN(1-e^{-\beta t_k^m})}}$$

에서 r을 구한다.

4) 로지스틱형

$$N=\frac{\sum_{k=1}^{n}\frac{W_k}{1+Ae^{-\alpha t_k}}}{\sum_{k=1}^{n}\frac{1}{\{1+Ae^{-\alpha t_k}\}^2}}$$

을 만족하는 N값을 구한다.

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

를 만족하는 A값을 구한다.

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

을 만족하는 α값을 구한다.

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

$$aN\left(\frac{1}{1+Ae^{-\alpha t_n}}-\frac{1}{1+A}\right)e^{-rN\left(\frac{1}{1+Ae^{-\alpha t_n}}-\frac{1}{1+A}\right)}$$

$$=\sum_{k=1}^{n}\{(m_k-m_{k-1})\cdot$$

$$\left[\left\{-N\left(\frac{1}{1+Ae^{-\alpha t_{k-1}}}-\frac{1}{1+A}\right)e^{-rN\left(\frac{1}{1+Ae^{-\alpha t_{k-1}}}-\frac{1}{1+A}\right)}\right\}\right.$$

$$+\left\{N\left(\frac{1}{1+Ae^{-\alpha t_k}}-\frac{1}{1+A}\right)e^{-rN\left(\frac{1}{1+Ae^{-\alpha t_k}}-\frac{1}{1+A}\right)}\right\}$$

$$\left.\div\left\{e^{-rN\left(\frac{1}{1+Ae^{-\alpha t_{k-1}}}-\frac{1}{1+A}\right)}-e^{-rN\left(\frac{1}{1+Ae^{-\alpha t_k}}-\frac{1}{1+A}\right)}\right\}\right]$$

에서 r을 구한다.

3.3 발행시각 결정

1) 지수함수형

발행시각을 $t=T$라 하고, 목표신뢰도를 $R(x|T)=R_0$라 하면

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

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

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

2) 레일레이형

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

3) 웨이불형

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

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

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

4) 로지스틱형

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

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

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

4. 결론

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, 웨이불형이며, 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, 웨이불형은 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 사료된다. 웨이불형은 앞의 두 가지 형에 대한 보완형으로 제안되었다고 할 만큼 융통성이 있어서 실제 현상과 유사한 특성을 나타낸다. 단, 웨이불형인 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.

참고문헌

[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.