

<분야 : 신뢰성 분석>

대규모 시스템의 신뢰도 분석을 위한 NHPP 모형
NHPP Model for Reliability Analysis in Large System

황 정 연*, 강 병 용, 임 주 환, 한 치 문
한국전자통신연구소

ABSTRACT

본 논문은 개발중인 대형 시스템의 소프트웨어 오류에 대한 확률적 신뢰도 모형을 적용하여 분석한다. 그리고 전체 시스템에 있어서 각 서브시스템의 고장 유형 (failure type)을 분석하기 위해 통계적 다변량 기법의 하나인 집락분석(cluster analysis)을 활용하였다. 신뢰도 모형으로는 고장 수 모형(failure count model)인 Goel-Okumoto의 NHPP(nonhomogeneous poisson process)모형을 적용했다. 그리고 각 서브시스템의 유형을 분석하기 위하여 집락분석을 하여 그 결과를 3차원 그래픽으로 표현한다. 그리하여 각 서브시스템별 특성을 자세히 설명한다. 따라서 전체시스템의 신뢰도 분석 결과 잔존 고장 수에 대한 예측을 제시하고, 서브시스템에 대한 분석으로 어느 서브시스템이 잔존 고장 수 퍼센트에 대하여 많은 비중을 차지하는지 알아본다. 본 연구에서 제시한 전체시스템에 대한 고장 발견 모형과 서브시스템별 유형에 대한 잔존 고장 분석은 유사한 기능과 구조를 가진 대규모 소프트웨어 시스템을 개발할 때 시스템 신뢰도 고장 발견 모형으로 사용할 수 있다.

Key Words : NHPP(Non-Homogeneous Poisson Process) Model, 적합성 검정 (good-ness of fit test), 집락분석(Cluster Analysis), 고장 발견 모형 (failure detection model)

대규모 시스템의 신뢰도 분석을 위한 NHPP 모형

NHPP Model for Reliability Analysis in Large System

황 정 연*, 강 병 용, 한 치 문, 임 주 환

한국전자통신연구소

ABSTRACT

본 논문은 개발중인 대형 시스템의 소프트웨어 오류에 대한 확률적 신뢰도 모형을 적용하여 분석한다. 그리고 전체 시스템에 있어서 각 서브시스템의 고장 유형 (failure type)을 분석하기 위해 통계적 다변량 기법의 하나인 집락분석(cluster analysis)을 활용하였다. 신뢰도 모형으로는 고장 수 모형(failure count model)인 Goel-Okumoto의 NHPP(nonhomogeneous poisson process)모형을 적용했다. 그리고 각 서브시스템의 유형을 분석하기 위하여 집락분석을 하여 그 결과를 3차원 그래픽으로 표현한다. 그리하여 각 서브시스템별 특성을 자세히 설명한다. 따라서 전체시스템의 신뢰도 분석결과 잔존 고장 수에 대한 예측을 제시하고, 서브시스템에 대한 분석으로 어느 서브시스템이 잔존 고장 수 퍼센트에 대하여 많은 비중을 차지하는지 알아본다. 본 연구에서 제시한 전체시스템에 대한 고장 발견 모형과 서브시스템별 유형에 대한 잔존 고장 분석은 유사한 기능과 구조를 가진 대형 시스템을 개발할 때 시스템 신뢰도 고장 발견 모형으로 사용할 수 있다.

Key Words : NHPP(Non-Homogeneous Poisson Process) Model, 적합성 검정 (good-ness of fit test), 집락분석(Cluster Analysis), 고장 발견 모형 (failure detection model)

1. 서론

소프트웨어 신뢰도는 매우 넓은 의미를 포함하고 있으며 소프트웨어 공학에서 다루는 모든 방법론과 모형들은 광의의 소프트웨어 신뢰도 분야로 볼 수 있다. 다 시말해 이들은 소프트웨어 신뢰도를 측정하기 위해 수많은 모형들이 제안되어 왔

는데 이 모형들은 크게 4가지로 분류한다.

시험중 개발관리자는 현재의 신뢰도(또는 남은 고장 수)와 원하는 신뢰도 수준이 주어질때, 언제까지 더 시험을 수행 해야 할지를(test stopping rule) 결정해야 한다. 이를 위해 시험 기간 동안 발생했던 고장 자료들을 바탕으로 몇가지 가정을 세우고 통계적 추정 방법들에 의해 신뢰도 추정을 해야한다. 이러한 통계적 추정 모형들은 미시적인 “소프트웨어 신뢰도 모형”이라 할 수 있다. 신뢰도 추정 방법들은 크게 2가지 기법으로 분류 하는데 이산적 기법과 연속적 기법으로 분류한다. 이때 이산적 기법으로는 입력 영역 중심 모형(input domain based model)과 고장 삽입 모형(fault seeding model)이 있다. 그리고 연속적 기법으로는 고장간 시간 모형(time between failure model)과 고장 수 모형(failure count model)으로 분류한다.

본 논문에서는 대규모 스위칭 시스템을 개발하면서 실제 시스템 개발상태의 소프트웨어 신뢰도 고장 발견 모형에 대한 분석 결과를 제시한다. 그리고 전체 시스템 뿐만 아니라 서브시스템에 대해서도 신뢰도 고장 발견 모형에 의한 분석을 한다. 아울러 서브시스템 간의 유형별 특성을 분석하기 위해서 통계적 다변량 기법의 하나인 집락분석을 하여 그 결과를 3차원 그래픽으로 표현한다. 그래서 각 서브시스템별 유형에 대해서 자세히 설명한다. 2장에서는 고장 발견 모형(failure detection model)에 대한 소개로 NHPP 모형에 대한 가정과 평균값 함수(mean value function), 고장율(failure rate), 적합성 검정(good-ness of fit test)에 대하여 설명한다. 3장에서는 실제 개발 경험으로 얻어진 시스템 소프트웨어 고장 자료에 대한 소개와 전체시스템에 대한 NHPP 모형에 의한 분석과 서브시스템의 유형 분석을 위한 집락분석 결과를 설명한다. 마지막으로 4장은 논문의 결론으로 전체 시스템의 신뢰도 분석 결과 잔존 고장 수는 어느 정도이며, 서브시스템에 대해서는 집락분석 결과를 3차원 그래픽으로 표현하여 어느 서브시스템에서 잔존 고장 수 퍼센트에 대하여 많은 비중을 차지하는지 설명한다.

2. 시스템 고장 발견 모형

고장 수 모형(failure count model)은 소프트웨어 신뢰도 추정 방법에 있어서 연속기법 중의 하나로 일정한 시험 기간중에 검출되어 발견된 고장(failure) 수를 확률적 모형화 한 것으로, 이 모형은 시스템으로 부터 고장이 제거 됨으로써 단위 시간당 검출될 고장 수가 감소 할것 이라는 가정을 기본으로 한다. 다음 기본 가정을 설명하면 다음과 같다.

- 고장 수 모형의 시험 구간은 서로 독립(independent)
- 각 시험 구간은 동질적(homogeneous)
- 구간 내의 고장 발생 수들은 상호 독립적

여기에 속한 대부분의 모형들은 poisson 분포를 갖는다고 가정하는 것이 기본 개념이다.

2.1 NHPP(Nonhomogeneous Poisson Process) Model

Goel-Okumoto의 NHPP 모형에 의한 시스템 신뢰도 추정은 추정만을 나타내지 않고 잔존 고장 수를 예측할 수 있는 모형이다. 따라서 시스템 팩키지 배포 시기를 결정하기 위해 언제까지 시험을 수행해야 할 지를(test stopping rule) 결정하도록 제시한다. 다음은 NHPP 모형에 대해서 설명한다.

$\{N(t), t > 0\}$ 을 counting process라 하면, $N(t)=y$ 는 시간 $(0, t]$ 사이에 일어난 고장 갯수를 나타낸다. 누적 고장수는 $N(t)$ 로 표기하고, $\lambda(t)$ 는 민감함수(intensity function) 또는 고장 발견율(failure detection rate)이라고 했을때 아래의 네가지 조건을 포함하면 NHPP(Non-homogeneous poisson process) 모형이라 부른다.

(1) $N(0)=0$

시험 구간 $t=0$ 에서는 고장이 발견되지 않는다.

(2) $\{N(t), t > 0\}$ 은 독립적인 증가를 한다.

다른 시간 구간에서 발견된 고장은 서로 독립이다.

(3) $P\{(t, t+h) \text{ 동안 두개 이상의 사건이 일어남}\} = 0(h)$

$P(t+h)$ 구간에서 2개 이상의 고장이 발견될 확률은 작으므로 무시한다.

(4) $P\{(t, t+h) \text{ 동안 하나의 사건이 일어남}\} = \lambda(t)h+o(h)$

$t+h$ 구간에서 1개의 고장이 발견될 확률은 $\lambda(t)$ 에 비례한다.

Poisson 분포에서 t 시점 까지 누적 고장 수가 y 개일 확률은

$$\Pr\{N(t)=y\} = \{m(t)^y/y!\} * e^{-m(t)}, \quad y=0, 1, 2, \dots \text{-----} (1)$$

식 (1)에서 $N(t)$ 는 누적된 고장 수이고, $m(t)$ 는 t 시점 까지 발견된 평균 고장 수(mean value function)라 부르며, 시간 t 에 대해서 비감소 함수, 증가 함수(non-decreasing function)로 Poisson 분포를 이루며 아래의 조건을 따른다.

$$m(t) = \begin{cases} 0, & t = 0 \\ a, & t \rightarrow \infty \end{cases} \text{-----} (2)$$

식 (2)에서 $t=0$ 은 처음 시간 t 가 0인 상태에서는 고장이 발견되지 않은 것으로

$m(t)=0$ 으로 나타내고, $m(t)=a$ 는 t 시간이 무한대(∞)로 갈수록 전체 발견된 누적 고장수를 의미한다. 여기서 a 는 발견된 전체 누적 고장 갯수이다. 우리가 시험을 무한대로 하더라도 발견될 수 있는 총 고장수는 a 임을 알수있다. 다음 $m(t)$ 의 조건식은 아래의 식 (3)과 같다.

$$m(t) = a(1 - e^{-bt}) \quad \text{----- (3)}$$

$m(t)$ 에 대하여 미분을 하여 구한 $\lambda(t)$ 의 조건식은

$$\lambda(t) = m'(t) = ab * e^{-bt} \quad \text{----- (4)}$$

식 (4)에서 $\lambda(t)$ 는 지수분포 형태를 따르며 민감함수(intensity function), 고장 발견율(failure detection rate)이라 부른다. 그리고 상수 b 는 어떤 시간 t 에 대하여 고장에 관한 발생율(occurrence rate of a failure)로써 비례상수로 본다. 다음 식 (1)에 식 (3)을 치환하면 새로운 식

$$P(N(t)=y) = \frac{\{a(1 - e^{-bt})\}^y}{y!} * e^{-a(1 - e^{-bt})}, \quad y > 0 \quad \text{----- (5)}$$

이 구해진다. 위 식에서 평균은 $m(t)=a(1 - e^{-bt})$ 를 갖는 poisson 분포를 따른다.

2.2 잔존 고장 수

만약, $\bar{N}(t)$ 는 t 시점까지 잔존 고장 수라 하자. 이때 $\bar{N}(t)$ 의 식은 아래와 같다.

$$\bar{N}(t) = N(\infty) - N(t) \quad \text{----- (6)}$$

이고, 잔존 고장 수 $\bar{N}(t)$ 는 모형에서 추정된 총 누적 고장 수(a)에서 t 시점까지의 누적 고장 수를 뺀 값이다. 그리고 잔존 고장 수 $\bar{N}(t)$ 의 평균 기대값은

$$E(\bar{N}(t)) = a e^{-bt} \quad \text{----- (7)}$$

가 된다. 왜냐하면, $N(t)$ 가 Poisson 분포를 식 (5)에서 평균이 $a(1 - e^{-bt})$ 를 갖기 때문이다. 만일, t 시점까지 총 발견된 고장 수를 y 라 하면, $N(t)=y$ 이다. 이때에 잔존 고장 수 $\bar{N}(t)=x$ 에 관한 조건부 분포 식은 다음과 같다.

$$P(\bar{N}(t)=x \mid N(t)=y) = P\{N(\infty)=y+x \mid a\} \quad \text{----- (8)}$$

$$P(\bar{N}(t)=x \mid N(t)=y) = \left[\frac{a^{y+x}}{(y+x)!} \right] e^{-a}, \quad x = 0, 1, 2 \quad \text{----- (9)}$$

$$E\{\bar{N}(t)=x \mid N(t)=y\} = a - y \text{ ----- (10)}$$

식 (10)은 시스템의 t시점에서 잔존 고장 수를 구하는 식으로써 시스템 패키지를 배포 할 것인지, 아니면 배포 하지 않을 것인지에 대한 중요한 결정을 내리게 한다. 따라서 남아있는 고장 수(x)는 소프트웨어 신뢰도에 있어서 매우 중요한 역할을 한다.

2.3 적합성 검정

적합성 검정(good-ness of fit test)이란 모수(parameter : a, b)를 추정(estimation)하기 위하여 주어진 고장 자료를 가지고서 통계적 모형을 설정 할 때, 주어진 자료가 설정한 통계적 확률 모형에 타당한지를 검토하는 통계적 검정(testing) 방법을 적합성 검정이라 한다. 이 적합성 검정에는 모수적 검정(parametric test) 방법으로는 χ^2 -검정이 있고, 비모수적 검정(nonparametric test) 방법에는 Kolmogorov-Smirnov 검정이 있다. 또한, 모형의 적합성 검정으로 F-검정이 있다. 이 검정은 분산분석표(anova table)에 의해서 검정한다. 다음은 적합성 검정 절차 순서에 대하여 설명한다.

1. 가설을 설정한다.

$$\begin{cases} H_0(\text{귀무가설}) : \text{설정된 NHPP 모형은 자료에 적합하지 않다} \\ H_1(\text{대립가설}) : \text{설정된 NHPP 모형은 자료에 적합하다} \end{cases}$$

2. 유의수준 $\alpha=(0.05, 0.01)$ 값을 설정한다.

$F(\theta_1, \theta_2 ; 0.05)$ 인 F-분포의 표값을 찾는다

3. 검정통계량 F값을 구한다.

F값은 $F=MSR/MSE$ 이다

4. 결론을 내린다.

유의수준에 의한 F-분포의 표값과 F-검정통계량 값을 서로 비교하여 검정한다.

3. 시스템 소프트웨어 고장 자료와 신뢰도 분석

대형 시스템을 개발하는 과정에 있어서 시스템의 현재 상태와 신뢰도를 평가하여 추후 개발 일정과 관리에 반영하는 일은 상당히 중요하다. 여러 개의 서브시스템이 결합하여 하나의 시스템을 형성할 때, 각 서브시스템의 개발 상태와 신뢰도를 평가하고 전체 시스템에 대한 서브시스템의 유형을 분석하는 것은 대형 시

시스템의 고 품질을 개발하는데 필수적이다.

3.1 시스템 고장 자료

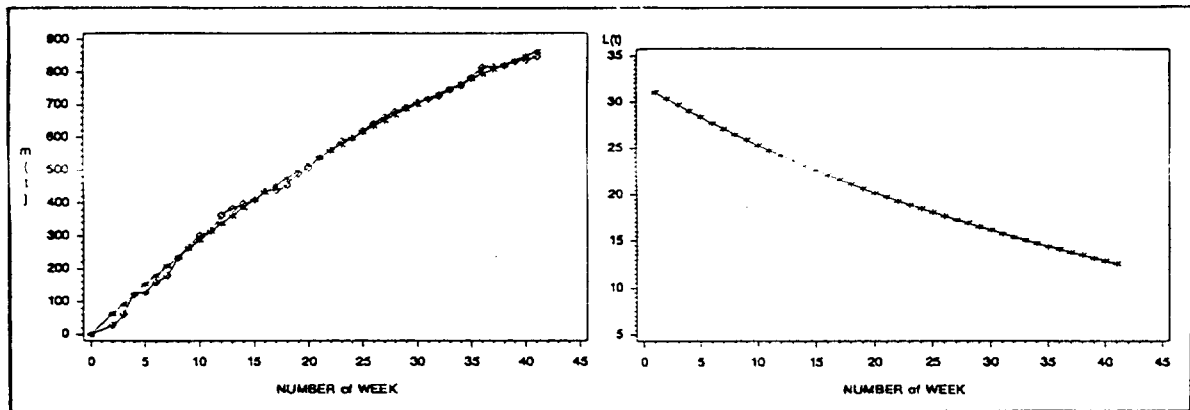
본 연구에 이용된 대형 시스템은 프로그램 소스코드가 약 1,500,000 라인으로 대규모 시스템이다. 이 시스템을 통합하고 시험하는 작업은 많은 자원과 시간이 요구되었다. 이 과정에서 새로운 요구사항의 발생으로 인하여 기능이 추가되거나 변경되기도 하였다. <표 1>은 대형 시스템의 시험 단계에서 검출된 시험 횟수별 고장 수와 누적 수를 나타내고 있다.

Time	Failure	Cumulative Failure	Time	Failure	Cumulative Failure
1	16	16	22	24	562
2	13	29	23	27	569
3	32	61	24	10	599
4	60	121	25	21	620
5	8	129	26	22	642
6	22	159	27	19	661
7	20	179	28	17	678
8	56	235	29	13	691
9	33	268	30	16	707
10	34	302	31	9	716
11	15	317	32	10	726
12	47	364	33	20	746
13	19	383	34	13	759
14	16	399	35	24	783
15	13	411	36	30	813
16	26	437	37	3	816
17	3	440	38	3	819
18	18	458	39	13	832
19	32	490	40	4	836
20	18	508	41	10	846
21	20	538			

<표 1 발견된 전체 시스템의 고장 자료>

3.2 전체 시스템과 서브시스템 분석

시스템 고장 자료에 대하여 NHPP 모형을 적용하여 분석하는데 SAS(statistics analysis system)라는 통계 패키지를 활용하였다. 아래의 <그림 1>에서 좌측은 전체 시스템에 대한 NHPP 모형의 평균값 함수 $m(t)$ 에 대한 그림이고, 우측은 고장 발견율 $\lambda(t)$ 에 대한 그림이다.



<그림 1 전체 시스템의 NHPP 모형에 의한 $m(t)$ 와 $\lambda(t)$ >

다음 아래의 <표 2>는 전체 시스템에 대해서 NHPP 모형을 적용하여 모수(a, b)를 추정할 때 얻어진 결과치 이다.

Source	DF	Sum of Squares	Mean Square	④ F
Model	2	12974568.614	6487284.307	31818.32
Residual	39	7951.386	203.882	
Total	41	12982520.000		

Parameter	Estimate	Asymptotic Std. Error	③ Asymptotic 95 % Confidence Interval	
			Lower	Upper
① a	1426.328630	11.396097013	1403.2779593	1449.3793001
② b	0.022596	0.000254679	0.0220805	0.0231108

<표 2 전체 시스템에 NHPP 모형을 적용한 모수 추정 결과>

위의 <표 2>에서 ①의 a값은 전체 시스템에 대한 NHPP 모형을 적용하여 구한 것으로 누적 고장 수의 값은 1426.329이다. 그리고 ②의 b값은 0.022596으로 어떤 시점 t에 대한 고장 발생율을 나타낸다. 다음은 식 (3)과 (4)에 의해서 평균 함수 값 m(t)와 고장 발견율 λ(t)의 값을 구하면 아래와 같다.

$$m(t) = 1426.329 * (1 - e^{-0.022596t}) \text{ 이고,}$$

$$\lambda(t) = 1426.329 * 0.022596 * e^{-0.022596t}$$

다음은 잔존 고장 수를 얻기위해 식 (10)에 의해서 구하면 $\bar{N}(t) = 1427 - 846 = 581$ 이다. 따라서 전체 시스템에 대하여 얻어진 잔존 고장 수는 581 임을 알 수 있다. 그리고 잔존 고장 수에 대한 95% 신뢰구간을 추정하기 위해 ③의 결과치에 의해서 구하여 보면 다음과 같다.

$$\bar{N}(t)_L = 1403.278 - 846 = 557.278 \text{ 이고,}$$

$$\bar{N}(t)_U = 1449.379 - 846 = 603.379$$

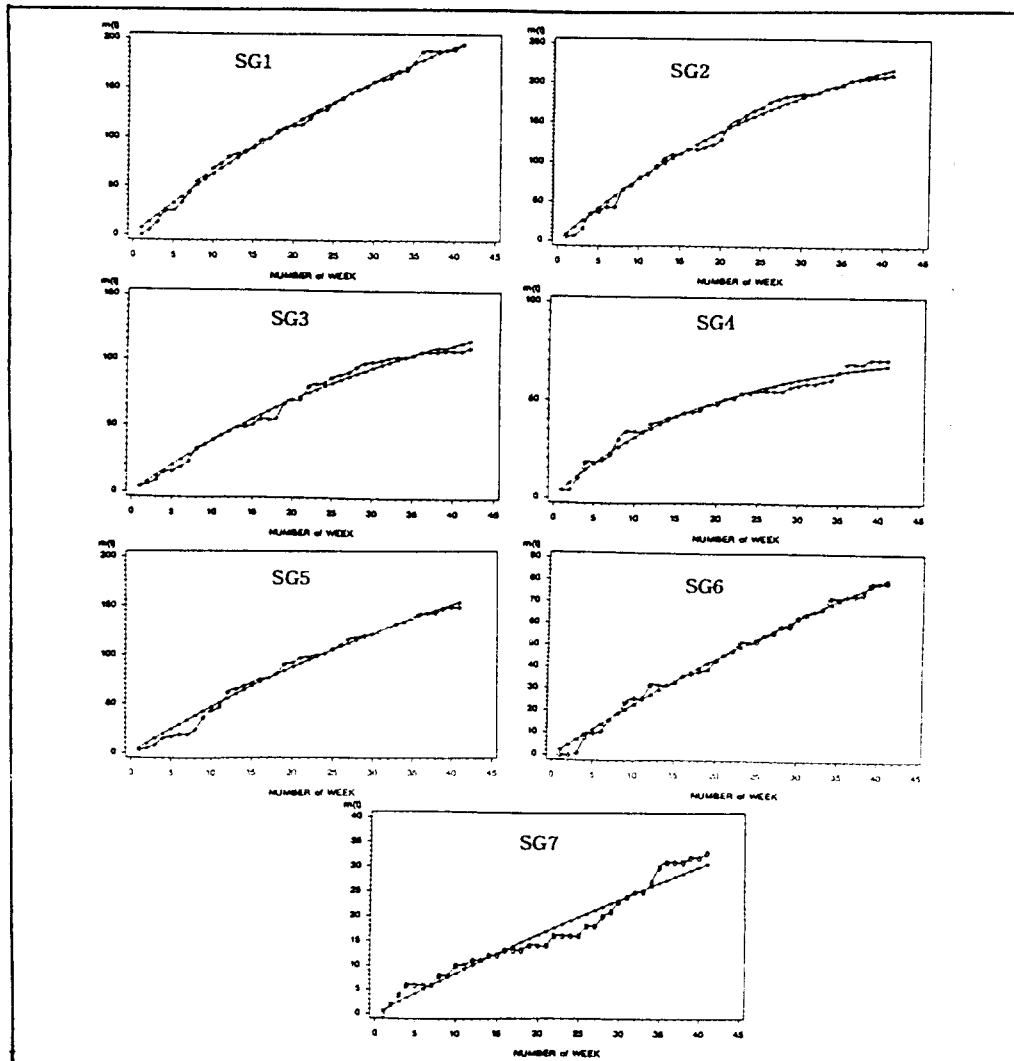
따라서 전체 시스템에 대한 잔존 고장 수의 95% 신뢰구간을 추정하면

$$\bar{N}(t)_L \leq \bar{N}(t) \leq \bar{N}(t)_U \text{ ----- (11)}$$

위의 식 (11)에 의해서 95% 신뢰구간을 추정하여 보면, 잔존 고장 수는 558 ~ 604 임을 알 수 있다. 다음은 NHPP 모형에 대한 적합성 검정(good-ness of fit test)을 살펴본다.

- (1) 가설 $\begin{cases} H_0(\text{귀무가설}) : \text{설정된 NHPP 모형은 자료에 적합하지 않다} \\ H_1(\text{대립가설}) : \text{설정된 NHPP 모형은 자료에 적합하다} \end{cases}$
- (2) 유의 수준 $\alpha=0.05$ 나 0.01 로 설정한다. 이때 F-분포의 표값은 $F(2, 39; 0.05)=3.24$ 와 $F(2, 39; 0.01)=5.195$ 이다.
- (3) 검정 통계량 F의 값은 ④인 $F=31818.32$ 이다.
- (4) 결론을 내리면, F-검정 통계량 $F=31818.32 > F(2, 39; 0.05)=3.24$ 보다 크므로 $H_0(\text{귀무가설})$ 를 기각한다.

따라서 $H_0(\text{귀무가설})$ 를 기각하고, $H_1(\text{대립가설})$ 을 채택하므로 설정한 NHPP 모형은 본 자료의 신뢰도 모형에 적합하다고 판단된다. 또한 <그림 1>에서 발견된 고장 자료와 NHPP 모형에 의한 $m(t)$ 는 동일하게 나타났다. 다음은 각 서브시스템별 NHPP 모형에 의한 신뢰도 분석 결과를 나타낸 평균값 함수 $m(t)$ 의 그림이다.



<그림 3 서브시스템별 NHPP 모형에 의한 $m(t)$ >

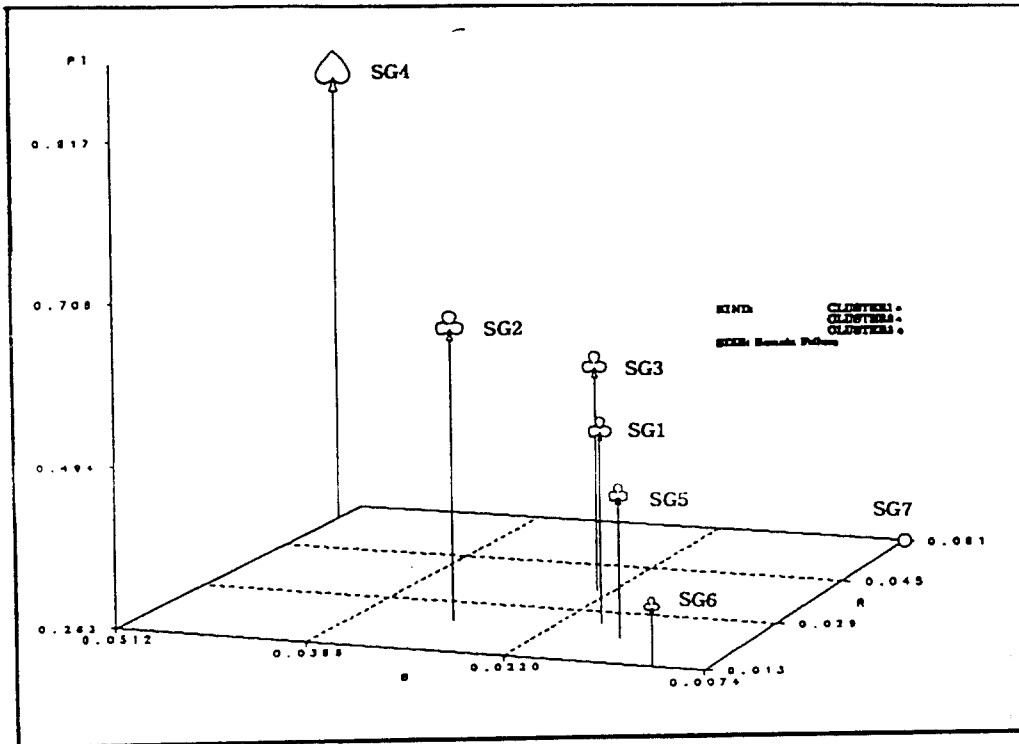
3.3 서브시스템 유형에 대한 집락분석

전체 시스템 고장 자료에서 서브시스템을 7가지로 분류하여 각 서브시스템에 대해서 각각 총 고장 수(a), 어느 한 시점에서의 고장 발견율(b), t시점까지 발견된 고장을 퍼센트(P_T), 평균 굴절율(r)을 구하여 최단 연결법(single linkage)을 이용하여 집락분석을 하였다. 그리고 그 결과는 아래와 같다.

SG	A	B	PT	R	Y	CLUSTER	CLUSNAME
SG1	360.803	0.018956	0.538	0.025	194	1	CL3
SG5	314.776	0.016250	0.470	0.021	148	1	CL3
SG2	311.125	0.029428	0.668	0.023	211	1	CL3
SG6	219.457	0.011146	0.360	0.013	79	1	CL3
SG3	188.279	0.023138	0.584	0.036	110	1	CL3
SG7	116.635	0.007421	0.283	0.061	33	2	SG7
SG4	77.467	0.051153	0.917	0.056	71	3	SG4

<표 3 서브시스템에 대한 집락분석 결과>

위의 집락분석 결과 SG1, SG5, SG2, SG6, SG3이 집락 1로 형성되었고, 집락 2에는 SG7이 형성 되었다. 그리고 집락 3에는 SG4가 각각 동일한 집단으로 형성 되었음을 알 수 있다. 다음 아래의 그래픽은 집락분석 결과를 3차원 그래픽으로 나타내어 각각 서브시스템별 특성을 나타낸다.



<그림 4 집락분석에 의한 3차원 그래픽>

위의 <그림 4>에서 크로바(club) 모양은 집락 1을 나타내고 원모양은 집락 2를 나타낸다. 그리고 스페이드(spade) 모양은 집락 3을 의미한다. 여기서 P_T 는 t 시점까지 발견된 고장 수의 퍼센트를 의미한다. 다음 B 는 어느 시점의 구간에 대한 고장 발견율을 의미하며 마지막으로 R 은 평균 굴절 비율에 의해서 표현된 것이다. 여기서 집락 2인 SG4의 경우는 t 시점까지 고장 수 퍼센트가 가장 많이 발견되었음을 알려준다. 반대로 발견된 고장 수 퍼센트가 가장 낮은 서브시스템의 그룹은 SG7임을 나타내고 있다. 다음 위의 그래픽에서 어느 시점의 구간에 대한 고장 발견율이 가장 높은 서브시스템은 SG4이다. 결국, 전체 시스템에 대한 잔존 고장 수 퍼센트에 대하여 많은 비중을 차지하는 서브시스템으로는 SG7, SG6, SG5, SG1, SG3 순으로 아직 잔존 고장 수가 많이 발견 되어야 함을 알 수 있다.

4. 결 론

본 연구는 개발중인 대형 시스템에 관한 고장 발견 모형으로 Goel-Okumoto의 고장 수 모형인 NHPP 모형을 적용하여 잔존 고장 수에 대하여 예측을 하였다. 그리고 기능별 7개 서브시스템에 대해서도 각각 고장 수 모형인 NHPP 모형을 적용하여 각 서브시스템별 5가지 특성을 각각 구하여 집락분석을 활용하였다. 그리하여 서브시스템에 대한 유형을 3차원 그래픽으로 표현해서 자세히 설명했다. 앞으로 시스템 소프트웨어 신뢰도에 관한 연구는 단순히 신뢰도 모형을 적용하여 잔존 고장 수를 예측하는 것에만 국한 할게 아니라, 폭 넓게 블럭별 또는 고장 발생 원인 등을 고려하여 다양한 측면에서 원인파악을 위한 시스템의 고장 유형 분석을 연구하여야 할 것이다. 이렇게 함으로써 대형 시스템에 대한 고 품질의 제품을 개발 할 수 있다.

[참 고 문 헌]

- [1] Amrit L. Goel and Kazuhira Okumoto. "Time-Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures," *IEEE Trans. Reliability*, Vol. R-28, 1979, pp. 206-211.
- [2] Amrit L. Goel. "Software Error Detection Model with Applications," *The Journal of system and software* 1. 1980, pp. 243-249.
- [3] Amrit L. Goel. "Software Reliability Models : Assumptions, Limitations, and Applicability," *IEEE Trans. Software Engineering*. Vol SE-11, 1985.
- [4] John D. Musa and Kazuhira Okumoto. "Software Reliability Measurement, Prediction, Application," 1987, McGraw-Hill Book Co.
- [5] Martin L. Shooman. "Software Engineering Design, Reliability and Management ," 1983 McGraw-Hill Book Co.
- [6] Norman F. Schneidewind. "Software Reliability Model with Optimal Selection of Failure Data," *IEEE Trans. Software Eng.*, Vol 19, pp. 1095-1993.
- [7] Ram Chillarege, Wei-Lun Kao and Richard G. Condit. "Defect Type and its Impact on the Growth Curve," *IEEE Trans. Reliability*, 1991, pp. 246-255.
- [8] Sheldon M. Ross. "Software Reliability: The Stopping Rule Problem," *IEEE Trans. Software Engineering*., 1985 Vol. SE-II, pp. 1472-1476.
- [9] A, A. Afifi. "Computer-Aided Multivariate Analysis," 2nd Edition., 1990, Van Nostrand Reinhold.