

고장진단 전문가시스템 지원에 편리한 신뢰도 계산방법

이 승 철

(중앙대학교 공과대학 전기공학과 교수)

1. 서 론

최근 수년간에 걸쳐 발전소 및 전력계통의 각종 고장진단을 실시간에 수행하기 위한 전문가시스템(expert system)들의 개발이 활발히 진행되어 왔고 이러한 개발 추세는 앞으로 더욱 가속될 전망이다. 그러나 이러한 대형 플랜트나 시스템용 고장진단 전문가시스템들이 본격적으로 실용화되기 위하여는 빠른 처리속도외에도 선행 해결되어야 할 두 가지의 중요한 과제가 있다.

그 첫째는, 전문가시스템들에 각종 계기측정치들의 정확한 상황을 알려주는 일이다. 즉 거의 실시간에 운전상황을 판단해야하는 전문가시스템에 있어서는 계기들로부터 들어오는 각종 측정치들이 운전상황판단과 고장진단의 궁극적인 근거가 된다. 그러나 실제 계기로 부터 들어오는 측정치들은 여러가지 원인으로 인하여 잡음이 섞이거나 바이어스를 갖기도하고 또는 계기자체의 고장으로 인하여 틀린값을 갖는 경우가 자주있다. 전문가시스템에 입력되는 계기측정치들 중 정확도가 매우 낮거나 틀리는 측정치들이 있을 경우 해당 측정치들이 틀렸다는 것을 발견하여 고장진단의 추론과정에서 제외시키지 않으면 대개는 측정치들 상호간에 존재하는 해석상의 논리적인 모순으로 인해 고장진단이 불가능하여 진다. 더욱이 바람직하지 않은 경우는 그러한 틀린 계기값들에 근거하여 우연히 어떤상황 판단의 결론에 이른 경우이며, 이경우에는 잘못된 상황판단에 근거한 오조작과 그에따른 더 큰 사고를 유발할 가능성도 있다.

따라서 계기측정치들이 나타내는 어떠한 고장의 징후도 먼저 계기값이 틀렸는지 실제 고장인지 아니면 드물게 양쪽 다 인지의 구별이 가능한 한 선행되어야 한다. 이와 같은 이유로 인하여 많은 고장진단 전문가 시스템의 경우 고장진단의 편의를 위하여 주어지는 계기값이 항상 정확하다는 다소 비현실적인 가정하에 고장진단을 수행하는 것을 종종 볼 수 있다. 따라서 고장진단을 정확히 하기 위하여는

먼저 전문가시스템들의 궁극적인 판단의 근거가되는 계기 측정값들의 정확한 상황을 해당 전문가시스템에 알려주어 전문가시스템이 입력되는 계기값들을 분별력(discretion)있게 사용할 수 있도록 해야한다.

둘째는, 고장진단의 대상이되는 각종기기들의 신뢰도를 전문가시스템에 알려주거나 또는 운전중 전문가시스템이 스스로 신뢰도 관련정보를 계속적으로 수집하여 update할 수 있도록 self learning능력을 갖추도록 하여야한다. 이는 현실적으로 발전소나 전력계통과 같은 대형플랜트나 시스템에서 모든 기기들의 모든고장을 발견할 수 있도록 계기들을 설치할 수는 없기 때문이다. 따라서 시스템에 이상이나 고장이 발생할 경우 대개는 설치되어 있는 계기들의 위치와 종류에 따라 고장구역과 고장의 종류에 대한 분별력(resolution)의 한계가 정해진다. 이경우 전문가시스템이 고장구역내에 있는 기기와 계기들의 고장종류별 신뢰도에 대한 정확한 정보를 가지고 있으면 보다 자세하고 고장복구에 실질적인 도움을 줄 수 있는 진단결과를 제시할 수 있다.

계기측정치들의 정확한 상황을 알기위하여는 각 계기들의 정밀도와 신뢰도에 근거한 계기값들의 유효성여부를 검증하여야 한다. 계기값들의 유효성 여부를 검증려면 계기값들에 어떠한 형태로든 redundancy가 존재해야하며 검증방법은 크게 temporal redundancy와 spatial redundancy에 근거하여 검증하는 방법으로 나눌 수 있다. 본 논문에서는 주로 고장진단 전문가시스템 지원에 편리한 신뢰도 계산방법을 소개하고자 하며 계기값들의 검증방법은 따로 [2, 7, 9]등을 참조하기 바란다.

발전소나 전력계통은 수많은 기기와 계기들로 구성이 되어 있으며 각 기기나 계기들의 고장발생빈도가 비교적 낮은 것이 현실이다. 또한 운전중인 기기나 계기들을 따로 떼어내어 여러 표본을 구성해서 신뢰도 계산을 위한 시험을 한다는 것은 현실성이 없다. 따라서 이들 기기나 계기들의 신뢰도 계산을 하기 위한 충분하고 체계적인 데이터의 수집이 어려워 각종기기의 신뢰도를 실제 고장진단에 반영하

는데에는 현실적으로 많은 어려움이 따르고 있다. 본 논문에서는 이와같이 표본수가 비교적 적고 표본별로 테스트 시간이 불규칙하거나 사건에 체계적인 통계분석 계획이 없이 수집된 미흡한 데이터들을 이용하여서도 신뢰도를 계산하기에 편리한 방법을 하나 소개하고자 한다. 또한 이방법은 미흡한 데이터들을 실제 운전을 해 가면서 계속적으로 보완해 나가기도 편리하다. 여기서 소개하고자 하는 방법은 주로 과거에 비행기, 로켓 또는 Space Shuttle의 터보엔진의 터빈날개의 crack이나 배어링 이상 여부, 또는 그에 설치된 각종계기들의 신뢰도를 계산하는데 쓰여져 온 방법으로 [1, 6], 표본 데이터 양이 부족하고 또 고장빈도도 매우 적거나 심지어 없는 경우, 그리고 표본들의 테스트 시간이 불규칙한 경우에도 적용이 가능한 이점이 있다.

2. 이론 배경

여기서 소개하는 신뢰도 계산방법은 Waloddi Weibull이 1951년 제안한 Weibull distribution에 기초한 방법으로 신뢰도 계산을 간단히 graph를 이용하여 수행할 수 있는 장점이 있다. Weibull distribution은 shape parameter와 특성수명(characteristic life or scale parameter) 이라고하는 2개의 parameter를 갖는데 그 shape parameter의 크기에 따라 다른 여러 distribution들을 유사하게 나타낼 수 있어 매우 융통성이 큰 distribution이다.

먼저 Weibull probability density function (pdf)은 다음과 같이 표현된다.

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (1)$$

여기서 β 는 shape parameter이고 η 는 특성수명이며 t 는 통상 기기의 사용시간이나 사용횟수(cycle)를 나타낸다.

상기 density function으로 부터 Weibull distribution은 다음과 같이 나타내어진다.

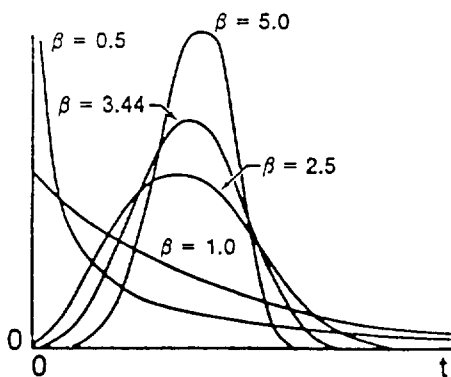


그림 1. 여러 Shape Parameter값에 따른 Weibull Probability Density Function

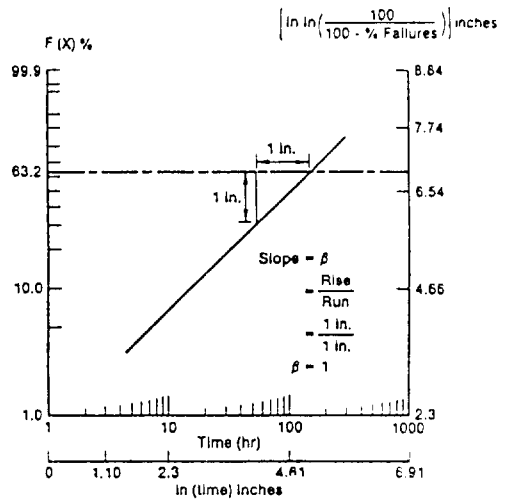


그림 2. Weibull Analysis Graph 용지의 구성

$$F(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (2)$$

위에서 $F(t)$ 는 결국 t 에서 고장날 확률 즉 unreliability가 된다. Weibull pdf를 β 와 η 값에 근거하여 그래프로 나타내면 그림1과 같다.

그림 1에서 알 수 있듯이 Weibull distribution은 β 값이 1 근처에서는 exponential, 그리고 3.5 근처에서는 normal distribution에 근접해 가는 것을 알 수 있다.

식 (2)로 부터 t 에서의 신뢰도(Reliability)는

$$R(t) = 1 - F(t) = e^{-\left(\frac{t}{\eta}\right)^\beta} \quad (3)$$

로 되고 따라서

$$\frac{1}{1 - F(t)} = e^{\left(\frac{t}{\eta}\right)^\beta} \quad (4)$$

상기 양변에 자연대수를 두번 취하면

$$\begin{aligned} \ln \ln \left(\frac{1}{1 - F(t)} \right) &= \beta \ln \left(\frac{t}{\eta} \right) \\ &= \beta \ln t - \beta \ln \eta \end{aligned} \quad (5)$$

따라서 상기 식은 다음과 같은 직선식으로 놓아질 수가 있다.

$$Y = \beta X - \beta \ln \eta \quad (6)$$

여기서

$$\begin{aligned} X &= \ln t \\ Y &= \ln \ln \left(\frac{1}{1 - F(t)} \right) \end{aligned} \quad (7)$$

상기 식 (7)은 그림 2와 같은 대수용지를 사용할 경우

직선으로 나타낼 수 있다.

식 (6)으로 부터 β 는 상기 graph상의 직선의 실제 기울기가 됨을 알 수 있다. 또한 식 (2)에서 $t = \eta$ 로 놓을 경우

$$F(t) = 1 - e^{-1} = 0.632 \quad (8)$$

가 되어, 특성수명 η 는 β 값에 관계없이 기기수명의 63.2%에 해당되는 시간임을 알 수 있다.

3. 신뢰도 및 특성수명 계산

실제 기기의 실험 또는 운전 데이터로 부터 전항의 그림 1에서 보인 Weibull analysis paper에 그래프를 그리기 위해서는 각각의 고장시간 t_j 에 해당하는 고장확률 $F(t_j)$ 를 구할 필요가 있다. 여기서 t_j 는 고장난 표본들을 sorting하여 고장시간의 길이 순으로 나열하였을 경우 j 번째의 고장시간이다. $F(t_j)$ 의 값은 Rank Regression방법을 사용하여 다음과 같이 구할 수 있다.

3.1 Median Rank값의 산출

고장데이터에서 n 개의 고장시간중 각각의 고장시간 t_j 에 대한 $F(t_j)$ 를 rank regression 방법에 의해 구하기 위하여는 먼저 데이터를 고장시간의 증가순서대로 sorting하여야 한다. 이와 같이 sorting한 n 개의 데이터상에서 j 번째의 고장시간 t_j 를 기준으로 생각할 때, $(j - 1)$ 개의 고장은 이미 발생하였고 한개의 고장이 t_j 의 순간에 발생하였으며 나머지 $(n - j)$ 고장은 t_j 이후에 발생하였으므로 각각의 결과가 발생할 확률을 distribution $F(t_j)$ 와 density $f(t_j)$ 를 사용하여 다음과 같이 나타낼 수 있다.

Number of Outcomes	Probability
$j - 1$	$F(t_j)$
1	$f(t_j)dt$
$n - j$	$1 - F(t_j)$

상기의 상황(event)은 multinomial pdf를 이용하여 다음 식으로 표시할 수 있다.

$$\frac{n!}{(j-1)! 1! (n-j)!} [F(t_j)]^{j-1} f(t_j) dt * [1 - F(t_j)]^{n-j} \quad (9)$$

상기 pdf의 면적적분이 50%가 되는 $F(t_j)$ 의 값을 구하여 이를 $F(t_j)$ 의 Median Rank값이라 하며 상기와 같은 event가 생길 가능성을 가능한 한 크게하는 $F(t_j)$ 의 값이라 하겠

다. 이러한 $F(t_j)$ 의 값은 n 과 j 의 값에따라 미리 계산되어 통상 median rank table이라고 불리는 표에 그값이 열거되어 있다. 실용적으로 $F(t_j)$ 는 다음과 같은 개략식을 사용하여 무방하다.

$$F(t_j) = \frac{j-0.3}{n+0.4} \quad (10)$$

여기서 n 는 전체 표본(sample)수 이다. (상기의 Median Rank값대신 $F(t_j)$ 의 기대값(Expected Value)을 사용 할 수도 있으며 이 경우는 다음식을 이용할 수 있다.

$$E[F(t_j)] = \frac{j}{n+1}$$

3.2 Shape Parameter와 특성수명의 산정

각 고장시간에 대하여 median rank값이 정해지면 해당되는 점들을 Weibull analysis paper에 plot한다. 다음에는 이들 plot된 점들을 통과하는 직선을 목적(目測)으로 긋는다. 이때 shape parameter β 는 그어진 직선의 기울기를 직접 측정함으로써 얻을 수 있고, 특성수명 η 는 이 직선이 종축인 $F(t)$ 축의 63.2%에 해당하는 값의 t 를 횡축으로부터 구함으로써 얻을 수 있다. 일단 이와같이 Weibull 직선이 얻어지면 임의의 시간에서의 신뢰도나 역으로 일정한 신뢰도에 이르는 시간을 간단하게 그래프로 부터 구할 수가 있다. 이와같이 목적으로 직선을 그어 β 와 η 를 구하는 것은 간단하면서도 실용적으로도 신뢰도계산에 별 문제가 없다. 만일 β 와 η 를 좀더 자세히 구하고 싶으면 최소자승 (least square) 법을 사용하여 각 plot된 점들을 전체적으로 가장 가깝게 지나가는 직선을 구하면 된다.

3.3 고장나지 않은 표본들이 섞여있을 경우

실제 현장에서 데이터들을 수집하다 보면 장시간 테스트나 실제 운전을 하였음에도 불구하고 고장이 발생치 않은 기기가 많이 있게 마련이다. 예를 들어 전체 n 개의 시험기기들중 r (여기서 $r < n$)개만 시험시간중 고장이 나고 $(n - r)$ 개는 고장이 발생치 않았다고 하자. 이때 고장이 발생치 않은 $(n - r)$ 개는 고장이 발생치 않은 시간이 충분히 길 경우, 이 기기들의 전체 신뢰도를 향상시키는 방향으로 기여하여야 할 것이다. 이와같이 테스트기간중 고장이 발생치 않았으나 여러사정으로 인하여 테스트를 중단하여야 했을 경우 일정시간 테스트에 견디어 낸 표본들을 suspended item들이라하며 그러한 suspended item들을 포함하고 있는 테스트데이터를 censored data라 한다. 이러한 suspended item을 감안하여 신뢰도 계산을 할 경우는 sorting시 이들 suspended item을 포함하여 ordering을 하며, 이경우 고장난 item들의 rank order number가 suspended item들의 영향으로 약간씩 증가하게되나 (즉 $F(t_j)$ 에서 j 의 값이 증가) 표본수의 증가에 따른 n 의 크기의 증가가 더 크기때문에 그결과로 $F(t_j)$ 의 median rank값들이 전체적으로 감소하게 된다 (식 (10) 참조). 따라서 Weibull 그래프상에서 전체 $F(t_j)$ 의 직선이 기울기는 일정한채 오른쪽으로 이동하게되어

특성수명이 늘어나는 결과를 초래하게 된다. Suspended item 을 포함할 경우의 rank order number의 조정은 다음 식에 의거하여 행하여 준다.

$$\text{Rank Increment} = ((n + 1) - (\text{previous sensor's adjusted rank order})) / (1 + (\text{number of items beyond present suspended item}))$$

3.4 Maximum Likelihood 방법을 이용한 Weibull Analysis

β 와 η 의 값은 상기와 같이 그래프를 사용하는 외에도 특정한 고장데이터가 얻어질 확률인 likelihood function을 최대화하는 Maximum Likelihood 방식에 의해 직접구할 수도 있다. 그러나 β 값을 구하기 위하여는 Newton_Raphson method등의 numerical iteration procedure를 사용해야하는 번거로움이 있고 또 고장 mode가 하나 이상일때는 기종기가 다른 2개 이상의 직선을 그어야 할 때도 있기 때문에 본 해설에서는 비교적 간단하면서도 실용적으로 이점이 있는 본 그래프방식만을 소개하였다.

3.5 고장난 item이 전혀 없을 경우의 접근방식

실험이나 운전을 하다 보면 때로는 기기들이 일정 테스트시간이 경과하여도 전혀 고장이 나지 않는 경우도 많이 있다. 이러한 경우는 소위 Weibayes 방법을 써서 신뢰도를 계산해 낼 수 있다. 근본적으로 Weibayes방법은 비록 현재까지 고장이 발생치 않았더라도 첫번째 고장이 임박하였다고 가정할 경우 conservative한 신뢰도의 추정이 될 수 있다는데 착안하여 고안되었다. 이 경우 failure mode를 나타내는 β 값은 표본들 또는 표본들과 유사한 기기들의 과거 사용경험에 기초한 engineering judgement에 의거하여 적절한 값을 상정할 필요가 있다.

이어서 η 는 상정한 β 값을 이용하여 maximum likelihood estimate를 통해 다음식에서 산출할 수가 있다.

$$\hat{\eta} = \left(\frac{\sum_{j=1}^N t_j^{\beta}}{r} \right)^{\frac{1}{\beta}}$$

상기식에서 β 는 상정한 β 값이고 $\hat{\eta}$ 는 η 의 maximum likelihood estimator이다. r 는 고장난 표본 수이나 이 경우 고장난 표본이 없으므로 r 는 1로 대입하면 된다. 이와같이 얻어진 $\hat{\eta}$ 는 참 η 값의 conservative한 63% lower confidence bound가 된다.

4. 적용 예

4.1 Complete Data의 경우

지금까지 설명한 신뢰도 계산 방법의 적용을 예를 들어 보이고자한다. 다음의 표 1을 어떤 기기들의 고장났을때까

표 1 고장 데이터 예

Serial No. of the Equip.	Failure Time (min)
1	90
2	96
4	30
5	49
8	82

표 2 정렬된 표 1과 Median Rank 값

Rank Order No.	Serial No.	Failure Time	Median Rank
1	4	30	12.9
2	5	49	31.3
3	8	82	50.0
4	1	90	68.6
5	2	96	87.0

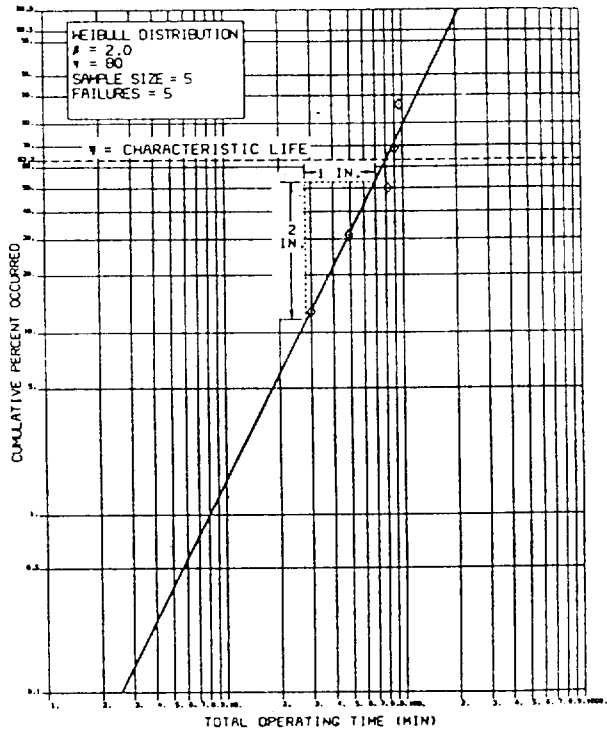


그림 3. 표 2의 결과로 부터 얻어진 Weibull Analysis Graph

지의 시간 (또는 사용 횟수로 하여도 무방함.) 이라고 가정 하자. (표본기기가 하나일 경우, 일단 수리하고 나면 새것과 다름없는 수명 특성을 나타낸다고 간주할 수 있을때에는, 즉 memoryless 특성일 경우, 기기의 수리후 다음 고장까지의 시간간격들로 취급할 수도 있을 것이다.)

상기 데이터를 고장시간의 증가순으로 나열하고 식 10에 의거하여 median rank의 값을 구하면 표 2와 같다.

표 2의 결과를 Weibull 대수용지에 plot하고 이들 plot된

점들을 가능한 한 가깝게 지나는 직선을 그리면 그림 3과 같다. 이와같이 그려진 그래프에서 β 가 2인 것을 직선의 기울기로 부터 구할 수 있고, 또 특성수명이 80인 것은 $F(t)$ 가 63.2%인 선과 직선이 만나는 점의 시간으로 부터 구할 수가 있다.

4.2 Censored Data의 경우

아래의 표 3은 전항의 표 2의 경우이다 시험을 종료했을 때까지 고장이 발생치 않은 3개의 데이터를 추가했을 경우이다. 고장이 발생치 않은 표본들의 시험결과는 시험시간의 길이 순서대로 표 2에 삽입하였다. 이 경우 rank order는 식 (11)에 의거하여 다시 조정하였으며 조정된 결과도 표 3에 함께 나타내었다.

표 3 Suspended Item들을 포함한 Censored Data

Rank Order No.	Serial No.	Failure Time(min)	Median Rank
susp.	7	10	
1.125	4	30	9.82
susp.	6	45	
2.438	5	49	25.45
3.751	8	82	41.08
5.064	1	90	56.71
6.377	2	96	72.35
susp.	3	100	

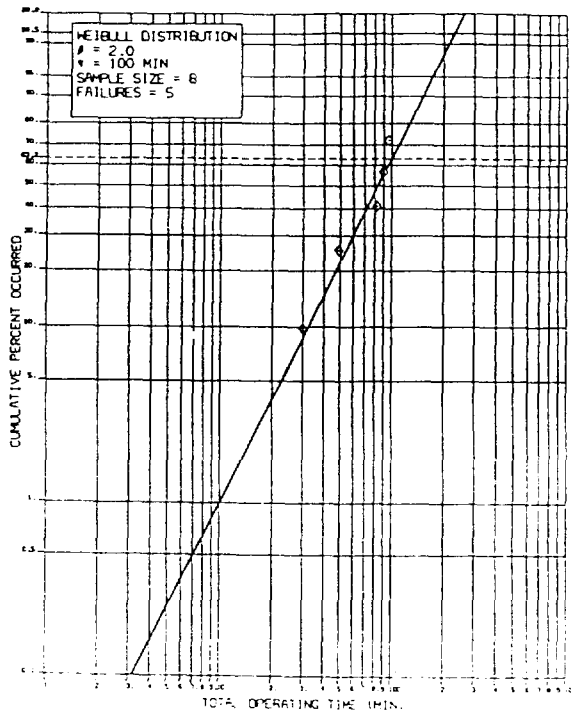


그림 4. 표 3의 결과로 부터 얻어진 Suspended Item들을 포함한 Weibull Analysis Graph

상기 표 3의 결과를 Weibull analysis paper에 plot하고 graph를 그리면 그림 4와 같다. 그림 3과 비교하여 기울기 즉 shape parameter β 는 같고 특성수명 η 만이 100으로 증가 되었음을 알 수 있다. 따라서 고장 데이터에 suspended item들이 추가 될 경우 고장 mode에는 영향이 없고 특성수명만이 늘어나 전체적으로 신뢰도가 증가되는 것을 알 수 있다.

5. 결 론

본 논문에서는 적은데이터나 censored 데이터도 비교적 유용하게 활용하여 기기의 신뢰도 계산을 할 수 있는 방법을 소개하였다. 이러한 방법은 기기의 테스트를 따로 시스템운전과 분리하여 수행하기 어렵거나 또는 여러 개의 시험표본을 준비하기 어려운경우, 또 기기들의 고장이 그다지 잦지 않아 충분한 고장데이터의 수집이 어려운 경우등에 적용하기 편리하며 발전소나 전력계통은 이러한 경우의 좋은 예가 될 수 있다고 생각한다. 또한 본 방법은 기기들 뿐만이 아니라 각종 측정계기류들의 신뢰도 계산에도 광범위하게 적용할 수 있다. 특히 suspended item들을 신뢰도 계산에 기여시킬 수 있기 때문에 실제로 운전을 해가면서도 끊임없이 기기들의 운전데이터를 수집하고 이들의 신뢰도를 update해 나갈 수 있다. 따라서 본 신뢰도 계산방법은 각종 고장진단 전문가시스템 개발시 그러한 전문가시스템들이 보다 상세하고 고장부위에 실질적인 도움을 줄 수 있는 진단결과를 제시할 수 있도록 지식 베이스를 보강하는데 이용할 수 있을 것이다.

Acknowledgement

본 해설의 그림 게재에 협조해 준 Dr. Abernethy께 감사사를 드린다.

참 고 문 헌

- [1] R. B. Abernethy, "Weibull Analysis Handbook," AFWAL-TR-83-2079, Pratt & Whitney Aircraft, Nov., 1983
- [2] J. J. Deyst Jr., R. M. Kanazawa, and J.P. Pasquenza, "Sensor validation: A method to enhance the quality of the man/machine interface in nuclear power station," IEEE Trans. on Nuclear Science, Vol. 28, No. 1, pp. 886-890, Feb. 1981.
- [3] W. W. Hines and D. C. Montgomery, "Probability and Statistics in Engineering and Management Science," 2nd Ed., John Wiley & Sons, 1980.
- [4] L. G. Johnson, "The Statistical Treatment of Fatigue Experiments," 1959.

- [5] K. C. Kapur and L. R. Lamberson, "Reliability in Engineering Design," John Wiley & Sons, 1977.
- [6] S. C. Lee, "Reliability Analysis for High Pressure Oxygen Turbo Pump Blade Downstream Shank Cracks for Space Shuttle Main Engine," UTSI Technical Report 87-14, NASA Contract No. NAS8-37104, 1987.
- [7] S. C. Lee, "Sensor Value Validation based on systematic exploration of the Sensor Redundancy for Fault Diagnosis Knowledge-based Systems," IEEE Transactions on Systems, Man and Cybernetics, Vol. 24, No. 4, pp. 594-605, April, 1994.
- [8] W. Nelson, "Applied Life Data Analysis," John Wiley and Sons, New York, 1982.
- [9] E. A. Scarl, J. R. Jamieson, and C. I. deLaune, "Diagnosis and Sensor Value Validation through Knowledge of Structure and Function," IEEE Transactions

on Systems, Man and Cybernetics, Vol. 17, No. 3, pp. 360-369, May/June 1987.

저 자 소 개



이승철(李承哲)

1946년 4월 16일생. 1969년 서울공대 전기공학과 졸업. 1974년 재무부 전매청 신탄진연 초제조창 전력과장. 1980년 현대엔지니어링 기전사업부 차장. 1985년 미국 University of Florida 전기공학과 졸업(공학박사). 1992년 미국 University of Tennessee, Space Institute 전산과 조교수. 1995년 상기 대 전기공학과 연구부교수. 현재 중앙대학교 전기공학과 부교수.